Contents

Chapter 1

Introduction

1.1 Background Information

The well-known theorem of Schur (see for example [38]) states that if A is a complex $m \times m$ matrix, then there exists a unitary $m \times m$ matrix U, such that $U^{-1}AU$ is an upper triangular matrix. In other words, each square complex matrix can be reduced to upper triangular form by a unitary similarity transformation. For *pairs* of square complex matrices, the following result was obtained by McCoy [42].

Theorem 1.1.1 *Let A, Z be a pair of complex $m \times m$ matrices. Then the following two statements are equivalent:*

1. *There exists an invertible $m \times m$ matrix S, such that both $S^{-1}AS$ and $S^{-1}ZS$ are upper triangular matrices.*
2. *For each polynomial $p(\lambda, \mu)$ in the non-commuting variables λ and μ, the $m \times m$ matrix $p(A, Z)(AZ - ZA)$ is nilpotent.*

A pair of $m \times m$ matrices A, Z which satisfies the statements of Theorem 1.1.1 is said to admit simultaneous reduction to upper triangular form. The proof of Theorem 1.1.1 in [42] is rather involved. Elementary proofs of this theorem have been obtained in [23] and [28]. The theorem is made more explicit for certain pairs of matrices in [34] and [35]. Further, a recent extension of Theorem 1.1.1 is given in [45]. The literature on this subject, which includes a paper of Frobenius [27] of almost a century ago, is extensive. For more information and references, see [36]. Generalizations of Theorem 1.1.1 to an infinite dimensional context has been obtained in [37] and [43]; see also Section 1.3.

This thesis also deals with simultaneous reduction to triangular forms. However, the emphasis is not on simultaneous reduction to "the same", e.g. upper, triangular form, but on simultaneous reduction to "complementary" triangular forms. As we shall see later on, the motivation for studying this issue comes from systems theory. But let us first give the definition and state two of the earliest results.

Let A and Z be $m \times m$ matrices. We say that the pair A, Z admits simultaneous reduction to complementary triangular forms, if there exists an invertible matrix S (not necessarily unitary) such that $S^{-1}AS$ is an upper triangular matrix and $S^{-1}ZS$ is a lower triangular matrix.

Here are two early results about this notion. The first result was implicit in the proof of Theorem 1.6 in [6]; for an explicit statement, see [5]. Theorem 1.1.2 below also appears in Chapter 2 as Theorem 2.2.1, as is indicated between brackets. The second result appeared in [7].

Theorem 1.1.2 (Theorem 2.2.1) *Let A, Z be a pair of $m \times m$ matrices. If either A or Z is diagonalizable, then the pair A, Z admits simultaneous reduction to complementary triangular forms.*

Theorem 1.1.3 *Let A, Z be a pair of $m \times m$ matrices, such that* $\operatorname{rank}(A - Z) = 1$ *and such that $\sigma(A) \cap \sigma(Z) = \emptyset$. Then the pair A, Z admits simultaneous reduction to complementary triangular forms.*

If a pair of $m \times m$ matrices A, Z admits simultaneous reduction to upper triangular form, we may assume without loss of generality, that the invertible $m \times m$ matrix S involved can be chosen to be unitary. For complementary triangular forms, this is not the case: There exist pairs of $m \times m$ matrices A, Z that admit simultaneous reduction to complementary triangular forms, for which the invertible $m \times m$ matrix S involved can not be taken unitary. This already indicates that simultaneous reduction to upper triangular form and simultaneous reduction to complementary triangular forms are quite different matters. Nevertheless, a connection between the two notions is formulated in the following proposition, which is proved in [16].

Proposition 1.1.4 *Let A and Z be $m \times m$ matrices. The pair A, Z admits simultaneous reduction to complementary triangular forms if and only if there exists a positive definite $m \times m$ matrix H, such that the pair $A, H^{-1}Z^*H$ admits simultaneous reduction to upper triangular form.*

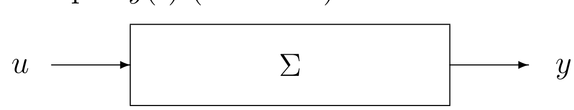
The collection of pairs of $m \times m$ matrices that admit simultaneous reduction to complementary triangular forms is denoted by $\mathcal{C}(m)$. At the end of Section 2.3, the collections of low order matrices $\mathcal{C}(2)$ and $\mathcal{C}(3)$ are described completely. For $m \geq 4$, a full description of $\mathcal{C}(m)$ is not known. One could hope for a reasonable general description by combining Proposition 1.1.4 and McCoy's theorem (Theorem 1.1.1). Unfortunately, the existence problem of the positive definite matrix H in Proposition 1.1.4 turns out to be as complicated as the study of simultaneous reduction to complementary triangular forms itself.

Altogether, there are no non-trivial results concerning the general case. On the other hand, quite some satisfactory results are obtained for pairs of matrices, taken from certain classes of matrices. Results in this direction have been obtained by several authors; see [7], [12],[13], [14], [15], and [26]. Some of these results are stated in Section 2.2. It should be mentioned that Theorem 2.2.2 from that section, which deals with pairs of first companion matrices, is connected to the Two Machine Flow Shop Problem from job scheduling theory; see the end of Section 2.4 for more details and references.

Now let us give a motivation for studying complementary triangular forms for pairs of matrices. We shall do this by making a connection with systems theory. Consider the linear dynamical system

$$(\Sigma) \begin{cases} x'(t) = Ax(t) + Bu(t) & \\ y(t) \;= Cx(t) + u(t) & t \geq 0 \\ x(0) \;= 0 & \end{cases} , \tag{1.1}$$

where $x(t)$ is an m-vector, $u(t)$ and $y(t)$ are n-vectors, and A, B and C are matrices of the appropriate sizes. The equations represent a system Σ with input $u(t)$ and output $y(t)$ (at time t) as illustrated below.



By taking the Laplace transform of (1.1), and by cancelling the Laplace transform $\hat{x}(\lambda)$ of $x(t)$, we obtain that the Laplace transforms $\hat{u}(\lambda)$ and $\hat{y}(\lambda)$ of, respectively, the input vector $u(t)$ and the output vector $y(t)$ are related as follows:

$$\hat{y}(\lambda) = W(\lambda)\hat{u}(\lambda).$$

Here $W(\lambda)$ is the so-called transfer function of Σ, given by

$$W(\lambda) = I_n + C(\lambda I_m - A)^{-1}B. \tag{1.2}$$

The $n \times n$ matrix function $W(\lambda)$ is rational, i.e., its entries are the quotient of two polynomials. Further, $W(\lambda)$ is analytic at infinity with value $W(\infty) = I_n$, the $n \times n$ identity matrix. In the following, we shall consider factorizations of such transfer functions which lead to cascade decompositions of the underlying systems. A more detailed survey of this material from [6] is presented in Section 2.1.

It can be shown that all rational $n \times n$ matrix functions $W(\lambda)$ analytic at infinity with $W(\infty) = I_n$ can be written in realization form (1.2) for certain matrices A, B and C. The smallest possible integer m for which a given rational matrix function $W(\lambda)$ admits a realization (1.2) is called the MacMillan degree of $W(\lambda)$ and is denoted by $\delta(W)$. One may interpret the integer $\delta(W)$ as a measure of complexity of the corresponding system Σ. In fact, it equals the number of poles of $W(\lambda)$ counted according to their pole multiplicities (cf. [6]).

To put factorizations of these rational matrix functions in perspective, we will first discuss the scalar case: The quotient $w(\lambda) = p(\lambda)/q(\lambda)$ of two polynomials of degree m is a scalar rational function. We will assume that the polynomials are monic, i.e., have leading coefficient equal to one, so that $w(\infty) = 1$. The scalar function $w(\lambda)$ has Macmillan degree $\delta(w) = m$ if and only if $p(\lambda)$ and $q(\lambda)$ have no roots in common. The Fundamental Theorem of Algebra, applied both to $p(\lambda)$ and $q(\lambda)$, then yields

$$w(\lambda) = \left(\frac{\lambda - \alpha_1}{\lambda - \beta_1}\right)\left(\frac{\lambda - \alpha_2}{\lambda - \beta_2}\right)\cdots\left(\frac{\lambda - \alpha_m}{\lambda - \beta_m}\right),$$

where $\alpha_1, \alpha_2, \ldots, \alpha_m$ are the roots of $p(\lambda)$, and $\beta_1, \beta_2, \ldots, \beta_m$ are the roots of $q(\lambda)$. Note that

$$w_k(\lambda) = \frac{\lambda - \alpha_k}{\lambda - \beta_k} = 1 + \frac{\beta_k - \alpha_k}{\lambda - \beta_k}, \qquad k = 1, \ldots, m$$

are scalar rational functions of MacMillan degree one. Therefore, each scalar rational function $w(\lambda)$ of Macmillan degree m, with $w(\infty) = 1$, is the product

$$w(\lambda) = w_1(\lambda) w_2(\lambda) \cdots w_m(\lambda)$$

of m scalar rational functions of MacMillan degree one.

We will now discuss this type of factorization for rational matrix functions. Rational matrix functions of MacMillan degree one are called elementary rational matrix functions. Such a function is of the form

$$W(\lambda) = I_n + \frac{1}{\lambda - \alpha} c b^T, \tag{1.3}$$

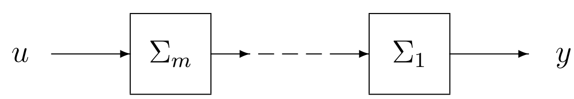
where α is a scalar and b and c are n-vectors. A complete factorization of a minimal realization (1.2) is a factorization

$$W(\lambda) = W_1(\lambda)W_2(\lambda)\cdots W_m(\lambda) \tag{1.4}$$

into $m = \delta(W)$ elementary rational matrix functions

$$W_k(\lambda) = I_n + \frac{1}{\lambda - \alpha_k} c_k b_k^T, \qquad k = 1, \ldots, m.$$

Each elementary rational matrix function $W_k(\lambda)$ corresponds to an "elementary" system Σ_k. In this manner, the complete factorization (1.4) corresponds to the cascade decomposition



of the system Σ into "elementary" systems $\Sigma_1, \ldots, \Sigma_m$. The question of which rational matrix functions admit a complete factorization is answered by the following theorem, which appeared in [5] and [7].

Theorem 1.1.5 (Theorem 2.1.2) *A rational matrix function $W(\lambda) = I_n + C(\lambda I_m - A)^{-1}B$ of MacMillan degree m admits a complete factorization if and only if the pair of $m \times m$ matrices $A, A - BC$ admits simultaneous reduction to complementary triangular forms.*

Not all rational matrix functions admit a complete factorization. For example, the rational 2×2 matrix function

$$W(\lambda) = \begin{pmatrix} 1 & \frac{-1}{\lambda^2} \\ 0 & 1 \end{pmatrix}$$

does not have this property; see also Example 2.4.3.

Up to so far, known material concerning complementary triangular forms and its connection with systems theory has been presented. In the next two sections, an outline is given of the new results in the thesis. Section 1.2 concerns the first part, and Section 1.3 concerns the second part of the thesis.

1.2 Finite Matrices

The first new result we shall discuss here is closely related to Theorem 1.1.5. As mentioned in the previous section, not all rational matrix functions admit a complete factorization. Recall that in the definition of a complete factorization, we required the elementary factors to be of a specific type, as in (1.3). If more general types of elementary factors are allowed (e.g. non-square ones), then the situation becomes different; see [52].

The factorization result we are about to present also involves elementary factors of the specific form (1.3), as in the case of complete factorization. However, we will allow the number of elementary factors to be larger than the MacMillan degree of the rational matrix function under consideration. In terms of [6], this means that the factorizations need not be minimal. The smallest number of factors that is needed to factorize a given rational matrix function $W(\lambda)$ into elementary factors is denoted by $\rho(W)$, and a factorization into $\rho(W)$ factors is called a quasicomplete factorization. We now state the factorization result.

Theorem 1.2.1 (Theorem 2.4.2) *Each non-trivial rational matrix function*

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

admits a quasicomplete factorization. In fact, the number of factors involved in such a factorization satisfies the estimates

$$\delta(W) \leq \rho(W) \leq 2\delta(W) - 1. \tag{1.5}$$

The first inequality in (1.5) is obvious, the second one requires a non-trivial proof that uses the Pole Assignment Theorem from systems theory. Another aspect of the proof is that it starts with a minimal realization $W(\lambda) = I_n + C(\lambda I_m - A)^{-1}B$, from which another realization $W(\lambda) = I_n + \tilde{C}(\lambda I_\rho - \tilde{A})^{-1}\tilde{B}$ is constructed. The matrices $\tilde{A}, \tilde{B}$ and $\tilde{C}$ are particular extensions of the matrices A, B and C, constructed in such a way that the pair $\tilde{A}$, $\tilde{A} - \tilde{B}\tilde{C}$ admits simultaneous reduction to complementary triangular forms.

A special type of extensions of those mentioned in the last paragraph are extensions with zeroes. This type of extensions leads to the following problem, which, by the way, also comes up in the study of complementary triangular forms in an infinite dimensional context (see Chapter 5).

Let A_1 and Z_1 be $m_1 \times m_1$ matrices. Does there exist a nonnegative integer m_2, such that the pair of $(m_1 + m_2) \times (m_1 + m_2)$ matrices $A_1 \oplus O_{m_2}$, $Z_1 \oplus O_{m_2}$ admits simultaneous reduction to complementary triangular forms? In other words, does the pair A_1, Z_1 admit simultaneous reduction to complementary triangular forms after extension with zeroes?

Note that we include the case when the pair A_1, Z_1 admits simultaneous reduction to complementary triangular forms without extension, i.e., the case when the integer m_2 can be taken zero.

In Section 3.2, an example is given of a pair of 4×4 matrices A_1, Z_1 which does not admit simultaneous reduction to complementary triangular forms, but obtains this property after extension with one zero. In short, $(A_1, Z_1) \notin \mathcal{C}(4)$, while $(A_1 \oplus 0, Z_1 \oplus 0) \in \mathcal{C}(5)$.

In Section 3.3 up to and including Section 3.8, it is shown that for all pairs of matrices, that are taken from classes of matrices for which a transparent description of simultaneous reduction to complementary triangular forms is known, e.g. for pairs of first companion matrices, the situation is different from the example in Section 3.2. For these pairs A_1, Z_1, the following will be shown: If m_2 is a nonnegative integer such that $(A_1 \oplus O_{m_2}, Z_1 \oplus O_{m_2}) \in \mathcal{C}(m_1+m_2)$, then also $(A_1, Z_1) \in \mathcal{C}(m_1)$. We may conclude that for such a pair, extending with zeroes does not produce the property of simultaneous reduction to complementary triangular forms, unless the pair had this property to begin with.

As was mentioned before, the problem of complementary triangular forms after extensions with zeroes comes up in the infinite dimensional setting. The following result is obtained there as a by-product.

Proposition 1.2.2 (Corollary 5.3.3) *If the pair of $m_1 \times m_1$ matrices A_1, Z_1 admits simultaneous reduction to complementary triangular forms after extension with zeroes, then $(A_1 \oplus O_{m_2}, Z_1 \oplus O_{m_2}) \in \mathcal{C}(m_1 + m_2)$ for a nonnegative integer m_2 that satisfies the estimate*

$$m_2 \leq 8m_1^2 - 3m_1.$$

Even in the general case, the estimate in Proposition 1.2.2 is probably not sharp. Before we turn to the infinite dimensional setting, one more point about quasicomplete factorization has to be made. In addition to an earlier remark with respect to the Two Machine Flow Shop Problem, we mention that there are strong indications that the concept of quasicomplete factorization is related to certain aspects of this topic from job scheduling theory.

1.3 Bounded Operators

In the second part of the dissertation, the notion of complementary triangular forms for pairs of bounded operators on an infinite dimensional Banach space is considered. A precise description of what is meant by a bounded linear operator in upper triangular form is given in terms of so-called maximal nests of invariant subspaces. A nest of subspaces is a collection of subspaces which is linearly ordered by inclusion. Maximal nests of subspaces are nests that are not properly contained in any other nest. Further, a subspace M is invariant for the linear operator A, if $x \in M$ implies $Ax \in M$, so $AM \subseteq M$.

A linear operator A acting on a Banach space X is called upper triangular with respect to a maximal nest of subspaces $\mathcal{M}$, if $\mathcal{M}$ consists of subspaces that are invariant for A. To illustrate the definition, we give two examples. The first example is very simple: Let

$$U = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

denote the backward shift on $l_2(\mathbf{Z}^+)$. If $\{e_k \mid k \in \mathbf{Z}^+\}$ denotes the standard basis in $l_2(\mathbf{Z}^+)$, then U is upper triangular with respect to the maximal nest of subspaces

$$\mathcal{M} = \{\text{span}\{e_1, \dots, e_k\} \mid k \in \mathbf{Z}^+\} \cup \{(0), l_2(\mathbf{Z}^+)\}.$$

The second example, which is quite different from the finite dimensional setting, is the Volterra operator of integration V, acting on $L_2(0,1)$. It is given by

$$Vf(x) = \int_0^x f(t)dt,$$

and it is upper triangular with respect to the continuous maximal nest of subspaces

$$\mathcal{M} = \{L_2(\tau, 1) \mid 0 \leq \tau \leq 1\}.$$

Indeed, the subspaces $(0 \leq \tau \leq 1)$

$$L_2(\tau, 1) = \{f \mid f \in L_2(0,1), f = 0 \text{ a.e. on } (0, \tau)\}$$

are invariant for V. In fact, this is the only maximal invariant nest of subspaces for V, hence V is unicellular; see for example [30] or [48].

It is well-known, that each compact operator on a Banach space has a maximal nest of invariant subspaces; see [1]. This result can be seen as an infinite dimensional analogue to Schur's theorem. Recall that an operator is compact, if it maps the unit ball into a compact set. Further, a bounded operator A is quasinilpotent, if $\sigma(A) = \{0\}$. The analogue to McCoy's theorem for pairs of compact operators on an infinite dimensional Banach space is provided by the following result from [37].

Theorem 1.3.1 *Let A and Z be compact operators acting on a Banach space X. Then A and Z have a common maximal nest of invariant subspaces if and only if for each polynomial $p(\lambda, \mu)$ in the non-commuting variables λ and μ, the compact operator $p(A, Z)(AZ - ZA)$ is quasinilpotent.*

Before we define complementary triangular forms for a pair of bounded operators on a Banach space, we return to the finite matrix case (see Section 3.1). A pair of $m \times m$ matrices A, Z admits simultaneous reduction to complementary triangular forms if and only if there exists a collection of projections $\mathcal{P} = \{P_k \mid 0 \leq k \leq m\}$, such that the collections $\{\text{Ran } P_k \mid 0 \leq k \leq m\}$ and $\{\text{Ker } P_k \mid 0 \leq k \leq m\}$ are maximal nests of invariant subspaces for A and Z respectively. With the natural ordering on projections, the collection of projections $\mathcal{P}$ is a maximal nest of projections.

In order to extend the notion of complementary triangular forms to pairs of bounded operators on an infinite dimensional Banach space X, we need to consider nests of projections on X. In Chapter 4, it is explained that the most straightforward definition that comes to mind, namely that of maximal nests of projections, does not really work. One needs the somewhat more restricted notion of a simple nest of projections, which is introduced in Section 4.1.

With this notion available to us, we introduce complementary triangular forms for pairs of bounded operators on a Banach space. It is convenient to give such a definition by describing the relevant collection of pairs of bounded operators.

The collection $\mathcal{C}(X)$ consists of pairs of bounded operators A, Z acting on X with the following property: There exists a simple nest of projections $\mathcal{P}$ on X, such that $AP = PAP$, and $PZ = PZP$ for all $P \in \mathcal{P}$.

Theorem 4.1.3 in Section 4.1 states that a nest of projections is simple if and only if the collections of subspaces $\{\text{Ran } P \mid P \in \mathcal{P}\}$ and $\{\text{Ker } P \mid P \in \mathcal{P}\}$ are maximal nests of subspaces. It follows that if $\mathcal{P}$ is a simple nest of projections, such that $AP = PAP$ and $PZ = PZP$ for all $P \in \mathcal{P}$, then $\{\text{Ran } P \mid P \in \mathcal{P}\}$ and $\{\text{Ker } P \mid P \in \mathcal{P}\}$ are maximal nests of invariant subspaces for A and Z respectively.

For examples of simple nests of projections, we refer to the end of Section 4.1. The aim of the subsequent sections in Chapter 4 is to put the definition of a simple nest of projections in a broader perspective. Section 4.2 considers nests of projections from a topological point of view. Section 4.3 studies nests of projections on reflexive Banach spaces. In this particular setting, converses are obtained to results from the preceding two sections.

In Chapter 5, complementary triangular forms for pairs of finite rank operators is considered. The following lemma, established in section 5.1, is useful to study such pairs.

Lemma 1.3.2 (Lemma 5.1.1) *Let A and Z be finite rank operators on a Banach space X. Then there exist subspaces $M, N \subseteq X$, such that $M \oplus N = X$, $\dim M = m < \infty$, and*

$$A = \begin{pmatrix} A_M & O \\ O & O_N \end{pmatrix}, \quad Z = \begin{pmatrix} Z_M & O \\ O & O_N \end{pmatrix},$$

where A_M and Z_M are the restrictions of A and Z to M, and where O_N denotes the zero operator on N.

Since A_M and Z_M are acting on the m-dimensional subspace M, they can be identified with $m \times m$ matrices by fixing a basis in M. We introduce the following terminology: The triple $(M, N, \{A_M, Z_M\})$ with the properties as in Lemma 1.3.2 is called a matrix reduction for the pair of finite rank operators A, Z. The following definition of complementary triangular forms uses the notion of a matrix reduction. Again, we give the definition by describing the relevant collection of pairs of finite rank operators.

The collection $\mathcal{C}_f(X)$ consists of those pairs of finite rank operators A, Z, such that there exists a matrix reduction $(M, N, \{A_M, Z_M\})$ for the pair, with $(A_M, Z_M) \in \mathcal{C}(m)$, where $m = \dim M$.

Theorem 5.3.1 in Section 5.3 shows that if the pair of finite rank operators A, Z satisfies $(A, Z) \in \mathcal{C}(X)$, then $(A, Z) \in \mathcal{C}_f(X)$. The other inclusion $\mathcal{C}_f(X) \subseteq \mathcal{C}(X)$ holds at least in the cases when X is a Hilbert space or when X is a Banach space with a Schauder basis; see Section 5.3.

Let a pair of finite rank operators A, Z be given, together with a matrix reduction $(M, N, \{A_M, Z_M\})$ for the pair. In principle, it is possible that $(A, Z) \in \mathcal{C}_f(X)$, but that on the other hand, $(A_M, Z_M) \notin \mathcal{C}(m)$, where $m = \dim M$. This is due to the fact that a matrix reduction for a pair of finite rank operators is not unique. It turns out that $(A, Z) \in \mathcal{C}_f(X)$ if and only if the pair of $m \times m$ matrices A_M, Z_M admits simultaneous reduction to complementary triangular forms after extension with zeroes; see Proposition 5.3.2.

One of the most basic results on complementary triangular forms for pairs of finite matrices is Theorem 1.1.2. It also underlies the proof of the fact that each rational matrix function admits a quasicomplete factorization. One might hope for an extension of this result to the infinite dimensional context. It has been an element of surprise (and to a certain extent of disappointment) that such a generalization does not hold true. Not even when one assumes both operators under consideration to be diagonalizable and compact.

Indeed, in Section 6.3, pairs of diagonalizable compact operators A, Z on $l_2(\mathbf{Z}^+)$ are constructed, for which there exists no invertible operator S, such that $S^{-1}AS$ is upper triangular, and $S^{-1}ZS$ is lower triangular with respect to the standard basis in $l_2(\mathbf{Z}^+)$. The construction of these pairs of diagonalizable compact operators uses a unitary infinite matrix, that does not admit lower-upper factorization, even after independently permuting rows and columns. Also, a pair of compact operators A, Z is presented, with Z diagonalizable, such that there exists no bounded nest of projections $\mathcal{P}$, with $AP = PAP$ and $PZ = PZP$ for all $P \in \mathcal{P}$. Finally, positive results are obtained for pairs of bounded operators, where one of the operators is of finite rank.

1.4 Notational Remarks

We will now clarify some notational conventions, used throughout this dissertation. The set of all integers is indicated by $\mathbf{Z}$. The set of all strictly positive (negative) integers is denoted by $\mathbf{Z}^+$ ($\mathbf{Z}^-$). If zero is included, write $\mathbf{Z}_0^+$ ($\mathbf{Z}_0^-$). The set of all complex numbers is denoted by $\mathbf{C}$. The symbol $\subseteq$ denotes inclusion, where equality may hold, the symbol $\subset$ denotes inclusion, where equality does not hold, i.e., proper inclusion.

If X is a vector space, then the extended integer $\dim X$ denotes its dimension. All vector spaces in this thesis are complex. By an operator between two vector spaces we will always mean a linear mapping. If T is an operator acting on X, then $\text{Ran } T = \{Tx \mid x \in X\}$, $\text{Ker } T = \{x \in X \mid Tx = 0\}$, and $\text{rank } T = \dim \text{Ran } T$.

If X is a Banach space, then a linear submanifold $Y \subseteq X$ is called a *subspace*, if it is closed in the norm topology. If X_1 and X_2 are subspaces in a Banach space X, then $X_1 + X_2 = \{x_1 + x_2 \mid x_1 \in X_1, x_2 \in X_2\}$ denotes the smallest linear submanifold in X, that contains both X_1 and X_2. In general, $X_1 + X_2$ need not be closed. In the case when $X_1 \cap X_2 = (0)$, we write $X_1 + X_2 = X_1 \oplus X_2$. Further, if $Y_1, Y_2 \subseteq X$ are subspaces, such that $Y_1 \oplus Y_2 = X$, then this direct sum is called a *decomposition of* X. If $Y \subseteq X$ is a subspace, the quotient Banach space is written as $X/Y = \{x + Y \mid x \in X\}$.

The *conjugate space* of a Banach space X, the Banach space of all bounded linear functionals on X, is denoted by X^*.

If the linear operator $T : X \longrightarrow X$ acts from X into X, and if $X = X_1 \oplus X_2$, then write

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

where $T_{kl} : X_k \longrightarrow X_l$ denotes the compression of T to the relevant subspaces. If $T_{12} = T_{21} = O$, then write $T = T_{11} \oplus T_{22}$. If $M \subseteq X$ is a subspace, then A_M denotes the compression of A to the subspace M. The identity operator and zero operator on the vector space X are denoted by I_X and O_X respectively. For a bounded operator T on X, let $\rho(T)$ and $\sigma(T)$ denote, respectively, the resolvent set and the spectrum of T.

Let $\mathcal{I}$ denote an arbitrary index set, and let X_i denote sets for $i \in \mathcal{I}$. Then write

$$\bigcap\{X_i \mid i \in \mathcal{I}\} = \{x \mid x \in X_i \text{ for all } i \in \mathcal{I}\},$$

$$\bigcup\{X_i \mid i \in \mathcal{I}\} = \{x \mid x \in X_i \text{ for some } i \in \mathcal{I}\}.$$

If $X_i \subseteq X$ are subsets in a Banach space, then $\operatorname{span}\{X_i \mid i \in \mathcal{I}\}$ denotes the closed linear hull of all X_i, i.e., the smallest subspace in X, that contains all X_i.

The index and glossary indicate on which page the specific notion or symbol is defined or explained. If a notion is defined on a certain page, it is printed in *emphasized* font. Numbering of theorems, propositions, etc. is related to the section in which they occur. For example, Theorem 2.2.1 can be found in Section 2.2.

Part I

Finite Matrices

Chapter 2

Quasicomplete Factorizations

The main topic of this chapter is factorization of rational matrix functions into elementary factors. The factorizations under consideration need not be minimal. Section 2.1 reviews minimal factorizations of rational matrix functions, and its connection with simultaneous reduction to complementary triangular forms. The latter notion is the main subject of Section 2.2, which gives an overview of known results, and of Section 2.3, where new results in this direction are presented. Finally, in Section 2.4, a new factorization theorem is proved.

2.1 Rational Matrix Functions

In this section, known material from systems theory is reviewed, from which the main subject of this dissertation originates. For more background information, the reader is referred to [6] and [7], and the references given there.

An $n \times n$ *rational matrix function* $W(\lambda) = (\ w_{ij}(\lambda)\ )_{i,j=1}^{n}$ is an $n \times n$ matrix with rational functions $w_{ij}(\lambda)$ as its entries. In this chapter, all rational $n \times n$ matrix functions $W(\lambda)$ are assumed to be analytic at ∞, with value $W(\infty) = I_n$, the $n \times n$ identity matrix. From systems theory it is known (see for example Theorem 2.2 in [6]), that such a matrix function $W(\lambda)$ can be written in the form

$$W(\lambda) = I_n + C(\lambda I_m - A)^{-1}B, \qquad \lambda \in \rho(A), \tag{2.1}$$

where the matrices A, B and C are of the appropriate sizes. An expression of the form (2.1) is called a *realization* of $W(\lambda)$. The smallest possible integer m for which a given rational matrix function $W(\lambda)$ admits a realization (2.1) is called the *MacMillan degree* of $W(\lambda)$, and is denoted by $\delta(W)$. If $m = \delta(W)$, then (2.1) is called a *minimal realization*. An equivalent requirement for (2.1) to be a minimal realization is that the pair of matrices A, B is *controllable*, i.e.,

$$\text{Ran } B + \text{Ran}(AB) + \cdots + \text{Ran}(A^{m-1}B) = \mathbf{C}^m,$$

and that the pair of matrices C, A is *observable* , i.e.,

$$\text{Ker } C \cap \text{Ker}(CA) \cap \cdots \cap \text{Ker}(CA^{m-1}) = (0).$$

The trivial rational matrix function $W(\lambda) = I_n$ satisfies $\delta(W) = 0$.

By the state space isomorphism theorem ([6], Theorem 3.1), to be stated below, all minimal realizations for a given rational matrix function are mutually similar.

Theorem 2.1.1 *Let $W(\lambda)$ be a rational matrix function, and let*

$$W(\lambda) = I_n + C(\lambda I_m - A)^{-1}B = I_n + K(\lambda I_m - F)^{-1}G$$

be two minimal realizations. Then there exists an invertible $m \times m$ matrix S, such that $CS = K$, $S^{-1}AS = F$, and $S^{-1}B = G$.

If (2.1) is a (not necessarily minimal) realization for $W(\lambda)$, then the inverse $W^{-1}(\lambda)$ is given by the realization

$$W^{-1}(\lambda) = I_n - C(\lambda I_m - A^{\times})^{-1}B, \quad \lambda \in \rho(A^{\times}),$$

where $A^{\times} = A - BC$. Note that $\delta(W^{-1}) = \delta(W)$.

If $W_1(\lambda)$, $W_2(\lambda)$ and $W(\lambda)$ are rational matrix functions, then $W(\lambda) = W_1(\lambda)W_2(\lambda)$ denotes a *factorization* of $W(\lambda)$. If the minimal realizations of the factors are given by

$$W_j(\lambda) = I_n + C_j(\lambda I_{m_j} - A_j)^{-1}B_j, \quad j = 1, 2,$$

then the realization

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

$$I_n + \begin{pmatrix} C_1 & C_2 \end{pmatrix} \left[\lambda I_{m_1+m_2} - \begin{pmatrix} A_1 & B_1C_2 \\ O & A_2 \end{pmatrix}\right]^{-1} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \tag{2.2}$$

is called a *product realization* of $W(\lambda)$. The MacMillan degree satisfies the sublogarithmic property

$$\delta(W) \leq \delta(W_1) + \delta(W_2).$$

In the case when $\delta(W) = \delta(W_1) + \delta(W_2)$, i.e., when the product realization (2.2) is a minimal realization of $W(\lambda)$, the factorization $W(\lambda) = W_1(\lambda)W_2(\lambda)$ is called a *minimal factorization.*

A rational matrix function of MacMillan degree one is called an *elementary rational matrix function.* A minimal realization of an elementary rational matrix function $W(\lambda)$ is of the form

$$W(\lambda) = I_n + \frac{1}{\lambda - \alpha} cb^T,$$

where α is a complex number, and cb^T is an $n \times n$ matrix of rank one (b, c are n-vectors here). The inverse of $W(\lambda)$ is given by

$$W(\lambda)^{-1} = I_n - \frac{1}{\lambda - \alpha^\times} cb^T,$$

where $\alpha^\times = \alpha - b^T c$.

Let $W(\lambda)$ be an $n \times n$ rational matrix function with minimal realization as in (2.1), so $\delta(W) = m$. If $W(\lambda)$ admits a factorization

$$W(\lambda) = W_1(\lambda) \cdots W_m(\lambda),$$

where $W_1(\lambda), \ldots, W_m(\lambda)$ are elementary rational matrix functions, we say that $W(\lambda)$ admits a *complete factorization.* Note that a complete factorization is minimal. A necessary and sufficient condition for complete factorization of a rational matrix function in terms of the realization matrices is given by the following theorem (Theorem 6.1 in [7]). We will recapitulate its instructive proof.

Theorem 2.1.2 *Let $W(\lambda)$ be a rational $n \times n$ matrix function with minimal realization (2.1). Then $W(\lambda)$ admits a complete factorization if and only if there exists an invertible $m \times m$ matrix S, such that $S^{-1}AS$ is an upper triangular matrix, and $S^{-1}A^\times S$ is a lower triangular matrix.*

Proof The only if part is proved as follows. Assume that the minimal realization $W(\lambda) = I_n + C(\lambda I_m - A)^{-1}B$ admits a complete factorization

$W(\lambda) = W_1(\lambda) \cdots W_m(\lambda)$, and let the factors be given by the minimal realizations

$$W_j(\lambda) = I_n + \frac{1}{\lambda - \alpha_j} c_j b_j^T, \quad j = 1, \ldots, m.$$

Then the product realization of $W(\lambda)$ is given by

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

where

$$\tilde{A} = \begin{pmatrix} \alpha_1 & b_1^T c_2 & \cdots & b_1^T c_m \\ 0 & \alpha_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & b_{m-1}^T c_m \\ 0 & \cdots & 0 & \alpha_m \end{pmatrix},$$

$$\tilde{B} = \begin{pmatrix} b_1^T \\ b_2^T \\ \vdots \\ b_m^T \end{pmatrix}, \quad \tilde{C} = \begin{pmatrix} c_1 & c_2 & \cdots & c_m \end{pmatrix}. \tag{2.3}$$

Computation yields

$$\tilde{A}^\times = \tilde{A} - \tilde{B}\tilde{C} = \begin{pmatrix} \alpha_1^\times & 0 & \cdots & 0 \\ b_2^T c_1 & \alpha_2^\times & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ b_m^T c_1 & \cdots & b_m^T c_{m-1} & \alpha_m^\times \end{pmatrix},$$

with $\alpha_j^\times = \alpha_j - b_j^T c_j$ for $1 \leq j \leq m$. This product realization is again minimal, so the state space isomorphism theorem (Theorem 2.1.1) provides that there exists an invertible $m \times m$ matrix S, such that

$$S^{-1}AS = \tilde{A}, \quad S^{-1}A^\times S = \tilde{A}^\times,$$

and in particular, that $S^{-1}AS$ is upper triangular, and $S^{-1}A^\times S$ is lower triangular.

To prove the if part, let S be an invertible $m \times m$ matrix, such that $S^{-1}AS$ is upper triangular, and $S^{-1}A^\times S$ is lower triangular. It is not difficult to see that $\tilde{A} = S^{-1}AS$, $\tilde{B} = S^{-1}B$ and $\tilde{C} = CS$ can be written in the form (2.3). The factorization $W(\lambda) = W_1(\lambda) \cdots W_m(\lambda)$ can now easily be derived from the product realization $W(\lambda) = I_n + \tilde{C}(\lambda I_m - \tilde{A})^{-1}\tilde{B}$. □

2.2 Complementary Triangular Forms

Theorem 2.1.2 in the previous section motivates the study of the following property:

A pair of $m \times m$ matrices A, Z admits *simultaneous reduction to complementary triangular forms*, if there exists an invertible $m \times m$ matrix S, such that $S^{-1}AS$ is an upper triangular matrix and $S^{-1}ZS$ is a lower triangular matrix.

Clearly, not all pairs of $m \times m$ matrices A, Z have this property; for example, consider the case when $A = Z$ is non-diagonalizable. The collection of all pairs of $m \times m$ matrices that admit simultaneous reduction to complementary forms, which we will denote by $\mathcal{C}(m)$, has been studied by various authors (cf. [5], [7], [12], [13], [15] and [26]).

In this section, a concise exposition of known results on simultaneous reduction to complementary triangular forms is given. Theorem 2.2.1 below first appeared in [6], Theorem 3.4, in terms of complete factorization of rational matrix functions. The result as stated below appeared in [5], Theorem 3.2. The proof of this theorem, as given in [7], Theorem 1.2, will be postponed until Section 6.2, where analogues of and counterexamples to this theorem for pairs of bounded operators acting on an infinite dimensional space are discussed.

Theorem 2.2.1 *Let A and Z be $m \times m$ matrices. If either A or Z is diagonalizable, then the pair A, Z admits simultaneous reduction to complementary triangular forms.*

Most results on simultaneous reduction to complementary triangular forms are concerned with pairs of matrices that belong to certain classes of matrices. First, we will state two results for pairs of companion matrices. Recall that a *first companion matrix* is of the form

$$C_a = \begin{pmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & 1 & \\ & & & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{m-2} & -a_{m-1} \end{pmatrix}, \tag{2.4}$$

where $a_0, \ldots, a_{m-1}$ are complex numbers. Note that there is a one-to-one correspondence between monic polynomials of degree m and first companion matrices, given by the equation

$$\det(\lambda - C_a) = a_0 + a_1\lambda + \cdots + a_{m-1}\lambda^{m-1} + \lambda^m.$$

The generalized eigenvectors of a first companion matrix are determined by the corresponding eigenvalues as follows. If α is an eigenvalue for the $m \times m$ first companion matrix C_a with algebraic multiplicity n, then the vectors

$$x_1(\alpha) = \left(\ 1, \alpha, \cdots, \alpha^{m-1} \ \right)^T,$$

$$x_k(\alpha) = \frac{1}{(k-1)!}(\frac{d}{d\alpha})^{k-1} x_1(\alpha), \quad k = 2, \ldots, n,$$

satisfy $(A - \alpha)x_1(\alpha) = 0$ and $(A - \alpha)x_k(\alpha) = x_{k-1}(\alpha)$ for $k = 2, \ldots, n$. In other words, the vectors $x_1(\alpha), \ldots, x_n(\alpha)$ form a Jordan chain of C_a.

At this point, we also introduce some general terminology concerning matrices. A vector $\beta = (\beta_1, \ldots, \beta_m)^T$ is called a *spectral vector* for an $m \times m$ matrix B, if $\beta_1, \ldots, \beta_m$ are the eigenvalues of B, counted according to their algebraic multiplicities. If $T = (\ T_{ij}\ )_{i,j=1}^m$ denotes a complex $m \times m$ matrix, then $\operatorname{diag} T = (T_{11}, \ldots, T_{mm})^T$ denotes the *diagonal vector* of T. In particular, the diagonal vector of an upper or lower triangular matrix is a spectral vector for that matrix.

A pair of $m \times m$ matrices A, Z is said to admit *simultaneous reduction to complementary triangular forms with diagonals*

$$\alpha = (\alpha_1, \ldots, \alpha_m)^T, \quad \zeta = (\zeta_1, \ldots, \zeta_m)^T, \tag{2.5}$$

whenever there exists an invertible $m \times m$ matrix S, such that $S^{-1}AS$ is upper triangular, $S^{-1}ZS$ is lower triangular, and

$$\operatorname{diag}(S^{-1}AS) = \alpha, \quad \operatorname{diag}(S^{-1}ZS) = (\zeta_m, \ldots, \zeta_1)^T.$$

In this case, we will write $(A, Z) \in \mathcal{C}(\alpha, \zeta)$.

In Chapter 3, where a geometrical description of simultaneous reduction to complementary triangular forms is given, it will be justified that the spectral vector ζ appears in reversed order on the diagonal of $S^{-1}ZS$.

The *reversed identity* or *rotation matrix* R, defined by $Re_k = e_{m-k+1}$ for $k = 1, \ldots, m$, transforms upper triangular matrices into lower triangular matrices and vice versa, i.e., T is an upper triangular $m \times m$ matrix if and only if $R^{-1}TR$ is a lower triangular $m \times m$ matrix. Also, $R^{-1} = R$. Using this matrix, it is immediate that $(A, Z) \in \mathcal{C}(\alpha, \zeta)$ if and only if $(Z, A) \in \mathcal{C}(\zeta, \alpha)$.

The following theorem, that describes simultaneous reduction to complementary triangular forms for pairs of first companion matrices, is taken from [7], Theorem 3.2.

Theorem 2.2.2 *Let A and Z be first companion $m \times m$ matrices. Then the pair A, Z admits simultaneous reduction to complementary triangular forms with diagonal vectors α and ζ as in (2.5), if and only if these vectors are spectral vectors for A and Z respectively, and satisfy*

$$\alpha_k \neq \zeta_l, \quad k + l \leq m.$$

The reversed identity R transforms the first companion matrix C_a, defined in (2.4), to the *third companion matrix*

$$\hat{C}_a = R^{-1} C_a R = \begin{pmatrix} -a_{m-1} & -a_{m-2} & \cdots & -a_1 & -a_0 \\ 1 & 0 & & & \\ & 1 & \ddots & & \\ & & \ddots & 0 & \\ & & & 1 & 0 \end{pmatrix}.$$

We state Theorem 3.2 from [15], that deals with pairs of matrices consisting of a first companion matrix and a third companion matrix.

Theorem 2.2.3 *Let A be a first companion $m \times m$ matrix and Z be a third companion $m \times m$ matrix. Then the pair A, Z admits simultaneous reduction to complementary triangular forms with diagonal vectors α and ζ as in (2.5), if and only if these vectors are spectral vectors for A and Z respectively, and satisfy*

$$\alpha_k \zeta_l \neq 1, \quad k + l \leq m.$$

The following result, Theorem 4.1 from [12], deals with certain pairs of nilpotent matrices that are called *sharply upper triangular matrices.* The symbol O_n denotes the $n \times n$ zero matrix.

Theorem 2.2.4 *Let $1 \leq \alpha, \omega \leq m-1$, and let A_{12} be an invertible upper triangular $(m-\alpha)\times(m-\alpha)$ matrix and Z_{12} be an invertible upper triangular $(m-\omega)\times(m-\omega)$ matrix. Let the $m \times m$ matrices A and Z be given by*

$$A = \begin{pmatrix} O & A_{12} \\ O_\alpha & O \end{pmatrix}, \quad Z = \begin{pmatrix} O & Z_{12} \\ O_\omega & O \end{pmatrix}.$$

Then the pair A, Z admits simultaneous reduction to complementary triangular forms, if and only if $\alpha + \omega > m$, α does not divide ω, and ω does not divide α.

We will now discuss pairs of nonderogatory Jordan Matrices. Recall that a matrix B is called *nonderogatory*, if each eigenvalue $\beta \in \sigma(B)$ has geometric multiplicity $\dim \operatorname{Ker}(B-\beta) = 1$. If α is a complex number, then the $n \times n$ matrix

$$J(\alpha, n) = \begin{pmatrix} \alpha & 1 & 0 & \cdots & 0 \\ 0 & \alpha & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \alpha & 1 \\ 0 & \cdots & \cdots & 0 & \alpha \end{pmatrix}$$

denotes the upper triangular $n \times n$ Jordan block with eigenvalue α. A matrix of the form $J(\alpha_1, n_1) \oplus \cdots \oplus J(\alpha_s, n_s)$ is called a *Jordan matrix*. This Jordan matrix is nonderogatory if and only if the eigenvalues $\alpha_1, \ldots, \alpha_s$ are mutually distinct. Consider the nonderogatory Jordan $m \times m$ matrices

$$J_\alpha = J(\alpha_1, k_1) \oplus \cdots \oplus J(\alpha_s, k_s), \quad J_\zeta = J(\zeta_1, l_1) \oplus \cdots \oplus J(\zeta_t, l_t), \tag{2.6}$$

with $k_1, \ldots, k_s$ and $l_1, \ldots, l_t$ non-zero positive integers, such that $k_1 + \cdots + k_s = l_1 + \cdots + l_t = m$, with $\alpha_1, \ldots, \alpha_s$ the distinct eigenvalues for J_α, and $\zeta_1, \ldots, \zeta_t$ the distinct eigenvalues for J_ζ. For $1 \leq \rho \leq s$ and $1 \leq \sigma \leq t$, we say that the Jordan blocks $J(\alpha_\rho, k_\rho)$ and $J(\zeta_\sigma, l_\sigma)$ have a *diagonal overlap* on more than one position, if the set

$$\left\{ 1 + \sum_{i=1}^{\rho-1} k_i, \ldots, \sum_{i=1}^{\rho} k_i \right\} \cap \left\{ 1 + \sum_{j=1}^{\sigma-1} l_j, \ldots, \sum_{j=1}^{\sigma} l_j \right\}$$

contains more than one element. We now state Theorem 4.1 in [13].

Theorem 2.2.5 *Let J_α and J_ζ be nonderogatory Jordan matrices. Then the pair J_α, J_ζ admits simultaneous reduction to complementary triangular forms, if and only if J_α and J_ζ contain no Jordan blocks, that have an overlap on more than one diagonal position.*

2.3 Spectral Polynomials

In this section, new results on simultaneous reduction to complementary triangular forms are obtained, using the following type of polynomial: Let B be an $m \times m$ matrix and let the mutually distinct eigenvalues of B be denoted by $\beta_1, \ldots, \beta_s$. Define the *spectral polynomial* of B by

$$p_B(\lambda) = (\lambda - \beta_1) \cdots (\lambda - \beta_s). \tag{2.7}$$

This polynomial is the monic polynomial of minimal degree vanishing on the spectrum of B. Note that the matrix $p_B(B)$ is nilpotent and that $p_B(B) = O_m$ if and only if B is diagonalizable. In fact, the subspace Ker $p_B(B)$ is the linear span of all eigenvectors of B.

First, we turn to simultaneous reduction to complementary triangular forms for pairs of nilpotent matrices. The following theorem is a generalization of Lemmas 1.1 and 1.2 in [12]. Note that, in this theorem, $Z \neq O_m$ and (2.8) imply $A \neq O_m$.

Theorem 2.3.1 *Let A and Z be nilpotent $m \times m$ matrices, $Z \neq O_m$, and assume that*

$$\text{Ker } A \subseteq \text{Ker } Z + \text{Ran } Z \tag{2.8}$$

and

$$\text{Ker } A \cap \text{Ran } A \subseteq \text{Ran } Z. \tag{2.9}$$

Then the pair A, Z does not admit simultaneous reduction to complementary triangular forms.

Proof Assume that (2.8) and (2.9) hold and that the pair A, Z admits simultaneous reduction to complementary triangular forms, i.e., there exists a basis $s_1, \ldots, s_m$ in $\mathbf{C}^m$, such that $(1 \leq k \leq m)$

$$As_k \in \text{span}\{s_1, \ldots, s_{k-1}\}, \quad Zs_k \in \text{span}\{s_{k+1}, \ldots, s_m\}. \tag{2.10}$$

For $m = 2$, a contradiction is immediate. For $m \geq 3$, we obtain a contradiction by induction. Indeed, the following hypothesis will be proved for $n = 1, \dots, m-1$:

$$(\Pi_n) \begin{cases} \text{span}\{s_1, \dots, s_n\} \subseteq \text{Ker } A \\ \text{Ran } Z \subseteq \text{span}\{s_{n+2}, \dots, s_m\} \end{cases} .$$

Step 1 Note first that

$$(\Pi_1) \begin{cases} s_1 \in \text{Ker } A \\ \text{Ran } Z \subseteq \text{span}\{s_3, \dots, s_m\} \end{cases} .$$

Indeed, $As_1 = 0$ by (2.10), so it remains to prove Ran $Z \subseteq \text{span}\{s_3, \dots, s_m\}$. Since $s_1 \in$ Ker A, (2.8) implies that $s_1 = x + y$, $x \in$ Ker Z, $y \in$ Ran Z. But then by (2.10), Ran $Z \subseteq \text{span}\{s_2, \dots, s_m\}$, so $y \in \text{span}\{s_2, \dots, s_m\}$. Further, $0 = Zx = Zs_1 - Zy$ and $Zs_1 = Zy \in \text{span}\{s_3, \dots, s_m\}$ follows from (2.10). Again by (2.10), $Zs_k \in \text{span}\{s_3, \dots, s_m\}$ for $k = 2, \dots, m$ and Step 1 is proved.

Step 2 Let $n \in \{1, \dots, m-2\}$. Then (Π_n) implies (Π_{n+1}).

Assume that (Π_n) holds. It then follows $As_{n+1} \in$ Ran $A \cap$Ker A, since $As_{n+1} \in \text{span}\{s_1, \dots, s_n\} \subseteq$ Ker A. For that reason, and by (2.9), $As_{n+1} \in$ Ran Z. But Ran $Z \subseteq \text{span}\{s_{n+2}, \dots, s_m\}$, so $As_{n+1} = 0$. Therefore $\text{span}\{s_1, \dots, s_{n+1}\} \subseteq$ Ker A. We need to prove that Ran $Z \subseteq \text{span}\{s_{n+3}, \dots, s_m\}$. Fix $1 \leq j \leq n+1$. Then $s_j \in$ Ker A and by (2.8), $s_j = x_j + y_j$, $x_j \in$ Ker Z and $y_j \in$ Ran Z. We obtain $y_j \in \text{span}\{s_{n+2}, \dots, s_m\}$. Further, $0 = Zx_j = Zs_j - Zy_j$ and for that reason, $Zs_j = Zy_j \in \text{span}\{s_{n+3}, \dots, s_m\}$. It follows that Ran $Z \subseteq \text{span}\{s_{n+3}, \dots, s_m\}$ and (Π_{n+1}) is satisfied.

Using Step 1 and Step 2, the statement (Π_{m-1}) follows by induction. But (Π_{m-1}) implies that Ran $Z = (0)$ or that $Z = O_m$, a contradiction. The theorem is proved. □

If $p(\lambda)$ and $q(\lambda)$ are polynomials, and A and Z are $m \times m$ matrices, then it is not difficult to see that $(A, Z) \in \mathcal{C}(m)$ implies $(p(A), q(Z)) \in \mathcal{C}(m)$. This argument provides the following corollary to Theorem 2.3.1.

Corollary 2.3.2 *Let A and Z be non-diagonalizable $m \times m$ matrices. If*

$$\text{Ker } p_A(A) \subseteq \text{Ker } p_Z(Z) + \text{Ran } p_Z(Z)$$

and

$$\text{Ker } p_A(A) \cap \text{Ran } p_A(A) \subseteq \text{Ran } p_Z(Z),$$

then the pair A, Z does not admit simultaneous reduction to complementary triangular forms.

Corollary 2.3.2 gives a *necessary* condition for simultaneous reduction to complementary triangular forms. Proposition 2.3.3 below provides a *sufficient* condition.

Proposition 2.3.3 *Let A and Z be $m \times m$ matrices. If either*

$$\text{Ker } p_A(A) + \text{Ker } p_Z(Z) = \mathbf{C}^m \tag{2.11}$$

or

$$\text{Ran } p_A(A) \cap \text{Ran } p_Z(Z) = (0), \tag{2.12}$$

then the pair A, Z admits simultaneous reduction to complementary triangular forms.

The subspace Ker $p_A(A)$+Ker $p_Z(Z)$ in (2.11) is the linear span of the eigenvectors of A and the eigenvectors of Z. Since the eigenvectors of a diagonalizable matrix span the whole space, it follows that Theorem 2.2.1 is a special case of Proposition 2.3.3. Before proving Proposition 2.3.3 we first state a lemma, the proof of which is straightforward and left to the reader.

Lemma 2.3.4 *Let A_1 and Z_1 be $m_1 \times m_1$ matrices, A_2 and Z_2 be $m_2 \times m_2$ matrices, A_{12} an $m_1 \times m_2$ matrix, Z_{21} an $m_2 \times m_1$ matrix. Define $m = m_1 + m_2$ and consider the $m \times m$ matrices*

$$A = \begin{pmatrix} A_1 & A_{12} \\ O & A_2 \end{pmatrix}, \quad Z = \begin{pmatrix} Z_1 & O \\ Z_{21} & Z_2 \end{pmatrix}.$$

If both pairs A_1, Z_1 and A_2, Z_2 admit simultaneous reduction to complementary forms, then the pair A, Z has the same property.

Proof of Proposition 2.3.3 We first prove that (2.11) implies that $(A, Z) \in \mathcal{C}(m)$. Write dim Ker $p_A(A) = k$ and dim Ker $p_Z(Z) = l$. There exist vectors $\phi_1, \ldots, \phi_k$ and complex numbers $\alpha_1, \ldots, \alpha_k$ such that

$$\text{Ker } p_A(A) = \text{span}\{\phi_1, \ldots, \phi_k\}, \quad A\phi_i = \alpha_i \phi_i, \quad i = 1, \ldots, k$$

and there exist vectors $\psi_1, \ldots, \psi_l$ and complex numbers $\zeta_1, \ldots, \zeta_l$ such that

$$\text{Ker } p_Z(Z) = \text{span}\{\psi_1, \ldots, \psi_l\}, \quad Z\psi_j = \zeta_j \psi_j, \quad j = 1, \ldots, l.$$

Then (2.11) implies that $k+l \geq m$ and that the vectors $\phi_1, \ldots, \phi_k, \psi_1, \ldots, \psi_l$ span the whole space $\mathbf{C}^m$. From this collection of vectors, a basis in $\mathbf{C}^m$ can be selected: There exist an integer $0 \leq s \leq m$, a strictly increasing mapping $\pi : \{1, \ldots, s\} \longrightarrow \{1, \ldots, k\}$ and a strictly increasing mapping $\rho : \{1, \ldots, m-s\} \longrightarrow \{1, \ldots, l\}$, such that $\phi_{\pi(1)}, \ldots, \phi_{\pi(s)}, \psi_{\rho(1)}, \ldots, \psi_{\rho(m-s)}$ is a basis in $\mathbf{C}^m$. With respect to this basis, A and Z assume the forms

$$A = \begin{pmatrix} D_1 & A_{12} \\ O & A_{22} \end{pmatrix}, \quad Z = \begin{pmatrix} Z_{11} & O \\ Z_{21} & D_2 \end{pmatrix},$$

where D_1 is an $s \times s$ diagonal matrix with diagonal $(\alpha_{\pi(1)}, \ldots, \alpha_{\pi(s)})^T$ and D_2 is an $(m-s) \times (m-s)$ diagonal matrix with diagonal $(\zeta_{\rho(1)}, \ldots, \zeta_{\rho(m-s)})^T$. By Theorem 2.2.1 and Lemma 2.3.4 it now follows that the pair A, Z admits simultaneous reduction to complementary triangular forms.

Second, we prove that (2.12) implies $(A, Z) \in \mathcal{C}(m)$. First note that (2.12) implies

$$\text{Ker } p_{A^*}(A^*) + \text{Ker } p_{Z^*}(Z^*) = \mathbf{C}^m.$$

By the argument given in the first part of the proof, it follows that the pair A^*, Z^* admits simultaneous reduction to complementary triangular forms. It easily follows that the pair A, Z has the same property. The proposition is proved. □

Corollary 2.3.2 and Proposition 2.3.3 lead to necessary and sufficient conditions for simultaneous reduction to complementary triangular forms on a special class of matrices; the almost diagonalizable matrices. An $m \times m$ matrix B is called *almost diagonalizable*, if rank $p_B(B) = 1$. In other words, the Jordan canonical form of an almost diagonalizable square matrix contains one Jordan block of size two, and all other blocks are of size one. The following theorem specifies Corollary 2.3.2 and Proposition 2.3.3 for almost diagonalizable matrices.

Theorem 2.3.5 *Let A and Z be almost diagonalizable $m \times m$ matrices. Then the following are equivalent:*

1. *The pair A, Z admits simultaneous reduction to complementary triangular forms.*
2. Ker $p_A(A) \neq$ Ker $p_Z(Z)$ *or* Ran $p_A(A) \neq$ Ran $p_Z(Z)$.
3. *$p_A(A)$ is not a scalar multiple of $p_Z(Z)$.*

The equivalence of the second and the third statement is contained in [13], Theorem 1.4. A somewhat different characterization of simultaneous reduction to complementary triangular forms for pairs of almost diagonalizable matrices is presented in Theorem 9 in [14], which extends the main result in [26].

Remark 2.3.6 We now present a full description of simultaneous reduction to complementary triangular forms for pairs of matrices of order less or equal to three.

A full description of $\mathcal{C}(2)$ is given in [7]: The pair of 2×2 matrices A, Z admits simultaneous reduction to complementary triangular forms if and only if $\text{Ker } p_A(A) + \text{Ker } p_Z(Z) = \mathbf{C}^2$.

A full description of $\mathcal{C}(3)$ is already given in the final section of [14]. We now present a somewhat different description below.

Let A and Z be 3×3 matrices. We may assume without loss of generality, that $\text{rank } p_A(A) \leq \text{rank } p_Z(Z)$. Note that $\text{rank } p_Z(Z) \leq 2$, since $p_Z(Z)$ is a nilpotent 3×3 matrix. We distinguish four cases.

1. $\text{rank } p_A(A) = 0$,
2. $\text{rank } p_A(A) = \text{rank } p_Z(Z) = 1$,
3. $\text{rank } p_A(A) = 1$, $\text{rank } p_Z(Z) = 2$,
4. $\text{rank } p_A(A) = \text{rank } p_Z(Z) = 2$.

The results presented in this chapter can be used to tackle these four cases as follows:

Case 1 In this case, A is diagonalizable by assumption, so Theorem 2.2.1 states that the pair A, Z admits simultaneous reduction to complementary triangular forms.

Case 2 Both A and Z are almost diagonalizable, and Theorem 2.3.5 can be applied. Hence the following are equivalent:

1. The pair A, Z admits simultaneous reduction to complementary triangular forms.
2. $\text{Ker } p_A(A) + \text{Ker } p_Z(Z) = \mathbf{C}^3$ or $\text{Ran } p_A(A) \cap \text{Ran } p_Z(Z) = (0)$.

Case 3 We claim that also in this case, the following are equivalent:

1. The pair A, Z admits simultaneous reduction to complementary triangular forms.

2. $\text{Ker } p_A(A) + \text{Ker } p_Z(Z) = \mathbf{C}^3$ or $\text{Ran } p_A(A) \cap \text{Ran } p_Z(Z) = (0)$.

Proposition 2.3.3 provides that the second statement implies the first one. As for the converse, assume that the invertible 3×3 matrix $S = \begin{pmatrix} s_1, s_2, s_3 \end{pmatrix}$ transforms A and Z into complementary triangular forms. Write

$$S^{-1}p_A(A)S = \begin{pmatrix} 0 & x_{12} & x_{13} \\ 0 & 0 & x_{23} \\ 0 & 0 & 0 \end{pmatrix}, \quad S^{-1}p_Z(Z)S = \begin{pmatrix} 0 & 0 & 0 \\ y_{21} & 0 & 0 \\ y_{31} & y_{32} & 0 \end{pmatrix}.$$

If $s_2 \notin \text{Ker } p_A(A)$, then $x_{12} \neq 0$. Since $\text{rank } p_A(A) = 1$, it follows that $x_{23} = 0$. Consequently, $\text{Ran } p_A(A) = \text{span}\{s_1\}$. Since $\text{Ran } p_Z(Z) = \text{span } \{s_2, s_3\}$, we get

$$\text{Ran } p_A(A) \cap \text{Ran } p_Z(Z) = (0).$$

On the other hand, if $s_2 \in \text{Ker } p_A(A)$, then $\text{Ker } p_A(A) = \text{span}\{s_1, s_2\}$. Since $\text{Ker } p_Z(Z) = \text{span}\{s_3\}$, it follows that

$$\text{Ker } p_A(A) + \text{Ker } p_Z(Z) = \mathbf{C}^3.$$

Case 4 In this case, both A and Z are similar to a 3×3 upper triangular Jordan block. It is not difficult to see that in this case, the following holds (see also Proposition 3.4 in [13]):

1. The pair A, Z admits simultaneous reduction to complementary triangular forms.

2. $\text{Ker } p_A(A) \oplus \text{Ker } p_Z(Z)^2 = \mathbf{C}^3$ and $\text{Ker } p_A(A)^2 \oplus \text{Ker } p_Z(Z) = \mathbf{C}^3$.

Example 2.3.7 This example illustrates, that $\mathcal{C}(4)$ cannot be fully described in terms of the spectral polynomials given by (2.7). Consider the 4×4 matrices

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then A and Z are transformed into complementary triangular forms by means of the invertible 4×4 matrix

$$S = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}$$

as follows:

$$S^{-1}AS = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad S^{-1}ZS = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence, $(A, Z) \in \mathcal{C}(4)$. Further, $B^2(B - I) = Z$ is a non-diagonalizable 4×4 matrix. Therefore, $(B, Z) \notin \mathcal{C}(4)$. On the other hand,

$$p_A(A) = p_B(B) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

so any description in terms of $p_A(A), p_B(B)$ and $p_Z(Z)$ only makes no distinction between the pairs A, Z and B, Z.

2.4 Factorizations into Elementary Factors

Scalar rational matrix functions always admit a complete factorization; see Section 7 in [7] and Section 1.2 in the introduction of this thesis.

In general, however, rational matrix functions need not admit a complete factorization. Indeed, for each pair of $m \times m$ matrices A, Z, one can construct a minimal realization

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

such that $Z = A^\times = A - BC$ (for details, see Theorem 5.1 in [7]). In particular, rational matrix functions that do not admit complete factorization can be constructed from pairs of $m \times m$ matrices, that do not admit simultaneous reduction to complementary triangular forms.

For this reason, more relaxed notions of factorization into elementary factors have been considered. In [52], it is shown that each rational matrix function admits a minimal factorization into nonsquare elementary rational matrix functions.

In this section, it will be proved that any rational matrix function

$$W(\lambda) = I_n + C(\lambda I_m - A)^{-1}B \tag{2.13}$$

is the product of elementary rational matrix functions, i.e., is of the form

$$W(\lambda) = W_1(\lambda)\cdots W_N(\lambda), \tag{2.14}$$

where $W_1(\lambda), \ldots, W_N(\lambda)$ again are elementary rational matrix functions, but where the number of factors N may exceed the MacMillan degree of $W(\lambda)$.

Before doing so in Theorem 2.4.1 below, we need to introduce the following. Let $W(\lambda)$ be a rational matrix function, with minimal realization (2.13). Let $N \geq m$, and consider two nonnegative integers m_1, m_3, such that $m_1 + m + m_3 = N$. Define the matrices

$$\hat{A} = \begin{pmatrix} A_1 & A_{12} & A_{13} \\ O & A & A_{23} \\ O & O & A_{33} \end{pmatrix}, \quad \hat{B} = \begin{pmatrix} B_1 \\ B \\ O \end{pmatrix}, \quad \hat{C} = \begin{pmatrix} O & C & C_3 \end{pmatrix}, \tag{2.15}$$

where A_1 is an $m_1 \times m_1$, A_{12} an $m_1 \times m$, A_{13} an $m_1 \times m_3$, A_{23} an $m \times m_3$, and A_{33} an $m_3 \times m_3$ matrix. Further, B_1 is an $m_1 \times n$, and C_3 an $n \times m_3$ matrix. A realization

$$W(\lambda) = I_n + \hat{C}(\lambda I_N - \hat{A})\hat{B}, \tag{2.16}$$

with $\hat{A}$,$\hat{B}$ and $\hat{C}$ as in (2.15) is called an *extended realization* of the minimal realization (2.13). We will allow ourselves to write the matrices of such an extended realization as

$$\hat{A} = \begin{pmatrix} * & * & * \\ O & A & * \\ O & O & * \end{pmatrix}, \quad \hat{B} = \begin{pmatrix} * \\ B \\ O \end{pmatrix}, \quad \hat{C} = \begin{pmatrix} O & C & * \end{pmatrix},$$

i.e., without specification of m_1 and m_3, and the corresponding block matrix entries.

Theorem 2.4.1 *Let $W(\lambda)$ be a rational $n \times n$ matrix function with minimal realization as in (2.13). Then $W(\lambda)$ admits a factorization as in (2.14), with*

$N \geq m$, if and only if there exists an extended realization (2.16) of the minimal realization (2.13), with the matrices

$$\hat{A} = \begin{pmatrix} * & * & * \\ O & A & * \\ O & O & * \end{pmatrix}, \quad \hat{B} = \begin{pmatrix} * \\ B \\ O \end{pmatrix}, \quad \hat{C} = \begin{pmatrix} O & C & * \end{pmatrix},$$

such that the pair of $N \times N$ matrices $\hat{A}$, $\hat{A}^\times = \hat{A} - \hat{B}\hat{C}$ admits simultaneous reduction to complementary triangular forms.

Proof To prove the only if part, assume that $W(\lambda)$ admits a factorization as in (2.14). Denote the product realization by

$$W(\lambda) = I_n + \tilde{C}(\lambda I_N - \tilde{A})^{-1}\tilde{B}.$$

Then $\tilde{A}$ is upper triangular, and $\tilde{A} - \tilde{B}\tilde{C}$ is lower triangular: See the first part of the proof of Theorem 2.1.2. By Theorem 3.2 in [6], there exists an invertible $N \times N$ matrix T, such that

$$T^{-1}\tilde{A}T = \begin{pmatrix} * & * & * \\ O & F & * \\ O & O & * \end{pmatrix}, \quad T^{-1}\tilde{B} = \begin{pmatrix} * \\ G \\ O \end{pmatrix}, \quad \tilde{C}T = \begin{pmatrix} O & H & * \end{pmatrix},$$

with $W(\lambda) = I_n + H(\lambda I_m - F)^{-1}G$ a minimal realization. The state space isomorphism theorem (Theorem 2.1.1) provides an invertible $m \times m$ matrix V, such that $V^{-1}FV = A$, $V^{-1}G = B$ and $HV = C$. Define the invertible $N \times N$ matrix (same block structure as $T^{-1}\tilde{A}T$) by

$$W = \begin{pmatrix} I & O & O \\ O & V & O \\ O & O & I \end{pmatrix},$$

and write $S = TW$. Then $\hat{A} = S^{-1}\tilde{A}S$, $\hat{B} = S^{-1}\tilde{B}$ and $\hat{C} = \tilde{C}S$ are of the form as described in the theorem. The if part is proved in the same fashion as the if part of Theorem 2.1.2. □

Let $W(\lambda)$ be a rational matrix function, and let $\rho(W)$ denote the infimum of all integers N, such that $W(\lambda)$ admits a factorization into N elementary rational matrix functions. In Theorem 2.4.2 below, it is shown that $\rho(W) < \infty$, i.e., that all rational matrix functions admit a factorization into a finite number of elementary rational matrix functions. The factorization involving the minimal number $\rho = \rho(W)$ elementary factors

$$W(\lambda) = W_1(\lambda) \cdots W_\rho(\lambda)$$

is called a *quasicomplete factorization*. Obviously, $\rho(W) \geq \delta(W)$.

For a square matrix B, we define the spectral polynomial $p_B(\lambda)$ as in (2.7).

Theorem 2.4.2 *Consider the minimal realization (2.13), and write $A^\times = A - BC$. Define the integer*

$$k(W) = m - \dim \left(\mathrm{Ker}\ p_A(A) + \mathrm{Ker}\ p_{A^\times}(A^\times) \right).$$

Then $\rho(W) \leq \delta(W) + k(W) \leq 2\delta(W) - 1$.

Proof The pair of matrices A, B in the minimal realization (2.13) is controllable. Therefore, according to the pole assignment theorem (Theorem 6.5.1 in [31]), there exists for each m-tuple of complex numbers $g_1, \ldots, g_m$ an $n \times m$ matrix K, such that $A + BK$ has eigenvalues $g_1, \ldots, g_m$. We will assume these eigenvalues to be distinct, so that $A + BK$ is diagonalizable. We will also assume that this set of complex numbers does not intersect $\sigma(A) \cup \sigma(A^\times)$.

Consider the subspace $M = \mathrm{Ker}\ p_A(A) + \mathrm{Ker}\ p_{A^\times}(A^\times)$. By definition, codim $M = k(W)$. Write $k = k(W)$, and note that $0 \leq k \leq m - 1$. There exist k eigenvectors $x_1, \ldots x_k$ for $A + BK$, such that $M \oplus \mathrm{span}\{x_1, \ldots, x_k\} = \mathbf{C}^m$. After renumbering the eigenvalues, we may write $(A + BK)x_j = g_j x_j$ for $j = 1, \ldots, k$. Define the $m \times k$ matrix $X = (x_1,\ x_2,\ \ldots,\ x_k)$. Then $(A + BK)X = XG$, where G is an $k \times k$ diagonal matrix with diagonal $\mathrm{diag}(G) = (g_1,\ \ldots,\ g_k)^T$. Define the $n \times k$ matrix $F = -KX$ to obtain $AX - XG = BF$. Consider the matrices

$$\hat{A} = \begin{pmatrix} A & BF \\ O & G \end{pmatrix}, \quad \hat{B} = \begin{pmatrix} B \\ O \end{pmatrix}, \quad \hat{C} = \begin{pmatrix} C & F \end{pmatrix}.$$

Further, write

$$\hat{A}^\times = \hat{A} - \hat{B}\hat{C} = \begin{pmatrix} A^\times & O \\ O & G \end{pmatrix},$$

where, as usual, $A^\times = A - BC$. Note that

$$\begin{pmatrix} I_m & -X \\ O & I_k \end{pmatrix} \begin{pmatrix} A & O \\ O & G \end{pmatrix} \begin{pmatrix} I_m & X \\ O & I_k \end{pmatrix} =$$

$$\begin{pmatrix} A & AX - XG \\ O & G \end{pmatrix} = \begin{pmatrix} A & BF \\ O & G \end{pmatrix} = \hat{A}.$$

Since $\sigma(A) \cap \sigma(G) = \emptyset$, we get $p_{\hat{A}}(\lambda) = p_A(\lambda) p_G(\lambda)$. Hence

$$p_{\hat{A}}(\hat{A}) = \begin{pmatrix} I_m & -X \\ O & I_k \end{pmatrix} \begin{pmatrix} p_G(A) & O \\ O & p_A(G) \end{pmatrix} \begin{pmatrix} p_A(A) & O \\ O & O_k \end{pmatrix} \begin{pmatrix} I_m & X \\ O & I_k \end{pmatrix},$$

where $p_G(A)$ and $p_A(G)$ are invertible matrices. We have also used $p_G(G) = O_k$. Therefore,

$$\text{Ker } p_{\hat{A}}(\hat{A}) = \begin{pmatrix} I_m & -X \\ O & I_k \end{pmatrix} \text{Ker } \begin{pmatrix} p_A(A) & O \\ O & O_k \end{pmatrix}.$$

Also,

$$\text{Ker } p_{\hat{A}^\times}(\hat{A}^\times) = \text{Ker } \begin{pmatrix} p_{A^\times}(A^\times) & O \\ O & O_k \end{pmatrix}.$$

It is not difficult to verify, that

$$\text{Ker } p_{\hat{A}}(\hat{A}) + \text{Ker } p_{\hat{A}^\times}(\hat{A}^\times) = \mathbf{C}^{m+k}$$

if and only if

$$\text{Ker } p_A(A) + \text{Ker } p_{A^\times}(A^\times) + \text{Ran } X = M + \text{Ran } X = \mathbf{C}^m.$$

By construction, the latter is the case. Proposition 2.3.3 then yields that the pair $\hat{A}$, $\hat{A}^\times$ admits simultaneous reduction to complementary triangular forms. Next, use Theorem 2.4.1 to obtain that $W(\lambda)$ admits a factorization as in (2.14), with $N = m + k = \delta(W) + k(W)$. The theorem is proved. □

Consider $W(\lambda)$ and its realization as in Theorem 2.4.2. Since $\delta(W) = \delta(W^*)$, and $\rho(W) = \rho(W^*)$, we can state a dual version of Theorem 2.4.2. Indeed,

$$k^*(W) = \text{codim}\left(\text{Ker } p_{A^*}(A^*) + \text{Ker } p_{(A^*)^\times}((A^*)^\times)\right) =$$

$$\dim\left(\text{Ran } p_A(A) \cap \text{Ran } p_{A^\times}(A^\times)\right)$$

leads to $\rho(W) \leq \delta(W) + k^*(W) \leq 2\delta(W) - 1$. Conclusively, it follows that $\rho(W) \leq \delta(W) + \min\{k(W), k^*(W)\}$.

To illustrate Theorem 2.4.2 and its proof, we will factorize a given rational matrix function, that does not admit a complete factorization, into elementary factors. Such a factorization of this particular rational matrix function was already known to Thijsse.

Example 2.4.3 In this example, we will compute a quasicomplete factorization for the rational matrix function

$$W(\lambda) = \begin{pmatrix} 1 & \frac{-1}{\lambda^2} \\ 0 & 1 \end{pmatrix},$$

with minimal realization

$$W(\lambda) = I_2 + C(\lambda I_2 - A)^{-1}B,$$

where the realization matrices are given by

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Compute

$$A^\times = A - BC = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Since $A = A^\times$ is non-diagonalizable, it follows that the pair A, $A^\times$ does not admit simultaneous reduction to complementary triangular forms. By Theorem 2.1.2, $W(\lambda)$ does not admit a complete factorization, so $\rho(W) > 2$. We will now follow the proof of Theorem 2.4.2. Note that with

$$K = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

we have

$$A + BK = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

so $A + BK$ is diagonalizable. Note that $k(W) = 1$, and take

$$X = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad G = 1, \quad F = \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$

so that $AX - XG = BF$. We now obtain the extended matrices

$$\hat{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \hat{B} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \hat{C} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

and

$$\hat{A}^\times = \hat{A} - \hat{B}\hat{C} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

By construction, the pair $\hat{A}$, $\hat{A}^\times$ admits simultaneous reduction to complementary triangular forms. In fact, the similarity

$$S = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

transforms $\hat{A}$ and $\hat{A}^\times$ into complementary triangular forms as follows:

$$S^{-1}\hat{A}S = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad S^{-1}\hat{A}^\times S = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

In addition,

$$S^{-1}\hat{B} = \begin{pmatrix} 0 & 1 \\ 0 & -1 \\ 0 & 1 \end{pmatrix}, \quad \hat{C}S = \begin{pmatrix} -1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}.$$

As explained in the proof of Theorem 2.1.2, we may now calculate the elementary factors explicitly and obtain $W(\lambda) = W_1(\lambda)W_2(\lambda)W_3(\lambda)$, with

$$W_1(\lambda) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{\lambda - 1}\begin{pmatrix} -1 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{-1}{\lambda-1} \\ 0 & \frac{\lambda}{\lambda-1} \end{pmatrix},$$

$$W_2(\lambda) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{\lambda}\begin{pmatrix} -1 \\ 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{\lambda} \\ 0 & 1 \end{pmatrix},$$

$$W_3(\lambda) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{\lambda - 1}\begin{pmatrix} 0 \\ -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\lambda-1}{\lambda} \end{pmatrix}.$$

We may conclude that $\rho(W) = 3$.

The following example concerns a particular type of extended realizations.

Example 2.4.4 Consider the 4×4 matrices

$$A_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad Z_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

In Example 3.2.1, it will be shown that the pair A_1, Z_1 does not admit simultaneous reduction to complementary triangular forms. The matrices

$$B_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad C_1 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

satisfy $A_1 - Z_1 = B_1 C_1$. Further, the pair A_1, B_1 is controllable, and the pair C_1, A_1 is observable, so the realization

$$W(\lambda) = I_3 + C_1(\lambda I_4 - A_1)^{-1} B_1 \tag{2.17}$$

is minimal. By Theorem 2.1.2, the rational matrix function $W(\lambda)$ does not admit a complete factorization. On the other hand, Example 3.2.1 shows that the pair of 5×5 matrices $A = A_1 \oplus 0$, $Z = Z_1 \oplus 0$ admits simultaneous reduction to complementary triangular forms. Define the matrices

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix},$$

which are obtained form B_1 and C_1 respectively by adding zero entries at the appropriate places, such that

$$W(\lambda) = I_3 + C(\lambda I_5 - A)^{-1} B$$

is an extended realization of (2.17), and such that $A - BC = Z$. The 5×5 invertible matrix

$$S = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

puts A and Z into complementary triangular forms. Therefore, we obtain a factorization into elementary factors

$$W(\lambda) = W_1(\lambda) \cdots W_5(\lambda),$$

where

$$W(\lambda) = \begin{pmatrix} 1 - \frac{1}{\lambda^3} & -\frac{1}{\lambda} & 0 \\ 0 & 1 & -\frac{1}{\lambda} \\ \frac{1}{\lambda} & 0 & 1 \end{pmatrix},$$

and

$$W_1(\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{\lambda} \\ 0 & 0 & 1 \end{pmatrix}, W_2(\lambda) = \begin{pmatrix} 1 & -\frac{1}{\lambda} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$W_3(\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{\lambda} \\ 0 & 0 & 1 \end{pmatrix}, W_4(\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{\lambda} & 0 & 1 \end{pmatrix},$$

$$W_5(\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{\lambda} \\ 0 & 0 & 1 \end{pmatrix}.$$

We conclude that $\delta(W) = 4$ and $\rho(W) = 5$.

As in Example 2.4.4, consider for a minimal realization (2.13) an extended realization (2.16) with matrices

$$\hat{A} = \begin{pmatrix} O & O & O \\ O & A & O \\ O & O & O \end{pmatrix}, \quad \hat{B} = \begin{pmatrix} O \\ B \\ O \end{pmatrix}, \quad \hat{C} = \begin{pmatrix} O & C & O \end{pmatrix},$$

where all extension blocks consist of zero entries. Write $m_2 = N - m$. Without violating the generality, we may assume that $\hat{A} = A \oplus O_{m_2}$ and $\hat{A}^\times = A^\times \oplus O_{m_2}$. Consider the set

$$\{m + m_2 \mid m_2 \in \mathbf{Z}_0^+, \quad (A \oplus O_{m_2}, A^\times \oplus O_{m_2}) \in \mathcal{C}(m + m_2)\},$$

and denote its infimum by $\rho_0(A, A^\times)$. As usual, the infimum over the empty set is defined $\inf \emptyset = +\infty$. First of all, note that $\rho_0(A, A^\times) = m$ if and only if

$W(\lambda)$ admits a complete factorization. Further, note that $\rho(W) \leq \rho_0(A, A^\times)$. The infimum $\rho_0(A, A^\times)$ will be studied in more detail in Chapter 3.

Let $W(\lambda)$ be an $n \times n$ rational matrix function, and let $\nu(W)$ denote the maximal number of non-trivial factors that can occur in a minimal factorization of $W(\lambda)$. Then $1 \leq \nu(W) \leq \delta(W)$. Write $\nu(W) = \nu$, and consider the minimal factorization

$$W(\lambda) = W_1(\lambda) \cdots W_\nu(\lambda).$$

By Theorem 2.4.2, applied on all factors separately, we get

$$\rho(W) \leq \sum_{j=1}^{\nu} \rho(W_j) \leq \sum_{j=1}^{\nu} [\delta(W_j) + k(W_j)] \leq$$

$$\sum_{j=1}^{\nu} [2\delta(W_j) - 1] = 2\delta(W) - \nu(W),$$

so $\rho(W) + \nu(W) \leq 2\delta(W)$. This inequality is not sharp. Indeed, the rational matrix function

$$W(\lambda) = I_3 + C(\lambda I_3 - A)^{-1}B,$$

with

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix},$$

satisfies $\delta(W) = 3$, $\rho(W) = 4$, and $\nu(W) = 1$.

A *companion based rational matrix function* is a rational matrix function, that admits a minimal realization (2.13), with A and $A^\times$ first companion matrices. Complementary triangular forms for a pair of first companion matrices is now well understood; see Theorem 2.2.2. This result states that a pair of first companion matrices admits simultaneous reduction to complementary triangular forms if and only if a combinatorial condition is met. In this manner, complete factorization for companion based rational matrix functions is characterized. Surprisingly enough, an important job scheduling problem, known as the Two Machine Flow Shop Problem, can be rewritten in terms of the combinatorial condition mentioned above. It turns out that there exists a feasible schedule for an instance of the Two Machine Flow shop Problem if and only if the associated companion based rational matrix function admits a complete factorization. Moreover, the minimal infeasibility of the instance of the Two Machine Flow Shop Problem is measured by the nonnegative integer $\delta(W) - \nu(W)$, which equals zero if there exists a feasible schedule. For details, see [9], [10] and [11].

The question of whether quasicomplete factorization for companion based rational matrix functions has a meaningful interpretation in job scheduling requires further research.

Chapter 3

Extensions with Zeroes

Let m_1 be a positive integer. A pair of $m_1 \times m_1$ matrices A_1, Z_1 admits *simultaneous reduction to complementary triangular forms after extension with zeroes*, if there exists a nonnegative integer m_2, such that the pair of matrices $A = A_1 \oplus O_{m_2}$, $Z = Z_1 \oplus O_{m_2}$ admits simultaneous reduction to complementary triangular forms.

Simultaneous reduction to complementary triangular forms after extensions with zeroes is the main subject of this chapter. First of all, we remark that if the pair of $m_1 \times m_1$ matrices A_1, Z_1 admits simultaneous reduction to complementary triangular forms, then so does the pair $A_1 \oplus O_{m_2}$, $Z_1 \oplus O_{m_2}$. Indeed, if S_1 is an invertible $m_1 \times m_1$ matrix that reduces the pair A_1, Z_1 to complementary triangular forms, then the invertible matrix $S = S_1 \oplus I_{m_2}$ reduces the pair $A = A_1 \oplus O_{m_2}$, $Z = Z_1 \oplus O_{m_2}$ to complementary triangular forms.

Recall that $\mathcal{C}(m)$ denotes the collection of all pairs of $m \times m$ matrices that admit simultaneous reduction to complementary triangular forms. We have just shown that if m_2 and m_3 are nonnegative integers, with $m_2 \leq m_3$, then $(A_1 \oplus O_{m_2}, Z_1 \oplus O_{m_2}) \in \mathcal{C}(m_1 + m_2)$ implies that $(A_1 \oplus O_{m_3}, Z_1 \oplus O_{m_3}) \in \mathcal{C}(m_1 + m_3)$. Therefore, it is appropriate to consider for a given pair of $m_1 \times m_1$ matrices A_1, Z_1 the infimum

$$\rho_0(A_1, Z_1) = \inf\{m_1 + m_2 \mid m_2 \in \mathbf{Z}_0^+, \quad (A_1 \oplus O_{m_2}, Z_1 \oplus O_{m_2}) \in \mathcal{C}(m_1 + m_2)\},$$

where the infimum over the empty set is defined as $\inf \emptyset = +\infty$.

It will turn out that the study of this infimum is difficult. This is due to the fact that the notion of simultaneous reduction to complementary triangular forms is not completely understood. Only partial results have been obtained in this direction (see Sections 2.2 and 2.3).

In this chapter, the infimum $\rho_0(A_1, Z_1)$ is studied for pairs of $m_1 \times m_1$ matrices A_1, Z_1.

In Section 3.2, an example is given of a pair of 4×4 matrices A_1, Z_1, for which the infimum $\rho_0(A_1, Z_1) = 5$, i.e., for which $m_1 < \rho_0(A_1, Z_1) < \infty$. In other words, this pair satisfies $(A_1, Z_1) \notin \mathcal{C}(4)$, while $(A_1 \oplus 0, Z_1 \oplus 0) \in \mathcal{C}(5)$.

In Section 3.3 up to and including Section 3.8, pairs of $m_1 \times m_1$ matrices A_1, Z_1 are discussed, for which $\rho_0(A_1, Z_1) = m_1$ or $\rho_0(A_1, Z_1) = \infty$. Note that for a pair of $m_1 \times m_1$ matrices A_1, Z_1, the infimum $\rho_0(A_1, Z_1) = m_1$ if and only if $(A_1, Z_1) \in \mathcal{C}(m_1)$, and that $\rho_0(A_1, Z_1) = \infty$ if and only if for any nonnegative integer m_2, one gets $(A_1 \oplus O_{m_2}, Z_1 \oplus O_{m_2}) \notin \mathcal{C}(m_1 + m_2)$.

In order to prove that the infimum $\rho_0(A_1, Z_1)$ is either m_1 or infinity for certain pairs of $m_1 \times m_1$ matrices A_1, Z_1, it is enough to prove that for any nonnegative integer m_2, $(A_1 \oplus O_{m_2}, Z_1 \oplus O_{m_2}) \in \mathcal{C}(m_1 + m_2)$ implies $(A_1, Z_1) \in \mathcal{C}(m_1)$. Section 3.3 up to and including Section 3.8 contain this type of results.

In Chapter 5, the infimum $\rho_0(A_1, Z_1)$ appears in the study of complementary triangular forms for pairs of finite rank operators acting on an infinite dimensional Banach space. Corollary 5.3.3 proves the following for a pair of $m_1 \times m_1$ matrices A_1, Z_1: If $\rho_0(A_1, Z_1) < \infty$, then actually $\rho_0(A_1, Z_1) \leq 2m_1(4m_1 - 1)$.

Section 3.1 discusses preliminaries that are used in this chapter.

3.1 Nests of Invariant Subspaces

In this preliminary section, a geometric description is given of triangular forms for matrices.

A complex $m \times m$ matrix A can be reduced to upper triangular form by means of a unitary transformation. This fact is known as Schur's theorem; see for example Theorem 5.2.2 in [38] or Theorem 1.9.1 in [31]. Let S be an invertible (e.g. unitary) $m \times m$ matrix, such that $S^{-1}AS$ is an upper triangular matrix. Write the matrix S in the form

$$S = (\ s_1,\ \ldots,\ s_m\ ) \tag{3.1}$$

i.e., as a row of m linear independent column vectors $s_1, \ldots, s_m$. These vectors define subspaces

$$M_k = \text{span}\{s_1, \ldots, s_k\}, \quad k = 0, \ldots, m, \tag{3.2}$$

which form the set $\mathcal{M} = \{M_k\}_{k=0}^{m}$. The set $\mathcal{M}$ is a *nest of subspaces*, i.e., is a set of subspaces that is linearly ordered by inclusion. Furthermore, $\mathcal{M}$ is not properly contained in any other nest of subspaces. Therefore, $\mathcal{M}$ is called a *maximal nest of subspaces*. As a matter of fact, *all* maximal nests of subspaces in $\mathbf{C}^m$ are of the form

$$\mathcal{M} = \{(0) = M_0 \subset M_1 \subset \cdots \subset M_m = \mathbf{C}^m\} = \{M_k\}_{k=0}^m,$$

where $\dim M_k = k$ for $k = 0, \ldots, m$. Finally, the maximal nest of subspaces as defined in (3.2) consists of invariant subspaces for A, and is called a *maximal invariant nest of subspaces* for A. Schur's Lemma thus says that each $m \times m$ matrix A has a maximal invariant nest of subspaces.

Conversely, let $\mathcal{M} = \{M_k\}_{k=0}^m$ be a maximal invariant nest of subspaces for A. For each $k \in \{1, \ldots, m\}$, there exists a unique complex number α_k, such that

$$(A - \alpha_k)M_k \subseteq M_{k-1}.$$

We define the diagonal of A with respect to the maximal invariant nest of subspaces $\mathcal{M}$ as

$$\operatorname{diag}(A; \mathcal{M}) = \alpha = (\alpha_1, \ldots, \alpha_m)^T.$$

Each set of vectors $s_1, \ldots, s_m$ that satisfies (3.2) defines an invertible $m \times m$ matrix S as in (3.1), such that $S^{-1}AS$ is upper triangular, with $\operatorname{diag}(S^{-1}AS) = \alpha$. Note that in particular, α is a spectral vector for A.

Let $\sigma \subseteq \sigma(A)$ be a non-empty subset of the spectrum of the $m \times m$ complex matrix A and define the linear subspace

$$N_\sigma(A) = \operatorname{span}\{\operatorname{Ker}(A - \alpha)^m \mid \alpha \in \sigma\}.$$

If $\sigma = \{\alpha\}$, we usually write $N_{\{\alpha\}}(A) = N_\alpha(A)$. The following lemma states that an invariant subspace of a matrix admits a decomposition, related to the spectrum of the matrix.

Lemma 3.1.1 *Let A be a complex $m \times m$ matrix and assume that the spectrum of A is the disjoint union of two non-empty subsets σ_1 and σ_2. Let M be an invariant subspace for A. Then M admits the decomposition*

$$M = M_1 \oplus M_2, \quad M_i = M \cap N_{\sigma_i}(A), \quad i = 1, 2.$$

A decomposition as in Lemma 3.1.1 is referred to as a *spectral decomposition* of M, associated with σ_1 and σ_2. For the proof of Lemma 3.1.1, we refer to Section 2.1 in [31].

We now state and prove two simple lemmas, that are used in the sequel.

Lemma 3.1.2 *Let $\mathcal{M} = \{M_k\}_{k=0}^m$ be a maximal invariant nest for the $m \times m$ matrix A and let N be an n-dimensional invariant subspace for A. Write $M_{k,1} = M_k \cap N$ for $k = 0, \ldots, m$, and denote the restriction of A to N by A_1. Put*

$$\tau(i) = \min \{k \mid 0 \le k \le m, \dim\, M_{k,1} = i\}, \quad i = 0, \ldots, n.$$

Then the nest $\mathcal{M}_1 = \{M_{\tau(i),1}\}_{i=0}^n$ is a maximal invariant nest in N for A_1.

Proof Fix $k \in \{0, \ldots, m\}$. Since M_k and N are invariant subspaces for A, the same is true for $M_{k,1}$. If $x \in M_{k,1}$, then $A_1 x = Ax \in M_{k,1}$, i.e., $M_{k,1}$ is an invariant subspace for A_1. Further note that $M_{0,1} = (0)$ and $M_{m,1} = N$. In addition it is easily verified that $\dim\, M_{k,1} - \dim\, M_{k-1,1} \le 1$. It follows that the integers $\tau(0), \tau(1), \ldots, \tau(n)$ are well-defined, that $\tau(0) < \tau(1) < \cdots < \tau(n)$ and that $\mathcal{M}_1$ is a maximal invariant nest in N for A_1. $\square$

Lemma 3.1.3 *Let A be an $m \times m$ matrix and let $\mathcal{M} = \{M_k\}_{k=0}^m$ be a maximal invariant nest for A with $\operatorname{diag}(A; \mathcal{M}) = (\alpha_1, \ldots, \alpha_m)^T$. Let $\alpha \in \sigma(A)$ and let N be an n-dimensional invariant subspace of A in $\operatorname{Ker}(A - \alpha)^m$. Define*

$$\tau(i) = \min \{k \mid 1 \le k \le m, \dim(M_k \cap N) = i\}, \quad i = 1, \ldots, n,$$

then $\alpha_{\tau(i)} = \alpha$.

Proof Fix $i \in \{1, \ldots, n\}$, and assume that $\alpha_{\tau(i)} = \hat{\alpha} \neq \alpha$. By definition, $\tau(i) \ge 1$, and $M_{\tau(i)-1} \cap N \subset M_{\tau(i)} \cap N$. Further recall that $(A - \hat{\alpha}) M_{\tau(i)} \subseteq M_{\tau(i)-1}$. Consequently,

$$(A - \hat{\alpha}) \left(M_{\tau(i)} \cap N \right) \subseteq M_{\tau(i)-1} \cap N \subset M_{\tau(i)} \cap N.$$

On the other hand, the restriction of $A - \hat{\alpha}$ to the subspace $M_{\tau(i)} \cap N \subseteq N_\alpha(A)$ is invertible. A contradiction is obtained and the lemma is proved. $\square$

To reformulate the notion of simultaneous reduction to complementary triangular forms in terms of maximal invariant nests of subspaces, we introduce the following terminology. Two maximal nests $\mathcal{M} = \{M_k\}_{k=0}^m$ and $\mathcal{N} = \{N_k\}_{k=0}^m$ in $\mathbf{C}^m$ are called *matching*, if

$$M_k \oplus N_{m-k} = \mathbf{C}^m, \quad k = 0, \ldots, m. \tag{3.3}$$

A set of projections $\mathcal{P} = \{P_k\}_{k=0}^m$ is called a *maximal nest of projections*, if the sets of subspaces $\{\text{Ran } P_k\}_{k=0}^m$ and $\{\text{Ker } P_k\}_{k=0}^m$ are (matching) maximal nests. For more details on nests of projections, we refer to Chapter 4. The following simple observation provides alternative descriptions of simultaneous reduction to complementary triangular forms.

Lemma 3.1.4 *Let A and Z be two $m \times m$ matrices, then the following are equivalent.*

1. *The pair A, Z admits simultaneous reduction to complementary triangular forms.*
2. *There exist matching maximal invariant nests $\mathcal{M}$ and $\mathcal{N}$ for A and Z respectively.*
3. *There exists a maximal nest of projections $\mathcal{P} = \{P_k\}_{k=0}^m$ such that $AP_k = P_kAP_k$ and $P_kZ = P_kZP_k$ for $k = 0, \ldots, m$.*

Proof To prove that the first statement implies the second one, let S be an invertible $m \times m$ matrix as in (3.1), such that $S^{-1}AS$ is upper triangular and $S^{-1}ZS$ is lower triangular. Then the maximal nests of subspaces $\mathcal{M} = \{M_k\}_{k=0}^m$ and $\mathcal{N} = \{N_k\}_{k=0}^m$, defined by

$$M_k = \text{span}\{s_1, \ldots, s_k\}, \quad N_{m-k} = \text{span}\{s_{k+1}, \ldots, s_m\}, \quad k = 0, \ldots, m,$$

are invariant for A and Z respectively, and are matching. Next, we prove that the second statement implies the first one. If $\mathcal{M} = \{M_k\}_{k=0}^m$ and $\mathcal{N} = \{N_k\}_{k=0}^m$ are matching maximal nests of subspaces, invariant for A and Z respectively, then the non-zero vectors $s_1, \ldots, s_m$, determined up to multiplicative constants by $s_k \in M_k \cap N_{m-k+1}$, define an invertible $m \times m$ matrix S as in (3.1), which reduces A and Z to complementary triangular forms. Here, we used

$$\dim(M_k \cap N_l) = \max\{k + l - m, 0\}.$$

The equivalence between the second and third statement is proved as follows. By definition, two matching maximal nests of subspaces give rise to a maximal nest of projections, and vice versa. In addition, $AP_k = P_kAP_k$ is equivalent to $A(\text{Ran } P_k) \subseteq \text{Ran } P_k$, and $P_kZ = P_kZP_k$ is equivalent to $Z(\text{Ker } P_k) \subseteq \text{Ker } P_k$. The proof is finished. □

Let S denote the invertible $m \times m$ matrix, obtained from the nests of subspaces $\mathcal{M}$ and $\mathcal{N}$ as in the proof of Lemma 3.1.4; note that S is determined up to multiplication to the right by an invertible diagonal matrix. Assume that

$$\text{diag}(A;\mathcal{M}) = (\alpha_1, \ldots, \alpha_m)^T = \alpha, \quad \text{diag}(Z;\mathcal{N}) = (\zeta_1, \ldots, \zeta_m)^T = \zeta.$$

Then

$$\text{diag}(S^{-1}AS) = (\alpha_1, \ldots, \alpha_m)^T = \alpha,$$

but

$$\text{diag}(S^{-1}ZS) = (\zeta_m, \ldots, \zeta_1)^T,$$

i.e., the entries of $\text{diag}(Z;\mathcal{N})$ appear in *reversed order* on the diagonal of the lower triangular matrix $S^{-1}ZS$.

Usually, we identify matrices with their action as a linear operator. This leads to the following convention, which is used several times in this chapter. Let B_1 be an $m_1 \times m_1$ matrix and let B_2 be an $m_2 \times m_2$ matrix, and define the $m \times m$ matrix $B = B_1 \oplus B_2$. In general, we make the identification that an $n \times n$ matrix acts as a linear operator on $\mathbf{C}^n$. Therefore, B acts on $\mathbf{C}^m$ and with respect to the decomposition $\mathbf{C}^m = \mathbf{C}^{m_1} \oplus \mathbf{C}^{m_2}$ it assumes the form

$$B = \begin{pmatrix} B_1 & O \\ O & B_2 \end{pmatrix}.$$

After making the identifications $\mathbf{C}^{m_1} = \mathbf{C}^{m_1} \oplus (0)$ and $\mathbf{C}^{m_2} = (0) \oplus \mathbf{C}^{m_2}$, we get that B_1 and B_2 denote the restrictions of B to the subspace $\mathbf{C}^{m_1}$ and $\mathbf{C}^{m_2}$ respectively.

3.2 An Example

The following example shows, that there exist pairs of matrices, that do not admit simultaneous reduction to complementary triangular forms, but obtain this property after extension with zeroes. Recall that an $m \times m$ matrix is *unicellular* or *uniserial*, i.e., has a unique maximal invariant nest of subspaces, if and only if it is similar to an $m \times m$ Jordan block.

Example 3.2.1 This example (see also Example 2.4.4) provides a pair of nilpotent 4×4 matrices A_1, Z_1, that does not admit simultaneous reduction to complementary triangular forms, while the the pair of 5×5 matrices $A_1 \oplus 0$, $Z_1 \oplus 0$ does have this property. Indeed, the pair of 4×4 matrices

$$A_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad Z_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

does not admit simultaneous reduction to complementary triangular forms (as will be shown later), while the pair of 5×5 matrices $A_1 \oplus 0$, $Z_1 \oplus 0$ is reduced to complementary triangular forms by the invertible 5×5 matrix

$$S = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

as follows:

$$S^{-1}(A_1 \oplus 0)S = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad S^{-1}(Z_1 \oplus 0)S = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

On the other hand, we will show that $(A_1, Z_1) \notin \mathcal{C}(4)$. Indeed, there exist no matching maximal invariant nests of subspaces for A_1 and Z_1 respectively. Note that Z_1 is unicellular, and that its unique maximal invariant nest of subspaces $\mathcal{N} = \{N_k\}_{k=0}^4$ is given by

$$\begin{aligned} &N_0 = (0), \\ &N_1 = \text{span}\{e_1\}, \\ &N_2 = \text{span}\{e_1, e_2\}, \\ &N_3 = \text{span}\{e_1, e_2, e_3\}, \\ &N_4 = \mathbf{C}^4. \end{aligned}$$

Assume there exists a maximal invariant nest of subspaces $\mathcal{M} = \{M_k\}_{k=0}^4$ for A_1, that matches $\mathcal{N}$. First of all, $M_1 \subseteq \text{Ker } A_1 = \text{span}\{e_3, e_4\}$. Further, to obtain that $M_1 \oplus N_3 = \mathbf{C}^4$, we have to take $M_1 = \text{span}\{e_4 + \alpha e_3\}$ for some complex number α, since $e_3 \in N_3$. Since $A_1(M_2) \subseteq M_1 \cap \text{Ran } A_1 = (0)$, we get

$M_2 = \text{Ker } A_1$. Since $A_1(M_3) \subseteq M_2 = \text{Ker} A_1$, it follows that $M_3 = \text{Ker } A_1^2 = \text{span}\{e_1, e_3, e_4\}$. But then $M_3 \cap N_1 = \text{span}\{e_1\} \neq (0)$, and a contradiction has been obtained. Therefore, $(A_1, Z_1) \notin \mathcal{C}(4)$.

3.3 Invertible Matrices

We start with the following simple observation.

Proposition 3.3.1 *Let A_i and Z_i be $m_i \times m_i$ matrices ($i = 1, 2$), and write $m = m_1 + m_2$. Define the $m \times m$ matrices $A = A_1 \oplus A_2$ and $Z = Z_1 \oplus Z_2$. Assume that $\sigma(A_1) \cap \sigma(A_2) = \sigma(Z_1) \cap \sigma(Z_2) = \emptyset$. If the pair A, Z admits simultaneous reduction to complementary triangular forms, then so do the pair A_1, Z_1 and the pair A_2, Z_2.*

Proof We write

$$A = \begin{pmatrix} A_1 & O \\ O & A_2 \end{pmatrix}, \qquad Z = \begin{pmatrix} Z_1 & O \\ O & Z_2 \end{pmatrix}$$

with respect to the decomposition $\mathbf{C}^m = \mathbf{C}^{m_1} \oplus \mathbf{C}^{m_2}$. Let $\mathcal{M} = \{M_k\}_{k=0}^m$ and $\mathcal{N} = \{N_k\}_{k=0}^m$ be matching maximal invariant nests of subspaces for A and Z respectively. Since $\sigma(A_1) \cap \sigma(A_2) = \emptyset$, we get

$$\mathbf{C}^{m_i} = \text{span}\{\text{Ker}(A - \alpha)^m \mid \alpha \in \sigma(A_i)\}, \qquad i = 1, 2,$$

and since $\sigma(Z_1) \cap \sigma(Z_2) = \emptyset$, we get

$$\mathbf{C}^{m_i} = \text{span}\{\text{Ker}(Z - \zeta)^m \mid \zeta \in \sigma(Z_i)\}, \qquad i = 1, 2.$$

We may apply Lemma 3.1.1 to obtain the decompositions

$$M_k = (M_k \cap \mathbf{C}^{m_1}) \oplus (M_k \cap \mathbf{C}^{m_2}),$$

$$N_{m-k} = (N_{m-k} \cap \mathbf{C}^{m_1}) \oplus (N_{m-k} \cap \mathbf{C}^{m_2}),$$

for $k = 0, \ldots, m$. The matching condition $M_k \oplus N_{m-k} = \mathbf{C}^m$ then gives

$$(M_k \cap \mathbf{C}^{m_i}) \oplus (N_{m-k} \cap \mathbf{C}^{m_i}) = \mathbf{C}^{m_i}, \tag{3.4}$$

$k = 0, \ldots, m$ and $i = 1, 2$. Fix $i \in \{1, 2\}$ and define the integers

$$\tau_i(s) = \min\{k \mid 0 \le k \le m, \dim(M_k \cap \mathbf{C}^{m_i}) = s\}, \quad s = 0, \ldots, m_i.$$

Lemma 3.1.2 provides that $\{M_{\tau_i(s)}\}_{s=0}^{m_i}$ is a maximal invariant nest for A_i. The matching condition (3.4) gives that $\{N_{m-\tau_i(s)}\}_{s=0}^{m_i}$ is a maximal invariant nest for Z_i which matches $\{M_{\tau_i(s)}\}_{s=0}^{m_i}$. By Lemma 3.1.4, the pair A_i, Z_i admits simultaneous reduction to complementary triangular forms. □

For a more elaborate proof of Proposition 3.3.1 in which certain aspects of the proof are written out in more detail, we refer to [54]. Theorem 3.3.2, which follows from Proposition 3.3.1, proves that for pairs of invertible matrices A_1, Z_1 the infimum $\rho_0(A_1, Z_1) \in \{m_1, \infty\}$.

Theorem 3.3.2 *Let A_1 and Z_1 be invertible $m_1 \times m_1$ matrices, and write $m = m_1 + m_2$. If the pair of $m \times m$ matrices $A = A_1 \oplus O_{m_2}$, $Z = Z_1 \oplus O_{m_2}$ admits simultaneous reduction to complementary triangular forms, then so does the pair A_1, Z_1.*

3.4 Unicellular and Nilpotent Matrices

Recall that an $m \times m$ matrix is unicellular if and only if it is similar to an $m \times m$ Jordan block. The following theorem shows that for pairs of unicellular matrices A_1, Z_1 the infimum $\rho_0(A_1, Z_1) \in \{m_1, \infty\}$.

Theorem 3.4.1 *Let A_1 and Z_1 be unicellular $m_1 \times m_1$ matrices and let $m = m_1 + m_2$. Define the $m \times m$ matrices $A = A_1 \oplus O_{m_2}$ and $Z = Z_1 \oplus O_{m_2}$. If the pair A, Z admits simultaneous reduction to complementary triangular forms then so does the pair A_1, Z_1.*

Proof The proof of the theorem is divided into three parts, corresponding to the following cases:

1. The matrices A_1 and Z_1 are invertible.
2. The matrices A_1 and Z_1 are singular, therefore nilpotent.
3. The matrix A_1 is singular, the matrix Z_1 is invertible. By a symmetry argument, the case A_1 invertible and Z_1 singular is also covered by this part.

Part 1: The matrices A_1 and Z_1 are invertible.

Apply Theorem 3.3.2.

Part 2: The matrices A_1 and Z_1 are nilpotent.

Assume there exist maximal invariant nests $\mathcal{M} = \{M_k\}_{k=0}^m$ and $\mathcal{N} = \{N_k\}_{k=0}^m$ for A and Z respectively, that are matching. Write $M_{k,i} = M_k \cap \mathbf{C}^{m_i}$ and $N_{k,i} = N_k \cap \mathbf{C}^{m_i}$ for $k = 0, \ldots, m$ and $i = 1, 2$. For $s = 0, \ldots, m_1$, introduce

$$\pi(s) = \min\{k \mid 0 \leq k \leq m, \dim M_{k,1} = s\},$$

and

$$\rho(s) = \min\{l \mid 0 \leq l \leq m, \dim N_{l,1} = s\}.$$

By Lemma 3.1.2, it follows that $\{M_{\pi(s),1}\}_{s=0}^{m_1}$ and $\{N_{\rho(s),1}\}_{s=0}^{m_1}$ are maximal invariant nest for A_1 and Z_1 respectively. It remains to prove that these nests are matching.

Since A_1 and Z_1 are unicellular nilpotent $m_1 \times m_1$ matrices, there exists a basis $\phi_1, \ldots, \phi_{m_1}$ for $\mathbf{C}^{m_1}$, such that

$$A_1\phi_1 = 0, \quad A_1\phi_s = \phi_{s-1}, \quad s = 2, \ldots, m_1.$$

There also exists a basis $\psi_1, \ldots, \psi_{m_1}$ in $\mathbf{C}^{m_1}$, such that

$$Z_1\psi_1 = 0, \quad Z_1\psi_s = \psi_{s-1}, \quad s = 2, \ldots, m_1.$$

We now have

$$M_{\pi(s),1} = \operatorname{span}\{\phi_1, \ldots, \phi_s\}, \quad N_{\rho(s),1} = \operatorname{span}\{\psi_1, \ldots, \psi_s\}, \quad s = 0, \ldots, m_1.$$

Claim For $s = 0, \ldots, m_1$, the following two identities hold:

$$M_{\pi(s)} = M_{\pi(s),1} \oplus M_{\pi(s),2}, \quad N_{\rho(s)} = N_{\rho(s),1} \oplus N_{\rho(s),2}.$$

To prove the first identity, fix $s \in \{1, \ldots, m_1\}$, the case $s = 0$ being trivial. Let $x \in M_{\pi(s)}$ and write $x = x_1 + x_2$, where $x_i \in \mathbf{C}^{m_i}$. We need to prove that $x_1 \in M_{\pi(s)}$. If $x_1 = 0$ this is trivial, so assume that $0 \neq x_1 = \gamma_1\phi_1 + \cdots + \gamma_p\phi_p$, where $\gamma_1, \ldots, \gamma_p$ are complex numbers and $p \in \{1, \ldots, m_1\}$ such that $\gamma_p \neq 0$. Then $Ax = A_1x_1 = \gamma_2\phi_1 + \cdots + \gamma_p\phi_{p-1}$. On the other hand,

$$AM_{\pi(s)} \subseteq M_{\pi(s-1),1} = \operatorname{span}\{\phi_1, \ldots, \phi_{s-1}\}.$$

Therefore $p-1 \leq s-1$, thus $p \leq s$ and hence $x_1 \in M_{\pi(s)}$. The second identity is dealt with similarly, so the claim is proved.

To finish the proof of Part 2, fix $s \in \{0, \ldots, m_1\}$ and distinguish two cases:

Case 1: $\pi(s) + \rho(m_1 - s) \leq m$.

In this case, $M_{\pi(s)} \cap N_{\rho(m_1-s)} = (0)$ and hence $M_{\pi(s),1} \oplus N_{\rho(m_1-s),1} \subseteq \mathbf{C}^{m_1}$. A dimension argument shows that equality holds.

Case 2: $\pi(s) + \rho(m_1 - s) > m$.

In this case, $M_{\pi(s)} + N_{\rho(m_1-s)} = \mathbf{C}^m$, and hence $M_{\pi(s),1} + N_{\rho(m_1-s),1} = \mathbf{C}^{m_1}$. A dimension argument shows that $M_{\pi(s),1} \cap N_{\rho(m_1-s),1} = (0)$.

In both cases, it is proved that

$$M_{\pi(s),1} \oplus N_{\rho(m_1-s),1} = \mathbf{C}^{m_1}.$$

Part 3: The matrix A_1 is nilpotent, the matrix Z_1 is invertible.

Assume there exist maximal invariant nests $\mathcal{M} = \{M_k\}_{k=0}^m$ and $\mathcal{N} = \{N_l\}_{l=0}^m$ for A and Z respectively, which are matching. Since A_1 is nilpotent, one may define -as in Part 2 of the proof- a strictly increasing mapping $\pi : \{0, \ldots, m_1\} \longrightarrow \{0, \ldots, m\}$ such that dim $M_{\pi(s),1} = s$, and $M_{\pi(s)} = M_{\pi(s),1} \oplus M_{\pi(s),2}$. Since Z_1 is invertible, it follows by the proof of Proposition 3.3.1, that $N_k = N_{k,1} \oplus N_{k,2}$ for $k = 0, \ldots, m$. In particular $(s = 0, \ldots, m_1)$,

$$M_{\pi(s)} \oplus N_{m-\pi(s)} = [M_{\pi(s),1} \oplus N_{m-\pi(s),1}] \oplus [M_{\pi(s),2} \oplus N_{m-\pi(s),2}] = \mathbf{C}^{m_1} \oplus \mathbf{C}^{m_2}.$$

Consequently, the maximal invariant nests $\{M_{\pi(s),1}\}_{s=0}^{m_1}$ and $\{N_{m-\pi(m_1-s),1}\}_{s=0}^{m_1}$ for A_1 and Z_1 are matching. This finishes the proof of the theorem. □

Proposition 3.4.2 below is an extension of Theorem 4.1 in [12], which deals with the case $m_2 = 0$. The latter result, Theorem 2.2.4 in this thesis, concerns pairs of sharply upper triangular matrices.

Proposition 3.4.2 *Let $1 \leq \alpha, \omega \leq m_1 - 1$, and let A_{12} be an invertible upper triangular $(m_1-\alpha) \times (m_1-\alpha)$ matrix and Z_{12} be an invertible upper triangular $(m_1 - \omega) \times (m_1 - \omega)$ matrix. Define the $m_1 \times m_1$ matrices*

$$A_1 = \begin{pmatrix} O & A_{12} \\ O_\alpha & O \end{pmatrix}, \quad Z_1 = \begin{pmatrix} O & Z_{12} \\ O_\omega & O \end{pmatrix}.$$

Let m_2 be any nonnegative integer, $m = m_1 + m_2$, and consider the $m \times m$ matrices $A = A_1 \oplus O_{m_2}$ and $Z = Z_1 \oplus O_{m_2}$. Then the pair A, Z admits simultaneous reduction to complementary triangular forms, if and only if $\alpha + \omega > m_1$, α does not divide ω, and ω does not divide α.

Proof By a symmetry argument, we may assume without loss of generality, that $\alpha \leq \omega$. First, we prove the only if part. Note that A and Z are non-zero nilpotent $m \times m$ matrices. If $\alpha + \omega \leq m_1$, then (2.8) and (2.9) hold and Theorem 2.3.1 implies that the pair A, Z does not admit simultaneous reduction to complementary triangular forms. Second, assume that $\alpha = \omega$. Then Ker A = Ker Z, and Ran A = Ran Z. Again, Theorem 2.3.1 implies that the pair A, Z does not admit simultaneous reduction to complementary triangular forms. The case when ω is a multiple of α is reduced to the case $\alpha = \omega$, by taking an appropriate power of A.

To prove the if part, note that the if part of Theorem 4.1 in [12] provides that the pair A_1, Z_1 admits simultaneous reduction to complementary triangular forms, and hence does the pair A, Z. □

Proposition 3.4.2 implies the following theorem on simultaneous reduction to complementary triangular forms after extension with zeroes for pairs of sharply upper triangular matrices. Hence for these pairs of $m_1 \times m_1$ matrices A_1, Z_1, we get $\rho_0(A_1, Z_1) \in \{m_1, \infty\}$.

Theorem 3.4.3 *Let A_1 and Z_1 be nilpotent matrices as in Proposition 3.4.2, and let $A = A_1 \oplus O_{m_2}$ and $Z = Z_1 \oplus O_{m_2}$. If the pair A, Z admits simultaneous reduction to complementary triangular forms, then so does the pair A_1, Z_1.*

3.5 First Companion Matrices

Theorem 3.5.1 *Let A_1 and Z_1 be first companion $m_1 \times m_1$ matrices, and define $m = m_1 + m_2$. Consider the $m \times m$ matrices $A = A_1 \oplus O_{m_2}$ and $Z = Z_1 \oplus O_{m_2}$. Assume that the pair A, Z admits simultaneous reduction to complementary triangular forms with diagonals*

$$(\alpha_1, \ldots, \alpha_m)^T, \quad (\zeta_1, \ldots, \zeta_m)^T. \tag{3.5}$$

Then there exist strictly increasing mappings $\pi, \rho : \{1, \ldots, m_1\} \to \{1, \ldots, m\}$, such that the pair A_1, Z_1 admits simultaneous reduction to complementary triangular forms with diagonals

$$(\alpha_{\pi(1)}, \ldots, \alpha_{\pi(m_1)})^T, \quad (\zeta_{\rho(1)}, \ldots, \zeta_{\rho(m_1)})^T. \tag{3.6}$$

Theorem 3.5.1 shows that for pairs of first companion $m_1 \times m_1$ matrices, the infimum $\rho_0(A_1, Z_1) \in \{m_1, \infty\}$. The transpose of a first companion matrix is called a *second companion matrix*, while a third companion matrix is obtained from a first companion matrix after transformation by means of the reversed identity (see Section 2.2). Finally, a *fourth companion matrix* is the transpose of a third companion matrix. In this manner, it is not difficult to see that Theorem 3.5.1 carries over to pairs of second, pairs of third and pairs of fourth companion matrices. Before proving Theorem 3.5.1, we present two lemmas.

Lemma 3.5.2 *Let m_1 be a positive integer, m_2 be a nonnegative integer, and let $m = m_1 + m_2$. Let $\pi, \rho : \{1, \ldots, m_1\} \longrightarrow \{1, \ldots, m\}$ and $\tau, \sigma : \{1, \ldots, m_2\} \longrightarrow \{1, \ldots, m\}$ be strictly increasing mappings, such that $\pi(s) \neq \tau(i)$ and $\rho(s) \neq \sigma(i)$ for all $s = 1, \ldots, m_1$ and $i = 1, \ldots, m_2$. Then $\sigma \leq \tau$ implies that $\pi \leq \rho$.*

Proof The mappings π and ρ are completely determined by the mappings τ and σ respectively. In fact,

$$\pi(s) = \begin{cases} s & 1 \leq s \leq \tau(1) - 1 \\ s + 1 & \tau(1) \leq s \leq \tau(2) - 2 \\ \vdots & \vdots \\ s + m_2 & \tau(m_2) - m_2 + 1 \leq s \leq m_1 \end{cases},$$

i.e., $\pi(s) = s + i$, if $\tau(i) - i + 1 \leq s \leq \tau(i+1) - i - 1$. For convenience, we write $\tau(0) = 0$ and $\tau(m_2 + 1) = m + 1$. In addition,

$$\rho(t) = \begin{cases} t & 1 \leq t \leq \sigma(1) - 1 \\ t + 1 & \sigma(1) \leq t \leq \sigma(2) - 2 \\ \vdots & \vdots \\ t + m_2 & \sigma(m_2) - m_2 + 1 \leq t \leq m_1 \end{cases},$$

thus $\rho(s) = s + j$, if $\sigma(j) - j + 1 \leq s \leq \sigma(j+1) - j - 1$. Again, we write $\sigma(0) = 0$ and $\sigma(m_2 + 1) = m + 1$.

Fix $s \in \{1, \ldots, m_1\}$. To show that $\pi(s) \leq \rho(s)$, let $i, j \in \{0, \ldots, m_2\}$ such that

$$\tau(i) - i + 1 \leq s \leq \tau(i+1) - i - 1, \quad \sigma(j) - j + 1 \leq s \leq \sigma(j+1) - j - 1.$$

We prove $i \leq j$: If we assume that $i \geq j + 1$, then, using $\sigma \leq \tau$ in the first inequality, we get

$$\sigma(i) - i + 1 \leq \tau(i) - i + 1 \leq s \leq \sigma(j+1) - j - 1 < \sigma(j+1) - (j+1) + 1 \leq \sigma(i) - i + 1,$$

a contradiction. Further, $i \leq j$ implies that $\pi(s) = s + i \leq s + j = \rho(s)$. □

Lemma 3.5.3 *Let* B_1 *be an* $m_1 \times m_1$ *matrix,* $\dim \operatorname{Ker} B_1 = 1$ *and* $B = B_1 \oplus O_{m_2}$*. If* $M \subseteq N_0(B)$ *is an invariant subspace for* B *and* $\operatorname{Ker} B_1 \not\subseteq BM$*, then* $M \subseteq \operatorname{Ker} B$

Proof Let $\dim N_0(B_1) = n$. Since $\dim \operatorname{Ker} B_1 = 1$, there exist a basis $y_1, \ldots, y_n$ in $N_0(B_1)$, such that $B_1 y_1 = 0$ and $B_1 y_{k+1} = y_k$ for $k = 1, \ldots, n-1$. Let $0 \neq x \in M$, and assume that $x \notin \operatorname{Ker} B$. Then there exists $p \in \{2, \ldots, n\}$ and complex numbers $\xi_1, \ldots, \xi_p$, with $\xi_p \neq 0$, and $u \in \mathbf{C}^{m_2}$, such that $x = \xi_1 y_1 + \cdots + \xi_p y_p + u$. Then $B^{p-1} x = \xi_p y_1 \in BM$, which contradicts the assumption $\operatorname{span}\{y_1\} = \operatorname{Ker} B_1 \not\subseteq BM$. □

Proof of Theorem 3.5.1 We will actually prove a slight generalization of the result as stated in the theorem. In the proof we shall consider, for a given complex number γ, the $m \times m$ matrices $A = A_1 \oplus \gamma I_{m_2}$ and $Z = Z_1 \oplus \gamma I_{m_2}$. The theorem corresponds to the case when $\gamma = 0$. The proof consists of three parts, dealing with the following cases:

1. $\gamma \notin \sigma(A_1) \cup \sigma(Z_1)$.
2. $\gamma \notin \sigma(A_1)$, $\gamma \in \sigma(Z_1)$. By a symmetry argument, the case $\gamma \in \sigma(A_1)$, $\gamma \notin \sigma(Z_1)$ is also covered here.
3. $\gamma \in \sigma(A_1) \cap \sigma(Z_1)$.

Part 1: $\gamma \notin \sigma(A_1) \cup \sigma(Z_1)$.

Apply Proposition 3.3.1.

Part 2: $\gamma \notin \sigma(A_1)$, $\gamma \in \sigma(Z_1)$.

Assume that the pair A, Z admits simultaneous reduction to complementary triangular forms, by means of an $m \times m$ similarity S, with diagonals α and ζ as in (3.5). Let the maximal invariant nests (defined by S) for A and Z be $\mathcal{M} = \{M_k\}_{k=0}^m$ and $\mathcal{N} = \{N_l\}_{l=0}^m$, respectively.

We have to prove that the pair A_1, Z_1 admits simultaneous reduction to complementary triangular forms with diagonals as in (3.6). By Theorem 2.2.2, it suffices to find strictly increasing mappings $\pi, \rho : \{1, \ldots, m_1\} \longrightarrow \{1, \ldots, m\}$ such that the vectors in (3.6) are spectral vectors for A_1 and Z_1, and such that

$$\alpha_{\pi(s)} \neq \zeta_{\rho(t)}, \quad s + t \leq m_1. \tag{3.7}$$

To define π and ρ, we will analyse the diagonals α and ζ. More specifically, we will determine restrictions on the positions of the eigenvalues on the diagonals. The first claim below concerns the eigenvalues different from γ.

Claim 1 If $\alpha_k = \zeta_l \neq \gamma$, then $k + l > m$.

To prove the claim, assume that $\alpha_k = \zeta_l = \beta \neq \gamma$. Then $\beta \in \sigma(A_1) \cap \sigma(Z_1)$ and

$$\operatorname{Ker}(\beta - A_1) = \operatorname{Ker}(\beta - Z_1) = \operatorname{span}\{x_1(\beta)\},$$

since A_1 and Z_1 are first companion matrices. Because $\beta \neq \gamma$, it holds that

$$\operatorname{Ker}(\beta - A) = \operatorname{Ker}(\beta - Z) = \operatorname{span}\begin{pmatrix} 1, & \beta, & \cdots, & \beta^{m_1 - 1}, & 0, & \cdots, & 0 \end{pmatrix}^T = L \subseteq \mathbf{C}^m.$$

On the other hand, $L \subseteq M_k$, since $\beta = \alpha_k$, and $L \subseteq N_l$, since $\beta = \zeta_l$. Therefore $(0) \neq L \subseteq M_k \cap N_l$, so $k + l > m$. The claim is proved.

To obtain restrictions on the positions of the eigenvalues equal to γ, the generalized eigenspaces $N_\gamma(A)$ and $N_\gamma(Z)$ are studied. Note that $N_\gamma(A) = N_\gamma(A_1) \oplus \mathbf{C}^{m_2} = \mathbf{C}^{m_2}$, and $N_\gamma(Z) = N_\gamma(Z_1) \oplus \mathbf{C}^{m_2}$; so in particular, $N_\gamma(A) \subseteq N_\gamma(Z)$. Note that $\dim\, N_\gamma(Z_1) = q$ is a strictly positive integer and that $\dim\, N_\gamma(Z) = q + m_2$. Define

$$\tau(i) = \min\left\{k \mid 1 \leq k \leq m, \dim\left(M_k \cap N_\gamma(A)\right) = i\right\}, \quad i = 1, \ldots, m_2,$$

$$\sigma(j) = \min\left\{l \mid 1 \leq l \leq m, \dim\left(N_l \cap N_\gamma(Z)\right) = j\right\}, \quad j = 1, \ldots, q + m_2.$$

Claim 2 It is immediate from Lemma 3.1.3 that

$$\alpha_{\tau(i)} = \gamma, \quad i = 1, \ldots, m_2,$$

and

$$\zeta_{\sigma(j)} = \gamma, \quad j = 1, \ldots, q + m_2. \tag{3.8}$$

Claim 3 If $i + j > q + m_2$, then $\tau(i) + \sigma(j) > m$.

To prove the claim, assume that $\tau(i) + \sigma(j) \leq m$. Then $M_{\tau(i)} \cap N_{\sigma(j)} = (0)$, and hence

$$\left(M_{\tau(i)} \cap N_\gamma(A)\right) \oplus \left(N_{\sigma(j)} \cap N_\gamma(Z)\right) \subseteq N_\gamma(A) + N_\gamma(Z) = N_\gamma(Z).$$

A dimension argument shows that $i + j \leq q + m_2$ and the claim is proved.

As a consequence of Claim 3, the following inequalities hold:

$$\tau(i) + \sigma(q + m_2 - i + 1) > m, \quad i = 1, \ldots, m_2. \tag{3.9}$$

It will be convenient to use the following notation:

$$\hat{\zeta}_l = \zeta_{m-l+1}, \quad l = 1, \dots, m, \tag{3.10}$$

and

$$\hat{\sigma}(j) = m - \sigma(q + m_2 - j + 1) + 1, \quad j = 1, \dots, q + m_2. \tag{3.11}$$

The expressions (3.8) and (3.9) are rewritten according to (3.10) and (3.11) as follows:

$$\hat{\zeta}_{\hat{\sigma}(j)} = \gamma, \quad j = 1, \dots, q + m_2, \tag{3.12}$$

and

$$\hat{\sigma}(i) \leq \tau(i), \quad i = 1, \dots, m_2. \tag{3.13}$$

Define the strictly increasing mappings $\pi, \hat{\rho} : \{1, \dots, m_1\} \longrightarrow \{1, \dots, m\}$, such that $\pi(s) \neq \tau(i)$ and $\hat{\rho}(s) \neq \hat{\sigma}(i)$ for all $s = 1, \dots, m_1$ and $i = 1, \dots, m_2$. By Lemma 3.5.2, inequality (3.13) implies that $\pi \leq \hat{\rho}$. Define the strictly increasing mapping $\rho : \{1, \dots, m_1\} \longrightarrow \{1, \dots, m\}$ by

$$\rho(t) = m - \hat{\rho}(m_1 - t + 1) + 1, \quad t = 1, \dots, m_1.$$

This equation and $\pi \leq \hat{\rho}$ together imply that

$$\pi(s) + \rho(m_1 - s + 1) \leq m + 1, \quad s = 1, \dots, m_1. \tag{3.14}$$

Note that the vectors in (3.6) are indeed spectral vectors for A_1 and Z_1, as they are obtained from the spectral vectors for A and Z by omitting m_2 eigenvalues γ. (Consider Claim 2 and the definition of π and ρ.) To prove that the spectral vectors in (3.6) satisfy the ordering condition (3.7), assume that $\alpha_{\pi(s)} = \zeta_{\rho(t)} = \beta$. Then $\beta \neq \gamma$, since $\gamma \notin \sigma(A_1)$. It follows from Claim 1 that $\pi(s)+\rho(t) > m$. Then (3.14) implies that $\rho(m_1 - s+1) \leq m+1-\pi(s) \leq \rho(t)$, so $m_1 - s+1 \leq t$ (ρ is strictly increasing) or $s+t > m_1$. Part 2 of the theorem is proved.

Part 3: $\gamma \in \sigma(A_1) \cap \sigma(Z_1)$.

Assume that the pair A, Z admits simultaneous reduction to complementary triangular with diagonals α and ζ as in (3.5). Notation will be consistent with Part 2 of the proof, unless explicitly stated otherwise.

As in Part 2, the main course of the proof of Part 3 will be as follows: Restrictions on the diagonals α and ζ, based on the matching condition on the maximal invariant nests $\mathcal{M}$ and $\mathcal{N}$ for A and Z respectively, are used to define strictly increasing mappings $\pi, \rho : \{1, \ldots, m_1\} \longrightarrow \{1, \ldots, m\}$, such that the vectors (3.6) are spectral vectors for A_1 and Z_1, and such that condition (3.7) is satisfied.

First of all, Claim 1 in Part 2 of the proof remains valid. The restrictions on the positions of the eigenvalues equal to γ on the diagonals α and ζ need more attention. Note that both $p = \dim N_\gamma(A_1)$ and $q = \dim N_\gamma(Z_1)$ are strictly positive integers and $\dim N_\gamma(A) = p + m_2$, $\dim N_\gamma(Z) = q + m_2$. By symmetry, we may assume without loss of generality that $p \le q$. Since A_1 and Z_1 are first companion matrices, it follows that $N_\gamma(A) \subseteq N_\gamma(Z)$.

Define

$$\tau(i) = \min\left\{k \mid 1 \le k \le m, \dim\left(M_k \cap N_\gamma(A)\right) = i\right\}, \quad i = 1, \ldots, p + m_2,$$

$$\sigma(j) = \min\left\{l \mid 1 \le l \le m, \dim\left(N_l \cap N_\gamma(Z)\right) = j\right\}, \quad j = 1, \ldots, q + m_2.$$

Claim 4 The following is an immediate consequence of Lemma 3.1.3.

$$\alpha_{\tau(i)} = \gamma, \quad i = 1, \ldots, p + m_2,$$

and

$$\zeta_{\sigma(j)} = \gamma, \quad j = 1, \ldots, q + m_2.$$

Since A_1 and Z_1 are first companion matrices, $L = \mathrm{Ker}(\gamma - A_1) = \mathrm{Ker}(\gamma - Z_1) = \mathrm{span}\{x_1(\gamma)\}$. Define the integers

$$\tau_* = \min\left\{k \mid 1 \le k \le m, L \subseteq M_k\right\}$$

and

$$\sigma_* = \min\left\{l \mid 1 \le l \le m\}, L \subseteq N_l\right\}.$$

There exist $i_* \in \{1, \ldots, p + m_2\}$, such that $\tau_* = \tau(i_*)$. Indeed, by definition of τ_*, we get $x_1(\gamma) \notin M_{\tau_*-1}$ and $x_1(\gamma) \in M_{\tau_*}$. Therefore, $M_{\tau_*} = M_{\tau_*-1} \oplus \mathrm{span}\{x_1(\gamma)\}$. It follows that $(A - \gamma)M_{\tau_*} \subseteq M_{\tau_*-1}$, so $\alpha_{\tau_*} = \gamma$. This gives $\tau_* \in \{\tau(1), \ldots, \tau(p + m_2)\}$.

Similarly, it is shown that there exists $j_* \in \{1, \ldots, q + m_2\}$, such that $\sigma_* = \sigma(j_*)$.

Apply Lemma 3.5.3 to $B_1 = A_1 - \gamma$, $B = A - \gamma$, $M = M_{\tau(i_*)} \cap N_\gamma(A)$, and to $B_1 = Z_1 - \gamma$, $B = Z - \gamma$, $M = N_{\sigma(j_*)} \cap N_\gamma(Z)$, to obtain

$$M_{\tau(i_*)} \cap N_\gamma(A), N_{\sigma(j_*)} \cap N_\gamma(Z) \subseteq L \oplus \mathbf{C}^{m_2}, \tag{3.15}$$

and by a dimension argument, $i_*, j_* \leq m_2 + 1$. In addition, $L \subseteq M_{\tau(i_*)} \cap N_{\sigma(j_*)}$ implies that

$$\tau(i_*) + \sigma(j_*) > m. \tag{3.16}$$

Claim 5 If $j \leq j_*$ and $i + j > p + m_2$, then

$$\tau(i) + \sigma(j) > m.$$

The claim is proved as follows: Assume that $j \leq j_*$, then

$$N_{\sigma(j)} \cap N_\gamma(Z) \subseteq N_{\sigma(j_*)} \cap N_\gamma(Z) \subseteq L \oplus \mathbf{C}^{m_2} \subseteq N_\gamma(A).$$

If in addition, $\tau(i) + \sigma(j) \leq m$, it follows that

$$\Big(M_{\tau(i)} \cap N_\gamma(A)\Big) \oplus \Big(N_{\sigma(j)} \cap N_\gamma(Z)\Big) \subseteq N_\gamma(A),$$

and a dimension argument provides $i + j \leq p + m_2$. The claim is proved.

In particular, Claim 5 implies that

$$\tau(p + m_2 - j + 1) + \sigma(j) > m, \quad j = 1, \ldots, j_*. \tag{3.17}$$

Claim 6 If $i \leq i_*$, $j \leq j_*$ and $i + j > m_2 + 1$, then

$$\tau(i) + \sigma(j) > m.$$

To prove the claim, let $i \leq i_*$ and $j \leq j_*$. If we assume that $\tau(i) + \sigma(j) \leq m$, then

$$\Big(M_{\tau(i)} \cap N_\gamma(A)\Big) \oplus \Big(N_{\sigma(j)} \cap N_\gamma(Z)\Big) \subseteq$$

$$\Big(M_{\tau(i_*)} \cap N_\gamma(A)\Big) + \Big(N_{\sigma(j_*)} \cap N_\gamma(Z)\Big) \subseteq L \oplus \mathbf{C}^{m_2},$$

and a dimension argument provides $i + j \leq m_2 + 1$. The claim is proved.

Claim 7 There exists a pair κ,λ such that

$$\kappa \in \{0, \ldots, i_* - 1\}, \quad \lambda \in \{0, \ldots, j_* - 1\}, \tag{3.18}$$

$$\kappa + \lambda \leq m_2, \tag{3.19}$$

$$\tau(\kappa + 1) + \sigma(\lambda + 1) > m. \tag{3.20}$$

To prove the claim, we consider two cases.

Case 1 $i_* + j_* \leq m_2 + 2$

In this case, put $\kappa = i_* - 1$ and $\lambda = j_* - 1$. Then it is immediate that (3.18) and (3.19) are satisfied. Further (3.16) implies (3.20).

Case 2 $i_* + j_* > m_2 + 2$

Define the integer $d = i_* + j_* - m_2 > 2$. In this case, define $\kappa = i_* - 1$ and $\lambda = j_* - d + 1$. Then it is easily verified that (3.18) and (3.19) are satisfied. Since $(\kappa + 1) + (\lambda + 1) = m_2 + 2$, Claim 6 implies (3.20). The claim is proved.

In the proof of Claim 7, the integers κ and λ were defined as follows (put $d = \max\{i_* + j_* - m_2, 2\}$)

$$\kappa = i_* - 1, \quad \lambda = j_* - d + 1.$$

In general, there may exist other pairs of integers κ, λ, which also satisfy the conditions of Claim 7. For the proof, it suffices to consider only this pair of integers.

Note that Claim 3 in Part 2 remains valid, and that we also obtain (3.9). As in Part 2, the notation introduced by the equations (3.10) and (3.11) is used to rewrite (3.8) and (3.9) as (3.12) and (3.13).

Further, equation (3.17) can be rewritten as

$$\hat{\sigma}(q + m_2 - j + 1) \leq \tau(p + m_2 - j + 1), \quad j = 1, \ldots, j_*, \tag{3.21}$$

and (3.20) as

$$\hat{\sigma}(q + m_2 - \lambda) \leq \tau(\kappa + 1). \tag{3.22}$$

Define the strictly increasing mappings $\tilde{\tau}, \tilde{\sigma} : \{1, \ldots, m_2\} \longrightarrow \{1, \ldots, m\}$ as follows:

$$\tilde{\tau}(i) = \begin{cases} \tau(i) & i = 1, \ldots, m_2 - \lambda \\ \tau(p+i) & i = m_2 - \lambda + 1, \ldots, m_2 \end{cases}, \tag{3.23}$$

and

$$\tilde{\sigma}(j) = \begin{cases} \hat{\sigma}(j) & j = 1, \ldots, \kappa \\ \hat{\sigma}(q+j) & j = \kappa + 1, \ldots, m_2 \end{cases}. \tag{3.24}$$

Note that the p integers

$$\tau(m_2 - \lambda + 1), \ldots, \tau(p + m_2 - \lambda) \tag{3.25}$$

are not in the range of $\tilde{\tau}$, and that the q integers

$$\hat{\sigma}(\kappa + 1), \ldots, \hat{\sigma}(q + \kappa) \tag{3.26}$$

are not in the range of $\tilde{\sigma}$.

Claim 8 $\tilde{\sigma} \leq \tilde{\tau}$

First, let $1 \leq i \leq \kappa$. Since $\kappa \leq i_* - 1 \leq m_2$, (3.13) implies that $\tilde{\sigma}(i) = \hat{\sigma}(i) \leq \tau(i) = \tilde{\tau}(i)$.

Second, let $\kappa + 1 \leq i \leq m_2 - \lambda$. Then (3.22) implies $\tilde{\sigma}(i) = \hat{\sigma}(q + i) \leq \hat{\sigma}(q + m_2 - \lambda) \leq \tau(\kappa + 1) \leq \tau(i) = \tilde{\tau}(i)$.

Finally, let $m_2 - \lambda + 1 \leq i \leq m_2$. Then (3.21) implies $\tilde{\sigma}(i) = \hat{\sigma}(q + i) \leq \tau(p + i) = \tilde{\tau}(i)$, and the claim is proved.

Define the strictly increasing mappings $\pi, \tilde{\rho} : \{1, \ldots, m_1\} \longrightarrow \{1, \ldots, m\}$, such that $\pi(s) \neq \tilde{\tau}(i)$ and $\tilde{\rho}(s) \neq \tilde{\sigma}(i)$ for all $s = 1, \ldots, m_1$ and $i = 1, \ldots, m_2$. By Lemma 3.5.2, the inequality in Claim 8 implies that $\pi \leq \tilde{\rho}$. Define the strictly increasing mapping $\rho : \{1, \ldots, m_1\} \longrightarrow \{1, \ldots, m\}$ by

$$\rho(t) = m - \tilde{\rho}(m_1 - t + 1) + 1, \quad t = 1, \ldots, m_1.$$

This equation and $\pi \leq \tilde{\rho}$ imply (3.14).

The vectors in (3.6) are spectral vectors for A_1 and Z_1 respectively, for the same reason as described in Part 2 of the proof. The integers in (3.25) are in the range of π. They indicate the positions of the eigenvalue γ in the spectral vector of A_1 given in (3.6) as follows: If $\pi(s)$ is one of the integers (3.25), then $\alpha_{\pi(s)} = \gamma$. The integers (3.26) are in the range of $\tilde{\rho}$. For that reason, the q integers

$$\sigma(m_2 - \kappa + 1), \ldots, \sigma(q + m_2 - \kappa) \tag{3.27}$$

are in the range of ρ and indicate the positions of the eigenvalues equal to γ in the spectral vector of Z_1 in (3.6).

We need to prove that condition (3.7) is satisfied for the spectral vectors of A_1 and Z_1, as given in (3.6).

If $\alpha_{\pi(s)} = \zeta_{\rho(t)} \neq \gamma$, then by the same argument as given in Part 2, it follows that $s + t > m_2$.

If $\alpha_{\pi(s)} = \zeta_{\rho(t)} = \gamma$, since $\gamma \in \sigma(A_1) \cap \sigma(Z_1))$, then $\pi(s)$ is one of the integers (3.25) and $\rho(t)$ is one of the integers (3.27). Therefore,

$$\pi(s) + \rho(t) \geq \tau(m_2 - \lambda + 1) + \sigma(m_2 - \kappa + 1) \geq \tau(\kappa + 1) + \sigma(\lambda + 1) > m.$$

Apply (3.14) as in Part 2, to obtain that $s + t > m_1$. This finally finishes the proof of the theorem. □

3.6 First and Third Companion Matrices

We now come to pairs of matrices, consisting of a first companion matrix and a third companion matrix. By taking transposes, the result also deals with pairs of matrices consisting of a second and a fourth companion matrix. For these pairs of $m_1 \times m_1$ matrices A_1, Z_1, we get $\rho_0(A_1, Z_1) \in \{m_1, \infty\}$.

Theorem 3.6.1 *Let A_1 be a first companion $m_1 \times m_1$ matrix, and Z_1 be a third companion $m_1 \times m_1$ matrix. Define the matrices $A = A_1 \oplus O_{m_2}$ and $Z = Z_1 \oplus O_{m_2}$, and let $m = m_1 + m_2$. Assume that the pair A, Z admits simultaneous reduction to complementary triangular forms with diagonals α and ζ as in (3.5). Then there exist strictly increasing mappings $\pi, \rho : \{1, \ldots, m_1\} \longrightarrow \{1, \ldots, m\}$, such that the pair A_1, Z_1 admits simultaneous reduction to complementary triangular forms with diagonals as in (3.6).*

Before we prove Theorem 3.6.1, the following technical lemma is proved.

Lemma 3.6.2 *Let A_1 and Z_1 be $m_1 \times m_1$ matrices, a, z be complex numbers, and define the matrices $A = A_1 \oplus aI_{m_2}$ and $Z = Z_1 \oplus zI_{m_2}$. Let $m = m_1 + m_2$. Assume that the pair A, Z admits simultaneous reduction to complementary triangular forms with diagonals α and ζ as in (3.5). Then there exist strictly increasing mappings $\gamma, \delta : \{1, \ldots, m_2\} \longrightarrow \{1, \ldots, m\}$, such that*

$$\alpha_{\gamma(i)} = a, \quad \zeta_{\delta(i)} = z, \quad i = 1, \ldots, m_2, \tag{3.28}$$

and

$$\gamma(i) + \delta(m_2 - i + 1) > m, \quad i = 1, \ldots, m_2. \tag{3.29}$$

Proof Let $\mathcal{M} = \{M_k\}_{k=0}^m$ and $\mathcal{N} = \{N_l\}_{l=0}^m$ be matching maximal invariant nests for A and Z respectively, and $\alpha = \operatorname{diag}(A; \mathcal{M})$ and $\zeta = \operatorname{diag}(Z; \mathcal{N})$. Define

$$\gamma(i) = \min\left\{k \mid 1 \le k \le m, \dim\left(M_k \cap \mathbf{C}^{m_2}\right) = i\right\}, \quad i = 1, \ldots, m_2$$

and

$$\delta(j) = \min\left\{l \mid 1 \le l \le m, \dim\left(N_l \cap \mathbf{C}^{m_2}\right) = j\right\}, \quad j = 1, \ldots, m_2.$$

Then by Lemma 3.1.3, (3.28) follows. To prove (3.29), fix $i \in \{1, \ldots, m_2\}$. It holds that

$$\dim\left(M_{\gamma(i)} \cap N_{\delta(m_2-i+1)} \cap \mathbf{C}^{m_2}\right) \ge$$

$$\dim\left(M_{\gamma(i)} \cap \mathbf{C}^{m_2}\right) + \dim\left(N_{\delta(m_2-i+1)} \cap \mathbf{C}^{m_2}\right) - m_2 = 1,$$

which implies that $\gamma(i) + \delta(m_2 - i + 1) > m$. The lemma is proved. □

Proof of Theorem 3.6.1 Assume that the pair A, Z admits simultaneous reduction to complementary triangular forms with diagonals α and ζ as in (3.5). Denote the matching maximal invariant nests for A and Z respectively by $\mathcal{M} = \{M_k\}_{k=0}^m$ and $\mathcal{N} = \{N_l\}_{l=0}^m$. As in the proof of Theorem 3.5.1, the matching of the nests $\mathcal{M}$ and $\mathcal{N}$ forces conditions on the diagonals α and ζ. Accordingly, mappings $\pi, \rho : \{1, \ldots, m_1\} \longrightarrow \{1, \ldots, m\}$ will be defined such that the condition

$$\alpha_{\pi(s)}\zeta_{\rho(t)} \ne 1, \quad s + t \le m_1 \tag{3.30}$$

is satisfied. Theorem 2.2.3 then yields the result. We now proceed with the proof.

Claim If $\alpha_k\zeta_l = 1$, then $k + l > m$.

The claim is proved as follows: If $\alpha_k\zeta_l = 1$, then $\beta = \alpha_k \ne 0$, and $\zeta_l = 1/\beta$. In particular, $\beta \in \sigma(A_1)$, and $1/\beta \in \sigma(Z_1)$. Therefore,

$$\operatorname{Ker}(\beta - A_1) = \operatorname{Ker}(1/\beta - Z_1) = \operatorname{span}\{x_1(\beta)\},$$

since A_1 is a first companion matrix, and Z_1 is a third companion matrix. Since $\beta \ne 0$, we get

$$\operatorname{Ker}(\beta - A) = \operatorname{Ker}(1/\beta - Z) = \operatorname{span}\{(1\ \beta\ \cdots\ \beta^{m_1-1}\ 0\ \cdots\ 0\ )^T\}.$$

By Lemma 3.6.2, where we take $a = z = 0$, there exist strictly increasing mappings $\gamma, \delta : \{1, \ldots, m_2\} \longrightarrow \{1, \ldots, m\}$ which satisfy (3.28) and (3.29).

Let $\hat{\delta}(i) = m - \delta(m_2 - i + 1) + 1$ for $i = 1, \ldots, m_2$. By (3.29), we get $\hat{\delta} \leq \gamma$. Define strictly increasing mappings $\pi, \hat{\rho} : \{1, \ldots, m_1\} \longrightarrow \{1, \ldots, m\}$, such that $\hat{\delta}(i) \neq \hat{\rho}(s)$ and $\gamma(i) \neq \pi(s)$ for $i = 1, \ldots, m_2$ and $s = 1, \ldots, m_1$. Lemma 3.5.2 implies $\pi \leq \hat{\rho}$. Define $\rho(s) = m - \hat{\rho}(m_1 - s + 1) + 1$ for $s = 1, \ldots, m_1$. Obviously $\pi(s) \leq m - \rho(m_1 - s + 1) + 1$, so $\pi(s) + \rho(m_1 - s + 1) \leq m + 1$.

Assume that $\alpha_{\pi(s)}\zeta_{\rho(t)} = 1$. By the claim, $\pi(s) + \rho(t) > m$. We get $\pi(s) + \rho(t) \geq m + 1 \geq \pi(s) + \rho(m_1 - s + 1)$, so $\rho(t) \geq \rho(m_1 - s + 1)$. Since ρ is increasing, it follows that $t \geq m_1 - s + 1$, thus $s + t > m_1$. The theorem is proved. □

The Proof of Theorem 3.5.1 is more complicated than the proof of Theorem 3.6.1. This is due to the fact that the zero entries in the spectral vectors (3.5) do not interfere with the ordering condition (3.30). Indeed, if $\alpha\zeta = 1$, then $\alpha \neq 0$ and $\zeta \neq 0$.

In the last two sections, we have not dealt with all pairs of companion matrices. For example, we have not described complementary triangular forms after extension with zeroes for pair of matrices, consisting of a first and a second companion matrix. As it seems, results in this direction are very hard to obtain.

3.7 Almost Diagonalizable Matrices

We first turn to pairs of matrices, where one of the matrices is of rank one. We may restrict ourselves to non-diagonalizable rank one matrices, since Theorem 2.2.1 deals with pairs of matrices, containing a diagonalizable matrix. Recall that a rank one matrix is non-diagonalizable if and only if it is nilpotent.

Proposition 3.7.1 *Let A and Z be complex $m \times m$ matrices, and assume that Z is a nilpotent rank one operator. Then the pair A, Z admits simultaneous reduction to complementary triangular forms if and only if there exist non-trivial subspaces M, N, with*

$$M \oplus N = \mathbf{C}^m, \quad AM \subseteq M, \quad \operatorname{Ran} Z \subseteq N \subseteq \operatorname{Ker} Z. \tag{3.31}$$

Proof Assume there exist subspaces M, N, such that (3.31) holds. Then with respect to $M \oplus N = \mathbf{C}^m$,

$$A = \begin{pmatrix} A_1 & A_{12} \\ O & A_2 \end{pmatrix}, \qquad Z = \begin{pmatrix} O & O \\ Z_{21} & O \end{pmatrix},$$

so the pair A, Z admits simultaneous reduction to complementary triangular forms, by Lemma 2.3.4.

On the other hand, if $\mathcal{M} = \{M_k\}_{k=0}^m$ and $\mathcal{N} = \{N_l\}_{l=0}^m$ are matching maximal invariant nests for A and Z respectively, then define

$$n = \min\{k \mid k = 0, \ldots, m, \text{ Ran } Z \subseteq N_k\}.$$

Since $(0) \neq \text{Ran } Z = Z(N_m) \subseteq N_{m-1}$, it follows that $1 \leq n \leq m-1$. Therefore, $N = N_n$ is a non-trivial subspace. Furthermore, since Z is nilpotent rank one and by the minimality of n, $ZN \subseteq N_{n-1} \cap \text{Ran } Z = (0)$. Therefore, $\text{Ran } Z \subseteq N \subseteq \text{Ker } Z$. Take $M = M_{m-n}$ to complete the proof. □

Proposition 3.7.2 *Let A_1 and Z_1 be $m_1 \times m_1$ matrices, Z_1 nilpotent and of rank one. If the pair $A_1 \oplus O_{m_2}$, $Z_1 \oplus O_{m_2}$ admits simultaneous reduction to complementary triangular forms, then so does the pair A_1, Z_1.*

Proof A matrix is nilpotent and of rank one, if and only if its extension with zeroes has this property. Therefore, we may assume that $m_2 = 1$. Write $m = m_1 + 1$. If the pair $A = A_1 \oplus 0$, $Z = Z_1 \oplus 0$ admits simultaneous reduction to complementary triangular forms, then, by Proposition 3.7.1, there exist subspaces $M, N \subseteq \mathbf{C}^m$, such that (3.31) is satisfied. Let P denote the projection onto $\mathbf{C}^{m_1}$ along $\mathbf{C}^1$, and write $M_1 = M \cap \mathbf{C}^{m_1}$, $N_1 = N \cap \mathbf{C}^{m_1}$.

Then both M_1 and PM are invariant subspaces for A_1, both N_1 and PN are invariant subspaces for Z_1, and $\text{Ran } Z_1 \subseteq N_1 \subseteq \text{Ker } Z \cap \mathbf{C}^{m_1} = \text{Ker } Z_1$. Also, $\text{Ran } Z_1 \subseteq PN \subseteq P\text{Ker } Z = \text{Ker } Z_1$. It is now sufficient to prove that either M_1 or PM matches with either N_1 or PN, since Proposition 3.7.1 can then be applied to obtain that the pair A_1, Z_1 admits simultaneous reduction to complementary triangular forms.

Since $M \oplus N = \mathbf{C}^m$, it follows that $M_1 \cap N_1 = (0)$, and that $PM + PN = \mathbf{C}^{m_1}$. If either $M_1 + N_1 = \mathbf{C}^{m_1}$ or $PM \cap PN = (0)$, we are done. Therefore, assume that both $M_1 \oplus N_1 \neq \mathbf{C}^{m_1}$ and that $PM \cap PN \neq (0)$.

Note that $PM = (M + \mathbf{C}^1) \cap \mathbf{C}^{m_1}$, so $\dim(PM/M_1) \leq 1$. In the same fashion, we get $\dim(PN/N_1) \leq 1$. Consequently,

$$1 \leq \dim(PM \cap PN) = \dim PM + \dim PN - \dim(PM + PN) \leq$$

$$\dim PM + \dim PN - m_1 \leq \dim M_1 + \dim N_1 + 2 - m_1 \leq$$

$$m_1 - 1 + 2 - m_1 = 1.$$

The inequalities turn out to be equalities, and we may conclude that

$$\dim(M_1 + N_1) = m_1 - 1, \quad \dim(PM \cap PN) = 1,$$

$$\dim(PM/M_1) = \dim(PN/N_1) = 1.$$

We distinguish two cases, that need not be exclusive, but cover all situations.

Case 1 $\quad PM \cap PN \not\subseteq M_1$

In this case, $M_1 \cap PN = (0)$, and by a dimension argument, $M_1 \oplus PN = \mathbf{C}^{m_1}$.

Case 2 $\quad PM \cap PN \not\subseteq N_1$

In this case, $PM \cap N_1 = (0)$, and by a dimension argument, $PM \oplus N_1 = \mathbf{C}^{m_1}$.

The proposition is proved. □

The following theorem deals with simultaneous reduction to complementary triangular forms after extension with zeroes for pairs of matrices that contain an almost diagonalizable matrix. For such pairs of $m_1 \times m_1$ matrices A_1, Z_1 we also have that $\rho_0(A_1, Z_1) \in \{m_1, \infty\}$.

Theorem 3.7.3 *Let Z_1 be an almost diagonalizable $m_1 \times m_1$ matrix, and let A_1 be any $m_1 \times m_1$ matrix. Let m_2 be a nonnegative integer, and define $m = m_1 + m_2$. If the the pair of $m \times m$ matrices $A = A_1 \oplus O_{m_2}$, $Z = Z_1 \oplus O_{m_2}$ admits simultaneous reduction to complementary triangular forms, then so does the pair A_1, Z_1.*

Proof It is easy to see that $(A, Z) \in \mathcal{C}(m)$ implies that $(A, p_{Z_1}(Z)) \in \mathcal{C}(m)$, where $p_{Z_1}(\lambda)$ is a spectral polynomial, as defined in (2.7). Since Z_1 is almost diagonalizable, the matrix $p_{Z_1}(Z) = p_{Z_1}(Z_1) \oplus O_{m_2}$ is nilpotent and of rank one. Proposition 3.7.2 provides that $(A_1, p_{Z_1}(Z_1)) \in \mathcal{C}(m_1)$. By Proposition 3.7.1, there exist $M, N \subseteq \mathbf{C}^{m_1}$, with $M \oplus N = \mathbf{C}^{m_1}$, and such that with respect to this decomposition,

$$A_1 = \begin{pmatrix} B_1 & B_{12} \\ O & B_2 \end{pmatrix}, \quad p_{Z_1}(Z_1) = \begin{pmatrix} O_n & O \\ Y_{21} & O_{m_1-n} \end{pmatrix},$$

where dim $M = n$. Recall that Ker $p_{Z_1}(Z_1)$ coincides with the span of all eigenvectors of Z_1. Since $p_{Z_1}(Z_1)$ is of rank one, the subspace Ran $p_{Z_1}(Z_1)$ is the span of one of these eigenvectors. Further,

$$\text{Ran } p_{Z_1}(Z_1) \subseteq N \subseteq \text{Ker } p_{Z_1}(Z_1),$$

so there exist $m_1 - n$ eigenvectors $\phi_1, \ldots, \phi_{m_1-n}$ of Z_1, such that Ran $p_{Z_1}(Z_1) = \text{span}\{\phi_1\}$, and

$$M \oplus \text{span}\{\phi_1, \ldots, \phi_{m_1-n}\} = \mathbf{C}^{m_1}.$$

Write $\hat{N} = \text{span}\{\phi_1, \ldots, \phi_{m_1-n}\}$, then with respect to the decomposition $M \oplus \hat{N} = \mathbf{C}^{m_1}$, we get

$$A_1 = \begin{pmatrix} \hat{B}_1 & \hat{B}_{12} \\ O & \hat{B}_2 \end{pmatrix}, \quad Z_1 = \begin{pmatrix} \hat{Z}_1 & O \\ \hat{Z}_{21} & \hat{Z}_2 \end{pmatrix}, \quad p_{Z_1}(Z_1) = \begin{pmatrix} \hat{O}_n & O \\ \hat{Y}_{21} & O_{m_1-n} \end{pmatrix}.$$

Since $p_{Z_1}(\hat{Z}_1) = O_n$, and $p_{Z_1}(\hat{Z}_2) = O_{m_1-n}$, it follows that $\hat{Z}_1$ and $\hat{Z}_2$ are diagonalizable. Lemma 2.3.4 now implies that $(A_1, Z_1) \in \mathcal{C}(m_1)$, and the theorem is complete. □

The following Theorem deals with pairs of matrices of low order.

Theorem 3.7.4 *Let $m_1 \leq 3$, and let A_1 and Z_1 be $m_1 \times m_1$ matrices. Let m_2 be a nonnegative integer, $m = m_1 + m_2$, and define the $m \times m$ matrices $A = A_1 \oplus O_{m_2}$, $Z = Z_1 \oplus O_{m_2}$. If the pair A, Z admits simultaneous reduction to complementary triangular forms, then so does the pair A_1, Z_1.*

Proof The case when $m_1 = 1$ is trivial. The case when $m_1 = 2$ follows from Theorem 2.2.1 and Theorem 3.7.3. Indeed, if the pair of matrices A_1, Z_1 contains a diagonalizable matrix, then Theorem 2.2.1 provides the result, and if both A_1 and Z_1 are non-diagonalizable, then they are almost diagonalizable, and Theorem 3.7.3 can be used.

The case when $m_1 = 3$ is dealt with as follows: If the pair A_1, Z_1 contains a diagonalizable matrix, then apply Theorem 2.2.1. If the pair A_1, Z_1 contains an almost diagonalizable matrix, use Theorem 3.7.3. The only case, that has not been dealt with, is the case when both A_1 and Z_1 are unicellular. In that case, apply Theorem 3.4.1. The theorem is proved. □

3.8 Jordan Matrices

A *strictly upper triangular matrix* is an upper triangular matrix with zero diagonal, i.e., a nilpotent upper triangular matrix. Recall that a nonderogatory nilpotent matrix is unicellular.

Lemma 3.8.1 *Let $2 \le \mu, \nu \le m_1$ be integers, A_1 a nonderogatory strictly upper triangular $\mu \times \mu$ matrix, Z_1 a nonderogatory strictly upper triangular $\nu \times \nu$ matrix. Write $m = m_1 + m_2$ and consider the $m \times m$ matrices*

$$A = \begin{pmatrix} A_1 & O & O \\ O & O_{m_1-\mu} & O \\ O & O & O_{m_2} \end{pmatrix}, \qquad Z = \begin{pmatrix} O_{m_1-\nu} & O & O \\ O & Z_1 & O \\ O & O & O_{m_2} \end{pmatrix}.$$

Then the pair A, Z admits simultaneous reduction to complementary triangular forms if and only if $\mu + \nu \le m_1 + 1$.

The case when A (or Z) is a zero matrix, corresponding to the case when $\mu = 1$ (or $\nu = 1$), is excluded in Lemma 3.8.1.

Proof To prove the only if part, note that $\mu + \nu > m_1 + 1$ implies that $\mathrm{Ker}\ Z \subseteq \mathrm{Ker}\ A + \mathrm{Ran}\ A$, and that $\mathrm{Ker}\ Z \cap \mathrm{Ran}\ Z \subseteq \mathrm{Ran}\ A$. By Theorem 2.3.1, the pair A, Z does not admit simultaneous reduction to complementary triangular forms.

To prove the if part, assume that $\mu+\nu \le m_1+1$. Then $\mathrm{Ker}\ A+\mathrm{Ker}\ Z = \mathbf{C}^m$, and Proposition 2.3.3 implies that the pair A, Z admits simultaneous reduction to complementary triangular forms. The lemma is proved. □

Consider the following special case of Lemma 3.8.1.

Corollary 3.8.2 *Let $m = m_1 + m_2$ and $2 \le \mu, \nu \le m_1$. The pair of $m \times m$ matrices*

$$\begin{pmatrix} J(0,\mu) & O & O \\ O & O_{m_1-\mu} & O \\ O & O & O_{m_2} \end{pmatrix}, \qquad \begin{pmatrix} O_{m_1-\nu} & O & O \\ O & J(0,\nu) & O \\ O & O & O_{m_2} \end{pmatrix}.$$

admits simultaneous reduction to complementary triangular forms if and only if $\mu + \nu \le m_1 + 1$.

Note that the condition $\mu + \nu \leq m_1 + 1$ in Corollary 3.8.2 is equivalent to the statement, that the Jordan blocks $J(0, \mu)$ and $J(0, \nu)$ have an overlap on the diagonal on at most one position. If A is a square matrix, then Inv A denotes the lattice of invariant subspaces for A. The following two lemmas are quite straightforward and the proofs are left to the reader.

Lemma 3.8.3 *Let A,$\hat{A}$ and Z,$\hat{Z}$ be square matrices, such that* Inv $A \subseteq$ Inv $\hat{A}$ *and* Inv $Z \subseteq$ Inv $\hat{Z}$*. If the pair A, Z admits simultaneous reduction to complementary triangular forms, then so does the pair $\hat{A}$, $\hat{Z}$.*

Lemma 3.8.4 *Let A and Z be square matrices, and assume that* Inv $A \subseteq$ Inv Z*. Then the pair A, Z admits simultaneous reduction to complementary triangular forms if and only if Z is diagonalizable.*

The following lemma will be used, together with Lemma 3.8.4, in Proposition 3.8.6.

Lemma 3.8.5 *Let A_1 and Z_1 be strictly upper triangular $m_1 \times m_1$ matrices, with A_1 nonderogatory. Let $A = A_1 \oplus O_{m_2}$ and $Z = Z_1 \oplus O_{m_2}$. Then* Inv $A \subseteq$ Inv Z*.*

Proof Write $m = m_1 + m_2$. First of all, there exists an invertible $m_1 \times m_1$ matrix R_1, such that $R_1^{-1} A_1 R_1 = J(0, m_1)$. It turns out, by Proposition 2.1 in [8], that R_1 is upper triangular. Consequently, if $R = R_1 \oplus I_{m_2}$, then $R^{-1} A R = J(0, m_1) \oplus O_{m_2}$ and $R^{-1} Z R$ is strictly upper triangular. For this reason, we may assume without loss of generality, that $A_1 = J(0, m_1)$.

Let $M \in$ Inv A be a non-trivial subspace. We need to prove that $M \in$ Inv Z. Let $x \in M$. If $x \in \mathbf{C}^{m_2}$, then $Zx = 0 \in M$. Assume that $x \notin \mathbf{C}^{m_2}$. Let $e_1, \ldots, e_m$ denote the standard basis in $\mathbf{C}^m = \mathbf{C}^{m_1} \oplus \mathbf{C}^{m_2}$. Write $x = \sum_{j=1}^{s} \zeta_j e_j + \sum_{j=m_1+1}^{m} \zeta_j e_j$, with $s \in \{1, \ldots, m_1\}$ and $\zeta_s \neq 0$. If $s = 1$, then $Zx = 0 \in M$. If $2 \leq s \leq m_1$, then $A^k x = (J(0, m_1) \oplus O_{m_2})^k x = \sum_{j=1}^{s-k} \zeta_{j+k} e_j \in M$, for $1 \leq k \leq s - 1$. Since $\zeta_s \neq 0$, it follows that $\text{span}\{e_1, \ldots, e_{s-1}\} \subseteq M$. Since Z is strictly upper triangular, $Zx \in \text{span}\{e_1, \ldots, e_{s-1}\} \subseteq M$. Therefore, $ZM \subseteq M$ and the lemma is proved. □

Proposition 3.8.6 *Let A and Z be non-zero strictly upper triangular $m_1 \times m_1$ matrices, with A_1 nonderogatory. Let $A = A_1 \oplus O_{m_2}$ and $Z = Z_1 \oplus O_{m_2}$. Then the pair A, Z does not admit simultaneous reduction to complementary triangular forms.*

Proof Assume that the pair A, Z admits simultaneous reduction to complementary triangular forms. Also, Lemma 3.8.5 implies that Inv $A \subseteq$ Inv Z. Therefore, by Lemma 3.8.4, the nilpotent matrix Z is diagonalizable, and hence equal to the zero matrix. A contradiction has been obtained and the lemma is proved. □

The following proposition contains the only if part of a result in [13], Theorem 2.2.5 in this thesis, which deals with the case $m_2 = 0$.

Proposition 3.8.7 *Let J_α and J_ζ be nonderogatory $m_1 \times m_1$ Jordan matrices as in (2.6), and assume that there exist Jordanblocks in J_α and J_ζ with a diagonal overlap on more than one position. Then the pair $J_\alpha \oplus O_{m_2}$, $J_\zeta \oplus O_{m_2}$ does not admit simultaneous reduction to complementary triangular forms.*

Proof Let $1 \leq \rho \leq s$ and $1 \leq \sigma \leq t$, such that the Jordan blocks $J(\alpha_\rho, k_\rho)$ and $J(\zeta_\sigma, l_\sigma)$ have a diagonal overlap on more than one position. It is immediate that the Jordan blocks with overlap on more than one diagonal position have size $k_\rho, l_\sigma > 1$.

There exists a polynomial $p(\lambda)$, such that $p(J(\alpha_i, k_i)) = O_{k_i}$, if $i \neq \rho$, and $p(J(\alpha_\rho, k_\rho)) = J(0, k_\rho)$. Indeed, if a polynomial $p(\lambda)$ satisfies the interpolation conditions

$$p^{(\nu)}(\alpha_i) = 0, \quad \nu = 0, \ldots, k_i - 1, \quad i \neq \rho,$$

$$p(\alpha_\rho) = 0, \quad p^{(1)}(\alpha_\rho) = 1, \quad p^{(\nu)}(\alpha_\rho) = 0, \quad \nu = 2, \ldots, k_\rho - 1,$$

it has the desired properties. Similarly, there exists a polynomial $q(\lambda)$, such that $q(J(\zeta_j, l_j)) = O_{l_j}$, if $j \neq \sigma$, and $q(J(\zeta_\sigma, l_\sigma)) = J(0, l_\sigma)$. Then

$$p(J_\alpha) = O_{k_1} \oplus \cdots \oplus O_{k_{\rho-1}} \oplus J(0, k_\rho) \oplus O_{k_{\rho+1}} \oplus \cdots \oplus O_{k_s},$$

$$q(J_\zeta) = O_{l_1} \oplus \cdots \oplus O_{l_{\sigma-1}} \oplus J(0, l_\sigma) \oplus O_{l_{\sigma+1}} \oplus \cdots \oplus O_{l_t}.$$

We may assume that $\sum_{i=1}^{\rho-1} k_i \leq \sum_{j=1}^{\sigma-1} l_j$, by interchanging the roles of J_α and J_ζ, if necessary. Further, $J(\alpha_\rho, k_\rho)$ and $J(\zeta_\sigma, l_\sigma)$ have an overlap on more than one position implies that $\sum_{i=1}^{\rho} k_i > \sum_{j=1}^{\sigma-1} l_j + 1$. There are two cases to distinguish.

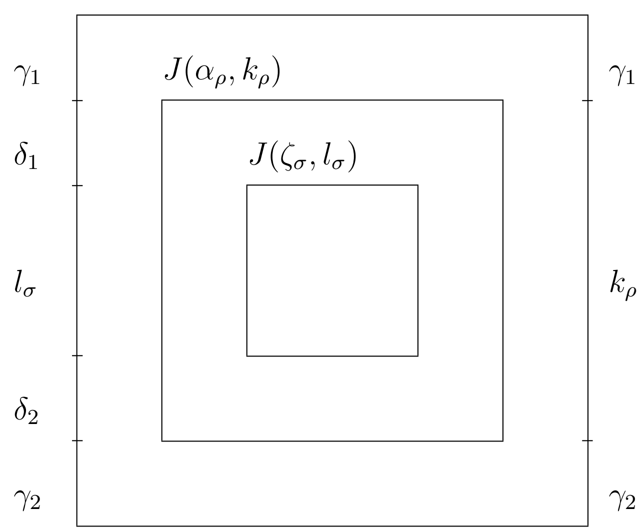
Case 1: $\sum_{i=1}^{\rho} k_i \geq \sum_{j=1}^{\sigma} l_j$, hence $\sum_{i=1}^{\rho-1} k_i \leq \sum_{j=1}^{\sigma-1} l_j < \sum_{j=1}^{\sigma} l_j \leq \sum_{i=1}^{\rho} k_i$.

In this case, the Jordan block $J(\alpha_\rho, k_\rho)$ covers the Jordan block $J(\zeta_\sigma, l_\sigma)$. Let $m = m_1 + m_2$ and consider the nonnegative integers

$$\gamma_1 = \sum_{i=1}^{\rho-1} k_i, \qquad \gamma_2 = m - \sum_{i=1}^{\rho} k_i,$$

$$\delta_1 = \sum_{j=1}^{\sigma-1} l_j - \gamma_1, \qquad \delta_2 = m - \gamma_2 - \sum_{j=1}^{\sigma} l_j.$$

The following picture illustrates the position of the Jordan blocks $J(\alpha_\rho, k_\rho)$ and $J(\zeta_\sigma, l_\sigma)$ relative to each other. The integer values $\gamma_1, \gamma_2, \delta_1, \delta_2$ and k_ρ, l_σ determine the measure and position of these Jordan blocks.



It follows that

$$p(J_\alpha \oplus O_{m_2}) = O_{\gamma_1} \oplus J(0, k_\rho) \oplus O_{\gamma_2},$$

$$q(J_\zeta \oplus O_{m_2}) = O_{\gamma_1} \oplus (O_{\delta_1} \oplus J(0, l_\sigma) \oplus O_{\delta_2}) \oplus O_{\gamma_2}.$$

Define the $m \times m$ permutation matrix

$$\Pi = \begin{pmatrix} O & I_{\gamma_1} & O \\ I_{k_\rho} & O & O \\ O & O & I_{\gamma_2} \end{pmatrix},$$

then

$$\Pi^{-1}p(J_\alpha \oplus O_{m_2})\Pi = J(0, k_\rho) \oplus O_{\gamma_1+\gamma_2},$$

and

$$\Pi^{-1}q(J_\zeta \oplus O_{m_2})\Pi = (O_{\delta_1} \oplus J(0, l_\sigma) \oplus O_{\delta_2}) \oplus O_{\gamma_1+\gamma_2}.$$

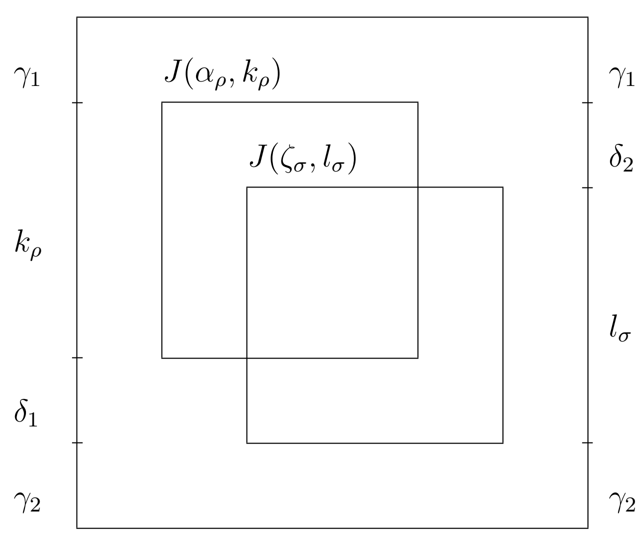
We may now apply Proposition 3.8.6 to the matrices $A_1 = J(0, k_\rho)$ and $Z_1 = O_{\delta_1} \oplus J(0, l_\sigma) \oplus O_{\delta_2}$, to obtain that the pair $A_1 \oplus O_{\gamma_1+\gamma_2}$, $Z_1 \oplus O_{\gamma_1+\gamma_2}$ does not admit simultaneous reduction to complementary triangular forms. Since $A_1 \oplus O_{\gamma_1+\gamma_2} = \Pi^{-1}p(J_\alpha \oplus O_{m_2})\Pi$, and $Z_1 \oplus O_{\gamma_1+\gamma_2} = \Pi^{-1}q(J_\zeta \oplus O_{m_2})\Pi$, we obtain that also the pair $J_\alpha \oplus O_{m_2}$, $J_\zeta \oplus O_{m_2}$ does not admit simultaneous reduction to complementary triangular forms.

Case 2: $\sum_{i=1}^{\rho} k_i < \sum_{j=1}^{\sigma} l_j$, hence $\sum_{i=1}^{\rho-1} k_i \leq \sum_{j=1}^{\sigma-1} l_j < \sum_{i=1}^{\rho} k_i < \sum_{j=1}^{\sigma} l_j$.

In this case, The Jordanblocks $J(\alpha_\rho, k_\rho)$ and $J(\zeta_\sigma, l_\sigma)$ have an overlap on more than one position, while $J(\zeta_\sigma, l_\sigma)$ is not covered by $J(\alpha_\rho, k_\rho)$. Define the nonnegative integers

$$\gamma_1 = \sum_{i=1}^{\rho-1} k_i, \quad \gamma_2 = m - \sum_{j=1}^{\sigma} l_j,$$

$$\delta_1 = m - \gamma_1 - \gamma_2 - k_\rho, \quad \delta_2 = m - \gamma_1 - \gamma_2 - l_\sigma.$$



The preceding picture illustrates the position of the Jordan blocks $J(\alpha_\rho, k_\rho)$ and $J(\zeta_\sigma, l_\sigma)$ relative to each other. The integer values $\gamma_1, \gamma_2, \delta_1, \delta_2$ and k_ρ, l_σ determine the measure and position of these Jordan blocks. Then

$$p(J_\alpha \oplus O_{m_2}) = O_{\gamma_1} \oplus (J(0, k_\rho) \oplus O_{\delta_1}) \oplus O_{\gamma_2},$$

$$q(J_\zeta \oplus O_{m_2}) = O_{\gamma_1} \oplus (O_{\delta_2} \oplus J(0, l_\sigma)) \oplus O_{\gamma_2}.$$

Defining the $m \times m$ permutation matrix ($n = m - \gamma_1 - \gamma_2$)

$$\Pi = \begin{pmatrix} O & I_{\gamma_1} & O \\ I_n & O & O \\ O & O & I_{\gamma_2} \end{pmatrix},$$

we get

$$\Pi^{-1} p(J_\alpha \oplus O_{m_2})\Pi = (J(0, k_\rho) \oplus O_{\delta_1}) \oplus O_{\gamma_1+\gamma_2},$$

and

$$\Pi^{-1} q(J_\zeta \oplus O_{m_2})\Pi = (O_{\delta_2} \oplus J(0, l_\sigma)) \oplus O_{\gamma_1+\gamma_2}.$$

We may now apply Corollary 3.8.2 to the matrices

$$\begin{pmatrix} J(0, k_\rho) & O & O \\ O & O_{\delta_1} & O \\ O & O & O_{\gamma_1+\gamma_2} \end{pmatrix}, \qquad \begin{pmatrix} O_{\delta_2} & O & O \\ O & J(0, l_\sigma) & O \\ O & O & O_{\gamma_1+\gamma_2} \end{pmatrix}.$$

and obtain that this pair of matrices does not admit simultaneous reduction to complementary triangular forms: Indeed, the overlap on more than one diagonal position is described by $\sum_{i=1}^{\rho} k_i > \sum_{j=1}^{\sigma-1} l_j + 1$, or $\gamma_1 + k_\rho > m - \gamma_2 - l_\sigma + 1$, or $k_\rho + l_\sigma > n + 1$, and we can apply Corollary 3.8.2. Since $J(0, k_\rho) \oplus O_{\delta_1} \oplus O_{\gamma_1+\gamma_2} = \Pi^{-1} p(J_\alpha \oplus O_{m_2})\Pi$, and $O_{\delta_2} \oplus J(0, l_\sigma) \oplus O_{\gamma_1+\gamma_2} = \Pi^{-1} q(J_\zeta \oplus O_{m_2})\Pi$, we obtain that also the pair $J_\alpha \oplus O_{m_2}$, $J_\zeta \oplus O_{m_2}$ does not admit simultaneous reduction to complementary triangular forms. The theorem is proved. □

The following theorem concerns simultaneous reduction to complementary triangular forms after extension with zeroes for pairs of nonderogatory Jordan matrices. For such pairs of $m_1 \times m_1$ matrices A_1, Z_1 it holds that $\rho_0(A_1, Z_1) \in \{m_1, \infty\}$.

Theorem 3.8.8 *Let A_1 and Z_1 be nonderogatory $m_1 \times m_1$ Jordan matrices, and define the matrices $A = A_1 \oplus O_{m_2}$ and $Z = Z_1 \oplus O_{m_2}$. If the pair A, Z admits simultaneous reduction to complementary triangular forms, then so does the pair A_1, Z_1.*

Proof If the pair A, Z admits simultaneous reduction to complementary triangular forms, then, by Proposition 3.8.7, the Jordan matrices A_1 and Z_1 have no Jordanblocks with diagonal overlap on more than one position. By Theorem 2.2.5, this implies that the pair A_1, Z_1 admits simultaneous reduction to complementary triangular forms. The theorem is proved. □

Part II

Bounded Operators

Chapter 4

Nests of Subspaces and Projections

In the first part of the thesis, Lemma 3.1.4 characterizes complementary triangular forms for pairs of finite matrices in terms of nests of subspaces and projections. In the literature, nests of subspaces have been used to define triangular forms for bounded operators on an infinite dimensional Hilbert or Banach space; see [20], [30] and [47].

In this Chapter, we introduce the notion of a simple nest of *projections*. We shall use this notion in Chapters 5 and 6 to study complementary triangular forms for pairs of bounded operators on an infinite dimensional Hilbert or Banach space.

Simple nests of *subspaces* were discussed in [47], and it was shown there, that these nest are exactly the maximal ones (Lemma 4.1.1). For nests of projections, the situation is more complicated, and this chapter discusses these matters in detail.

In Section 4.1, simple nests of projections are defined, and some well-known examples of simple nests of projections are discussed.

In Section 4.2, the strong operator topology closure of nests of projections is considered.

Section 4.3 studies simple nests of projections versus maximal nests of projections in the case when the underlying Banach space is reflexive.

4.1 Simple Nests

This section defines simple nests of projections, and characterizes simple nests of projections in terms of their nests of ranges and kernels (Theorem 4.1.3).

At the end of this section, examples of simple nests are given.

A linear manifold M in a Banach space X is called a *subspace* , if it is closed in the norm topology of X. If $U \subseteq X$ is a subset, let *cl* U denotes its closure in the norm topology of X and let span U denotes its closed linear span in X with respect to the same topology. If $\{x_n\}$ is a sequence in X that converges to $x \in X$ in norm, we write $x_n \overset{\|\cdot\|}{\to} x$. Further, if $\{T_n\}$ is a sequence of bounded operators, then uniform convergence of the sequence to the bounded operator T is denoted by $T_n \overset{\|\cdot\|}{\to} T$. Convergence with respect to the strong operator topology, i.e., pointwise convergence in norm, is denoted by $T_n \overset{s}{\to} T$.

The collection of subspaces in X is partially ordered by inclusion. To fix notation, "$\subseteq$" denotes inclusion, where equality may hold, and "$\subset$" denotes proper inclusion, i.e., where equality does not hold. Let $\mathcal{S}$ be a non-empty collection of subspaces in X. A *lower bound* for $\mathcal{S}$ is a subspace L, such that $L \subseteq M$ for all $M \in \mathcal{S}$. For a non-empty collection of subspaces $\mathcal{S}$, the *greatest lower bound* or *infimum* always exists and is given by $\bigcap\{M \mid M \in \mathcal{S}\}$. We will denote this subspace by $\bigwedge \mathcal{S}$. An *upper bound* of the collection $\mathcal{S}$ is a subspace N, such that $M \subseteq N$ for all $M \in \mathcal{S}$. The *least upper bound* or *supremum* of $\mathcal{S}$ always exists, and is given by span$\{M \mid M \in \mathcal{S}\}$, the closed linear hull of all subspaces in $\mathcal{S}$. This subspace will be denoted by $\bigvee \mathcal{S}$.

A collection of subspaces $\mathcal{M}$ is called a *nest* of subspaces , if it is linearly ordered, i.e., if $M_1, M_2 \in \mathcal{M}$, then either $M_1 \subseteq M_2$ or $M_2 \subseteq M_1$. A subset of a nest is called a *subnest*. Let $\mathcal{M}$ be a nest of subspaces and $M_1, M_2 \in \mathcal{M}$ with $M_1 \subseteq M_2$. Define the interval

$$[M_1, M_2) = \{M \mid M \in \mathcal{M}, M_1 \subseteq M \subset M_2\}.$$

The intervals (M_1, M_2), $(M_1, M_2]$ and $[M_1, M_2]$ are defined in the same fashion. Note that these intervals also depend on the underlying nest $\mathcal{M}$. If the open interval $(M_1, M_2) = \emptyset$, then the pair of subspaces M_1, M_2 is called a *gap* in $\mathcal{M}$. The *dimension of the gap* is then defined as the dimension of the quotient space M_2/M_1.

A nest of subspaces is called *maximal*, if it is not properly contained in any other nest. In [47], Ringrose proposed the following definition: A nest of subspaces $\mathcal{M}$ is called *simple*, if it satisfies the following three conditions:

1. The trivial subspaces (0) and X are in $\mathcal{M}$.

2. If $\mathcal{M}_1$ is a non-empty subnest of $\mathcal{M}$, then the infimum $\bigwedge \mathcal{M}_1$ and supremum $\bigvee \mathcal{M}_1$ are in $\mathcal{M}$.

3. All gaps in $\mathcal{M}$ (if there are any) are one-dimensional.

If a nest of subspaces satisfies the first two conditions only, then it is called a *complete* nest. The following lemma is taken from [47] (Lemma 1, p.369).

Lemma 4.1.1 *A nest of subspaces is maximal if and only if it is simple.*

Note that by Zorn's lemma, maximal nests of subspaces in a Banach space do exist.

We now turn to nests of projections. A *projection* P acting on a Banach space X is a idempotent bounded linear operator (i.e., $P^2x = Px$ for all $x \in X$). One can define a partial ordering on the set of all projections on X as follows: Given two projections P_1, P_2, stipulate $P_1 \leq P_2$ if $P_1P_2 = P_2P_1 = P_1$, and $P_1 < P_2$ if in addition, $P_1 \neq P_2$. Note that $P_1 \leq P_2$ if and only if Ran $P_1 \subseteq$ Ran P_2 and Ker $P_2 \subseteq$ Ker P_1. A collection of projections $\mathcal{P}$ is called a *nest*, if it is linearly ordered, i.e., if $P_1, P_2 \in \mathcal{P}$, then either $P_1 \leq P_2$ or $P_2 \leq P_1$. An upper bound of a collection of projections and related notions are defined in a straightforward manner.

If P_1 and P_2 are two commuting projections, then define the *infimum* and *supremum* of P_1 and P_2 respectively as

$$P_1 \wedge P_2 = P_1P_2, \quad P_1 \vee P_2 = P_1 + P_2 - P_1P_2.$$

In this manner, we may define the infimum and supremum of any finite set of commuting projections. If $\mathcal{S}$ is a finite set of commuting projections, then the supremum P_0 of $\mathcal{S}$ satisfies

$$\text{Ran } P_0 = \text{span}\{\text{Ran } P \mid P \in \mathcal{S}\}, \quad \text{Ker } P_0 = \bigcap\{\text{Ker } P \mid P \in \mathcal{S}\}.$$

A nest of projections is a commuting set of projections. The following questions now arise: Does each nest of projections have a least upper bound? And if a nest of projections $\mathcal{P}$ has a least upper bound P_0, does

$$\text{Ran } P_0 = \text{span}\{\text{Ran } P \mid P \in \mathcal{P}\}, \quad \text{Ker } P_0 = \bigcap\{\text{Ker } P \mid P \in \mathcal{P}\}. \tag{4.1}$$

hold? Both questions are answered in the negative, as will be shown in Examples 4.3.5 and 4.3.7 respectively. For these reasons, we give the following definition. An upper bound P_0 for a nest of projections $\mathcal{P}$ is called a *strong supremum*, if (4.1) is satisfied. It is immediate, that a strong supremum of a nest of projections is also the least upper bound of the nest. If the strong supremum of a nest $\mathcal{P}$ exists, it is denoted by $\bigvee \mathcal{P}$. A projection Q_0 is called a *strong infimum* of the nest of projections $\mathcal{P}$, if

$$\text{Ran } Q_0 = \bigcap\{\text{Ran } P \mid P \in \mathcal{P}\}, \quad \text{Ker } Q_0 = \text{span}\{\text{Ker } P \mid P \in \mathcal{P}\}.$$

A strong infimum of a nest of projections is also the greatest lower bound for the nest. The strong infimum of a nest of projections $\mathcal{P}$, if it exists, will be denoted by $\bigwedge \mathcal{P}$. A projection Q is a lower bound of a nest of projections, if and only if $I - Q$ is an upper bound for the nest $\mathcal{P}^c = \{I - P \mid P \in \mathcal{P}\}$. In this fashion, results on least upper bounds and strong suprema have analogues for greatest lower bounds and strong infima respectively.

A nest of projections is *maximal*, if it is not properly contained in any other nest of projections. A nest of projections is called *bounded*, if it is a bounded set with respect to the operator norm. Proposition 4.2.3 below shows that an upper bound for a bounded nest of projections is a strong supremum if and only if it is in the strong operator closure of the nest. Example 4.2.2 gives a strong supremum of an unbounded nest, that is not in the strong operator closure of the nest. For these and other reasons, we will in general assume nests of projections to be bounded. Note that by Zorn's lemma, maximal nests of projections on a Banach space exist. The existence of a bounded maximal nest of projections is not assured by Zorn's lemma, however.

A pair of projections P_1, P_2 in the nest of projections $\mathcal{P}$ is a *gap* in $\mathcal{P}$, if $P_1 < P_2$ and

$$(P_1, P_2) = \{P \mid P \in \mathcal{P}, P_1 < P < P_2\} = \emptyset.$$

The *dimension of the gap* P_1, P_2 is defined as $\text{rank}(P_2 - P_1)$. A nest of projections $\mathcal{P}$ is a *simple* nest of projections , if it satisfies the following three conditions:

1. The trivial projections O, I are in $\mathcal{P}$.

2. For any non-empty subnest $\mathcal{P}_1 \subseteq \mathcal{P}$, the strong supremum $\bigvee \mathcal{P}_1$ and strong infimum $\bigwedge \mathcal{P}_1$ exist, and are in $\mathcal{P}$.

3. If P_1, P_2 is a gap in $\mathcal{P}$, then $\text{rank}(P_2 - P_1) = 1$.

A nest of projections, which satisfies the first two conditions only, is called a *complete* nest of projections. Having defined maximal and simple nests of projections, the question arises whether the two notions are equivalent, as for nests of subspaces (Lemma 4.1.1). The following lemma is a result in one direction.

Lemma 4.1.2 *A simple nest of projections on a Banach space X is maximal.*

Proof Assume that the nest of projections $\mathcal{P}$ is simple, but not maximal: There exists a projection $Q \notin \mathcal{P}$, such that $\mathcal{P} \cup \{Q\}$ is again a nest of projections. Define the sets

$$\mathcal{P}_+ = \{P \mid P \in \mathcal{P}, P > Q\}$$

and

$$\mathcal{P}_- = \{P \mid P \in \mathcal{P}, P < Q\}.$$

Then $\mathcal{P} = \mathcal{P}_+ \cup \mathcal{P}_-$ and $\mathcal{P}_+ \cap \mathcal{P}_- = \emptyset$. Note that $O \in \mathcal{P}_-$ and $I \in \mathcal{P}_+$. Define $P_+ = \bigwedge \mathcal{P}_+$ and $P_- = \bigvee \mathcal{P}_-$. Then $P_- \leq Q \leq P_+$. Since $P_-, P_+ \in \mathcal{P}$ and $Q \notin \mathcal{P}$, it follows that $P_- < Q < P_+$. Consequently, $\text{rank}(P_+ - P_-) > 1$. On the other hand, the pair P_-, P_+ is a gap in $\mathcal{P}$. A contradiction has been obtained and the lemma is proved. □

The converse of Lemma 4.1.2 does not hold in general. In Section 4.3, some counterexamples are given: See Examples 4.3.5 and 4.3.6. The next proposition characterizes a simple nest of projections in terms of its ranges and kernels.

Theorem 4.1.3 *A nest of projections $\mathcal{P}$ is simple if and only if the nests of subspaces*

$$\{\text{Ran } P \mid P \in \mathcal{P}\}, \quad \{\text{Ker } P \mid P \in \mathcal{P}\}$$

both are simple.

Proof To prove the only if part, assume that the nest of projections $\mathcal{P}$ is simple. We will prove here that the nest of subspaces $\mathcal{M} = \{\text{Ran } P \mid P \in \mathcal{P}\}$ is simple. The same argument, applied to the simple nest $\mathcal{P}^c = \{I - P \mid P \in \mathcal{P}\}$, proves that the nest of subspaces $\{\text{Ker } P \mid P \in \mathcal{P}\}$ is simple.

First, $(0), X \in \mathcal{M}$, since $O, I \in \mathcal{P}$. Second, if M_1, M_2 form a gap in $\mathcal{M}$, then there exist two projections $P_1, P_2 \in \mathcal{P}$, such that $M_1 = \text{Ran } P_1$ and $M_2 = \text{Ran } P_2$. It follows that P_1, P_2 is a gap in $\mathcal{P}$, so $\text{rank}(P_2 - P_1) = 1$. Therefore, the quotient space M_2/M_1 is one-dimensional. Finally, let $\mathcal{M}_1 \subseteq \mathcal{M}$ be a non-empty subnest. To prove that $\bigvee \mathcal{M}_1 \in \mathcal{M}$, define

$$\mathcal{P}_1 = \{P \mid P \in \mathcal{P}, \text{Ran } P \in \mathcal{M}_1\}.$$

Then $\mathcal{P}_1 \subseteq \mathcal{P}$ is a non-empty subnest. Therefore, $\bigvee \mathcal{P}_1 \in \mathcal{P}$, and it follows that $\bigvee \mathcal{M}_1 = \mathrm{Ran}\,(\bigvee \mathcal{P}_1) \in \mathcal{M}$. In the same fashion, one proves that $\bigwedge \mathcal{M}_1 \in \mathcal{M}$.

The if part is proved as follows. Assume that $\mathcal{P}$ is a nest of projections, such that the nests of its ranges and its kernels are simple. It follows, by reverting the relevant arguments given above, that $O, I \in \mathcal{P}$ and that all gaps in $\mathcal{P}$ (if there are any) are one-dimensional. Let $\mathcal{P}_1 \subseteq \mathcal{P}$ be a non-empty subnest, and let

$$M_1 = \mathrm{span}\{\mathrm{Ran}\ P \mid P \in \mathcal{P}_1\}, \qquad N_2 = \bigcap\{\mathrm{Ker}\ P \mid P \in \mathcal{P}_1\}.$$

There exist $Q_1, Q_2 \in \mathcal{P}$, such that Ran $Q_1 = M_1$ and Ker $Q_2 = N_2$. If $P \in \mathcal{P}_1$, then Ran $P \subseteq M_1$, so $P \leq Q_1$. In other words, Q_1 is an upper bound for $\mathcal{P}_1$ and hence Ker $Q_1 \subseteq N_2$. It follows that $Q_2 \leq Q_1$. On the other hand, if $P \in \mathcal{P}_1$, then $N_2 \subseteq$ Ker P and $P \leq Q_2$. Therefore, $M_1 \subseteq$ Ran Q_2 and $Q_1 \leq Q_2$ now follows. Hence, $Q_1 = Q_2 = \bigvee \mathcal{P}_1$ exists. Analogously, one proves that $\bigwedge \mathcal{P}_1$ exists. The theorem is proved. □

We will now discuss some well-known examples of simple nests of projections. First of all, in Hilbert space, a maximal nest of subspaces $\mathcal{M}$ induces a set of orthogonal (self-adjoint) projections

$$\mathcal{P} = \{P \mid \mathrm{Ran}\ P = M,\ \ M \in \mathcal{M}\}.$$

Note that the nests of subspaces

$$\{\mathrm{Ran}\ P \mid P \in \mathcal{P}\} = \mathcal{M}, \qquad \{\mathrm{Ker}\ P \mid P \in \mathcal{P}\} = \{M^{\perp} \mid M \in \mathcal{M}\},$$

are maximal, so by Theorem 4.1.3, $\mathcal{P}$ is simple.

Other simple nests are those induced by bases on a Banach or Hilbert space. Let X be a (separable) Banach space with *Schauder basis* $\{x_n\}_{n=1}^{\infty}$, i.e., a set of non-zero vectors $\{x_n\}_{n=1}^{\infty}$, such that each vector $x \in X$ corresponds to a unique sequence of scalars $\{\zeta_n\}_{n=1}^{\infty}$, with

$$x = \sum_{n=1}^{\infty} \zeta_n x_n,$$

where the infinite sum converges in the norm topology. A Schauder basis induces a bounded simple nest of projections in X. Indeed, for given $n \in \mathbf{Z}^+$, consider the projection P_n of rank n defined as

$$P_n \sum_{k=1}^{\infty} \zeta_k x_k = \sum_{k=1}^{n} \zeta_k x_k.$$

The set $\mathcal{P} = \{P_n \mid n \in \mathbf{Z}^+\} \cup \{O, I\}$ is a bounded nest of projections on X, that strongly converges to the identity ($P_n \stackrel{s}{\longrightarrow} I$); see for example [40], Proposition 1.a.2. Obviously, the trivial projections O and I are elements in $\mathcal{P}$. It is rather straightforward to verify that $I = \bigvee\{P_n \mid n \in \mathbf{Z}^+\}$. Further, it is easy to see that all gaps in the nest are one-dimensional. It follows that $\mathcal{P}$ is simple.

An *orthonormal basis* in a Hilbert space H is an orthonormal sequence of vectors $\{e_n\}_{n=1}^\infty$, that spans the whole space: $\text{span}\{e_n \mid n \in \mathbf{Z}^+\} = H$. The Fourier expansion

$$x = \sum_{n=1}^{\infty} (x, e_n) e_n, \qquad x \in H,$$

shows that an orthonormal basis is a Schauder basis. In fact, the simple nest of projections induced by an orthonormal basis consists of orthogonal projections.

A *Riesz basis* in a separable Hilbert space H is the image of an orthonormal basis in H under an invertible operator. In other words, if $\{x_n\}_{n=1}^\infty$ is a Riesz basis, and $\{e_n\}_{n=1}^\infty$ is an orthonormal basis, then there exists an invertible operator T, such that $Te_n = x_n$ for $n \in \mathbf{Z}^+$. For equivalent definitions of a Riesz basis, we refer to [53], Section 1.8. Also, a Riesz basis is a Schauder basis. As a matter of fact, it is rather difficult to find a bounded Schauder basis in a Hilbert space, that is not a Riesz basis. For examples of so-called conditional bases on a Hilbert space, we refer to [40], Proposition 2.b.11.

We state some additional examples of simple nests. In the following, fix $1 \leq p < \infty$. On $L_p(0,1)$, the projections

$$P_t f(x) = \begin{cases} f(x), & 0 \leq x \leq t \\ 0, & t < x \leq 1 \end{cases}$$

for $0 \leq t \leq 1$, define a simple continuous nest $\mathcal{P} = \{P_t \mid 0 \leq t \leq 1\}$. On $L_p(0,1)^2$, the projections

$$Q_t f(x,y) = \begin{cases} f(x,y), & 0 \leq x \leq t \\ 0, & t < x \leq 1 \end{cases}$$

for $0 \leq t \leq 1$, define a simple continuous nest $\mathcal{Q} = \{Q_t \mid 0 \leq t \leq 1\}$. The nest $\mathcal{P}$ defined above is essentially different from the nest $\mathcal{Q}$; see [20]. On $l_p(\mathbf{Z})$ with standard Schauder basis $\{e_n\}_{n \in \mathbf{Z}}$, the projections

$$P_n \left(\sum_{k=-\infty}^{\infty} \alpha_k e_k \right) = \sum_{k=-\infty}^{n} \alpha_k e_k$$

define the simple nest $\{P_n \mid n \in \mathbf{Z}\} \cup \{O, I\}$.

4.2 Closure of Nests

Proposition 4.2.1 below shows that a bounded and complete nest is a closed set of operators in the strong operator topology. Before we state and prove the proposition, we briefly present some topological prerequisites. Let (Y, τ) be a topological space, such that each point of Y is a closed set (so it is a T_1-space).

A *directed set* A is a set with a partial ordering $\preceq$, such that for all $\alpha, \beta \in A$, there exists $\gamma \in A$ with $\alpha \preceq \gamma$ and $\beta \preceq \gamma$. The set $\{y_\alpha \mid \alpha \in A\} \subseteq Y$ is called a *net*, if the set A is directed. A net $\{y_\alpha \mid \alpha \in A\}$ *converges* to $y \in Y$, if for each open neighbourhood $\mathcal{O}$ of y, there exists $\alpha \in A$, such that $x_\beta \in \mathcal{O}$, for all $\alpha \preceq \beta$. Let $\mathcal{S} \subseteq Y$ be an arbitrary subset, and let $\overline{\mathcal{S}}$ denotes its closure with respect to the topology τ. Note that $y \in \overline{\mathcal{S}}$ if and only if there exists a net in $\mathcal{S}$ that converges to y.

Let $\{y_\alpha \mid \alpha \in A\}$ be a net, that converges to $y \in Y$. Assume that the directed set $A = A_1 \cup A_2$ is the union of two sets. Then either A_1 or A_2 is directed, and it follows that either $\{y_\alpha \mid \alpha \in A_1\}$ or $\{y_\alpha \mid \alpha \in A_2\}$ contains a net that converges to y.

In the proposition below, we will consider the space of all bounded operators on a Banach space, equipped with the strong operator topology. The strong operator closure of a set of bounded operators $\mathcal{S}$ is denoted by $\overline{\mathcal{S}}$.

Proposition 4.2.1 *A bounded and complete nest of projections on a Banach space X is closed with respect to the strong operator topology.*

Proof Let $\mathcal{P}$ be a bounded and complete nest of projections on X. We need to prove that $\mathcal{P} = \overline{\mathcal{P}}$, i.e., that the nest equals its closure in the strong operator topology. Clearly, the nest is contained in its closure. To prove the other inclusion, let $T \in \overline{\mathcal{P}}$. By the discussion before the proposition, there exists a net $\mathcal{P}_1 = \{P_\alpha \mid \alpha \in A\} \subseteq \mathcal{P}$, that converges to T in the strong operator topology. Since $\mathcal{P}$ is bounded, it follows that T is a projection which commutes with all $P \in \mathcal{P}$. We need to prove that in fact $T \in \mathcal{P}$. If $T = O$ or $T = I$, this is immediate, so we may assume that T is a non-trivial projection. Fix $P \in \mathcal{P}$ and define the following two sets

$$\mathcal{Q}_+ = \{P_\alpha \mid \alpha \in A, P_\alpha \geq P\},$$

$$\mathcal{Q}_- = \{P_\alpha \mid \alpha \in A, P_\alpha \leq P\}.$$

Since $\mathcal{Q}_+ \cup \mathcal{Q}_- = \mathcal{P}_1$, at least one of the subsets $\mathcal{Q}_+$ or $\mathcal{Q}_-$ contains a net that converges to T. If $\mathcal{Q}_-$ contains a net $\{P_\alpha\}$ convergent to T, then $P = PP_\alpha \xrightarrow{s} PT$, so $T \geq P$. If $\mathcal{Q}_+$ contains a net $\{P_\alpha\}$ convergent to T, then $I - P = (I-P)(I-P_\alpha) \xrightarrow{s} (I-P)(I-T)$, so $T \leq P$. Assume that $T \notin \mathcal{P}$ and define the subnests

$$\mathcal{P}_+ = \{P \mid P \in \mathcal{P}, P > T\},$$

$$\mathcal{P}_- = \{P \mid P \in \mathcal{P}, P < T\}.$$

Note that $O \in \mathcal{P}_-$ and that $I \in \mathcal{P}_+$. For each $P \in \mathcal{P}$, either $T \leq P$ or $T \geq P$ hence $T < P$ or $T > P$, so $\mathcal{P}_- \cup \mathcal{P}_+ = \mathcal{P}$. Let $P_+ = \bigwedge \mathcal{P}_+$ and $P_- = \bigvee \mathcal{P}_-$. Then $P_-, P_+ \in \mathcal{P}$ and $P_- \leq T \leq P_+$. Since $T \notin \mathcal{P}$, we obtain that $P_- < T < P_+$. In other words, Ran $P_- \subset$ Ran $T \subset$ Ran P_+ and Ker $P_- \supset$ Ker $T \supset$ Ker P_+. There exist $x, y \in X$, $\|x\| = \|y\| = 1$, such that

$$\begin{array}{lll} P_- x = 0, & Tx = x, & P_+ x = x, \\ P_- y = 0, & Ty = 0, & P_+ y = y. \end{array}$$

Since $(\mathcal{P}_+ \cap \mathcal{P}_1) \cup (\mathcal{P}_- \cap \mathcal{P}_1) = \mathcal{P}_1$, at least one of the sets $\mathcal{P}_+ \cap \mathcal{P}_1$ or $\mathcal{P}_- \cap \mathcal{P}_1$ contains a net that converges to T. Note that if $P_\alpha \in \mathcal{P}_-$, then $P_\alpha \leq P_-$, and if $P_\alpha \in \mathcal{P}_+$, then $P_\alpha \geq P_+$. But, if $\mathcal{P}_+ \cap \mathcal{P}_1$ contains a convergent net, then $y = P_\alpha y \xrightarrow{\|\cdot\|} Ty = 0$ ($P_\alpha \in \mathcal{P}_+$), a contradiction. On the other hand, if $\mathcal{P}_- \cap \mathcal{P}_1$ contains a convergent net, it follows that $0 = P_\alpha x \xrightarrow{\|\cdot\|} Tx = x$ ($P_\alpha \in \mathcal{P}_-$), which also leads to a contradiction. Therefore, $T \in \mathcal{P}$ and the proposition is proved. □

In general, a closed and bounded nest of projections in a Banach space need not be complete, as Example 4.2.2 below shows. In a reflexive Banach space however, closed and bounded nests of projections, that contain the trivial projections O and I, are complete. This is shown in the next section, Corollary 4.3.3.

Example 4.2.2 Let $X = L_\infty[0,1]$ and for $0 \leq \tau \leq 1$, define the projections

$$P_\tau f(x) = \begin{cases} f(x) & 0 \leq x \leq \tau \\ 0 & \tau < x \leq 1 \end{cases}.$$

Note that if $0 \leq \sigma < \tau \leq 1$ and $e(x) = 1$ for almost all $x \in [0,1]$, then $\|P_\tau e - P_\sigma e\|_\infty = 1$. Also, $\|P_\tau\| = 1$ for $0 < \tau \leq 1$. Let $\tau_k = \frac{1}{2} - (\frac{1}{2})^k$, then

$\tau_k \to \frac{1}{2}$ as $k \to \infty$, while the sequence $\{P_{\tau_k} \mid k \in \mathbf{Z}^+\}$ does not have a Cauchy subsequence. For this reason, the nest

$$\mathcal{P} = \{P_{\tau_k} \mid k \in \mathbf{Z}^+\}$$

is closed, although it is not complete; the strong supremum $\bigvee \mathcal{P}$ does not exist: Indeed, we have the proper inclusion

$$\text{span}\{\text{Ran } P_{\tau_k} \mid k \in \mathbf{Z}^+\} \subset L_\infty[0, \frac{1}{2}],$$

while on the other hand,

$$\bigcap\{\text{Ker } P_{\tau_k} \mid k \in \mathbf{Z}^+\} = L_\infty[\frac{1}{2}, 1].$$

In Example 4.2.2, the strong supremum $\bigvee \mathcal{P}$ did not exist. Proposition 4.2.3 below shows that this is due to the fact that no upper bound for this nest is in the strong operator closure of the nest.

Proposition 4.2.3 *Let $\mathcal{P}$ be a bounded nest of projections and let P_0 be an upper bound for $\mathcal{P}$. Then the following are equivalent:*

1. $P_0 x \in cl\{Px \mid P \in \mathcal{P}\}$ *for all* $x \in X$,
2. $P_0 = \bigvee \mathcal{P}$,
3. $P_0 \in \overline{\mathcal{P}}$.

Before proving this proposition, we make a few observations. A net of projections $\{P_\alpha \mid \alpha \in A\}$ is called *directed upwards*, if for $\alpha, \beta \in A$, $\alpha \preceq \beta$ if and only if $P_\alpha \leq P_\beta$. If an upper bound P_0 for the nest of projections $\mathcal{P}$ is contained in the strong operator closure, then $\mathcal{P}$ itself is an upwards directed net of projections, that converges to P_0 in the strong operator norm. This will be shown in the proof of Proposition 4.2.3. A result in the same direction as Proposition 4.2.3 is Theorem 1 in [4].

Proof To prove that the first statement implies the second one, let $y \in \text{Ran } P_0$. By assumption, $\{Py \mid P \in \mathcal{P}\}$ has $P_0 y = y$ as a limit point. Therefore, $y \in M = cl\{\text{Ran } P \mid P \in \mathcal{P}\}$, and $\text{Ran } P_0 \subseteq M$ follows. Since P_0 is an upper bound for $\mathcal{P}$, also $M \subseteq \text{Ran } P_0$, and we get $M = \text{Ran } P_0$. Let

$x \in N = \bigcap\{\text{Ker } P \mid P \in \mathcal{P}\}$. Then $P_0 x = 0$, since $P_0 x$ is a limit point of $\{Px \mid P \in \mathcal{P}\} = \{0\}$. Therefore, $x \in \text{Ker } P_0$, and $N \subseteq \text{Ker } P_0$ follows. Again, since P_0 is an upper bound for $\mathcal{P}$, we obtain $\text{Ker } P_0 = N$. We conclude that $P_0 = \bigvee \mathcal{P}$.

To prove that the second statement implies the first one, assume that $\|P\| \leq C$ for all $P \in \mathcal{P}$. Let $x \in X$, then $y = P_0 x \in \text{Ran } P_0$. Let $\varepsilon > 0$. There exist $P_\varepsilon \in \mathcal{P}$, and $y_\varepsilon \in \text{Ran } P_\varepsilon$, such that $\|y - y_\varepsilon\| < \varepsilon$. Consequently,

$$P_0 x - Px = P_0 x - y_\varepsilon + y_\varepsilon - Px = (y - y_\varepsilon) + P(y_\varepsilon - P_0 x) =$$

$$(y - y_\varepsilon) + P(y_\varepsilon - y) = (I - P)(y - y_\varepsilon).$$

Therefore,

$$\|P_0 x - Px\| \leq (1 + C)\|y - y_\varepsilon\| \leq (1 + C)\varepsilon.$$

Since $\varepsilon > 0$ was taken arbitrary, $P_0 x \in cl\{Px \mid P \in \mathcal{P}\}$. Further, $x \in X$ was taken arbitrary, so the first statement now follows.

It is immediate, that the third statement implies the first one. We now prove that the first statement implies the third one. We prove that each open neighbourhood (in the strong operator topology) of P_0 of the form

$$\mathcal{O} = \bigcap_{k=1}^{n} \{S \mid \|P_0 x_k - S x_k\| < \varepsilon\}$$

for arbitrary $n \in \mathbf{Z}^+$, $x_1, \ldots, x_n \in X$, $\varepsilon > 0$ contains all projections in $\mathcal{P}$, that are greater or equal to a certain projection in $\mathcal{P}$. Let $\mathcal{O}$ be a neighbourhood of P_0 as above, and let $k \in \{1, \ldots, n\}$. Since $P_0 x_k \in cl\{Px \mid P \in \mathcal{P}\}$, there exists $P_k \in \mathcal{P}$, such that $\|P_0 x_k - P_k x_k\| \leq \varepsilon/(1 + C)$. Now let $P = \bigvee\{P_1, \ldots, P_n\}$. For each $Q \geq P$, we get $(I - Q)(P_0 - P_k) = P_0 - Q$, so

$$\|P_0 x_k - Q x_k\| \leq (1 + C)\|P_0 x_k - P_k x_k\| < \varepsilon.$$

Consequently, $Q \in \mathcal{O}$ and we have proved the proposition. In fact, the nest $\mathcal{P}$ itself is an upwards directed net that converges to its strong supremum. □

The following example shows, that a simple (complete) nest of projections acting on a Banach space need not be bounded or closed.

Example 4.2.4 Consider the Hilbert space $X = l_2(\mathbf{Z}^+)$, with standard orthonormal basis $\{e_k\}_{k=1}^{\infty}$. For each $n \in \mathbf{Z}^+$, consider the projection P_n of rank n, defined by

$$P_n e_k = \begin{cases} e_k + (-1)^{n-k} e_{n+1}, & 1 \le k \le n \\ 0, & k > n \end{cases}.$$

Then

$$\text{Ran } P_n = \text{span}\{e_1 + e_2, \ldots, e_n + e_{n+1}\},$$

$$\text{Ker } P_n = \text{span}\{e_1, \ldots, e_n\}^{\perp}.$$

It follows that $P_n < P_{n+1}$ for all $n \in \mathbf{Z}^+$, and that

$$\text{span}\{\text{Ran } P_n \mid n \in \mathbf{Z}^+\} = l_2(\mathbf{Z}^+),$$

$$\bigcap\{\text{Ker } P_n \mid n \in \mathbf{Z}^+\} = (0).$$

For this reason, $\bigvee\{P_n \mid n \in \mathbf{Z}^+\} = I$ exists. Moreover,

$$\mathcal{P} = \{P_n \mid n \in \mathbf{Z}^+\} \cup \{O, I\}$$

is a simple nest of projections. On the other hand, if we fix $n \in \mathbf{Z}^+$, and write $x = (\ (-1)^{n-1},\ (-1)^{n-2},\ \cdots,\ -1,\ 1,\ 0,\ \cdots\ )^T$, then $P_n x = x + n e_{n+1}$, so $\|P_n x\|^2 = \|x\|^2 + n^2$. Since $\|x\| = \sqrt{n}$, we get

$$\|P_n\| \ge \frac{\|P_n x\|}{\|x\|} = \sqrt{n+1}.$$

Thus the simple nest of projections $\mathcal{P}$ is not bounded. Note that Proposition 4.2.3 now fails; the subnest $\mathcal{P}_1 = \{P_n \mid n \in \mathbf{Z}^+\}$ has strong supremum $\bigvee \mathcal{P}_1 = I$, but $I \notin \overline{\mathcal{P}_1}$: Indeed, $P_n e_1 = e_1 + (-1)^{n-1} e_{n+1}$ does not converge to e_1 in norm.

A *Boolean algebra* $\mathcal{B}$ of projections is a commuting set of projections, that satisfies the following condition: $I \in \mathcal{B}$, and if $P, Q \in \mathcal{B}$, then $P \wedge Q \in \mathcal{B}$, $P \vee Q \in \mathcal{B}$, and $I - P \in \mathcal{B}$. One could, for example, consider Boolean algebras generated by a nest of projections. In contrast to the situation for nests (see Example 4.2.4), a complete Boolean algebra is necessarily bounded (Theorem 2.2 in [3]). On the other hand, the completeness result Proposition 4.2.3 is analogous to Lemma 2.3 in [3].

4.3 Nests in Reflexive Banach Spaces

In this section, we will discuss a converse of Lemma 4.1.2 in a reflexive Banach space: Theorem 4.3.4 below states that a bounded and maximal nest of projections acting on a reflexive Banach space is simple. Even in Hilbert space, there exist nests of projections (Example 4.3.5) which are maximal, but not simple. However, these nests of projections are not bounded. On the other hand, there exist nests of projections on a nonreflexive Banach space, which are bounded and maximal, but not simple (Example 4.3.6). Corollary 4.3.3 shows that in reflexive Banach spaces, a closed and bounded nest is complete. This is a converse to Proposition 4.2.1 in a reflexive Banach space. For preliminaries concerning the conjugate space of a Banach space, we refer to [2].

First, we fix some notation. Let the conjugate space of a Banach space X be denoted by X^*. The conjugate operator of a bounded operator T on X is denoted by T^*. If $U \subseteq X$ is a non-empty subset, its annihilator is given by

$$U^\perp = \{f \mid f \in X^*, f(x) = 0 \ \text{ for all } \ x \in U\}.$$

If $U = \emptyset$, define $U^\perp = X^*$. If $V \subseteq X^*$ is a non-empty subset, then its inverse annihilator is defined as

$$^\perp V = \{x \mid x \in X, f(x) = 0 \ \text{ for all } \ f \in X^*\}.$$

If $V = \emptyset$, define $^\perp V = X$.

In [41], the main ingredients are given to prove that maximal, bounded nests on a reflexive Banach space are simple. The following lemma is well-known and can be found in e.g. [33] (Corollary 1.8.8).

Lemma 4.3.1 *Two closed subspaces M, N in a Banach space have zero intersection and closed sum if and only if there exists $k > 0$, such that for all $x \in M$ and $y \in N$,*

$$\|x + y\| \geq k\|x\|.$$

The following theorem is contained in [41]. The proof of the theorem is instructive and therefore included here.

Theorem 4.3.2 *If $\mathcal{P}$ is a bounded nest of projections on a reflexive Banach space X, then its strong supremum $\bigvee \mathcal{P}$ and its strong infimum $\bigwedge \mathcal{P}$ exist.*

Proof Consider the bounded nest $\mathcal{P}$ and let $0 < C < \infty$, such that $\|P\| \leq C$ for all $P \in \mathcal{P}$. We first prove that $\bigvee \mathcal{P}$ exists. To this end, we will show that the subspaces

$$M = \text{span}\{\text{Ran } P \mid P \in \mathcal{P}\}$$

and

$$N = \bigcap\{\text{Ker } P \mid P \in \mathcal{P}\}$$

satisfy $M \oplus N = X$. Indeed, let $x \in M$ and $y \in N$. Then $Py = 0$ for all $P \in \mathcal{P}$ and there exist $P_n \in \mathcal{P}$ and $x_n \in \text{Ran } P_n$, such that $x_n \xrightarrow{\|\cdot\|} x$, as $n \to \infty$. Therefore,

$$\|x_n\| = \|P_n(x_n + y)\| \leq C\|x_n + y\|.$$

If we let n tend to infinity at both sides, we get $\|x\| \leq C\|x + y\|$ or

$$\|x + y\| \geq k\|x\|,$$

with $k = 1/C > 0$. Lemma 4.3.1 implies that the subspaces M and N have zero intersection and have closed sum. In other words, $M \oplus N \subseteq X$ is a closed subspace.

Consider the adjoint nest of projections $\mathcal{P}^* = \{P^* \mid P \in \mathcal{P}\}$, and the subspaces

$$M^* = \text{span}\{\text{Ran } P^* \mid P \in \mathcal{P}\}$$

and

$$N^* = \bigcap\{\text{Ker } P^* \mid P \in \mathcal{P}\}.$$

Note that $\mathcal{P}^*$ is again a bounded nest of projections, so a similar argument can be applied to obtain that the subspaces M^* and N^* have zero intersection and closed sum. Further, $X = {}^{\perp}(0) = {}^{\perp}(M^* \cap N^*)$, and reflexivity of X gives

$${}^{\perp}(M^* \cap N^*) = cl({}^{\perp}M^* + {}^{\perp}N^*).$$

We find ${}^{\perp}M^* = N$, and by reflexivity of X, we get ${}^{\perp}N^* = M$. Since $M + N$ is closed, it follows that $M + N = X$. Apply the argument presented here to $\mathcal{P}^c$ to obtain that $\bigwedge \mathcal{P}$ exists. $\square$

The following corollary to Theorem 4.3.2 is a converse to Proposition 4.2.1 in a reflexive Banach space.

Corollary 4.3.3 *Each closed and bounded nest of projections on a reflexive Banach space containing the trivial projections O and I is complete.*

Proof Let $\mathcal{P}$ be a closed and bounded nest of projections acting on the reflexive Banach space X, and assume that $\mathcal{P}$ contains the trivial projections O and I. Let $\mathcal{P}_1 \subseteq \mathcal{P}$ be a non-empty subnest. We need to prove that $\bigvee \mathcal{P}_1$ and $\bigwedge \mathcal{P}_1$ are in $\mathcal{P}$. By Theorem 4.3.2, these two projections exist. By Proposition 4.2.3, the projections are in the strong closure $\overline{\mathcal{P}_1}$. Since $\mathcal{P}_1 \subseteq \mathcal{P}$, and since $\mathcal{P}$ is closed, the projections $\bigvee \mathcal{P}_1$ and $\bigwedge \mathcal{P}_1$ are in $\mathcal{P}$. □

We now state a converse to Lemma 4.1.2 in a reflexive Banach space.

Theorem 4.3.4 *A maximal, bounded nest of projections in a reflexive Banach space is simple.*

Proof Let $\mathcal{P}$ be a maximal, bounded nest in a reflexive Banach space. It follows that $O, I \in \mathcal{P}$, since enlarging $\mathcal{P}$ with these trivial projections would otherwise result in a nest of subspaces properly containing $\mathcal{P}$, contrary to its maximality. If $\mathcal{P}_1 \subseteq \mathcal{P}$ is a subnest, then the strong infimum $\bigwedge \mathcal{P}_1$ and the strong supremum $\bigvee \mathcal{P}_1$ are in $\mathcal{P}$, since Theorem 4.3.2 implies that the collection $\mathcal{P} \cup \{\bigwedge \mathcal{P}_1, \bigvee \mathcal{P}_1\}$ is again a nest. If P_1, P_2 form a gap in $\mathcal{P}$ of dimension greater than one, then choose $0 \neq x \in \text{Ran}\ (P_2 - P_1)$ and the subspace $M \subset \text{Ran}\ (P_2 - P_1)$ such that $\text{span}\{x\} \oplus M = \text{Ran}\ (P_2 - P_1)$. Notice that $M \neq (0)$, since $\text{rank}(P_2 - P_1) > 1$. Define the mapping R on X by $Ry = y$ for $y \in \text{Ran}\ P_1$, $Rx = x$, $Rm = 0$ for $m \in M$, and $Rz = 0$ for $z \in \text{Ker}\ P_2$. Since $X = \text{Ran}\ P_1 \oplus \text{span}\{x\} \oplus M \oplus \text{Ker}\ P_2$, it follows that R is a projection and that $P_1 < R < P_2$. The collection $\mathcal{P} \cup \{R\}$ is a nest of projections properly containing $\mathcal{P}$, a contradiction. Therefore, $\text{rank}(P_2 - P_1) = 1$. The nest $\mathcal{P}$ is simple by definition and the theorem is proved. □

We will now present an example of a nest of projections in a Hilbert space which is maximal, but not simple. Since it is not bounded, the example does not contradict the statement of Theorem 4.3.4.

Example 4.3.5 Consider the Hilbert space $X = l_2(\mathbf{Z}^+)$ and let $\{e_n\}_{n=1}^{\infty}$ denote its standard orthonormal basis. Define the vectors

$$x_n = \tfrac{3}{5}e_n + \tfrac{4}{5}e_{n+1}, \quad n \in \mathbf{Z}^+.$$

It is not difficult to see that

$$\text{span}\{x_1, \ldots, x_n\} \oplus \text{span}\{e_1, \ldots, e_n\}^{\perp} = X,$$

for all $n \in \mathbf{Z}^+$. Fix $n \in \mathbf{Z}^+$. The subspaces

$$M = \text{span}\{x_1, x_3, \ldots, x_{2n-1}\},$$

and

$$N = \text{span}\{x_2, x_4, \ldots, x_{2n}\} \oplus \text{span}\{e_1, e_2, \ldots, e_{2n}\}^\perp$$

complement each other. Let P_n denote the projection onto M along N. In the same fashion, for fixed $m \in \mathbf{Z}^+$, consider the projection Q_m onto

$$\text{span}\{x_2, x_4, \ldots, x_{2m}\}$$

along

$$\text{span}\{x_1, x_3, \ldots, x_{2m-1}\} \oplus \text{span}\{e_1, e_2, \ldots, e_{2m}\}^\perp.$$

It follows that Ran $Q_m \subseteq$ Ker P_n and that Ran $P_n \subseteq$ Ker Q_m, so $P_n Q_m = Q_m P_n = O$. In other words,

$$P_n < I - Q_m, \quad n, m \in \mathbf{Z}^+.$$

Consider the nest of projections

$$\mathcal{P} = \{O, I\} \cup \{P_n \mid n \in \mathbf{Z}^+\} \cup \{I - Q_m \mid m \in \mathbf{Z}^+\}.$$

We will show that $\mathcal{P}$ is maximal, not bounded and not simple.

Assume there exists a projection R, such that $P_n \leq R$ for $n \in \mathbf{Z}^+$ and such that $R \leq I - Q_m$ for all $m \in \mathbf{Z}^+$. Then

$$\text{span}\{\text{Ran } P_n \mid n \in \mathbf{Z}^+\} \subseteq \text{Ran } R \subseteq \bigcap\{\text{Ker } Q_m \mid m \in \mathbf{Z}^+\}$$

and

$$\text{span}\{\text{Ran } Q_m \mid m \in \mathbf{Z}^+\} \subseteq \text{Ker } R \subseteq \bigcap\{\text{Ker } P_n \mid n \in \mathbf{Z}^+\}.$$

Since

$$V = \text{span}\{\text{Ran } P_n \mid n \in \mathbf{Z}^+\} = \bigcap\{\text{Ker } Q_m \mid m \in \mathbf{Z}^+\} =$$

$$\text{span}\{x_{2k-1} \mid k \in \mathbf{Z}^+\}$$

and

$$W = \{\text{Ran } Q_m \mid m \in \mathbf{Z}^+\} = \bigcap\{\text{Ker } P_n \mid n \in \mathbf{Z}^+\} =$$

$$\text{span}\{x_{2k} \mid k \in \mathbf{Z}^+\},$$

it follows that Ran $R = V$ and that Ker $R = W$. On the other hand, $V + W =$ span$\{x_k \mid k \in \mathbf{Z}^+\} \neq X$, a contradiction. Indeed,

$$y = \left(1, \ -\frac{3}{4}, \ (-\frac{3}{4})^2, \ \cdots \right)^T \in (V+W)^\perp.$$

Consequently, for each upper bound P_+ of $\mathcal{P}_1 = \{P_n \mid n \in \mathbf{Z}^+\}$, there exists $m \in \mathbf{Z}^+$, such that $I - Q_m < P_+$. Since $I - Q_m$ is itself an upper bound of $\mathcal{P}_1$, it follows that there exists no least upper bound of the subnest $\mathcal{P}_1$. In the same fashion one proves that there exists no greatest lower bound for $\{I - Q_m \mid m \in \mathbf{Z}^+\}$. We may conclude that the nest $\mathcal{P}$ is not simple. On the other hand, the nest $\mathcal{P}$ is maximal. Indeed, let R be a projection, such that $\mathcal{P} \cup \{R\}$ is a nest which properly contains $\mathcal{P}$. It follows that $R < I - Q_m$ for all $m \in \mathbf{Z}^+$ and that $P_n < R$ for all $n \in \mathbf{Z}^+$. We have already proved that such a projection R does not exist. Further, it is not difficult to see that for $k \in \mathbf{Z}^+$,

$$e_1 = \left(\sum_{j=1}^{k} \frac{5}{3}(-\frac{4}{3})^{j-1} x_j \right) + (-\frac{4}{3})^k e_{k+1}.$$

For $n \in \mathbf{Z}^+$, define the projection

$$R_n = P_n \vee Q_n = P_n + Q_n - P_n Q_n = P_n + Q_n,$$

where we used that $P_n Q_n = Q_n P_n = O$. Then rank $R_n = 2n$, and

$$\text{Ran } R_n = \text{span}\{x_1, x_2, \ldots, x_{2n}\},$$

$$\text{Ker } R_n = \text{span}\{e_1, e_2, \ldots, e_{2n}\}^\perp.$$

Consequently,

$$(I - R_n)e_1 = e_1 - \sum_{j=1}^{2n} \frac{5}{3}(-\frac{4}{3})^{j-1} x_j = (-\frac{4}{3})^{2n} e_{2n+1},$$

so $\|I - Q_n\| + \|P_n\| \geq \|I - R_n\| \geq \|(I - R_n)e_1\| = (\frac{4}{3})^{2n}$. It follows that the nest $\mathcal{P}$ is not bounded.

The following example is a maximal, bounded nest of projections in a nonreflexive Banach space, which is not simple.

Example 4.3.6 The sequence space

$$c = \{(\xi_1, \xi_2, \xi_3, \ldots) \mid \lim_{k\to\infty} \xi_k \in \mathbf{C}\}$$

is a nonreflexive Banach space with the supremum-norm. The subset of all zero-sequences

$$c_0 = \{x = (\xi_1, \xi_2, \xi_3, \ldots) \mid \lim_{k\to\infty} \xi_k = 0\}$$

is a closed subspace of c of codimension one. If $x = (\xi_1, \xi_2, \xi_3, \ldots) \in c_0$, then $x = \sum_{k=1}^{\infty} \zeta_k e_k$ converges in the supremum-norm. The set $\{e_k \mid k \in \mathbf{Z}^+\}$ is a Schauder basis in c_0. If $x = (\xi_1, \xi_2, \xi_3, \ldots) \in c$ and $\lim_{k\to\infty} \xi_k = \xi_0$, then $x - (\xi_0, \xi_0, \xi_0, \ldots) = \sum_{k=1}^{\infty}(\xi_k - \xi_0)e_k \in c_0$. Note that $e_0 = (1, 1, 1, \ldots) \in c$, but $e_0 \notin c_0$. Further,

$$x = \xi_0 e_0 + \sum_{k=1}^{\infty}(\xi_k - \xi_0)e_k$$

converges in the supremum-norm. It follows that the set $\{e_k \mid k \in \mathbf{Z}_0^+\}$ is a Schauder basis in the whole space c.

The construction of the example here is analogous to the construction in Example 4.3.5. For $n, m \in \mathbf{Z}^+$, the projections

$$P_n(\xi_1, \xi_2, \ldots,) = (\xi_1, 0, \xi_3, 0, \ldots, 0, \xi_{2n-1}, 0, 0, \ldots),$$

and

$$Q_m(\xi_1, \xi_2, \ldots,) = (0, \xi_2, 0, \xi_4, 0, \ldots, 0, \xi_{2m}, 0, 0, \ldots)$$

satisfy

$$P_n < I - Q_m, \quad n, m \in \mathbf{Z}^+.$$

Define the nest of projections

$$\mathcal{P} = \{P_n \mid n \in \mathbf{Z}^+\} \cup \{I - Q_m \mid m \in \mathbf{Z}^+\} \cup \{O, I\}.$$

We will show that $\mathcal{P}$ is maximal and bounded, but not simple. As in Example 4.3.5, there exists no projection R, such that $P_n \leq R$ for $n \in \mathbf{Z}^+$ and such that $R \leq I - Q_m$ for all $m \in \mathbf{Z}^+$. Again, we may conclude that the nest $\mathcal{P}$ is not simple, since there exist subnests of $\mathcal{P}$ which neither have greatest lower bound nor least upper bound. Following the same type of argument as in Example 4.3.5, the nest $\mathcal{P}$ is maximal. Finally, it is not difficult to see that

$$\sup\{\|P\| \mid P \in \mathcal{P}\} = 1.$$

In the following example, we present a nest of projections on a Hilbert space, that has a least upper bound, which is not its strong supremum.

Example 4.3.7 As in Example 4.3.5, consider the Hilbert space $X = l_2(\mathbf{Z}^+)$ and let $\{e_n\}_{n=1}^{\infty}$ denote its standard orthonormal basis. Define the vectors

$$x_n = \tfrac{3}{5}e_n + \tfrac{4}{5}e_{n+1}, \quad n \in \mathbf{Z}^+.$$

For $n \in \mathbf{Z}^+$, the projections P_n onto $\text{span}\{x_1, \ldots, x_n\}$ along $\text{span}\{e_1, \ldots, e_n\}^\perp$ define a nest of projections $\mathcal{P} = \{P_n \mid n \in \mathbf{Z}^+\}$. Since

$$\text{span}\{\text{Ran } P_n \mid n \in \mathbf{Z}^+\} = \text{span}\left\{\sum_{k=1}^{\infty}(-\frac{3}{4})^{k-1}e_k\right\}^\perp$$

and

$$\bigcap\{\text{Ker } P_n \mid n \in \mathbf{Z}^+\} = (0),$$

it follows that $\mathcal{P}$ has least upper bound I, which is not a strong supremum of $\mathcal{P}$.

Chapter 5

Finite Rank Operators

In this chapter, complementary triangular forms for pairs of finite rank operators acting on an infinite dimensional Banach space are studied. It will be shown that the study of complementary triangular forms for pairs of finite rank operators reduces more or less to the matrix problem studied in Chapter 3; see Theorem 5.3.1 and Proposition 5.3.2. In fact, the approach presented in this chapter sheds new light on the matters discussed in Chapter 3; see Corollary 5.3.3.

5.1 Matrix Reductions

In this section, we introduce the notion of a matrix reduction for a set of finite rank operators. This notion is based on the fact, that a finite rank operator is, roughly speaking, the direct sum of a finite matrix and a zero operator.

Let A be a finite rank operator acting on the Banach space X, i.e., A is a bounded operator on X, such that the linear manifold $\text{Ran } A = \{Ax \mid x \in X\}$ is finite dimensional. In that case, since $\text{Ran } A$ and $X/\text{Ker } A$ are linearly isomorphic, it follows that

$$\text{rank } A = \dim \text{Ran } A = \dim(X/\text{Ker } A).$$

A bounded operator A on X is *reduced* by a projection P on X, if $AP = PA$. Assume, in addition, that $AP = PA = A$, hence $A = PAP$. Write $M = \text{Ran } P$ and $N = \text{Ker } P$, then with respect to the decomposition $X = M \oplus N$, the operator A assumes the form

$$A = \begin{pmatrix} A_M & O \\ O & O_N \end{pmatrix}.$$

If $A = PAP$ and rank $P < \infty$, then obviously, rank $A < \infty$. The following lemma states a converse to this fact. Actually, the lemma involves a finite set of finite rank operators.

Lemma 5.1.1 *Let $\mathcal{A}$ be a finite set of finite rank operators. Then there exists a finite rank projection P, such that $A = PAP$ for all $A \in \mathcal{A}$.*

Proof Write $M_0 = \text{span}\{\text{Ran } A \mid A \in \mathcal{A}\}$, and $N_0 = \bigcap\{\text{Ker } A \mid A \in \mathcal{A}\}$. The subspace $M_0 + N_0 \subseteq X$ is closed and of finite codimension. Therefore, there exists a finite dimensional subspace R, such that $(M_0 + N_0) \oplus R = X$. Observe that $M = M_0 \oplus R$ is also a finite dimensional subspace, and that $M + N_0 = X$. Let $N \subseteq N_0$, such that $M \oplus N = X$. This defines the finite rank projection P onto M along N, i.e., Ran $P = M$ and Ker $P = N$. This establishes the lemma. □

Let A be a finite rank operator acting on X. By lemma 5.1.1, there exists a finite rank projection P, such that $A = PAP$. The restriction of A to $M = \text{Ran } P$, written A_M, can be identified with a finite matrix by fixing a basis in the finite dimensional subspace M. This matrix is determined up to similarity. Therefore, we can compute the Jordan Canonical form of A_M. Indeed, for $\alpha \in \sigma(A_M)$, compute $d_k(\alpha) = \text{dimKer}(\alpha I_M - A_M)^k$ for $1 \leq k \leq m$, and define $d_0(\alpha) = 0$. It is well-known (see for example [38]), that the number of α-Jordan blocks of size k is given by

$$n_k(\lambda) = [d_{k+1}(\lambda) - d_k(\lambda)] - [d_k(\lambda) - d_{k-1}(\lambda)]\,. \tag{5.1}$$

It is our aim to describe the Jordan canonical form of A_M -as far as possible- in terms of the original finite rank operator A. Since Ker $A = \text{Ker } A_M \oplus N$, we know beforehand, that the number $d_1(0) = \dim \text{Ker } A_M$ will not only depend on A, but also on the dimension of M. Therefore, we will treat the case $\lambda = 0$ separately.

We will first show that the number of λ-Jordanblocks of A_M for $\lambda \neq 0$ only depends on A, and not on M. Indeed, let λ be a non-zero complex number. Then $\text{Ker}(\lambda I_M - A_M)^k = \text{Ker}(\lambda I_X - A)^k$, hence $d_k(\lambda) = \dim \text{Ker}(\lambda I_X - A)^k$. Therefore, the number of λ-Jordanblocks for A_M of a given size is determined by the finite rank operator A through formula (5.1).

We now treat the case $\lambda = 0$. Note that Ker $A^k = \text{Ker } A_M^k \oplus N$. Therefore,

$$d_{k+1}(0) - d_k(0) = \dim\left(\text{Ker}(\lambda I_M - A_M)^{k+1}/\text{Ker}(\lambda I_M - A_M)^k\right) =$$

$$\dim\left(\text{Ker}(\lambda I_X - A)^{k+1}/\text{Ker}(\lambda I_X - A)^k\right).$$

It follows that the numbers $n_k(0)$ are determined by A for $k \geq 2$, again by formula (5.1). The number $n_1(0)$, the number of 0-Jordanblocks of A_M of size one, needs some special attention. Formula (5.1) gives

$$n_1(0) = [d_2(0) - d_1(0)] - [d_1(0) - d_0(0)] \,.$$

Note that $d_1(0) - d_0(0) = \dim \text{ Ker } A_M = \dim M - \text{rank } A$. Thus

$$n_1(0) = \dim \left(\text{Ker } A^2/\text{Ker } A\right) + \dim M - \text{rank } A$$

does not only depend on A, but also on $\dim M$. In particular, if $\dim M = \text{rank } A - \dim (\text{Ker } A^2/\text{Ker } A)$, then $n_1(0) = 0$.

A triple $(M, N, \mathcal{A}_M)$ is a *matrix reduction* for a collection $\mathcal{A}$ of finite rank operators, if $M, N \subseteq X$ are subspaces, and $\mathcal{A}_M$ is a set of finite matrices, such that

1. $\dim M < \infty$, and $M \oplus N = X$,
2. $\text{Ran } A \subseteq M$ and $N \subseteq \text{Ker} A$ for all $A \in \mathcal{A}$,
3. $\mathcal{A}_M = \{A_M \mid A \in \mathcal{A}\}$.

To clarify the notion of a matrix reduction, we give an example.

Example 5.1.2 Consider the set of finite rank operators $\mathcal{A} = \{A, B\}$ acting on the separable Hilbert space $l_2(\mathbf{Z}^+)$. With respect to the orthonormal basis $\{e_k\}_{k=1}^{\infty}$, let A and B be given by the infinite matrices

$$A = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ \frac{1}{2} & 0 & 0 & \cdots \\ \frac{1}{4} & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

We will construct a matrix reduction for the set $\mathcal{A}$, using Lemma 5.1.1. Let $v = \begin{pmatrix} 0, & 1, & \frac{1}{2}, & \frac{1}{4}, & \cdots \end{pmatrix}^T$. Note that $M_0 = \text{Ran } A + \text{Ran } B = \text{span}\{e_2, v\}$. Further, $N_0 = \text{Ker } A \cap \text{Ker } B = \text{span}\{e_1\}^{\perp}$. It follows that $M_0 + N_0 = \text{span}\{e_1\}^{\perp}$. Following the proof of Lemma 5.1.1, choose $R = \text{span}\{e_1\}$, to obtain $M = \text{span}\{e_1, e_2, v\}$. Next, choose $N = M^{\perp}$. This provides a matrix reduction $(M, N, \{A_M, B_M\})$ for $\{A, B\}$, where

$$A_M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad B_M = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Note that both A_M and B_M have the Jordan canonical form

$$J = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

It should be noted that a matrix reduction $(M, N, \mathcal{A}_M)$ for a given set of finite rank operators $\mathcal{A}$ is not unique. On the other hand, it appears that all matrix reductions for a given set of finite rank operators are closely related. This is described in the following proposition. The proof contains elements of the discussion on "Angular subspaces and angular operators" in [6], Section 5.1.

Proposition 5.1.3 *Let $\mathcal{A}$ be a set of finite rank operators with two matrix reductions $(M, N, \mathcal{A}_M)$ and $(M_1, N_1, \mathcal{A}_{M_1})$. Write $\dim M = m$, $\dim M_1 = m_1$, and assume that $m - m_1 = m_2 \geq 0$. Let $N_2 \subseteq N_1$ be an m_2-dimensional subspace. Then there exists an invertible operator $Q : M \longrightarrow M_1 \oplus N_2$, such that $QA_MQ^{-1} = A_{M_1} \oplus O_{N_2}$, for all $A \in \mathcal{A}$.*

To prove this proposition, we will need the following technical lemma.

Lemma 5.1.4 *Let $V_1, V_2 \subseteq X$ be subspaces in a Banach space X, such that*

$$\dim(X/V_1) = \dim(X/V_2) = k < \infty.$$

Then there exists a k-dimensional subspace $R \subseteq X$, such that

$$V_1 \oplus R = V_2 \oplus R = X. \tag{5.2}$$

Proof We start with the following simple remark: If $W_1, W_2 \subset X$ are proper subspaces, then $W_1 \cup W_2$ is also properly contained in the whole space X. Indeed, to avoid trivialities, we may assume that $W_1 \not\subseteq W_2$, and that

$W_2 \not\subseteq W_1$. Let $w_1 \in W_1 \backslash W_2$ and $w_2 \in W_2 \backslash W_1$. If $W_1 \cup W_2 = X$, then $w_1 + w_2 \in W_1$ or $w_1 + w_2 \in W_2$. Thus $w_1 \in W_2$ or $w_2 \in W_1$, a contradiction.

We prove the lemma by induction on the integer k. Indeed, if $k = 0$, then the statement is trivial. Assume that the statement of the lemma holds for $k = 0, \ldots, n-1$, with n a fixed positive integer. Let $V_1, V_2 \subset X$ be subspaces such that $\dim(X/V_1) = \dim(X/V_2) = n$. By the remark at the beginning of the proof, there exists $r \in X$ such that $r \notin V_1 \cup V_2$. Let $\hat{V}_\nu = V_\nu \oplus \text{span}\{r\}$ for $\nu = 1, 2$. Then $\hat{V}_1, \hat{V}_2 \subseteq X$ are closed subspaces such that $\dim(X/\hat{V}_1) = \dim(X/\hat{V}_2) = n - 1$. We may apply the induction hypothesis to obtain an $n-1$ dimensional subspace $\hat{R}$, such that $\hat{V}_\nu \oplus \hat{R} = X$ for $\nu = 1, 2$. The n dimensional subspace $R = \hat{R} \oplus \text{span}\{r\}$ complements the subspaces V_1, V_2, and the lemma is proved. $\square$

Proof of Proposition 5.1.3 First of all, write $M_3 = M_1 \oplus N_2$, and let $N_3 \subseteq N_1$ satisfy $N_3 \oplus N_2 = N_1$. Then $(M_3, N_3, \mathcal{A}_{M_3})$ is also a matrix reduction for $\mathcal{A}$, and $\dim M_3 = \dim M$. Let $\tilde{M} = M \cap M_3$, and define $V = \tilde{M} \oplus N$ and $V_3 = \tilde{M} \oplus N_3$. Then $\dim(X/V) = \dim(X/V_3) < \infty$, so by Lemma 5.1.4, there exists a finite dimensional subspace R, such that $V \oplus R = V_3 \oplus R = X$. Let P, P_3 be finite rank projections on X, with Ran $P = M$, Ker $P = N$ and Ran $P_3 = M_3$, Ker $P_3 = N_3$. Then (see Section 5.1 in [6])

$$T = P|_{\tilde{M} \oplus R} : \tilde{M} \oplus R \longrightarrow M, \quad T_3 = P_3|_{\tilde{M} \oplus R} : \tilde{M} \oplus R \longrightarrow M_3,$$

are invertible operators, since $(\tilde{M} \oplus R) \cap \text{Ker } P = (\tilde{M} \oplus R) \cap \text{Ker } P_3 = (0)$. Therefore, the operator

$$Q = T_3 T^{-1} : M \longrightarrow M_3$$

is invertible. We will prove that $QA_M = A_{M_3} Q$ for all $A \in \mathcal{A}$. First of all, note that $AP = PA = AP_3 = P_3 A = A$. Let $x \in M$, then $M = P(\tilde{M} \oplus R) = \tilde{M} \oplus P(R)$ implies that there exist $\tilde{m} \in \tilde{M}$ and $r \in R$, such that $x = \tilde{m} + Pr$. Then $A_M x = A_M \tilde{m} + A_M Pr = A\tilde{m} + APr = A\tilde{m} + Ar = A(\tilde{m} + r)$. Further, $T^{-1} A_M x = A_M x$, since $A_M x \in \text{Ran } A \subseteq \tilde{M}$, and $QA_M x = T_3 T^{-1} A_M x = T_3 A_M x = A_M x$, for the same reason.

On the other hand, $A_{M_3} Q x = A_{M_3} T_3 T^{-1}(\tilde{m} + Pr) = A_{M_3} T_3(\tilde{m} + r) = A_{M_3}(\tilde{m} + P_3 r) = A\tilde{m} + AP_3 r = A(\tilde{m} + r)$. Hence, $QA_M = A_{M_3} Q$. $\square$

5.2 Finite Rank Operators and Nests

In this section, we study the action of a finite rank operator on a (simple) nest of subspaces or projections. First, we introduce the notions of an invariant subspace and an invariant projection. For an introduction to the theory of invariant subspaces, see [17].

Let A be a bounded operator acting on a Banach space X. A subspace $M \subseteq X$ is an *invariant subspace* for A, if $x \in M$ implies that $Ax \in M$. Write $AM \subseteq M$. A nest of subspaces is called an *invariant nest* for A, if the nest consists of invariant subspaces for A. In general, a bounded operator acting on a Banach space need not have an invariant subspace, other than the trivial subspaces; (0) and X (cf. [25]). By the well-known result of Aronszajn and Smith [1] however, each compact operator acting on a Banach space of dimension greater than one has a non-trivial invariant subspace. It even follows (see for example Theorem 1 in [47]) that a compact operator has a simple nest of invariant subspaces. A projection P is *invariant* for the bounded operator A, if $AP = PAP$, or equivalently, if Ran P is an invariant subspace for A. Recall that only in a Banach space isomorphic to a Hilbert space, all subspaces are complemented, i.e., are the range of some bounded projection (cf. [39]). The existence of a Banach space on which only bounded projections act that are either finite rank or Fredholm [32], leads to the existence of a compact operator on that space, that has no invariant projections other than the trivial ones; O and I [44].

A nest of projections is called *invariant* for a bounded operator A, if all projections in the nest are invariant for A. The bounded operator A is called *upper triangular* with respect to a nest of projections or subspaces, if the nest is invariant for A. A bounded operator is called *lower triangular* with respect to a nest of projections $\mathcal{P}$, if it is upper triangular with respect to $\{I - P \mid P \in \mathcal{P}\}$, or in other words, upper triangular with respect to $\{\text{Ker } P \mid P \in \mathcal{P}\}$. Finally, a bounded operator is called *diagonal* with respect to a nest of projections, if it is both upper triangular and lower triangular with respect to the nest.

Let $\mathcal{M}$ be a complete nest of subspaces. For $M \in \mathcal{M}$, define

$$M_- = \text{span}\,\{L \mid L \in \mathcal{M}, L \subset M\}$$

and

$$M_+ = \bigcap \{N \mid N \in \mathcal{M}, M \subset N\}\,.$$

If there exists no $L \in \mathcal{M}$ such that $L \subset M$, then define $M_- = (0)$, and if there exists no $N \in \mathcal{M}$ such that $M \subset N$, then define $M_+ = X$. Further, if $M_- \neq M$, then $M_- \subset M$ is called the *immediate predecessor* of M in $\mathcal{M}$. In that case, $(M_-)_+ = M$. Indeed, $M_- \subset M$ gives

$$(M_-)_+ = \bigcap\{N \mid N \in \mathcal{M}, M_- \subset N\},$$

so $(M_-)_+ \subseteq M$. If $(M_-)_+ \subset M$ would hold, then there would exist $L \in \mathcal{M}$, with $M_- \subset L \subset M$, a contradiction. Analogously, if $M_+ \neq M$, then $M_+ \supset M$ is called the *immediate successor* of M, and $(M_+)_- = M$.

We now introduce some terminology and results from [47] or [48]. Consider a compact operator A acting on a Banach space, and let $\mathcal{M}$ be a simple invariant nest of subspaces for A. For each $M \in \mathcal{M}$, the quotient space M/M_- is at most one-dimensional. This follows from the fact that M_-, M is a gap in the simple nest $\mathcal{M}$, whenever $M_- \subset M$. If $M \in \mathcal{M}$ and $M_- \subset M$, then there exists a unique complex number α_M, such that $(A - \alpha_M)M \subseteq M_-$. On the other hand, if $M_- = M$, then $\alpha_M = 0$ gives $(A - \alpha_M)M = AM \subseteq M = M_-$.

The complex number α_M is called the *diagonal coefficient* of A at M. If α is a non-zero complex number, then the number of elements in $\{M \mid M \in \mathcal{M}, \quad \alpha_M = \alpha\}$ is called the *diagonal multiplicity* of α. We now state Theorem 2 in [47]; see also Theorem 4.3.10 in [48].

Proposition 5.2.1 *Let A be a compact operator, upper triangular with respect to a simple nest of subspaces $\mathcal{M}$, and let α be a non-zero complex number. Then α is a diagonal coefficient of A with respect to $\mathcal{M}$ if and only if it is an eigenvalue of A. Moreover, the diagonal multiplicity of α, as a diagonal coefficient of A with respect to $\mathcal{M}$, is finite and coincides with the algebraic multiplicity of α, as an eigenvalue of A.*

Let A be an operator of finite rank acting on the Banach space X and let $\mathcal{M}$ be a complete nest of subspaces. For the moment, we do *not* assume that $\mathcal{M}$ is invariant for A. Along with the finite rank operator A, define the mapping $d_A : \mathcal{M} \longrightarrow \{0, \ldots, \text{rank}A\}$ as

$$d_A(M) = \dim(AM), \quad M \in \mathcal{M}. \tag{5.3}$$

A mapping $d : \mathcal{M} \longrightarrow [-\infty, \infty]$ is called *monotonically increasing*, if $M_1, M_2 \in \mathcal{M}$ and $M_1 \subseteq M_2$ imply that $d(M_1) \leq d(M_2)$. The monotonically increasing mapping $d : \mathcal{M} \longrightarrow [-\infty, \infty]$ is called *left continuous* at $M \in \mathcal{M}$, if for each subnest $\mathcal{M}_1 \subseteq \mathcal{M}$, with $M = \bigvee \mathcal{M}_1$, it follows

$$\sup\{d(L) \mid L \in \mathcal{M}_1\} = d(M).$$

Lemma 5.2.2 *Let $\mathcal{M}$ be a complete nest of subspaces in the Banach space X, and let A be an operator of finite rank on X. The mapping d_A, as defined in (5.3), is monotonically increasing and left continuous on $\mathcal{M}$.*

Proof It is immediate, that d_A is monotonically increasing on $\mathcal{M}$. To prove that d_A is left continuous at $M \in \mathcal{M}$, let $\mathcal{M}_1 \subseteq \mathcal{M}$ be a non-empty subnest, such that $M = \bigvee \mathcal{M}_1$. If $x \in M$, there exist $M_n \in \mathcal{M}_1$ and $x_n \in M_n$ for $n \in \mathbf{Z}^+$, such that $x_n \xrightarrow{\|\cdot\|} x$. Consequently, $Ax_n \xrightarrow{\|\cdot\|} Ax$, so

$$AM \subseteq cl\left(\bigcup\{AL \mid L \in \mathcal{M}_1\}\right).$$

Since $\dim(AL) \in \{0, \dots, \text{rank } A\}$ for $L \in \mathcal{M}_1$, the set $\{AL \mid L \in \mathcal{M}_1\}$ has a finite number of elements. Therefore,

$$cl\left(\bigcup\{AL \mid L \in \mathcal{M}_1\}\right) = AL_1$$

for some $L_1 \in \mathcal{M}_1$. Since $L_1 \subseteq M$, we get $AL_1 \subseteq AM$, so $AL_1 = AM$. The lemma is proved. □

Write $d_A(\mathcal{M}) = \{d_0, d_1, \dots, d_r\}$, with $d_{j-1} < d_j$ for $j = 1, \dots, r$. In particular, $d_0 = 0$ and $d_r = \text{rank } A$. Define

$$\mathcal{M}_j = \{M \mid M \in \mathcal{M}, d_A(M) = d_j\}, \quad j = 0, \dots, r. \tag{5.4}$$

Note that $\mathcal{M}_j \neq \emptyset$, and that $\mathcal{M}_i \cap \mathcal{M}_j = \emptyset$, if $i \neq j$. We state some additional properties of this partition of $\mathcal{M}$.

Lemma 5.2.3 *Let $\mathcal{M}$ be a complete nest in X, A a finite rank operator on X, and $\mathcal{M}_j$ for $j = 0, \dots, r$, as defined in (5.4). Then*

1. *$\bigvee \mathcal{M}_j \in \mathcal{M}_j$.*
2. *If $\bigwedge \mathcal{M}_j \notin \mathcal{M}_j$, then $\bigwedge \mathcal{M}_j = \bigvee \mathcal{M}_{j-1}$.*
3. *The pair $\bigvee \mathcal{M}_{j-1}, \bigwedge \mathcal{M}_j$ is a gap in $\mathcal{M}$.*

Moreover, if the nest $\mathcal{M}$ is simple, then the quotient spaces $\bigwedge \mathcal{M}_j / \bigvee \mathcal{M}_{j-1}$ are at most one-dimensional for $j = 1, \dots, r$.

Proof The first statement follows from the continuity to the left of d_A on $\mathcal{M}$. Assume that $\bigwedge \mathcal{M}_j \notin \mathcal{M}_j$. Then $j > 0$. Since $\bigvee \mathcal{M}_{j-1} \subset M$ for all $M \in \mathcal{M}_j$, $\bigvee \mathcal{M}_{j-1} \subseteq \bigwedge \mathcal{M}_j$. Consequently, $d_{j-1} \leq d_A(\bigwedge \mathcal{M}_j) < d_j$, so $\bigwedge \mathcal{M}_j \in \mathcal{M}_{j-1}$. But then $\bigwedge \mathcal{M}_j \subseteq \bigvee \mathcal{M}_{j-1}$, and the second statement follows. To prove the third statement, let $M \in \mathcal{M}$ be a subspace, such that $\bigvee \mathcal{M}_{j-1} \subset M \subset \bigwedge \mathcal{M}_j$. Then $\bigvee \mathcal{M}_{j-1} \subset M$ implies that $d_A(M) > d_{j-1}$ and $M \subset \bigwedge \mathcal{M}_j$ implies that $d_A(M) < d_j$. Therefore, $d_A(M) \notin d_A(\mathcal{M})$, a contradiction. The lemma is proved. □

Theorem 5.2.4 *Let A be an operator of finite rank acting on the Banach space X. Let $\mathcal{M}$ be a simple invariant nest for A. Then there exists a finite subnest $\{M_k \mid k = 0, \ldots, n\}$ of $\mathcal{M}$, and complex numbers $\alpha_1, \ldots, \alpha_n$, such that the following hold:*

1. $M_0 = (0), \quad M_n = X, \quad M_{k-1} \subset M_k, \quad k = 1, \ldots, n.$
2. $n \leq 1 + 2 \text{ rank } A.$
3. $$(A - \alpha_k)M_k \subseteq M_{k-1}, \quad k = 1, \ldots, n. \tag{5.5}$$
4. *If* $\dim(M_k/M_{k-1}) \geq 2$, *then* $\alpha_k = 0$.

Proof Let $\tilde{\mathcal{M}}$ be the finite subnest of $\mathcal{M}$ defined by

$$\tilde{\mathcal{M}} = \left\{ \bigwedge \mathcal{M}_j, \bigvee \mathcal{M}_j \mid j = 0, \ldots, r \right\}.$$

It is sufficient to prove the following. If $(0) \neq M \in \tilde{\mathcal{M}}$, then there exist $\hat{M} \in \tilde{\mathcal{M}}$, with $\hat{M} \subset M$, and $\alpha_M \in \mathbf{C}$, such that $(A - \alpha_M)M \subseteq \hat{M}$. If in addition, $\alpha_M \neq 0$, then $\dim(M/\hat{M}) = 1$. To prove this, we distinguish several cases.

Case 1 $M = \bigvee \mathcal{M}_j = \bigwedge \mathcal{M}_j$. We will assume that $M \neq (0)$ and hence that $j > 0$. By the remark made before Proposition 5.2.1, there exists $\alpha_M \in \mathbf{C}$, such that

$$(A - \alpha_M)M = (A - \alpha_M) \bigvee \mathcal{M}_j = (A - \alpha_M) \bigwedge \mathcal{M}_j \subseteq$$

$$\left(\bigwedge \mathcal{M}_j \right)_{-} \subseteq \bigvee \mathcal{M}_{j-1} \in \mathcal{M}_{j-1}.$$

Define $\hat{M} = \bigvee \mathcal{M}_{j-1}$. Then $\hat{M} \subset M$, and $(A - \alpha_M)M \subseteq \hat{M}$. Furthermore, by the second part of Lemma 5.2.3, $\dim(M/\hat{M}) = 1$.

Case 2 $M = \bigvee \mathcal{M}_j \supset \bigwedge \mathcal{M}_j$. Then

$$AM = A \bigvee \mathcal{M}_j \subseteq \bigwedge \mathcal{M}_j,$$

since $A(\bigvee \mathcal{M}_j) = AL \subseteq L$ for all $L \in \mathcal{M}_j$. Take $\hat{M} = \bigwedge \mathcal{M}_j$.

Case 3 $M = \bigwedge \mathcal{M}_j = \bigvee \mathcal{M}_{j-1}$. This case is dealt with under Case 1 or Case 2, since $M = \bigvee \mathcal{M}_{j-1}$.

Case 4 $M = \bigwedge \mathcal{M}_j \supset \bigvee \mathcal{M}_{j-1}$. There exists $\alpha_M \in \mathbf{C}$, such that

$$(A - \alpha_M)M = (A - \alpha_M) \bigwedge \mathcal{M}_j \subseteq (\bigwedge \mathcal{M}_j)_- = \bigvee \mathcal{M}_{j-1}.$$

Take $\hat{M} = \bigvee \mathcal{M}_{j-1}$. The second part of Lemma 5.2.3 gives $\dim(M/\hat{M}) = 1$.

The number of elements in $\tilde{\mathcal{M}}$ does not exceed $2(r + 1) \leq 2(\text{rank } A + 1)$. Therefore, if $\tilde{\mathcal{M}} = \{M_k\}_{k=0}^{n}$, then

$$n \leq 1 + 2 \text{ rank } A.$$

The theorem is established. □

The following example illustrates the proof of Theorem 5.2.4. It also follows from the example, that the number $2(\text{rank } A + 1)$ in some cases is the minimal number of subspaces required for a finite subnest to have the properties mentioned in the theorem.

Example 5.2.5 Consider the $(2m + 1) \times (2m + 1)$ upper triangular matrix

$$A = \left(\begin{array}{ccccccccc}
0 & 1 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\
 & 1 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\
 & & 0 & 1 & & & \vdots & \vdots & \vdots \\
 & & & 1 & & & \vdots & \vdots & \vdots \\
 & & & & \ddots & & \vdots & \vdots & \vdots \\
 & & & & & \ddots & 0 & 0 & 0 \\
 & & & & & & 0 & 1 & 0 \\
 & & & & & & & 1 & 0 \\
 & & & & & & & & 0
\end{array}\right),$$

and let M_k denote the span of the first k standard basis vectors. Since A is upper triangular, $AM_k \subseteq M_k$ for $0 \leq k \leq 2m+1$. Note that rank $A = m$, and define, as in (5.4),

$$\mathcal{M}_j = \{M_{2j}, M_{2j+1}\}, \quad j = 0, \ldots, m.$$

The finite subnest $\tilde{\mathcal{M}}$, as indicated in the proof, coincides here with the original maximal nest of subspaces $\{M_k\}_{k=0}^{2m+1}$. Note that

$$AM_{2j+1} \subseteq M_{2j}, \quad AM_{2j+1} \not\subseteq M_{2j-1},$$

and

$$(A - I)M_{2j} \subseteq M_{2j-1}, \quad (A - I)M_{2j} \not\subseteq M_{2j-2}, \quad j = 1, \ldots, m.$$

Therefore, the nest $\tilde{\mathcal{M}}$ does not properly contain a subnest that has the required properties.

Proposition 5.2.6 *Let A be a bounded operator, and let $\mathcal{M}_1 = \{M_k\}_{k=0}^n$ be a finite nest of subspaces, with $M_{k-1} \subset M_k$ for $k = 1 \ldots, n$. Assume there exist complex numbers $\alpha_1, \ldots, \alpha_n$, such that*

$$(A - \alpha_k)M_k \subseteq M_{k-1}, \quad k = 1, \ldots, n.$$

If $\mathcal{M}$ is any nest of subspaces which contains $\mathcal{M}_1$, then for each non-zero $M \in \mathcal{M}$, there exist a complex number α_M and an $\hat{M} \in \mathcal{M}_1$, with $\hat{M} \subset M$, such that $(A - \alpha_M)M \subseteq \hat{M}$. In particular, the nest $\mathcal{M}$ is invariant for A.

Proof Let $\mathcal{M}$ be a nest of subspaces which contains $\mathcal{M}_1$, and let $M \in \mathcal{M}$, $M \neq (0)$. Let $1 \leq k \leq n$, such that $M_{k-1} \subset M \subseteq M_k$. It follows that

$$(A - \alpha_k)M \subseteq (A - \alpha_k)M_k \subseteq M_{k-1},$$

so let $\hat{M} = M_{k-1}$ and $\alpha_M = \alpha_k$. □

We reformulate Theorem 5.2.4 in term of nests of projections (instead of subspaces).

Proposition 5.2.7 *Let A be an operator of finite rank acting on the Banach space X. Let $\mathcal{P}$ be a simple invariant nest of projections for A. Then there exists a finite subnest $\{P_k \mid k = 0, \ldots, n\}$ of $\mathcal{P}$, and complex numbers $\alpha_1, \ldots, \alpha_n$, such that the following hold:*

1. $P_0 = O_X, \quad P_n = I_X, \quad P_{k-1} < P_k, \quad k = 1, \ldots, m.$
2. $n \leq 1 + 2 \text{ rank } A.$
3. $$(A - \alpha_k)P_k = P_{k-1}(A - \alpha_k)P_k, \quad k = 1, \ldots, n. \tag{5.6}$$
4. *If* $\text{rank}(P_k - P_{k-1}) \geq 2$, *then* $\alpha_k = 0$.

Proof Apply Theorem 5.2.4 on the finite rank operator A and its nest of invariant subspaces $\{\text{Ran } P \mid P \in \mathcal{P}\}$, which is simple also, according to Theorem 4.1.3. We obtain a finite nest of subspaces $(0) = M_0 \subset M_1 \subset \cdots \subset M_n = X$, and complex numbers $\alpha_1, \ldots, \alpha_n$ such that (5.5) holds. For each $k \in \{0, \ldots, n\}$, there exists exactly one projection $P_k \in \mathcal{P}$, such that $\text{Ran } P_k = M_k$. In this manner, equation (5.5) implies equation (5.6). The proposition is proved. □

An operator of finite rank n, acting on a Banach space, and upper triangular with respect to a complete nest $\mathcal{M}$ of subspaces, is the sum of n operators of rank one, all upper triangular with respect to $\mathcal{M}$. This result, due to Spanoudakis [51], is related to Theorem 5.2.4.

5.3 Complementary Triangular Forms

In this section, we will study complementary triangular forms for pairs of finite rank operators acting on a Banach space X. We first present a definition of complementary triangular forms for pairs of bounded operators in general.

The collection $\mathcal{C}(X)$ consists of pairs of bounded operators A, Z, such that there exists a simple nest of projections $\mathcal{P}$, with $AP = PAP$ and $PZ = PZP$, for all $P \in \mathcal{P}$.

If A and Z are finite rank operators acting on X, then there exist several matrix reductions for this pair (see Section 5.1). Proposition 5.1.3 in that section describes how all these matrix reductions are related. We will now define complementary triangular forms for a pair of finite rank operators, using a matrix reduction.

The collection $\mathcal{C}_f(X)$ consists of pairs of finite rank operators A, Z, such that there exists a matrix reduction $(M, N, \{A_M, Z_M\})$ for the pair, with $(A_M, Z_M) \in \mathcal{C}(M)$.

The crux of this definition is that M is finite dimensional, so that the pair A_M, Z_M can be identified with a pair of finite matrices. Lemma 3.1.4 then shows that the collections $\mathcal{C}(M)$ and $\mathcal{C}(m)$ can be identified, where $m = \dim M$.

Theorem 5.3.1 *Let A and Z be operators of finite rank acting on the Banach space X. Then $(A, Z) \in \mathcal{C}(X)$ implies that $(A, Z) \in \mathcal{C}_f(X)$.*

Proof Assume that $(A, Z) \in \mathcal{C}(X)$, i.e., there exists a simple nest of projections $\mathcal{P}$, such that $AP = PAP$ and $PZ = PZP$ for all $P \in \mathcal{P}$. If we apply Proposition 5.2.7 to A and $\mathcal{P}$, we obtain a finite subnest $\mathcal{P}_A \subseteq \mathcal{P}$, which is invariant for A as in (5.6), such that $\mathcal{P}_A$ contains at most $2 + 2$ rank A elements. In the same fashion we obtain a finite subnest $\mathcal{P}_Z \subseteq \mathcal{P}^c$ which is invariant for Z as in (5.6), containing at most $2 + 2$ rank A elements. Consider the finite subnest

$$\mathcal{P}_1 = \mathcal{P}_A \cup \mathcal{P}_Z^c \subseteq \mathcal{P}.$$

Write $\mathcal{P}_1 = \{P_0, P_1, \ldots, P_n\}$, with $O_X = P_0 < P_1 < \cdots < P_n = I_X$. By Proposition 5.2.6, applied to A and $\{\text{Ran } P_k \mid k = 0, \ldots, n\}$, and to Z and $\{\text{Ker } P_k \mid k = 0, \ldots, n\}$, it follows that there exist complex numbers $\alpha_1, \ldots, \alpha_n$, and $\zeta_1, \ldots, \zeta_n$, such that

$$(A - \alpha_k)P_k = P_{k-1}(A - \alpha_k)P_k, \quad P_k(Z - \zeta_k) = P_k(Z - \zeta_k)P_{k-1},$$

where the integer $n \leq 1 + 2 \text{ rank } A + 2 \text{ rank } Z$. Write $\Delta P_k = P_k - P_{k-1}$ for $k = 1, \ldots, n$. Define the finite dimensional subspace

$$M_k = \Delta P_k \left(\text{Ran } A + \text{Ran } Z\right), \quad k = 1, \ldots, n,$$

and define the space of finite codimension

$$N_k = \text{Ker}(A\Delta P_k) \cap \text{Ker } (Z\Delta P_k), \quad k = 1, \ldots, n.$$

Then $N_k = \text{Ker } \Delta P_k \oplus \tilde{N}_k$, where $\tilde{N}_k = N_k \cap \text{Ran } \Delta P_k$. Note that $M_k + \tilde{N}_k$ is a subspace of finite codimension in Ran ΔP_k. Therefore, there exists a finite dimensional subspace $R_k \subseteq \text{Ran } \Delta P_k$, such that

$$\left(M_k + \tilde{N}_k\right) \oplus R_k = \text{Ran } \Delta P_k.$$

In addition, let $\hat{N}_k \subseteq N_k$, such that

$$\hat{M}_k \oplus \hat{N}_k = \text{Ran } \Delta P_k,$$

where $\hat{M}_k = M_k \oplus R_k$. Define

$$M = M_1 \oplus \cdots \oplus M_n, \quad R = R_1 \oplus \cdots \oplus R_n,$$

$$N = \tilde{N}_1 \oplus \cdots \oplus \tilde{N}_n, \quad \hat{N} = \hat{N}_1 \oplus \cdots \oplus \hat{N}_n.$$

As a matter of fact,

$$N = \bigcap_{k=1}^{n} N_k.$$

Indeed, if $x = x_1 + \cdots + x_n$ with $x_k \in \tilde{N}_k$, then $m_k = x - x_k \in \text{Ker } \Delta P_k$ and hence $x = x_k + m_k \in \tilde{N}_k + \text{Ker } \Delta P_k$. On the other hand, if $x = x_k + m_k$ with $x_k \in \tilde{N}_k$ and $m_k \in \text{Ker } \Delta P_k$, then $x = (\sum_{k=1}^{n} \Delta P_k)x = \sum_{k=1}^{n} x_k \in \tilde{N}_1 \oplus \cdots \oplus \tilde{N}_n$. Define $\hat{M} = M \oplus R$, then

$$\text{Ran } A + \text{Ran } Z \subseteq \hat{M}.$$

Since $\hat{N} \subseteq N = \bigcap_{k=1}^{n} \left[\text{Ker}(A\Delta P_k) \cap \text{Ker}(Z\Delta P_k)\right]$, we get

$$\text{Ker } A \cap \text{Ker } Z \supseteq \hat{N}.$$

Also,

$$\hat{M} \oplus \hat{N} = X.$$

Since $A\Delta P_k = \sum_{j=1}^{k-1} \Delta P_j A \Delta P_k + \alpha_k \Delta P_k$, it follows that

$$A\hat{M}_k \subseteq \text{Ran}(A\Delta P_k) \subseteq \sum_{j=1}^{k-1} \hat{M}_j \oplus \text{Ran}(\alpha_k \Delta P_k).$$

In addition, $Z\Delta P_k = \sum_{j=k+1}^{n} \Delta P_j Z \Delta P_k + \zeta_k \Delta P_k$ implies

$$Z\hat{M}_k \subseteq \text{Ran}(Z\Delta P_k) \subseteq \sum_{j=k+1}^{n} \hat{M}_j \oplus \text{Ran}(\zeta_k \Delta P_k).$$

With respect to the decomposition

$$X = \hat{M}_1 \oplus \cdots \hat{M}_n \oplus \hat{N},$$

the finite rank operators A and Z assume the forms

$$A = \begin{pmatrix} \alpha_1 \hat{I}_1 & * & * & \cdots & * & O \\ O & \alpha_2 \hat{I}_2 & * & \cdots & * & O \\ O & O & \alpha_3 \hat{I}_3 & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \ddots & * & O \\ \vdots & & & \ddots & \alpha_n \hat{I}_n & O \\ O & \cdots & \cdots & \cdots & O & O_{\hat{N}} \end{pmatrix},$$

and

$$Z = \begin{pmatrix} \zeta_1 \hat{I}_1 & O & O & \cdots & \cdots & O \\ * & \zeta_2 \hat{I}_2 & O & & & \vdots \\ * & * & \zeta_3 \hat{I}_3 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ * & * & \cdots & * & \zeta_n \hat{I}_n & O \\ O & O & \cdots & O & O & O_{\hat{N}} \end{pmatrix}.$$

Here $\hat{I}_k$ denotes the identity operator on the (finite dimensional) subspace $\hat{M}_k$. We shall estimate the dimension of $\hat{M}$ from above. First of all,

$$\dim M = \sum_{k=1}^{n} \dim M_k \le n \dim(\text{Ran } A + \text{Ran } Z).$$

Further,

$$\text{codim } N \le \sum_{k=1}^{n} \text{codim } N_k \le n \text{ codim}\,(\text{Ker } A \cap \text{Ker } Z)\,.$$

Then $(M + N) \oplus R = X$ implies $\dim R \le \text{codim } N$. Therefore,

$$\dim \hat{M} = \dim M + \dim R \le$$

$$n \ \left(\dim[\text{Ran } A + \text{Ran } Z] + \text{codim}\,(\text{Ker } A \cap \text{Ker } Z)\right).$$

Finally, use that $n \le 1 + 2 \text{ rank } A + 2 \text{ rank } Z$. The theorem is proved. □

The other inclusion $\mathcal{C}_f(X) \subseteq \mathcal{C}(X)$ holds true, if the underlying Banach space X has the following geometric property: On each subspace $Y \subset X$ of finite co-dimension acts a simple nest of projections. It is not difficult to see that Hilbert spaces and Banach spaces with a Schauder basis have this geometric property.

Let m_1 be a positive integer and let m_2 be a nonnegative integer. Recall (see Chapter 3) that a pair of $m_1 \times m_1$ matrices A_1, Z_1 admits simultaneous reduction to complementary triangular forms after extension with m_2 zeroes, if the pair $A_1 \oplus O_{m_2}, Z_1 \oplus O_{m_2}$ admits simultaneous reduction to complementary triangular forms. If such a nonnegative integer m_2 exists for the pair A_1, Z_1, we say that A_1 and Z_1 admits simultaneous reduction to complementary triangular forms after extension with zeroes. The following proposition shows that this notion and the collection $\mathcal{C}_f(X)$ are closely related.

Proposition 5.3.2 *Let A and Z be finite rank operators acting on the infinite dimensional Banach space X. Let $(\{A_M, Z_M\}, M, N)$ be a matrix reduction for A and Z. Then the following are equivalent:*

1. *$(A, Z) \in \mathcal{C}_f(X)$.*
2. *The pair A_M, Z_M admits simultaneous reduction to complementary triangular forms after extension with zeroes.*

Proof To prove that the first statement implies the second one, consider a matrix reduction $(\{A_{M_0}, Z_{M_0}\}, M_0, N_0)$ for A and Z, such that the pair A_{M_0}, Z_{M_0} admits simultaneous reduction to complementary triangular forms. We may assume without loss of generality, that dim $M_0 \geq$ dim M. Indeed, if this is not the case, choose a positive integer m_2, such that $m_2 + \dim M_0 \geq$ dim M. Then choose an m_2-dimensional subspace $N_2 \subseteq N_0$. Further, let $N_1 \subseteq N_0$, such that $N_1 \oplus N_2 = N_0$, and define $M_1 = M_0 \oplus N_2$. In this manner, we have obtained a matrix reduction $(\{A_{M_1}, Z_{M_1}\}, M_1, N_1)$ for A and Z, with dim $M_1 \geq$ dim M, and such that the pair A_{M_1}, Z_{M_1} admits simultaneous reduction to complementary triangular forms.

Write dim $M = m$, dim $M_0 = m_0$, and $m_0 - m = m_2$. Then m_2 is a nonnegative integer. Let $N_2 \oplus N_0 = N$ be a decomposition of N, with dim $N_2 = m_2$. By Proposition 5.1.3, there exists an invertible operator $Q : M_0 \longrightarrow M \oplus N_2$, such that

$$QA_{M_0}Q^{-1} = A_M \oplus O_{N_2}, \quad QZ_{M_0}Q^{-1} = Z_M \oplus O_{N_2}.$$

It follows that the pair A_M, Z_M admits simultaneous reduction to complementary triangular forms after extension with m_2 zeroes.

To prove the converse, assume that m_2 is a nonnegative integer, and that the pair $A_M \oplus O_{m_2}, Z_M \oplus O_{m_2}$ admits simultaneous reduction to complementary triangular forms. Let $N_2 \oplus N_1 = N$ be a decomposition of N, with dim $N_2 = m_2$. Define $M_1 = M \oplus N_2$. Note that

$$A_{M_1} = A_M \oplus O_{N_2}, \quad Z_{M_1} = Z_M \oplus O_{N_2}.$$

It follows that $(\{A_{M_1}, Z_{M_1}\}, M_1, N_1)$ is a matrix reduction for A and Z, and that the pair A_{M_1}, Z_{M_1} admits simultaneous reduction to complementary triangular forms. □

If a pair of $m_1 \times m_1$ matrices A_1, Z_1 admits simultaneous reduction to complementary triangular forms after extension with zeroes, it is interesting to know the required number of extended zeroes. The following result, a corollary to Theorem 5.3.1, provides an estimate of this number from above.

Corollary 5.3.3 *Let m_1 be a positive integer, and let A_1 and Z_1 be $m_1 \times m_1$ matrices. Assume that the set*

$$\{m_1 + m_2 \mid m_2 \in \mathbf{Z}_0^+, \quad (A_1 \oplus O_{m_2}, Z_1 \oplus O_{m_2}) \in \mathcal{C}(m_1 + m_2)\}$$

is non-empty. Then the infimum $\rho_0(A_1, Z_1)$ taken over this set satisfies

$$\rho_0(A_1, Z_1) \leq 2m_1(4m_1 - 1).$$

Proof Let m_2 be a nonnegative integer such that the pair $A = A_1 \oplus O_{m_2}, Z = Z_1 \oplus O_{m_2}$ admits simultaneous reduction to complementary triangular forms. If both A_1 and Z_1 are invertible, Theorem 3.3.2 gives $(A_1, Z_1) \in \mathcal{C}(m_1)$, so $\rho_0(A_1, Z_1) = m_1$, and we are done. Therefore, we may assume that rank $A +$ rank $Z \leq 2\ m_1 - 1$.

Write $m = m_1 + m_2$ and apply Theorem 5.3.1 on the $m \times m$ matrices A and Z, acting on $X = \mathbf{C}^{m_1} \oplus \mathbf{C}^{m_2}$. It follows, that there exists a matrix reduction $(\{A_M, Z_M\}, M, N)$ for A and Z, such that $(A_M, Z_M) \in \mathcal{C}(M)$. We may assume without loss of generality, that dim $M \geq m_1$. Let $\nu =$ dim $M - m_1$.

The triples $(\{A_1, Z_1\}, \mathbf{C}^{m_1}, \mathbf{C}^{m_2})$ and $(\{A_M, Z_M\}, M, N)$ both are matrix reductions for A and Z. Note that dim $M - m_1 = \nu \geq 0$. By Proposition 5.1.3, there exists an invertible operator $Q : M \longrightarrow \mathbf{C}^{m_1} \oplus \mathbf{C}^{\nu}$, such that

$$QA_MQ^{-1} = A_1 \oplus O_\nu, \quad QZ_MQ^{-1} = Z_1 \oplus O_\nu.$$

It follows that $(A_1 \oplus O_\nu, Z_1 \oplus O_\nu) \in \mathcal{C}(m_1 + \nu)$, since $(A_M, Z_M) \in \mathcal{C}(M)$. Therefore,

$$\rho_0(A_1, Z_1) \leq m_1 + \nu = \text{dim } M.$$

The proof of Theorem 5.3.1 provides

$$\text{dim } M \leq 2m_1(4m_1 - 1),$$

where we used that

$$1 + 2 \text{ rank } A + 2 \text{ rank } Z \leq 4m_1 - 1,$$

$$\text{dim}(\text{Ran } A + \text{Ran } Z) \leq m_1, \quad \text{codim}(\text{Ker } A \cap \text{Ker} Z) \leq m_1.$$

This proves the corollary. □

Chapter 6

Diagonalizable Operators

In this chapter, analogues to Theorem 2.2.1 are discussed for pairs of bounded operators acting on the separable Hilbert space $l_2(\mathbf{Z}^+)$. Indeed, if a pair of bounded operators on this space contains a finite rank operator and a diagonalizable operator, then the pair admits simultaneous reduction to complementary triangular forms, whenever an obvious necessary condition is met. This is shown by Propositions 6.4.1 and 6.4.2, and the remarks after the respective proofs.

On the other hand, there exist pairs of diagonalizable operators that do not admit simultaneous reduction to complementary triangular forms (Theorem 6.3.1). The pairs of operators under consideration can even be taken self-adjoint and of trace-class. The construction of these examples is based on the existence of a unitary infinite matrix that does not admit lower-upper factorization, even after independently permuting rows and columns (Theorem 6.2.2). Another counterexample, that contradicts even a wider analogue of Theorem 2.2.1, is presented in Theorem 6.3.4.

6.1 Preliminaries

Consider the separable Hilbert space $l_2(\mathbf{Z}^+)$ with standard orthonormal basis $\{e_k\}_{k=1}^{\infty}$. Denote the orthoprojector (of rank m) on $\text{span}\{e_1, \ldots, e_m\}$ by E_m. Let

$$\mathcal{E}_+ = \{E_m \mid m \in \mathbf{Z}^+\} \cup \{O, I\}$$

denote the standard nest of orthoprojectors in $l_2(\mathbf{Z}^+)$.

A bounded operator T on $l_2(\mathbf{Z}^+)$ can be represented as an infinite matrix $(T_{kl})_{k,l=1}^{\infty}$, with respect to $\{e_k\}_{k=1}^{\infty}$, where $T_{kl} = e_k^* T e_l$. Of course, not every infinite matrix represents a bounded operator on $l_2(\mathbf{Z}^+)$. The *diagonal* of T is given by

$$\text{diag } T = (\ T_{11},\ T_{22},\ T_{33},\ \cdots\ )^T .$$

A bounded operator T is upper triangular with respect to $\mathcal{E}_+$, if $TE_m = E_mTE_m$ for all $m \in \mathbf{Z}^+$, i.e., if its infinite matrix as defined above is upper triangular. In the same fashion, lower triangular and diagonal operators (with respect to $\mathcal{E}_+$) are described. An operator is called *upper triangularizable*, if there exists an invertible operator S on $l_2(\mathbf{Z}^+)$, such that $S^{-1}AS$ is upper triangular with respect to $\mathcal{E}_+$. In the same fashion, *lower triangularizable* and *diagonalizable* operators are defined. Not all operators on a separable Hilbert space are upper or lower triangularizable. Indeed, both the Volterra operator of integration V on $L_2(0,1)$, given by

$$Vf(x) = \int_0^x f(t)dt,$$

and its adjoint V^* have no eigenvalues. Therefore, it is neither upper nor lower triangularizable. We will now define standard complementary triangular forms with respect to $\mathcal{E}_+$.

The collection $\mathcal{C}_+$ consists of those pairs of bounded operators A, Z on $l_2(\mathbf{Z}^+)$, such that there exists an invertible operator S on $l_2(\mathbf{Z}^+)$, with $S^{-1}AS$ upper triangular and $S^{-1}ZS$ lower triangular with respect to $\mathcal{E}_+$.

If τ is a permutation (i.e., bijection) on $\mathbf{Z}^+$, then the *permutation operator* U_τ is defined as $U_\tau e_k = e_{\tau(k)}$ for $k \in \mathbf{Z}^+$. This notion is analogous to the notion of a permutation matrix, as defined in [38], p. 64. The following lemma is easy to prove.

Lemma 6.1.1 *Let A and Z be bounded operators acting on $l_2(\mathbf{Z}^+)$, and let T be an invertible operator on $l_2(\mathbf{Z}^+)$. Then the following are equivalent:*

1. $(A, Z) \in \mathcal{C}_+$,
2. $(Z^*, A^*) \in \mathcal{C}_+$,
3. $(T^{-1}AT, T^{-1}ZT) \in \mathcal{C}_+$.

In contrast to the finite matrix situation, $(A, Z) \in \mathcal{C}_+$ does not imply that $(Z, A) \in \mathcal{C}_+$, since there exist upper triangular operators A, which are not lower triangularizable; see Proposition 6.3.5.

6.2 Lower-Upper Factorization

In this section, we discuss lower-upper factorization of invertible operators and explain how this notion is used in the study of simultaneous reduction to complementary triangular forms. This approach is motivated by a proof of Theorem 2.2.1. As promised in Section 2.2, we now present this proof (see also [7]).

Proof of Theorem 2.2.1 We may assume without loss of generality, that A is a diagonalizable $m \times m$ matrix, and Z is any $m \times m$ matrix. Let T, V be invertible $m \times m$ matrices, such that $T^{-1}AT$ is a diagonal matrix, and $V^{-1}ZV$ is a lower triangular matrix.

It is well-known that for the invertible $m \times m$ matrix $V^{-1}T$, there exists an $m \times m$ permutation matrix U_τ, such that $V^{-1}TU_\tau = LR^{-1}$, where L is an invertible lower triangular matrix, and R is an invertible upper triangular matrix. Define $S = VL = TU_\tau R$. Then

$$S^{-1}AS = R^{-1}(U_\tau^{-1}T^{-1}ATU_\tau)R$$

is upper triangular, since $U_\tau^{-1}T^{-1}ATU_\tau$ is a diagonal matrix, and R, R^{-1} are upper triangular matrices. Further,

$$S^{-1}ZS = L^{-1}(V^{-1}ZV)L$$

is the product of lower triangular matrices, and hence lower triangular. □

The proof of Theorem 2.2.1 suggests a connection between lower-upper factorization and simultaneous reduction to complementary triangular forms. Before we study this connection in more detail in the infinite dimensional context, we give a definition of lower-upper factorization in the general setting. Note that triangular forms with respect to an arbitrary nest of projections are defined at the beginning of Section 5.2.

An invertible operator S on a Banach space X admits a *lower-upper factorization* along a nest of projections $\mathcal{P}$ (e.g. $\mathcal{P} = \mathcal{E}_+$), if there exist invertible operators L and R, such that $S = LR$, where L, L^{-1} are lower triangular and R, R^{-1} are upper triangular operators with respect to $\mathcal{P}$.

The argument of the proof of Theorem 2.2.1 suggests the following proposition for pairs of bounded operators acting on $l_2(\mathbf{Z}^+)$.

Proposition 6.2.1 *Let A and Z be bounded operators acting on the Hilbert space $l_2(\mathbf{Z}^+)$. Then $(A, Z) \in \mathcal{C}_+$ if and only if there exist invertible operators T and V, such that $T^{-1}AT$ is upper triangular, and $V^{-1}ZV$ is lower triangular with respect to $\mathcal{E}_+$, and $V^{-1}T$ admits a lower-upper factorization along $\mathcal{E}_+$.*

Although Proposition 6.2.1 is formulated for operators on a separable infinite dimensional Hilbert space, a straightforward analogue of Theorem 2.2.1 does not follow. The reason for this is that, contrary to the finite matrix situation, not every infinite matrix which represents an invertible bounded operator admits a lower-upper factorization after permutation of rows and columns.

Indeed, fix $n \in \mathbf{Z}^+$ and consider the $n \times n$ matrix

$$U(n) = \frac{\sqrt{n}}{n}\left(e^{\frac{2\pi i(u-1)(v-1)}{n}}\right)_{u,v=1}^{n}.$$

This matrix, known as the *Fourier matrix* , is unitary: If $\theta = e^{\frac{2\pi i}{n}}$, then

$$\{U(n)U(n)^*\}_{kl} = \frac{1}{n}\sum_{j=1}^{n} U(n)_{kj}\overline{U(n)_{lj}} =$$

$$\frac{1}{n}\sum_{j=1}^{n}\theta^{(k-1)(j-1)}\theta^{-(l-1)(j-1)} = \frac{1}{n}\sum_{j=1}^{n}(\theta^{k-l})^{j-1}.$$

If $k = l$, then $\{U(n)U(n)^*\}_{kl} = 1$. If $k \neq l$, then $\lambda = \theta^{k-l}$ satisfies $\lambda^n - 1 = 0$ and $\lambda \neq 1$. Hence

$$\frac{\lambda^n - 1}{\lambda - 1} = \sum_{j=1}^{n}\lambda^{j-1} = 0.$$

Therefore, $\{U(n)U(n)^*\}_{kl} = 0$. We conclude that $U(n)U(n)^* = I_n$. Further, $U(n)$ diagonalizes $n \times n$ so-called circulant matrices; see [21], Section 3.2. Using Fourier matrices of increasing sizes, we define the unitary operator U acting on $l_2(\mathbf{Z}^+)$ with respect to the standard basis $\{e_k \mid k \in \mathbf{Z}^+\}$ as the infinite diagonal block-matrix

$$U = U(1) \oplus U(2) \oplus U(3) \oplus \cdots. \tag{6.1}$$

Theorem 6.2.2 *The unitary operator U on $l_2(\mathbf{Z}^+)$ as defined in (6.1) satisfies the following: For all permutation operators U_ρ and U_σ, the operator $U_\rho^* U U_\sigma$ does not admit a lower-upper factorization along $\mathcal{E}_+$.*

Before we give the proof of theorem 6.2.2, we state the following auxiliary lemma. The lemma is based on techniques used in [30], Chapter 4.

Lemma 6.2.3 *If an invertible operator S admits a lower-upper factorization along a bounded nest $\mathcal{P}$ of projections, then for each $P \in \mathcal{P}$, the operator $PSP + I - P$ is invertible. Moreover,*

$$\sup_{P \in \mathcal{P}} \|(PSP + I - P)^{-1}\| < \infty.$$

Proof Write $S = LR$, with L, L^{-1} lower triangular and R, R^{-1} upper triangular with respect to $\mathcal{P}$. Fix a projection $P \in \mathcal{P}$ and write

$$PSP + (I - P) = PLRP + (I - P) = PLPPRP + (I - P) =$$

$$(PLP + I - P)(PRP + I - P).$$

Note that $PLP + I - P$ has inverse $PL^{-1}P + I - P$, and that $PRP + I - P$ has inverse $PR^{-1}P + I - P$. Consequently,

$$(PSP + I - P)^{-1} = (PR^{-1}P + I - P)(PL^{-1}P + I - P).$$

Let $C > 0$ such that $\|P\|, \|I - P\| \leq C$ for all $P \in \mathcal{P}$. Since

$$\|PL^{-1}P + (I - P)\| \leq C^2\|L^{-1}\| + C$$

and

$$\|PR^{-1}P + (I - P)\| \leq C^2\|R^{-1}\| + C,$$

we get

$$\|(PSP + I - P)^{-1}\| \leq \|PR^{-1}P + (I - P)\|\|PL^{-1}P + (I - P)\| \leq$$

$$C^4\|R^{-1}\|\|L^{-1}\| + C^3\left(\|R^{-1}\| + \|L^{-1}\|\right) + C^2 < \infty.$$

The lemma has been proved. □

Proof of Theorem 6.2.2 Let U be as in (6.1) and let ρ, σ be permutations on $\mathbf{Z}^+$. Further, let U_ρ, U_σ be the corresponding permutation operators. Define the unitary operator

$$V = U_\rho^* U U_\sigma.$$

It follows that

$$V_{kl} = U_{\rho(k),\sigma(l)}, \quad U_{kl} = V_{\rho^{-1}(k),\sigma^{-1}(l)}, \quad k, l \in \mathbf{Z}^+.$$

We need to investigate whether V admits a lower-upper factorization. If $E_m V E_m + I - E_m$ is not invertible for certain $m \in \mathbf{Z}^+$, then, by Lemma 6.2.3, V does not admit a lower-upper factorization. For that reason, we may assume that all the operators $E_m V E_m + I - E_m$ are invertible ($m \in \mathbf{Z}^+$). We claim that this fact leads to the following restriction on the permutations ρ and σ. The operator

$$W = U_\rho U_\sigma^* = W(1) \oplus W(2) \oplus W(3) \oplus \cdots$$

is of the same block-diagonal form as U with respect to the standard basis $\{e_k\}_{k=1}^\infty$, and each matrix $W(n)$ is an $n \times n$ permutation matrix. As we shall see, this is more than we need for the proof of the theorem.

Fix $n \in \mathbf{Z}^+$ and write $\kappa_n = \frac{n(n-1)}{2}$. Let $\mathcal{I}_n = \{\kappa_n + 1, \ldots, \kappa_n + n\}$. To prove the theorem, it suffices to show that $\min \rho^{-1}(\mathcal{I}_n) = \min \sigma^{-1}(\mathcal{I}_n)$, but we will even prove that

$$\rho^{-1}(\mathcal{I}_n) = \sigma^{-1}(\mathcal{I}_n). \tag{6.2}$$

Write

$$\rho^{-1}(\mathcal{I}_n) = \{k_1, \ldots, k_n\}, \quad \sigma^{-1}(\mathcal{I}_n) = \{l_1, \ldots, l_n\},$$

with $k_{u-1} < k_u$ and $l_{u-1} < l_u$ for $u = 2, \ldots, n$. Put $k = \kappa_n + u$ and $l = \kappa_n + v$ with $1 \le u, v \le n$. Then

$$V_{\rho^{-1}(k),\sigma^{-1}(l)} = U_{kl} = \frac{\sqrt{n}}{n} e^{\frac{2\pi i(u-1)(v-1)}{n}}.$$

Further, $V_{\rho^{-1}(k),s} = 0$ if $s \notin \sigma^{-1}(\mathcal{I}_n)$, and $V_{r,\sigma^{-1}(l)} = 0$ if $r \notin \rho^{-1}(\mathcal{I}_n)$.

To prove (6.2), we will show that $k_u = l_u$ for $u = 1, \ldots, n$.

As a first step, we prove that $k_1 = l_1$. Indeed, write $k = k_1$ and $l = l_1$. Then $|V_{kl}| = \sqrt{n}/n$. If $s < l$, we obtain that $s \notin \rho^{-1}(\mathcal{I}_n)$ and hence that $V_{ks} = 0$. In the same manner it follows that $V_{rl} = 0$ if $r < k$. Assume that $k < l$. Then the k-th row of $E_k V E_k + I - E_k$ consists of zero elements only, since $V_{ks} = 0$ for $1 \le s \le k < l$. Therefore, the operator $E_k V E_k + I - E_k$ is not invertible, a contradiction, so $k \ge l$. Next, assume that $k > l$. The l-th column of $E_l V E_l + I - E_l$ consists of zero elements only, since $V_{rl} = 0$ for $1 \le r \le l < k$. The operator $E_l V E_l + I - E_l$ is not invertible and again a contradiction has been obtained. We conclude that $k = l$.

Second, fix $2 \le p \le n$ and assume that $k_u = l_u$ for $u = 1, \dots, p-1$. We will prove that $k_p = l_p$. First assume that $k = k_p < l_p$. The p rows of the operator $E_k V E_k + I - E_k$ labelled $k_1, \dots, k_p$ have nonzero entries exactly at the $p-1$ column positions $l_1, \dots, l_{p-1}$. This shows that these rows are linear dependent and so the operator $E_k V E_k + I - E_k$ is not invertible. On the other hand, if $k_p > l_p = l$, the operator $E_l V E_l + I - E_l$ is not invertible. Indeed, the p columns of this operator labelled $l_1, \dots, l_p$ have nonzero entries exactly at the $p-1$ row positions $k_1, \dots, k_{p-1}$. Again the columns are linear dependent. It follows that $k_p = l_p$ and by induction on p, we get (6.2).

To prove the theorem, fix $n \in \mathbf{Z}^+$ and let

$$k = \min \rho^{-1}(\mathcal{I}_n) = \min \sigma^{-1}(\mathcal{I}_n).$$

Note that $|V_{kk}| = \sqrt{n}/n$ and that $V_{kj} = V_{jk} = 0$ for $j = 1, \dots, k-1$. Identify $E_k V E_k$ with the $k \times k$ matrix of the operator with respect to the first k standard basis vectors, to obtain

$$E_k V E_k = \begin{pmatrix} E_{k-1} V E_{k-1} & O \\ O & V_{kk} \end{pmatrix}.$$

This $k \times k$ matrix is invertible with inverse

$$(E_k V E_k)^{-1} = \begin{pmatrix} (E_{k-1} V E_{k-1})^{-1} & O \\ O & V_{kk}^{-1} \end{pmatrix}.$$

It follows that

$$\|(E_k V E_k)^{-1}\| \ge |V_{kk}^{-1}| = \sqrt{n}.$$

But $n \in \mathbf{Z}^+$ was chosen arbitrary, so

$$\sup_{k \in \mathbf{Z}^+} \|(E_k V E_k + I - E_k)^{-1}\| \ge \sup_{k \in \mathbf{Z}^+} \|(E_k V E_k)^{-1}\| = \infty.$$

The theorem now follows from Lemma 6.2.3, applied to V and $\mathcal{E}_+$. □

6.3 Two Counterexamples

In this section, counterexamples to the matrix result Theorem 2.2.1 are given. First, we will construct a pair of diagonalizable bounded operators A, Z, which does not admit simultaneous reduction to standard complementary triangular forms, i.e., $(A, Z) \notin \mathcal{C}_+$. In the construction, it is possible to take A and Z self-adjoint and of trace-class. Second, a counterexample is given to Theorem 2.2.1 in terms of a larger class than $\mathcal{C}_+$, which nevertheless can be viewed as a adequate generalization of $\mathcal{C}(m)$.

Theorem 6.3.1 *Let D_1 and D_2 be bounded operators, diagonal with respect to $\mathcal{E}_+$, each with mutually distinct diagonal elements, and let U be the unitary operator as defined in (6.1). Let $A = D_1$ and $Z = U^* D_2 U$. Then $(A, Z) \notin \mathcal{C}_+$. Moreover, the operators A and Z can be chosen self-adjoint and trace-class.*

The proof of this theorem requires the following two lemmas.

Lemma 6.3.2 *Let D be a bounded diagonal operator acting on $l_2(\mathbf{Z}^+)$, given by*

$$De_k = \delta_k e_k, \quad k \in \mathbf{Z}^+,$$

with mutually distinct diagonal elements: $\delta_k \neq \delta_l$ if $k \neq l$. If M is an m-dimensional invariant subspace for D, then there exist distinct positive integers $\tau(1), \ldots, \tau(m)$, such that

$$M = \mathrm{span}\{e_{\tau(1)}, \ldots, e_{\tau(m)}\}.$$

Proof First of all, note that $\mathrm{Ker}(D - \delta) \neq (0)$ if and only if $\delta = \delta_k$ for some $k \in \mathbf{Z}^+$. In this case, k is uniquely defined and $\mathrm{Ker}(D - \delta) = \mathrm{span}\{e_k\}$. Let M be an m-dimensional invariant subspace for D, and let D_M denote the restriction of D to M. If $\delta \in \sigma(D_M)$, then $\mathrm{Ker}(D_M - \delta) \neq (0)$. Since $\mathrm{Ker}(D_M - \delta) = \mathrm{Ker}(D - \delta) \cap M$, it follows that $\mathrm{Ker}(D - \delta) \neq (0)$, so $\delta = \delta_k$ for precisely one $k \in \mathbf{Z}^+$. This proves that $\sigma(D_M) \subseteq \{\delta_k \mid k \in \mathbf{Z}^+\}$.

To prove that D_M has m distinct eigenvalues, assume by way of contradiction that there exists $k \in \mathbf{Z}^+$, such that $\dim\mathrm{Ker}(D_M - \delta_k)^2 \geq 2$. It then follows that $\dim\mathrm{Ker}(D - \delta_k)^2 \geq 2$, which violates $\mathrm{Ker}(D - \delta_k)^2 = \mathrm{span}\{e_k\}$. Let $\tau(1), \ldots, \tau(m)$ denote the distinct positive integers, such that

$$\sigma(D_M) = \{\delta_{\tau(1)}, \ldots, \delta_{\tau(m)}\}.$$

Note that $\mathrm{Ker}(D_M - \delta_{\tau(j)}) = \mathrm{Ker}(D - \delta_{\tau(j)}) = \mathrm{span}\{e_{\tau(j)}\}$. We may conclude that

$$M = \mathrm{span}\{e_{\tau(1)}, \ldots, e_{\tau(m)}\},$$

and the lemma is proved. □

Lemma 6.3.3 *Let D be a bounded diagonal operator acting on $l_2(\mathbf{Z}^+)$, given by*

$$De_k = \delta_k e_k, \quad k \in \mathbf{Z}^+,$$

with mutually distinct diagonal elements: $\delta_k \neq \delta_l$ if $k \neq l$. Let S be an invertible operator, such that $C = S^{-1}DS$ is upper triangular with respect to $\mathcal{E}_+$, with diag *$C = (\gamma_1, \gamma_2, \gamma_3, \ldots)^T$. Then there exists a permutation τ on $\mathbf{Z}^+$, such that $\gamma_k = \delta_{\tau(k)}$, and such that the invertible operator $U_\tau^* S$ is upper triangular with respect to $\mathcal{E}_+$.*

Proof Fix a positive integer m. The subspace $M_m = \text{span}\{e_1, \cdots, e_m\}$ satisfies $DSM_m \subseteq SM_m$. By Lemma 6.3.2, there exist distinct integers $\tau_m(1), \ldots, \tau_m(m)$, such that

$$SM_m = \text{span}\{e_{\tau_m(1)}, \ldots, e_{\tau_m(m)}\}.$$

The proof of the lemma implies $\sigma(D\,|_{SM_m}) = \{\delta_{\tau_m(1)}, \ldots, \delta_{\tau_m(m)}\}$. Since the proper inclusion $SM_m \subset SM_{m+1}$ holds for $m \in \mathbf{Z}^+$, an injective mapping τ on $\mathbf{Z}^+$ can be defined, such that

$$SM_m = \text{span}\{e_{\tau(1)}, \ldots, e_{\tau(m)}\}, \quad m \in \mathbf{Z}^+.$$

Indeed, let $\tau(1) = \tau_1(1)$ and let $\tau(m+1)$ be the unique element in the set $\{\tau_{m+1}(1), \ldots, \tau_{m+1}(m+1)\}$ which is not an element in $\{\tau_m(1), \ldots, \tau_m(m)\}$. Furthermore, $\bigcup_{m=1}^{\infty} SM_m \subseteq l_2(\mathbf{Z}^+)$ is dense, so the mapping τ is also surjective, and hence a permutation of the positive integers. Now

$$\sigma(D\,|_{SM_m}) = \sigma(C\,|_{M_m}) = \{\gamma_1, \ldots, \gamma_m\},$$

implies $\delta_{\tau(m)} = \gamma_m$ for $m \in \mathbf{Z}^+$. Further, $U_\tau^* SM_m = M_m$ for all $m \in \mathbf{Z}^+$, so $U_\tau^* S$ is upper triangular. □

If T is an invertible bounded operator on a Hilbert space, we shall use the short-hand notation $T^{-*} = (T^{-1})^*$ to indicate the adjoint of the inverse.

Proof of Theorem 6.3.1 Assume there exists an invertible operator S acting on $l_2(\mathbf{Z}^+)$, such that $S^{-1}AS$ is upper triangular and $S^{-1}ZS$ is lower triangular. Apply Lemma 6.3.3 to $A = D_1$ to obtain a permutation operator U_σ such that $R_1 = U_\sigma^* S$ is upper triangular. Further use that $(S^{-1}ZS)^* = (S^{-1}U^*D_2US)^* = S^*U^*D_2^*US^{-*}$ is upper triangular, and apply Lemma 6.3.3 to D_2^*, to obtain a permutation operator U_ρ such that $R_2 = U_\rho^* US^{-*}$ is upper triangular. Note that $R = R_1^{-1} = S^{-1}U_\sigma$ is upper triangular and $L = R_2^{-*} = U_\rho^* US$ is lower triangular. Then

$$LR = U_\rho^* USS^{-1}U_\sigma = U_\rho^* UU_\sigma,$$

i.e., the unitary operator U admits a lower-upper factorization after permutations of rows and columns. Here we use that an invertible operator is upper triangular with respect to $\mathcal{E}_+$ if and only if its inverse has this property. A contradiction has been obtained and the first part of the theorem is proved.

To prove the second part of the theorem, i.e., to obtain that D_1 and D_2 are self-adjoint trace-class operators, let the diagonals of both D_1 and D_2 be l_1-sequences of positive numbers. □

The second counterexample concerns a type of complementary triangular forms, which is less restrictive than $\mathcal{C}_+$. The notion of a simple nest is explained in Section 4.1. We will provide a counterexample to Theorem 2.2.1, involving the following definition of complementary triangular forms:

Let $\mathcal{C}_B(X)$ denote the collection of pairs of bounded operators A, Z on the Banach space X, such that there exists a bounded, simple nest of projections $\mathcal{P}$ on X with $AP = PAP$ and $PZ = PZP$ for all $P \in \mathcal{P}$.

The following operator, known as the *Donoghue shift*, is one of the operators featuring in the counterexample. Consider the compact weighted shift C, acting on $l_2(\mathbf{Z}^+)$, given by

$$Ce_1 = 0, \quad Ce_{k+1} = \gamma_k e_k, \quad k \in \mathbf{Z}^+, \tag{6.3}$$

with $|\gamma_k| \geq |\gamma_{k+1}| > 0$ for $k \in \mathbf{Z}^+$, and $\sum_{k=1}^\infty |\gamma_k|^2 < \infty$. We now state the counterexample.

Theorem 6.3.4 *Let C denote the Donoghue shift as in (6.3), and let D denote a diagonal operator, with mutually distinct diagonal elements, as in Lemma 6.3.2. Let $A = C$, and $Z = U^*DU$, where U is the unitary operator, defined in (6.1). Then $(A, Z) \notin \mathcal{C}_B(H)$.*

To prove this theorem, some preparations are required. The proposition below concerning the Donoghue shift is based on material in [22]. The result as stated here is taken from [46] (Theorem 4.12, p.67).

Proposition 6.3.5 *The only non-trivial invariant subspaces M for C as defined in (6.3) are the finite dimensional subspaces*

$$M = \operatorname{span}\{e_1, \ldots, e_m\}, \quad m \in \mathbf{Z}^+.$$

Proposition 6.3.6 *Let Q be a bounded projection on the Banach space X, and let $M \subseteq X$ be a closed subspace. Further, let T denote the restriction of Q to M. Then $M \oplus \mathrm{Ker}\ Q = X$ if and only if*

$$T : M \longrightarrow \mathrm{Ran}\ Q$$

is invertible. Moreover, in the case when these conditions are satisfied, the projection P onto M along Ker Q *is well-defined and satisfies*

$$\|T^{-1}\| \leq \|P\| \leq \|Q\|\ \|T^{-1}\|.$$

In particular, if $\|Q\| = 1$, then $\|P\| = \|T^{-1}\|$.

Proof The first part of the proposition is proved by Proposition 5.2 in [6]. We now prove the second part. Note first that for all $m \in M$, $T^{-1}Qm = m = Pm$. Further, $P(I - Q) = O$ implies $P = PQ$. Therefore,

$$\|T^{-1}\| = \sup_{0\neq m\in M} \frac{\|T^{-1}Qm\|}{\|Qm\|} = \sup_{0\neq m\in M} \frac{\|Pm\|}{\|Qm\|} =$$

$$\sup_{0\neq m\in M} \frac{\|PQm\|}{\|Qm\|} \leq \|P\|.$$

The reversed inequality is obtained as follows. First note that for each $x \in X$, there exists $m \in M$ such that $Qx = Qm$. Since $Qx = 0$ if and only if $Px = 0$, we may write

$$\|P\| = \sup_{x\neq 0} \frac{\|Px\|}{\|x\|} = \sup_{Qx\neq 0} \frac{\|Px\|}{\|Qx\|} \frac{\|Qx\|}{\|x\|} \leq$$

$$\|Q\|\ \sup_{Qx\neq 0} \frac{\|PQx\|}{\|Qx\|} = \|Q\|\ \sup_{0\neq m\in M} \frac{\|PQm\|}{\|Qm\|} = \|Q\|\ \|T^{-1}\|.$$

This proves the proposition. □

Proof of Theorem 6.3.4 Assume that $(A, Z) \in \mathcal{C}_B(H)$, so let $\mathcal{P}$ be a bounded, simple nest of projections, such that $AP = PAP$ and $PZ = PZP$ for all $P \in \mathcal{P}$. Fix a nontrivial $P \in \mathcal{P}$. By Proposition 6.3.5, $AP = PAP$ implies that there exists a positive integer m, such that rank $P = m$. Write $P = P_m$. In addition,

$$\mathrm{Ran}\ P_m = \mathrm{span}\{e_1, \ldots, e_m\} = \mathrm{Ran}\ E_m.$$

Since $\mathcal{P}$ is a simple nest of projections on a Hilbert space, $\{\text{Ran } P \mid P \in \mathcal{P}\}$ is a simple nest of subspaces, by Theorem 4.1.3. In fact, we have obtained that

$$\{\text{Ran } P \mid P \in \mathcal{P}\} = \{\text{Ran } E_m \mid m \in \mathbf{Z}^+\} \cup \{O, I\}.$$

Let $m \in \mathbf{Z}^+$, and consider the equation $P_m Z = P_m Z P_m$. Since $Z = U^* D U$, we get $U P_m U^* D = U P_m U^* D U P_m U^*$. Taking adjoints at both sides, we obtain $D^* Q_m = Q_m D^* Q_m$, with $Q_m = U P_m^* U^*$. Lemma 6.3.2, applied to D^* and Ran Q_m, implies that there exist distinct positive integers $\tau_m(1), \ldots, \tau_m(m)$, such that

$$\text{Ran } Q_m = \text{span}\{e_{\tau_m(1)}, \ldots, e_{\tau_m(m)}\}.$$

Since Ran $Q_m = U(\text{Ker } P_m)^\perp$, it follows that

$$\{\text{Ran } Q_m \mid m \in \mathbf{Z}^+\} \cup \{O, I\}$$

is a simple nest of subspaces. For that reason, there exists a permutation τ on $\mathbf{Z}^+$, such that for all $m \in \mathbf{Z}^+$,

$$\{\tau(1), \ldots, \tau(m)\} = \{\tau_m(1), \ldots, \tau_m(m)\},$$

which implies that Ran $Q_m = U_\tau$Ran E_m. Therefore,

$$\text{Ker } P_m = U^*(\text{Ran } Q_m)^\perp = U^* U_\tau \text{Ker } E_m, \qquad m \in \mathbf{Z}^+.$$

Proposition 6.3.6 can now be used to calculate $\|P_m\|$. Write $V = U_\tau^* U$. Since Ran P_m = Ran E_m and Ker $P_m = \text{Ker}(V^* E_m V)$, it follows that the operator

$$T_m = V^* E_m V \mid_{\text{Ran } E_m} : \text{Ran } E_m \longrightarrow V^* \text{Ran } E_m$$

is invertible. Therefore, the operator $\hat{T}_m = V^*(E_m V E_m + I - E_m)$ is invertible, and its inverse is given by

$$\hat{T}_m^{-1} = [(E_m V E_m)^{-1} + I - E_m] V.$$

Let $x \in$ Ran E_m and $y \in$ Ker E_m. Then

$$\hat{T}_m^{-1} V^*(x + y) = (E_m V E_m)^{-1} x + y = T_m^{-1} V^* x + y.$$

Here we used that for $x \in$ Ran E_m, one gets

$$(T_m^{-1} V^*)^{-1} x = V T_m x = V V^* E_m V x = E_m V x = (E_m V E_m) x.$$

Further,

$$\|\hat{T}_m^{-1}V^*(x+y)\|^2 = \|T_m^{-1}V^*x\|^2 + \|y\|^2 =$$

$$\|T_m^{-1}V^*x\|^2 + \|V^*y\|^2 \leq \max\{1, \|T_m^{-1}\|^2\}\|V^*(x+y)\|^2.$$

It follows that

$$\|P_m\| = \|T_m^{-1}\| \geq \|\hat{T}_m^{-1}\| = \|(E_mU_\tau^*UE_m + I - E_m)^{-1}\|.$$

The nest $\mathcal{P}$ is bounded, so

$$\sup_{m\in\mathbf{Z}^+} \|(E_mU_\tau^*UE_m + I - E_m)^{-1}\| \leq \sup_{m\in\mathbf{Z}^+} \|P_m\| < \infty,$$

contrary to elements in the proof of Theorem 6.2.2. The theorem is established. □

6.4 Pairs with a Finite Rank Operator

In this section, analogues of Theorem 2.2.1 are studied for a pair of bounded operators acting on the separable Hilbert space $l_2(\mathbf{Z}^+)$, where at least one of the operators is of finite rank. The results in this section agree up to a certain extent with the finite dimensional case.

Proposition 6.4.1 *Let Z be a bounded diagonalizable operator, and let A be an operator of finite rank, both acting on $l_2(\mathbf{Z}^+)$. Then $(A, Z) \in \mathcal{C}_+$.*

Proof By Lemma 6.1.1, we may assume without loss of generality that Z is diagonal; say $Ze_k = \zeta_k e_k$ for $k \in \mathbf{Z}^+$. Write rank $A = m$, and let $M = \text{Ran } A$. There exist vectors $b_1, \ldots, b_m \in M$, and integers $n_1 < \cdots < n_m$, such that

$$b_k = \sum_{j=n_k}^{\infty} \beta_{kj} e_j, \qquad \beta_{kn_k} \neq 0,$$

for $k = 1, \ldots, m$. The vectors are linearly independent and hence form a basis in M; $M = \text{span}\{b_1, \ldots, b_m\}$. We claim that if $N = \text{span}\{e_{n_1}, \ldots, e_{n_m}\}^\perp$, then

$$M \oplus N = l_2(\mathbf{Z}^+). \tag{6.4}$$

It suffices to prove that $M \cap N = (0)$. Let $x = \sum_{k=1}^{m} \xi_k b_k \in M \cap N$, and assume that $x \neq 0$. Then there exists an integer $1 \leq p \leq m$, such that $\xi_1 = \ldots = \xi_{p-1} = 0$, and $\xi_p \neq 0$. Then $e_{n_p}^* x = \beta_{n_p} \xi_p \neq 0$. On the other hand, $x \in N$, so $e_{n_p}^* x = 0$. A contradiction has been obtained. Therefore, the decomposition (6.4) indeed holds. With respect to this decomposition,

$$A = \begin{pmatrix} A_1 & A_{12} \\ O & O \end{pmatrix}, \quad Z = \begin{pmatrix} Z_1 & O \\ Z_{21} & Z_2 \end{pmatrix}.$$

Let $\hat{M} = \text{span}\{e_{n_1}, \ldots, e_{n_m}\}$. Since Z is diagonal with respect to $\{e_k\}_{k=1}^{\infty}$, we get

$$Z = \begin{pmatrix} \hat{Z}_1 & O \\ O & D_2 \end{pmatrix} : \hat{M} \oplus N \longrightarrow \hat{M} \oplus N.$$

By Lemma 1.4 in [6], the $m \times m$ matrices D_1 and $\hat{Z}_1$ are similar. In particular, Z_1 is diagonalizable.

It follows, by Theorem 2.2.1, that A_1 and Z_1 admit simultaneous reduction to complementary triangular forms: There exists a basis $s_1, \ldots, s_m$ for M, such that

$$A_1 s_k \in \text{span}\{s_1, \ldots, s_k\}, \quad Z_1 s_k \in \text{span}\{s_k, \ldots, s_m\}, \quad k = 1, \ldots, m.$$

Define the invertible operator S on $l_2(\mathbf{Z}^+)$ as

$$Se_j = \begin{cases} s_j, & j = 1, \ldots, m \\ e_{\pi(j)}, & j = m+1, \ldots, n_m \\ e_j, & j > n_m \end{cases},$$

where $\pi : \{m+1, \ldots, n_m\} \longrightarrow \{1, \ldots, n_m\} \backslash \{n_1, \ldots, n_m\}$ is any bijection. Write $L = \text{span}\{e_1, \ldots, e_m\}$. Then

$$S = \begin{pmatrix} S_1 & O \\ O & S_2 \end{pmatrix} : L \oplus L^{\perp} \longrightarrow M \oplus N.$$

Further, with respect to $L \oplus L^{\perp}$, we get

$$S^{-1}AS = \begin{pmatrix} S_1^{-1} A_1 S_1 & S_1^{-1} A_{12} S_2 \\ O & O \end{pmatrix}, \quad S^{-1}ZS = \begin{pmatrix} S_1^{-1} Z_1 S_1 & O \\ S_2^{-1} Z_{21} S_1 & S_2^{-1} Z_2 S_2 \end{pmatrix},$$

where $S_1^{-1} A_1 S_1$ is an upper triangular $m \times m$ matrix, $S_1^{-1} Z_1 S_1$ is a lower triangular $m \times m$ matrix, and $S_2^{-1} Z_2 S_2$ is a diagonal operator with respect to $\mathcal{E}_+$. Therefore, $(A, Z) \in \mathcal{C}_+$. The proposition is proved. □

The case when A is a bounded diagonalizable operator and Z is of finite rank is dealt with as follows: Apply Proposition 6.4.1 to the operators Z^* and A^* and use Lemma 6.1.1.

Proposition 6.4.2 *Let A be a bounded diagonalizable operator of finite rank, and let Z be a bounded operator, which is lower triangularizable. Then $(A, Z) \in \mathcal{C}_+$.*

Proof By Lemma 6.1.1, we may assume without loss of generality, that Z is lower triangular. To avoid trivialities, we assume that $A \neq O$. By assumption, there exists an invertible operator V, such that $V^{-1}AV$ is diagonal. Let the diagonal of $V^{-1}AV$ be given by

$$\mathrm{diag}(V^{-1}AV) = (\ \alpha_1, \alpha_2, \alpha_3, \dots\ )^T.$$

Since A is of finite rank, we get $m = \max\{k \mid k \in \mathbf{Z}^+, \alpha_k \neq 0\} < \infty$. Write $V(\mathrm{Ran}\ E_m) = M$, and $V(\mathrm{Ker}\ E_m) = N$. Define $d(t) = \dim(M \cap \mathrm{Ker}\ E_t)$ for $t \in \mathbf{Z}^+$, then $d : \mathbf{Z}^+ \longrightarrow \{0, \dots, m\}$ is decreasing, and $\lim_{t\to\infty} d(t) = 0$. Indeed, if $\lim_{t\to\infty} d(t) > 0$, there exists $0 \neq x \in M$, such that $x \in \mathrm{Ker}\ E_t$ for all $t \in \mathbf{Z}^+$, a contradiction. Let $\tau \in \mathbf{Z}^+$, such that $d(\tau) = 0$. Since $M \cap \mathrm{Ker}\ E_\tau = (0)$, and $M + \mathrm{Ker}\ A = l_2(\mathbf{Z}^+)$, there exists a finite dimensional subspace $R \subseteq \mathrm{Ker}\ A$ with $M \oplus R \oplus \mathrm{Ker}\ E_\tau = l_2(\mathbf{Z}^+)$. The vectors $y_k = Ve_k$ for $k = 1, \dots, m$ form a basis in M. In addition, let $y_{m+1}, \dots, y_\tau$ be a basis in R. Note that $Ay_k \in \mathrm{span}\{y_k\}$ for $k = 1, \dots, \tau$. Therefore, the restriction of A to $M_\tau = M \oplus R$ is diagonalizable. With respect to the decomposition $M_\tau \oplus \mathrm{Ker}\ E_\tau = l_2(\mathbf{Z}^+)$, we get

$$A = \begin{pmatrix} A_1 & A_{12} \\ O & O \end{pmatrix}, \qquad Z = \begin{pmatrix} Z_1 & O \\ Z_{21} & Z_2 \end{pmatrix},$$

where A_1 is a diagonalizable $\tau \times \tau$ matrix and Z_2 is a lower triangular operator. By Theorem 2.2.1, there exists a basis $s_1, \dots s_\tau$ in M_τ, such that $A_1 s_k \in \mathrm{span}\{s_1, \dots, s_k\}$ and $Z_1 s_k \in \mathrm{span}\{s_k, \dots, s_\tau\}$ for $k = 1, \dots, \tau$. The invertible operator S, defined by

$$Se_j = \begin{cases} s_j, & 1 \leq j \leq \tau \\ e_j, & j > \tau \end{cases},$$

puts A and Z into complementary triangular forms. The proposition is proved. □

If A is a bounded operator, that is upper triangularizable, and Z is diagonalizable and of finite rank, then apply Proposition 6.4.2, to obtain that $(Z^*, A^*) \in \mathcal{C}_+$. Next, apply Lemma 6.1.1 to obtain $(A, Z) \in \mathcal{C}_+$. We have now dealt with all pairs of operators A, Z, that contain a diagonalizable operator and a finite rank operator, and that satisfy the obvious necessary condition that A is upper triangularizable and Z is lower triangularizable.

Index

Glossary

$A^{\times}$ 16

A_M 12

$\mathbf{C}$ 11

C_a 19

$\hat{C}_z$ 21

$\mathcal{C}(m)$ 19

$\mathcal{C}(\alpha, \zeta)$ 20

$\mathcal{C}_+$ 114

$\mathcal{C}(X)$ 106

$\mathcal{C}_B(X)$ 122

$\mathcal{C}_f(X)$ 106

$cl\ U$ 76

$\overline{\mathcal{S}}$ 82

$\delta(W)$ 16

$\mathrm{diag}(A; \mathcal{M})$ 43

$\mathrm{diag}\ T$ 20, 114

$\dim X$ 11

E_m 113

$\mathcal{E}_+$ 113

I_n 15

I_X 12

$[M_1, M_2)$ 76

$\mathrm{Inv}\ A$ 67

$J(\alpha, n)$ 22

$k(W)$ 32

$k^*(W)$ 33

$\mathrm{Ker}\ A$ 11

M_+ 100

M_- 100

$N_\alpha(A)$ 43

$N_\sigma(A)$ 43

$\nu(W)$ 38

O_n 21

O_X 12

$X \oplus Y$ 11

$\mathcal{P}^c$ 78

$\mathcal{P}^*$ 88

$p_B(\lambda)$ 23

Ran A 11

rank A 11

$\rho(W)$ 31

$\rho_0(A_1, Z_1)$ 41

span U 76

$\subseteq$ 11

$\subset$ 11

T^{-*} 121

$U^\perp$ 87

${}^\perp V$ 87

$\bigvee \mathcal{P}$ 77

$\bigwedge \mathcal{P}$ 78

$\bigvee \mathcal{S}$ 76

$\bigwedge \mathcal{S}$ 76

X/Y 11

X^* 12

$x_k(\alpha)$ 20

$\mathbf{Z}$ 11

$\mathbf{Z}^\pm$ 11

$\mathbf{Z}_0^\pm$ 11

$T_n \xrightarrow{\|\cdot\|} T$ 76

$T_n \xrightarrow{s} T$ 76

$x_n \xrightarrow{\|\cdot\|} x$ 76

Bibliography

[1] N. Aronszajn, K.T. Smith, Invariant Subspaces of Completely Continuous Operators, *Ann. Math.* 60: 345-350 (1954).

[2] G. Bachman, L. Narici, Functional Analysis, Textbooks in Mathematics, Academic Press, New York (1966).

[3] W.G. Bade, On Boolean Algebras of Projections and Algebras of Operators, *Trans. A.M.S.* 80: 345-360 (1955).

[4] J.Y. Barry, On the Convergence of Ordered Sets of Projections, *Proc. A.M.S.* 5: 313-314 (1954).

[5] H. Bart, Transfer Functions and Operator Theory, *Lin. Alg. Appl.* 84: 33-61 (1986).

[6] H. Bart, I. Gohberg, M.A. Kaashoek, Minimal Factorization of Matrix and Operator Functions, Operator Theory: Adv. Appl. 1, Birkhäuser Verlag, Basel (1979).

[7] H. Bart, H. Hoogland, Complementary Triangular Forms of Pairs of Matrices, Realizations with Prescribed Main Matrices, and Complete Factorization of Rational Matrix Functions, *Lin. Alg. Appl.* 103: 193-228 (1988).

[8] H. Bart, P.S.M. Kop Jansen, Upper Triangularization of Matrices by Lower Triangular Similarities, *Lin. Alg. Appl.* 103: 229-248 (1988).

[9] H. Bart, L.G. Kroon, Companion Based Matrix Functions: Description and Factorization, *to appear in Lin. Alg. Appl.*

[10] H. Bart, L.G. Kroon, Factorization and Job Scheduling: a connection via Companion Based Rational Matrix Functions, *to appear.*

[11] H. Bart, L.G. Kroon, Variants of the Two Machine Flow Shop Problem, *to appear.*

[12] H. Bart, G.Ph.A. Thijsse, Complementary Triangular Forms of Upper Triangular Toeplitz Matrices, *Operator Theory: Adv. Appl.* 40: 133-149 (1989).

[13] H. Bart, G.Ph.A. Thijsse, Complementary Triangular Forms of Nonderogatory, Jordan and Rank One Matrices, Report 9003/B, Econometric Institute, Erasmus University Rotterdam (1990).

[14] H. Bart, G.Ph.A. Thijsse, Eigenspace and Jordan-Chain Techniques for the Description of Complementary Triangular Forms, Report 9353/B, Econometric Institute, Erasmus University Rotterdam (1993).

[15] H. Bart, H.K. Wimmer, Simultaneous Reduction to Triangular and Companion Forms of Pairs of Matrices: The Case rank$(I - AZ) = 1$, *Lin. Alg. Appl.* 150: 443-461 (1991).

[16] H. Bart, R.A. Zuidwijk, Triangular Forms after Extensions with Zeroes, *in preparation.*

[17] B. Beauzamy, Introduction to Operator Theory and Invariant Subspaces, North-Holland Math. Library 42, North-Holland Publ. Co., Amsterdam (1988).

[18] M.S. Brodskii, Triangular and Jordan Representations of Linear Operators, Transl. Math. Monographs 32, A.M.S., Providence, RI (1971).

[19] D. Carlson, S. Pierce, Common Eigenvectors and Quasicommutativity of Sets of Simultaneously Triangularizable Matrices, *Lin. Alg. Appl.* 71: 49-55 (1985).

[20] K.R. Davidson, Nest Algebras, Pitman Research Notes in Math. 191, Longman House, Burnt Mill, Harlow (1988).

[21] P.J. Davis, Circulant Matrices, Wiley and Sons, New York (1979).

[22] W.F. Donoghue, The Lattice of Invariant Subspaces of a Completely Continuous Quasinilpotent Transformation, *Pac. J. Math.* 7: 1031-1035 (1957).

[23] M.P. Drazin, J.W. Dungey, K.W. Gruenberg, Some Theorems on Commutative Matrices, *J. London Math. Soc.* 26: 221-228 (1951).

[24] P. Enflo, A Counterexample to the Approximation Property in Banach Spaces, *Acta Math.* 130: 309-317 (1973).

[25] P. Enflo, On the Invariant Subspace Problem in Banach Spaces, *Acta Math.* 158: 213-313 (1987).

[26] S. Friedland, Pairs of Matrices Which Do Not Admit a Complementary Triangular Form, *Lin. Alg. Appl.* 150: 119-123 (1990).

[27] G. Frobenius, Über vertauschbare Matrizen, *Sitz.-Ber. Akad. Wiss. Berlin*, 26: 601-614 (1896).

[28] F.J. Gaines, R.C. Thompson, Sets of Nearly Triangular Matrices, *Duke Math. J.* 35: 441-453 (1968).

[29] I. Gohberg, S. Goldberg, M.A. Kaashoek, Classes of Linear Operators I, Operator Theory: Adv. Appl. 49, Birkhäuser Verlag, Basel (1990).

[30] I. Gohberg, M.G. Krein, Theory and Applications of Volterra Operators in Hilbert Space, Transl. Math. Monographs 24, A.M.S, Providence, RI (1969).

[31] I. Gohberg, P. Lancaster, L. Rodman, Invariant Subspaces of Matrices with Applications, J. Wiley and Sons, New York (1986).

[32] W.T. Gowers, B. Maurey, The Unconditional Basic Sequence Problem, *preprint* (1992).

[33] R.V. Kadison, J.R. Ringrose, Fundamentals of the Theory of Operator Algebras 1 (Elementary Theory), Academic Press, New York (1983).

[34] T.J. Laffey, Simultaneous Triangularization of a Pair of matrices, *J. Alg.* 44: 550-557 (1977).

[35] T.J. Laffey, Simultaneous Triangularization of Matrices-Low Rank Cases and the Nonderogatory Case, *Lin. Multilin. Alg.* 6: 269-305 (1978).

[36] T.J. Laffey, Simultaneous Reduction of Sets of Matrices under Similarity, *Lin. Alg. Appl.* 84: 123-138 (1986).

[37] C. Laurie, E. Nordgren, H. Radjavi, P. Rosenthal, On Triangularization of Algebras of Operators, *J. Reine Angew. Math.* 327: 143-155 (1981).

[38] P. Lancaster, M. Tismenetsky, The Theory of Matrices, Second Edition with Applications, Academic Press, Orlando, Fl. (1985).

[39] J. Lindenstrauss, L. Tzafriri, On the Complemented Subspaces Problem, *Israel J. Math.* 9: 263-269 (1971).

[40] J. Lindenstrauss, L. Tzafriri, Classical Banach Spaces I (Sequence Spaces), Ergebnisse der Math. und ihre Grenzgebiete 92, Springer Verlag, Berlin (1977).

[41] E.R. Lorch, On a Calculus of Operators in Reflexive Vector Spaces, *Trans. A.M.S.* 45: 217-234 (1939).

[42] N.H. McCoy, On the Characteristic Roots of Matric Polynomials, *Bull. A.M.S.* 42: 592-600 (1936).

[43] G.J. Murphy, Triangularizable Algebras of Compact Operators, *Proc. A.M.S.* 84: 354-356 (1982).

[44] A. Pelczynski, *personal communication.*

[45] H. Radjavi, A Trace Condition Equivalent to Simultaneous Triangularizability, *Can. J. Math.* 38: 376-386 (1986).

[46] H. Radjavi, P. Rosenthal, Invariant Subspaces, Ergebnisse der Math. und ihrer Grenzgebiete 77, Springer-Verlag, Berlin (1973).

[47] J.R. Ringrose, Super-Diagonal Forms for Compact Linear Operators, *Proc. London Math. Soc.* 12: 367-384 (1962).

[48] J.R. Ringrose, Non-Self-Adjoint Compact linear Operators, van Nostrand, New York (1971).

[49] W. Rudin, Real and Complex Analysis, McGraw-Hill, New York (1986).

[50] W. Rudin, Functional Analysis, McGraw-Hill, New Dehli (1986).

[51] N.K. Spanoudakis, Generalizations of Certain Nest Algebra Results, *Proc. A.M.S.* 115: 711-723 (1992).

[52] S.H. Tan, J. Vandewalle, On Factorizations of Rational Matrices, *IEEE Transactions on Circuits and Systems* 35: 1179-1181 (1988).

[53] R.M. Young, An Introduction to Nonharmonic Fourier Series, Pure and Appl. Math., Academic Press, New York (1980).

[54] R.A. Zuidwijk, Triangular Forms for Pairs of Matrices and Zero Extensions, Report 9317/B Erasmus University Rotterdam (1993).

[55] R.A. Zuidwijk, Complementary Triangular Forms for Infinite Matrices, *submitted.*

[56] R.A. Zuidwijk, Quasicomplete Factorizations, *in preparation.*

Nederlandse Samenvatting - Dutch Summary

In deze samenvatting, die met name bedoeld is voor de lezer zonder wiskundige voorkennis, zal aan de hand van een analogie het onderwerp van het proefschrift worden uitgelegd. Verder wordt enige aandacht besteed aan de motivatie achter het onderzoek. Tenslotte worden enkele resultaten uit het proefschrift kort toegelicht.

Stelt u zich voor dat u door een beeldentuin van een museum voor moderne kunst loopt. In die tuin staan twee zwarte beelden, die zeer grillig van vorm zijn. Voor elk beeld apart is het echter zo, dat vanuit een bepaald perspectief het silhouet van het beeld zich aftekent als een onmiskenbare geometrische vorm: een zuivere driehoek. u kunt zich nu voorstellen dat, vanuit bepaalde plaatsen in de tuin bezien, het silhouet van het ene beeld zich inderdaad als een driehoek aftekent, terwijl het silhouet van het andere beeld een of andere grillige vorm is. De vraag die in het proefschrift centraal staat laat zich naar de analogie van de beeldentuin als volgt vertalen: Bestaat er een plek in de tuin van waaruit de silhouetten van beide beelden zich aftekenen als zuivere driehoeken?

De zwarte beelden in de analogie van de beeldentuin zijn in het proefschrift zogeheten vierkante matrices. Dit zijn vierkante tabellen met getallen, waar voor het gemak haken omheen zijn gezet. Een voorbeeld van een 3-bij-3 matrix is

$$\begin{pmatrix} \pi & 6 & 0 \\ 2 & 0 & \frac{3}{4} \\ \sqrt{2} & 0 & 12 \end{pmatrix}.$$

De driehoekige silhouetten van de beelden in de beeldentuin zijn in het proefschrift driehoeksvormen van matrices. Bij een matrix in driehoeksvorm staan op bepaalde plaatsen enkel nullen. We onderscheiden bovendriehoeksmatrices van de vorm

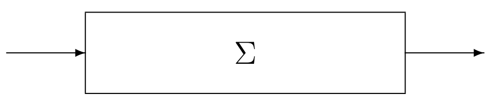
$$\begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \end{pmatrix},$$

en onderdriehoeksmatrices van de vorm

$$\begin{pmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{pmatrix}.$$

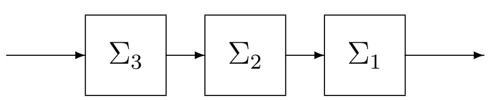
In deze laatste twee figuren duiden de asteriksen getallen aan, waarvan de precieze waarde nu niet van belang is. In dit proefschrift luidt de centrale vraag: Kan een gegeven tweetal vierkante matrices simultaan worden omgevormd tot een tweetal matrices, waarvan de een een bovendriehoeksmatrix is en de ander een onderdriehoeksmatrix is? Als dit mogelijk is, zeggen we dat het betreffende paar vierkante matrices een simultane reductie tot complementaire driehoeksvormen toestaat.

Waarom zijn we eigenlijk zo geinteresseerd in simultane reductie tot complementaire driehoeksvormen? Het antwoord op deze vraag moet worden gezocht in de systeemtheorie. Met deze theorie kan worden gekeken naar een systeem, dat we met de griekse letter sigma (Σ) zullen aanduiden, als hieronder schematisch weergegeven.



De figuur suggereert dat er iets in het systeem wordt gebracht en dat er iets uit komt. De systemen die in het proefschrift bekeken worden, laten in zekere zin een rechtlijnig verband zien tussen datgene wat er in wordt gebracht en datgene wat er vervolgens uit komt.

Nu kan zo'n systeem behoorlijk ingewikkeld zijn. Voor een beter begrip is het nuttig het systeem stap voor stap te analyseren. In de systeemtheorie is men dan ook geinteresseerd in de vraag of een systeem in een aantal eenvoudiger systemen kan worden opgedeeld. Zo zou men kunnen proberen een systeem Σ van complexiteitsgraad drie op te delen in drie systemen Σ_1, Σ_2 en Σ_3, elk van complexiteitsgraad één. (Een systeem van complexiteitsgraad één wordt een elementair systeem genoemd.) Deze opdeling is in de onderstaande figuur schematisch weergegeven.



Rest de vraag, of een systeem Σ wel op deze manier kàn worden geanalyseerd. Bestaande theorie geeft aan, dat een systeem Σ kan worden opgedeeld in elementaire systemen, waarvan het aantal gelijk is aan de graad van het systeem, precies als een tweetal vierkante matrices dat met het systeem wordt geassocieerd (we behandelen hier niet hoe dit precies gebeurt) simultane reductie tot complementaire driehoeksvormen toestaat. Het blijkt echter, dat niet alle systemen op deze wijze kunnen worden opgedeeld. Zo kan niet elk systeem van complexiteitsgraad drie worden opgedeeld in drie elementaire systemen (van complexiteitsgraad één).

Nu volgt een kort overzicht van de in het proefschrift bewezen resultaten.

In het tweede hoofdstuk van dit proefschrift wordt aangetoond dat elk systeem kan worden opgedeeld in elementaire systemen. Soms is daarvoor echter een groter aantal elementaire systemen nodig dan de complexiteitsgraad van het systeem. Zo kan een systeem van complexiteitsgraad drie altijd in minder dan zes elementaire systemen worden opgedeeld. Deze stelling is gebaseerd op deels bekende resultaten over simultane reductie tot complementaire driehoeksvormen voor paren van matrices. Zo maakt een stap in de bewijsgang gebruik van het in dit proefschrift verkregen feit, dat elk paar van matrices kan worden uitgebreid tot een paar van grotere matrices, dat simultane reductie tot complementaire driehoeksvormen toestaat.

In het derde hoofdstuk komt een zeer specifiek type van uitbreidingen van paren van matrices aan de orde. Het betreft hier zogeheten uitbreidingen met nullen. De vraag is of elk paar van matrices kan worden uitgebreid met enkel nullen tot een paar van grotere matrices, dat simultane reductie tot complementaire driehoeksvormen toestaat. Uit de beschouwingen in het derde hoofdstuk blijkt dat deze kwestie bijzonder ingewikkeld is.

In het tweede deel van dit proefschrift wordt gekeken naar complementaire driehoeksvormen voor paren van zogenaamde operatoren. Eindige matrices, die we tot nu toe hebben besproken, behoren tot die operatoren, maar er zijn veel meer soorten operatoren. Zo kan bijvoorbeeld de oneindig grote matrix

$$\begin{pmatrix} \frac{1}{2} & 1 & 0 & 0 & \cdots \\ 0 & \frac{1}{4} & \frac{1}{2} & 0 & \cdots \\ 0 & 0 & \frac{1}{8} & \frac{1}{4} & \ddots \\ 0 & 0 & 0 & \frac{1}{16} & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

ook als operator optreden. De stippellijntjes in de figuur suggereren een oneindige voortzetting van het regelmatige patroon dat zich laat raden. Over complementaire driehoeksvormen voor paren van eindige matrices is al het een en ander bekend. De vraag is nu, of voor paren van operatoren in het algemeen vergelijkbare resultaten kunnen worden verkregen.

In de hoofdstukken vier en vijf worden werkzame en zeer algemene definities van complementaire driehoeksvormen voor paren van operatoren ontwikkeld, die consistent zijn met die voor paren van eindige matrices. Desalniettemin blijkt in hoofdstuk zes aan de hand van voorbeelden dat in het algemeen voor paren van operatoren geen gelijksoortige resultaten kunnen worden verwacht als voor paren van eindige matrices.

Tenslotte verdient opmerking dat er sterke aanwijzingen zijn dat de eerder genoemde stelling uit hoofdstuk twee in verband kan worden gebracht met bepaalde besliskundige problemen.