Schur Complement Algebra and Operations
with Applications in Multivariate
Functions, Realizations, and Representations

by
Anthony Dean Stefan

A thesis
submitted to the College of Engineering and Science
at Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Master of Science
in
Applied Mathematics

Melbourne, Florida
May, 2021

We the undersigned committee
hereby approve the attached thesis

Schur Complement Algebra and Operations
with Applications in Multivariate
Functions, Realizations, and Representations

by Anthony Dean Stefan

Aaron Welters, Ph.D.
Assistant Professor
Mathematical Sciences
Committee Chair

Philip Bernhard, Ph.D.
Associate Professor
Computer Engineering and Sciences
Outside Committee Member

Vladislav Bukshtynov, Ph.D.
Assistant Professor
Mathematical Sciences
Committee Member

Munevver Mine Subasi, Ph.D.
Associate Professor and Department Head
Mathematical Sciences

ABSTRACT

Title:
Schur Complement Algebra and Operations
with Applications in Multivariate
Functions, Realizations, and Representations

Author:
Anthony Dean Stefan

Major Advisor:
Aaron Welters, Ph.D.

We provide a new approach to the following multidimensional realizability problem: Can an arbitrary square matrix, whose entries are from the field of multivariate rational functions over the complex numbers, be realized as a Schur complement of a linear matrix pencil with symmetries? To answer this problem, we prove the main theorem of M. Bessmertnyĭ, "On realizations of rational matrix functions of several complex variables," in Vol. 134 of *Oper. Theory Adv. Appl.*, pp. 157-185, Birkhäuser Verlag, Basel, 2002 and have included additional symmetries as an extension to his results. Furthermore, we were so thorough in our constructive approach that we also prove every real multivariate polynomial has a symmetric determinantal representation, which was first proved in J. W. Helton, S. A. McCullough, and V. Vinnikov, *Noncommutative convexity arises from linear matrix inequalities*, J. Funct. Anal. 240, 105-191 2006, as a direct application of our techniques. Our perspective is from a more "natural" and algorithmic approach using Schur complement algebra and operations with algebraic-functional symmetries. To further motivate the use of the Schur complement in realizability theory and its applications to synthesis and inverse problems, we give three quintessential examples of the Schur complement with symmetries in multivariate applied models: impedance matrices, Dirichlet-to-Neumann map, and effective conductivity tensor in the theory of composites. We then conclude with a discussion on the open problems related to and future directions of our study.

Table of Contents

Acknowledgements

The author would like to thank all of the thesis committee and Florida Tech for the support in my ambitions and the research program. I would like to especially thank Dr. Aaron Welters for his motivation and mentorship in my life, education, and research. I would like recognize the prolific experts M.F. Bessmertnyĭ, Graeme W. Milton, Mihai Putinar, Joseph A. Ball, and Victor Vinnikov for their significant contributions to the topics in this thesis, without them, we would not be here.

Chapter 1

Introduction

The core of this thesis is a study on the Schur complements of 2×2 block matrices from an algebraic and operator theoretical perspective in which the entries of the matrices come from the field of multivariate rational functions over the complex numbers. The question that becomes paramount is how to preserve symmetries of these block matrices both algebraic (such as real and symmetric or Hermitian) as well as functional (such as linear pencils or homogeneous) under algebraic operations on Schur complements. Specifically, these operations include the vector operations of addition and scalar multiplication, the matrix algebra operations of matrix products and Kronecker products, and the functional operation of composition.

Our motivation and the importance of this study is due to the fact that the Schur complement naturally arises in applications from many areas of mathematics as well as other areas of science such as physics and engineering. The consistent recurrence happens because one of the fundamental properties of the Schur complement is that it is part of the parameterization of the solution of certain constrained linear systems (see Proposition 3 of Chapter 2). Some examples of this are considered in Chapter 3 which includes the impedance matrix of an electrical network, the Dirichlet-to-Neumann map of a resistor network, and the effective conductivity tensor in the theory of multi-phase composites. In particular, all of these are important examples of Schur complements of linear matrix pencils with certain symmetries and are matrices with multivariate rational functions over the complex numbers as entries.

These examples motivate our studies as they pertain to realizability theory and its applications to synthesis and inverse problems. In regard to this theory, Chapter 5 represents an important contribution, and, in particular, our extension and approach to the Bessmertnyĭ realizability theorem (i.e., Theorem 77) which proves realizability for rational functions of several complex variables with functional symmetries as a Schur complement of linear matrix pencils. Our extension to this theorem [namely, the extension part (d) of Theorem 77] allows us to construct a Hermitian matrix pencil (i.e., all the matrices in the linear pencil are Hermitian matrices) if the function $f(z)$ is "Hermitian" as well, [i.e., has the functional property $f(\overline{z})^* = f(z)$].

Now from a mathematical perspective, Chapter 4 is arguably the most important chapter in this thesis. This is because we need it not only for proving the Bessmertnyĭ realizability theorem but to also lay the foundations for other applications in this thesis (see Chapter 6) as well as for future applications we have in mind (see Chapter 7). In Chapter 4, we provide explicit formulas for computing algebraic-functional operations on Schur complements. In fact, we are so thorough in our constructive approach to the algebra of operations on Schur complements in Chapter 4 that all we need to do to prove the Bessmertnyĭ realizability theorem in Chapter 5 is to construct realizations of squares and simple products.

Another instance of the applicability from the results of Chapter 4 and of the realizations of squares and simple products from Chapter 5 is the short and elementary proof we give in Chapter 6 on the symmetric determinantal representation of real multivariate polynomials. The first proof of this theorem, which we will call the *HMV Theorem*, was in 2006 by J. Helton, S. McCullough, and V. Vinnikov in [06HMV], and alternative proofs were given shortly thereafter in [11BG, 12RQ] along with extensions to arbitrary fields with characteristic different from two. The HMV Theorem generated quite a bit of interest in the scientific community especially given its relationship to hyperbolic multivariate polynomials and the Lax Conjecture (now a theorem in the case of three variables, see [05LPR]). The approaches taken in these papers [06HMV, 11BG, 12RQ] are based on various advanced methods outside the scope of elementary linear algebra and the basic theory of polynomials. In Chapter 6, however, we take an approach within this scope in proving the HMV Theorem (Theorem 84). In particular, the only results we need are those in Chapter 4, realization of squares and simple products in Chapter 5, and the elementary Schur complement determinate formula [(9) in Chapter 2]. In addition, our approach allows us to easily prove the extension of the HMV Theorem for arbitrary fields with characteristic not two and explicitly show why our method fails in fields of characteristic two (by relating it to the known result that HMV Theorem is false in fields of characteristic two, as was first shown in [13GMT]).

As we can see, much has been accomplished already in this thesis, but there is still a lot of work to be done, this is discussed in more detail along with open problems in Chapter 7. In particular, we discuss the open problems associated with the Bessmertnyĭ class of functions and the associated Milton class of functions which involve Bessmertnyĭ realizations of homogeneous rational matrix functions with positive-semidefinite matrix pencils and their relationship to Schur complements and effective operators in the theory of composites. Then, we conclude with future directions of our study such as: extending the Bessmertnyĭ Realization Theorem to include additional symmetries that arise in applications like the ones in Chapter 3, developing computational methods to implement the Schur complement algebra and operations from Chapter 4 to construct realizations with symmetries when they exist and test conjectures or hypothesises for problems with multivariate rational matrix functions, extending the Bessmertnyĭ Realizability Theorem in Chapter 5 over any arbitrary field (like the field extension of HMV theorem in Chapter 6), and finally, investigating recent developments on the

Bessmertnyĭ class of functions and use the insights obtained to study the Milton class of functions along with its application in the theory of composites (e.g., effective operators and their bounds, limitations, as well as realizability problems).

The structure of the thesis will proceed as follows and can be visualized in the flow diagram of Figure 1.1. In the remainder of Chapter 1, we will briefly summarize the contents of each chapter and provide a discussion on how results from previous work (specifically [20SW] and [21SW]) have been improved or modified. Next, in Chapter 2 we establish the notation, definitions, and preliminary results. Then, in Chapter 3 we motivate the study of the Schur complement by providing applications of how the Schur complement helps parameterize the solution of certain constrained linear systems and we point out how the applied models influence the algebraic-functional symmetries of the Schur complement. Afterward, in Chapter 4 (which is the cornerstone of the thesis) we develop the theory of elementary algebraic-functional operations on Schur complements and provide explicit examples of each operation (see List 4.1 for a complete list of the operations). Subsequently, in Chapter 5 we state and prove the Bessmerntyĭ Realization Theorem including our extension. We also provide several examples in this chapter to illustrate our methods of *realizing* rational functions of several complex variables. Next, in Chapter 6 we prove our HMV Theorem, provide an example using our approach, and discuss the extension of the HMV Theorem to arbitrary fields different from characteristic two. Finally, in Chapter 7 we conclude with a discussion on the open problems and the future research directions.

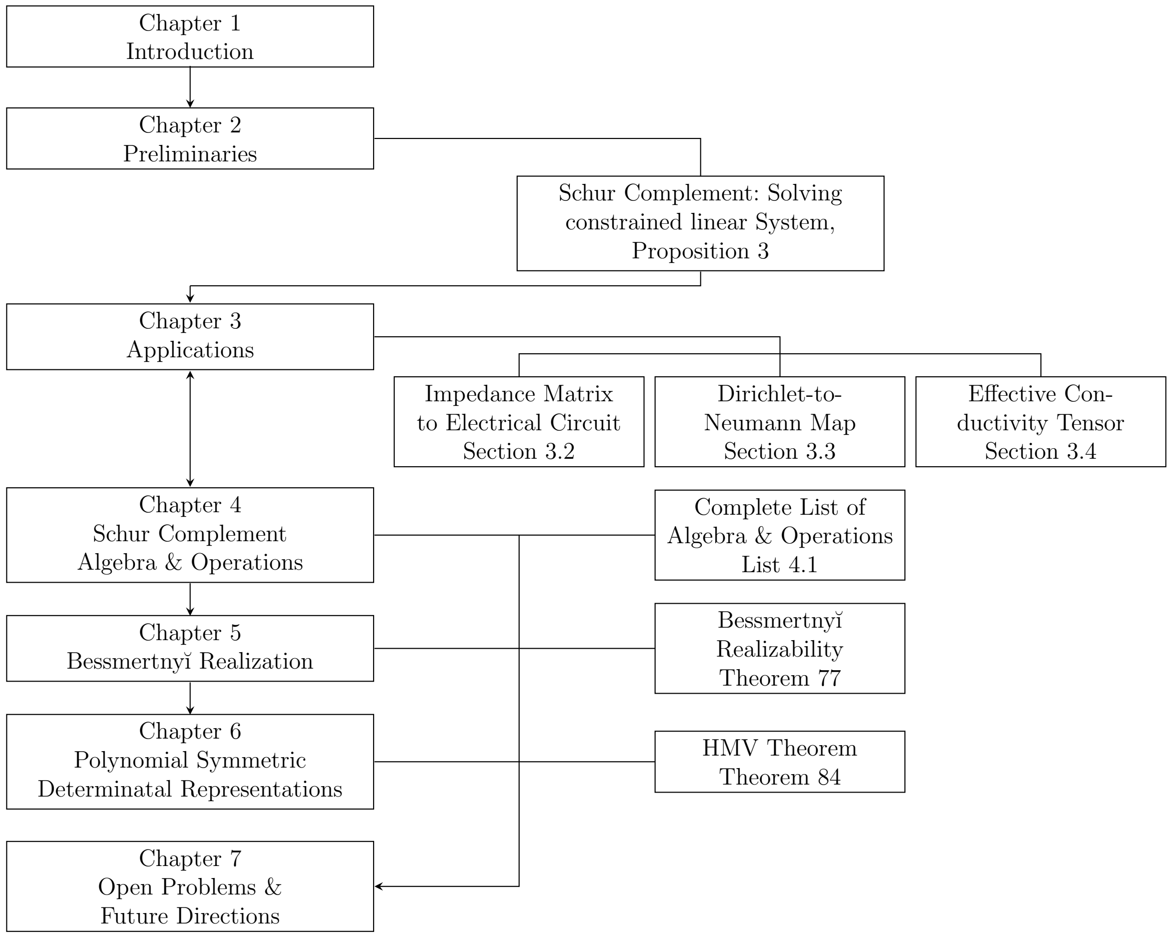


Figure 1.1: Flow diagram for thesis.

1.1 Summary of Chapters

1.2 Chapter 2: Preliminaries

Chapter 2 proves in general that the Schur complement (Definition 2) is part of the parameterization to the solution of a constrained linear system (Lemma 3). As shown in Figure 1.1, Lemma 3 is an essential result of the thesis that motivates our study and, in fact, is called into each section of Chapter 3 for parameterizing the solutions to our examples. Although all chapters contain fundamental chapter-specific preliminaries, Chapter 2 contains basic notation and operations of the entire thesis. In Section 2.1, the Schur complement and elementary properties are introduced, along with the block Schur complement decomposition (Lemma 4) and the Schur complement inverse formulas (Corollary 5), both of these elementary results are used to prove Lemma 41 in Chapter 4. Next, we provide notation and elementary properties of the Kronecker product (Section 2.2) which is not only defined as a Schur complement matrix operation itself in Section 4.4 of Chapter 4 but is also used in Chapter 5 for realizations of linear pencils (Section 5.2.1) and to realize multivariate rational functions (Examples 78, 79). Then, we provide the notation and definition of the direct sum (Section 2.3) of two matrices which is also defined as a Schur complement operation in Section 4.2 of Chapter 4. Finally, we conclude the preliminaries with the direct sum (closely related to shorting matrices), the determinant of a direct sum (8), and Schur's determinate formula (9) which are heavily used throughout Chapter 6 in proving symmetric determinantal representations of polynomials, (see Example 85).

1.3 Chapter 3: Applications

In Chapter 3 we provide explicit applications of the Schur complement parametrizing the solving of constrained linear systems. We show that each application encompasses the Schur complement with algebra-functional symmetries (e.g., Schur complement is real and symmetric and is a linear pencil in the field of rational multivariate functions over the complex numbers) is related to applied models. These examples include the impedance matrix to an RLC electric circuit (Section 3.2), a Dirichlet-to-Neumann map of a resistor only network (Section 3.3), the effective conductivity tensor in the theory of multi-phase composites (Section 3.4).

To do so, we provide a basic introduction to the topology, constraints, and constitutive relations of our models (Section 3.1). The study of electrical circuit theory for this thesis is primarily from [69BB] and [60PB], however, our study of the Dirichlet-to-Neumann map is primarily from [00CM]. We include in our theorems of proving the source impedance matrix Z for an RLC network Γ (Theorem 15) and the Dirichlet-to-Neumann matrix Λ (Theorem 20) is a Schur complement with algebraic-functional symmetries.

Although the Schur complement in the first two sections of Chapter 3 are 2×2 block matrices, in Section 3.4, the Schur complement is a block operator matrix (only occurrence in this thesis when the Schur complement is a block operator matrix). We still include this example because the theory of composites is the most quintessential model for our methods, research, and motivation for this thesis. Furthermore, the overlay of our definitions, theorems, and techniques in our study on Schur complements with more difficult problems, as the one in Section 3.4, illustrates the potential of solving synthesis and inverse problems in realizability theory in a more operator theory and algorithmic approach. Moreover, the open problems discussed in Chapter 7 directly relate to Section 3.4 (in particular Proposition 88) in the theory of composites. Hence, we solve the Z-problem (Definition 22) for the effective conductivity tensor in the theory of multi-phase composites in finite dimensions and solve for the effective operator.

1.4 Chapter 4: Schur Complement Algebra & Operations

As mentioned, the core of this thesis is a study on the Schur complements of 2×2 block matrices from an algebraic and operator theoretical perspective in which the entries of the matrices come from the field of multivariate rational functions over the complex numbers. Thus, Chapter 4 is the center point of this thesis as it provides a complete foundation of algebraic-functional operations for Schur complements over the complex numbers and we give explicit formulas to compute each operation. Furthermore, our conceptualization of the perspective to using the Schur complement in a more algorithmic algebraic approach is from the approach of solving synthesis problems in electrical network theory using elementary algebraic operations by R. Duffin in [55RD]. Hence, we provide an introduction to the *principal pivot transform* (PPT) which is an important transformation in the context of network synthesis problems (see, for instance, [65DHM, 66DHM, 00MT]) and may also play an important role in the theory of composites.

In addition, expanding on the work previously done in [20SW], we provide concrete examples for each of the operations and relations of the Schur complement. Moreover, there have been a few changes in the formulas to preserve symmetries, for instance, the formula for the C matrix of Proposition 37 was originally non-symmetric, but now after a switch in columns we were able to provide a symmetric representation when $A = B^T$ or $A = B^*$. Another example of an improvement to the Schur complement algebra and operations from [20SW] is how we introduce rank factorization of matrices (Lemma 54) for treating non-invertible matrices which is a result that is valid over any field, and was actually a result from later work in [21SW] now incorporated in Chapter 5.

One of the merits of this thesis is that it outlines the natural progression to proving the *Bessmertnyĭ realizability theorem* in Chapter 5 and then using techniques from both chapters (Chapter 4 and 5) to prove HMV theorem in Chapter 6. More importantly, because of this, the thesis identifies the individual aspects (by chapter) of these theorems that are needed [whether it be additional algebraic-functional symmetries of the Schur complement or different types of Schur complements needed (such as simple products or squares)]. This helps narrow down exactly where and what results should be focused on for future research directions.

1.5 Chapter 5: Bessmertnyĭ Realizability

As part of M. Bessmertnyĭ's 1982 Ph. D. thesis (in Russian) [82MB], he proved that every rational matrix-valued function of several variables could be written as the Schur complement of a linear matrix pencil (i.e., a *Bessmertnyĭ long resolvent representation*). As mentioned in [04KV], this theorem of his (a.k.a, the *Bessmertnyĭ realizability theorem*), was unknown to Western readers until parts of it were translated into English beginning in 2002 with [02MB]. The main theorem of M. Bessmertnyĭ in [02MB, Theorem 1.1] solves the realization problem, and his construction of the linear matrix pencil

$A(z) = A_0 + z_1 A_1 + \cdots + z_n A_n$ from $f(z)$ involves solving large systems of constrained linear equations in such a way that $A(z)$ inherits certain real, symmetric, or homogeneity properties from $f(z)$. Given the potential applications of this important theorem to realization problems in multivariate systems theory, electric network theory, and the theory of composites, we consider this theorem, its proof and its extensions, from a different viewpoint and method of approach than that of M. Bessmertnyĭ. In fact, we were very inspired and motivated by the abstract theory of composites approach to a similar realization problem developed by G. Milton in [16GM7], as well as the approach of R. Duffin in [55RD] on synthesis problems in electrical network theory that can be solved using elementary algebraic operations. Our main theorem (i.e., Theorem 77) provides an extension of M. Bessmertnyĭ's Theorem [02MB, Theorem 1.1] as mentioned above. In addition, other than taking a different approach to the presentation, we provide an example for the realization of Kronecker products: Part I (Example 70), which is not included in [20SW]. Also, we include the appropriate formula for realizations of simple products (Lemma 68), which was not recognized until later in [21SW] when proving our extension to arbitrary fields not characteristic two for our HMV theorem.

1.6 Chapter 6: Symmetric Determinant Representation

In this thesis, "*HMV Theorem*" is referring to the theorem that any real polynomial $p(z) \in \mathbb{R}[z]$ in n variables $[z = (z_1, \ldots, z_n)]$ has a *symmetric determinant representation*, i.e., there exists a linear matrix pencil $A_0 + \sum_{i=1}^{n} z_i A_i$ with symmetric matrices $A_0, \ldots, A_n \in \mathbb{R}^{m \times m}$ such that

$$p(z) = \det \left(A_0 + \sum_{i=1}^{n} z_i A_i \right). \tag{1.1}$$

This was first proven in 2006 by J. W. Helton, S. McCullough, and V. Vinnikov [06HMV]. In Chapter 6, we provide a new proof to the HMV theorem (see Theorem 84 and its proof) as a direct corollary of the results established by the thesis. For instance: the elementary theory of determinants (Lemma 9 and 8 of Chapter 2), the sum of Schur complements (Lemma 30 in Chapter 4), the realization of a simple products (Lemma 68 in Chapter 5). One of the merits of our proof that makes it short and elementary is that it requires no prior knowledge of multidimensional systems theory (as compared to [06HMV] and [12RQ]) or advanced representation theory for multivariate polynomials (as compared to [11BG]). In Chapter 6 we provide an example (that has been corrected from [21SW]) that not only demonstrates our approach but also illustrates the known problem of the HMV theorem not being true in a field of characteristic 2 (e.g., $\mathbb{F}_2$). As such, we provide a discussion on this and then extend our results to arbitrary fields different from characteristic 2. This discussion helped us pinpoint (Lemma 68) the exact reason there exist field constraints (field not characteristic two) for not only HMV theorem but also the Bessmertnyĭ realizability problem. In addition, Theorem 84 gave insight into how treating non-invertibility of matrices has a more general approach in Chapter 4 by use of Lemma 54.

1.7 Chapter 7: Open Problems and Future Directions

As mentioned, we discuss the open problems associated with the Bessmertnyĭ class of functions (Proposition 87) and the associated Milton class of functions Proposition 88) which involve Bessmertnyĭ realizations of homogeneous rational matrix functions with positive-semidefinite matrix pencils and their relationship to Schur complements and effective operators in the theory of composites. The open problem of the Bessmertnyĭ class of functions is detailed in [04KV] and recent developments have been made in [21MB]. In Chapter 7, we discuss this along with our intentions to incorporate and develop these methods in our techniques in progressing the associated Milton class of functions.

We conclude with a discussion on future directions of our study that include extending the Bessmertnyĭ Realization Theorem to include additional symmetries that arise in applications like the ones in Chapter 3. We are interested in developing Schur complement algebra for reactance functions or characterize other potential functional symmetries associated with rotational symmetries. We are also interested in developing computational methods to implement the Schur complement algebra and operations from Chapter 4 to construct realizations with symmetries when they exist and to test conjectures or hypothesises for problems with multivariate rational matrix functions. Our research program includes future work in coding a program for computing Bessmertnyĭ realizations by using symbolic algebraic functions that are the algebra and operations of Chapter 4 for a more algorithmic approach to realizability [as can be seen in the proof flow diagram of the Bessmertnyĭ Realizability Theorem (Figure 5.2) and the flow diagram of the HMV Theorem (Figure 6.1) of Chapter 5]. We are able to identify the exact source of failure to the the Bessmertnyĭ realizability theorem in fields characteristic two, hence extending the Bessmertnyĭ Realizability Theorem in Chapter 5 over any arbitrary field not characteristic two (like the field extension of HMV theorem in Chapter 6 is also future research). In addition, investigating the realizability of simple products failing in fields characteristic two and conjecturing the lack of existence or existence of realizations over fields characteristic two. Finally, investigating recent developments on the Bessmertnyĭ class of functions and use the insights obtained to study the Milton class of functions along with its application in the theory of composites (e.g., effective operators and their bounds, limitations, as well as realizability problems).

Chapter 2

Preliminaries

Let $\mathbb{F}$ denote an arbitrary field (e.g., $\mathbb{F} = \mathbb{C}$ or $\mathbb{F} = \mathbb{R}$) and its characteristic, n, be denoted by $\mathrm{char}(\mathbb{F}) = n$ or $\mathbb{F}_n$. Let the set of n-tuples be denoted by $\mathbb{F}^n$, where $z = (z_1, \ldots, z_n)$ denotes a point in $\mathbb{F}^n$, the polynomials $p(z) = p(z_1, \ldots, z_n)$ in n-variables (n indeterminants) with coefficients in a field $\mathbb{F}$ be denoted by $\mathbb{F}[z] = \mathbb{F}[z_1, \ldots, z_n]$, and the field of all rational functions in the n-variables with coefficients in $\mathbb{F}$, be denoted by $\mathbb{F}(z) = \mathbb{F}(z_1, \ldots, z_n)$ {its elements are the fractions of the form $\frac{p(z)}{q(z)}$, where $p(z), q(z) \in \mathbb{F}[z]$ with $q(z) \not\equiv 0$ [i.e., $q(z)$ is not the zero polynomial]}. We will denote the set of all $m \times n$ matrices with entries in $\mathbb{F}$ or $\mathbb{F}(z)$ as $\mathbb{F}^{m\times n}$ or $\mathbb{F}(z)^{m\times n}$, respectively. Complex conjugate of a complex number c will be denoted as $\overline{c}$. The transpose and conjugate transpose of a matrix A will be denoted by A^T and A^* (i.e., $A^* = \overline{A}^T$), respectively, and the inverse of an invertible matrix A will be denoted by A^{-1}. The symbol $\det A$ will denote the determinant of a square matrix A. The $m \times m$ identity matrix and the $m \times m$ zero matrix will be denoted by I_m and 0_m, respectively.

We will denote any matrix $A \in \mathbb{F}^{m\times n}$ that is partitioned in 2×2 block matrix form as

$$A = [A_{ij}]_{i,j=1,2} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{c:c} A_{11} & A_{12} \\ \hdashline A_{21} & A_{22} \end{array}\right], \tag{2.1}$$

where the matrix $A_{ij} \in \mathbb{F}^{m\times n}$ is called the (i, j)-block of A. If, in addition, the matrices A_k are all 2×2 block matrices partitioned conformally, then we can partition the matrix function $A(z) \in \mathbb{F}(z)^{m\times m}$, $z = (z_1, \ldots, z_n)$, conformally as a 2×2 block matrix and denote this by

$$A(z) = [A_{ij}(z)]_{i,j=1,2} = \begin{bmatrix} A_{11}(z) & A_{12}(z) \\ A_{21}(z) & A_{22}(z) \end{bmatrix} = \left[\begin{array}{c:c} A_{11}(z) & A_{12}(z) \\ \hdashline A_{21}(z) & A_{22}(z) \end{array}\right],$$

in other words, this block structure of $A(z)$ is independent of the variable z. For any rational $\mathbb{F}^{m\times m}$-valued function $A(z)$ (i.e., $A(z) \in \mathbb{F}(z)^{m\times m}$), we will write $\det A(z) \not\equiv 0$

whenever $\det A(z)$ is not identically equal to 0 as a rational function.

Definition 1 (Linear matrix pencil) *Let $A(z)$ be a linear matrix function in n-variables $z_1, \dots, z_n$ with coefficients in $\mathbb{F}$. A **linear matrix pencil** is of the form*

$$A(z) = A_0 + z_1 A_1 + \cdots + z_n A_n, \tag{2.2}$$

*where $A_0, A_1, \dots, A_n \in \mathbb{F}^{m\times m}$. If, in addition, A_j are real, symmetric, Hermitian, real and symmetric matrices for all $j = 0, 1, \dots, n$ then $A(z)$ is called a **linear real, symmetric, Hermitian, real and symmetric matrix pencil**, respectively.*

2.1 The Schur complement

Definition 2 (Schur complement) *Let $A = [A_{ij}]_{i,j=1,2} \in \mathbb{F}^{m\times m}$ be a 2×2 block matrix. The **Schur Complement** of a matrix $A = [A_{ij}]_{i,j=1,2}$ with respect to A_{22} [i.e., with respect to its $(2,2)$-block A_{22}], will be denoted as A/A_{22} and defined by*

$$A/A_{22} = A_{11} - A_{12}A_{22}^{-1}A_{21}, \tag{2.3}$$

whenever A_{22} is invertible. Similarly, the Schur complement of A with respect to A_{11}, denoted by A/A_{11}, is defined as

$$A/A_{11} = A_{22} - A_{21}A_{11}^{-1}A_{12},$$

whenever A_{11} is invertible. Furthermore, let $A = [A_{ij}(z)]_{i,j=1,2} \in \mathbb{F}(z)^{m\times m}$ be a 2×2 block matrix-function, we may write the Schur complement with respect to $A_{22}(z)$ as

$$A(z)/A_{22}(z) = A_{11}(z) - A_{12}(z)A_{22}(z)^{-1}A_{21}(z), \tag{2.4}$$

whenever $\det A_{22}(z) \not\equiv 0$, and treat it as rational matrix-valued function of z.

Proposition 3 *Let $[A_{ij}]_{i,j=1,2}$ be a 2×2 block matrix. The system of linear equations, $Av = b$, where*

$$b = \begin{bmatrix} w \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

in block from is

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} w \\ 0 \end{bmatrix}$$

which is equivalent to the linear system of equations

$$v_2 = -A_{22}^{-1}A_{21}v_1$$
$$A/A_{22}v_1 = w$$

provided A_{22} is invertible.

Proof. Let $[A_{ij}]_{i,j=1,2}$ be a 2×2 block matrix, with A_{22} invertible, that satisfies the system of equations $Av = b$, where

$$b = \begin{bmatrix} w \\ 0 \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

. Then the linear system of equations in block form is

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} A_{11}v_1 + A_{12}v_2 \\ A_{21}v_1 + A_{22}v_2 \end{bmatrix} = \begin{bmatrix} w \\ 0 \end{bmatrix},$$

which is equivalent to the system of equations

$$A_{11}v_1 + A_{12}v_2 = w, \tag{2.5}$$
$$A_{21}v_1 + A_{22}v_2 = 0. \tag{2.6}$$

Solving for v_2 in (2.6), we have $A_{22}v_2 = -A_{21}v_1$, which implies $v_2 = -(A_{22})^{-1}A_{21}v_1$. By back substitution in (2.5) it follows that,

$$\begin{aligned} A_{11}v_1 + A_{12}v_2 &= w \\ A_{11}v_1 + A_{12}(-A_{22}^{-1}A_{21}v_1) &= w \\ (A_{11} - A_{12}A_{22}^{-1}A_{21})v_1 &= w \\ A/A_{22}v_1 &= w, \end{aligned}$$

which is what we wanted to show. ∎

Some key elementary properties of the Schur complement, under the assumption that A_{22}^{-1} exists, are

$$(\lambda A) / (\lambda A)_{22} = \lambda (A/A_{22}), \tag{2.7}$$
$$\overline{(A/A_{22})} = (\overline{A}) / (\overline{A})_{22}, \tag{2.8}$$
$$(A/A_{22})^T = (A^T) / (A^T)_{22}, \tag{2.9}$$
$$(A/A_{22})^* = (A^*) / (A^*)_{22}, \tag{2.10}$$

for every $\lambda \in \mathbb{F} \setminus \{0\}$.

Lemma 4 (Schur complement decomposition) *If $A \in \mathbb{F}^{m\times m}$ is a 2×2 block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{2.11}$$

with Schur complement A/A_{22} then

$$A = \begin{bmatrix} I_{m-p} & A_{12}A_{22}^{-1} \\ 0 & I_p \end{bmatrix} \begin{bmatrix} A/A_{22} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I_{m-p} & 0 \\ A_{22}^{-1}A_{21} & I_p \end{bmatrix}. \tag{2.12}$$

Proof. Let $A \in \mathbb{F}^{m\times m}$ is a 2×2 block matrix of the form (2.11) and $A_{22} \in \mathbb{F}^{p\times p}$ is invertible, then the factorization of A in (2.12) follows immediately from block multiplication,

$$\begin{aligned}
\begin{bmatrix} I & A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A/A_{22} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ A_{22}^{-1}A_{21} & I \end{bmatrix} &= \begin{bmatrix} I & A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A/A_{22} & 0 \\ A_{22}A_{22}^{-1}A_{21} & A_{22} \end{bmatrix} \\
&= \begin{bmatrix} A/A_{22} + A_{12}A_{22}^{-1}(A_{22}A_{22}^{-1}A_{21}) & A_{12}A_{22}^{-1}A_{22} \\ A_{22}A_{22}^{-1}A_{21} & A_{22} \end{bmatrix} \\
&= \begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} + A_{12}A_{22}^{-1}A_{21} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \\
&= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \\
&= A.
\end{aligned}$$

■

Corollary 5 (Inverse formulas) *Let $A \in \mathbb{F}^{m\times m}$ is a 2×2 block matrix of the form (2.11) and $A_{22} \in \mathbb{F}^{p\times p}$ is invertible, then*

$$\begin{bmatrix} I_{m-p} & A_{12}A_{22}^{-1} \\ 0 & I_p \end{bmatrix}^{-1} = \begin{bmatrix} I_{m-p} & -A_{12}A_{22}^{-1} \\ 0 & I_p \end{bmatrix},$$

$$\begin{bmatrix} I_{m-p} & 0 \\ A_{22}^{-1}A_{21} & I_p \end{bmatrix}^{-1} = \begin{bmatrix} I_{m-p} & 0 \\ -A_{22}^{-1}A_{21} & I_p \end{bmatrix}.$$

Proof. Let $A \in \mathbb{F}^{m\times m}$ is a 2×2 block matrix with the factorization (2.11), with A_{22}^{-1} exists, consider

$$\begin{bmatrix} I & A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & A_{12}A_{22}^{-1} - A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

Similarly,

$$\begin{bmatrix} I & 0 \\ A_{22}^{-1}A_{21} & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ A_{22}^{-1}A_{21} - A_{22}^{-1}A_{21} & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

This proves they are inverses. ■

2.2 Kronecker Product Properties

Definition 6 (Kronecker product) *The **Kronecker product** of the two matrices $A = [a_{ij}]_{i,j=1}^{m,n} \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{p\times q}$ is the matrix $A \otimes B \in \mathbb{F}^{mp\times qn}$ defined as*

$$A \otimes B = [a_{ij}B]_{i,j=1}^{m,n} \in \mathbb{F}^{mp\times nq}. \tag{2.13}$$

Some key elementary properties of the Kronecker products are

$$\overline{A \otimes B} = \overline{A} \otimes \overline{B}, \tag{2.14}$$

$$(A \otimes B)^T = A^T \otimes B^T, \tag{2.15}$$

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}, \tag{2.16}$$

$$(A \otimes B)^* = A^* \otimes B^*. \tag{2.17}$$

2.3 Direct Sum & Determinants

Definition 7 (Direct sum) *The **direct sum** $A \oplus B$ of two matrices $A \in \mathbb{F}^{k\times l}$ and $B \in \mathbb{F}^{p\times q}$ is defined to be the matrix $A \oplus B \in \mathbb{F}^{(k+p)\times(l+q)}$ with the 2×2 block matrix form*

$$A \oplus B = \begin{bmatrix} A & 0_{k\times p} \\ 0_{l\times q} & B \end{bmatrix}.$$

The following result is well-known, see [02FIS, Sec. 4.3, Exercise 21]).

Lemma 8 (Determinant of a direct sum) *If $A \in \mathbb{F}^{m\times m}$ and $B \in \mathbb{F}^{k\times k}$ are symmetric matrices then the direct sum $A \oplus B$ is a symmetric matrix and*

$$\det(A \oplus B) = (\det A)(\det B). \tag{2.18}$$

The following result is also well-known, see [05FZ, p. 19, Theorem 1.1]).

Lemma 9 (Schur's determinant formula) *If $A = [A_{ij}]_{i,j=1,2} \in \mathbb{F}^{N\times N}$ and A_{22} is invertible then*

$$\det A = \det(A_{22}) \det(A/A_{22}). \tag{2.19}$$

Proof. The proof is a direct result of Lemma 4. ■

Chapter 3

Applications

In this chapter we introduce a few applications in order to motivate the study of Schur complements (Definition 2). These applications will illustrate how the Schur complement naturally arises as part of the parameterization of the solution of certain constrained linear systems (Proposition 3) with algebraic-functional symmetries related to Chapters 4 and 5 [e.g., real and symmetric linear matrix pencil (Definition 1)].

Our main motivation for the results of this thesis is its potential applications to problems associated with realizability, synthesis, and inverse problems involving either the impedance and Dirichlet-to-Neumann matrices in electrical networks or the effective tensor in the theory of composites. Here we give several examples which show how our main objects of study, namely, Schur complements of linear matrix pencils, arise as impedance matrices of electrical networks, Dirichlet-to-Neumann matrices of resistor networks, and as effective conductivity tensors in the theory of composites.

The rest of the chapter will proceed as follows. In Section 3.1, we begin with a basic introduction to electric circuit theory, which is primarily from [69BB] and [60PB]. In Section 3.2, we show how the impedance matrix (in the loop equations of the electric network) is a Schur complement of a real and symmetric linear matrix pencil, which maps a source current to the source voltage. In Section 3.3, we give an introduction to the Dirichlet-to-Neumann map from electric circuit theory, which is primarily from [00CM]. We show how the Dirichlet-to-Neumann matrix is a Schur complement of a real and symmetric linear matrix pencil which maps the voltage potentials on the boundary to the currents into the boundary. Finally, in Section 3.4 we include an introduction to the effective conductivity tensor in the theory of composites, which is primarily from [16GM]. We show how the effective conductivity tensor is a Schur complement of a real and symmetric linear operator pencil, which maps the average electric field to the average electric current in the conductivity equation with periodic coefficients.

3.1 Electrical Circuit Theory

In this section we introduce the basic background on electrical network theory (our primary source on this is [69BB]) which we need for Sections 3.2 and 3.3. All vector spaces considered here are spaces of column vectors over the field $\mathbb{R}$ (or $\mathbb{C}$ when convenient).

The fundamental problem to solve in any electric network Γ is to solve for the voltage and current in the network. In the next two sections we will consider this problem in more detail. In this section though we just treat the definition of an electric network Γ as follows.

An *electric network* Γ will consist of:

1. a finite linear directed graph G (i.e., the topology of the network: nodes, edges, and the directed connections between them);
2. constraints (i.e., Kirchhoff's laws: KCL, KVL);
3. constitutive relations [e.g., Ohm's law: $v(t) = Ri(t)$].

The *graph* G associated with Γ is a finite (i.e., finite set of nodes and edges) linear directed (i.e., edges are oriented) graph such that every edge (i.e., "branch") has a unique pair of distinct nodes incident on it (these nodes are its endpoints; no edge is itself a loop, i.e., no self-loops) and such that every node (i.e., "vertex") is incident on at least one edge (allows the possibility that two edges have the same pair of nodes incident on them, i.e., multi-edges are possible, but no node is an isolated point) [60PB, p. 216], [61SR, p. 9], [69BB, p. 70]. We will denote by p_i the ith node of the graph for $i = 1, \ldots, \kappa$ (where G has κ nodes) and e_j denote the jth edge for $j = 1, \ldots, \mu$ of the graph. Next, the topology of the network with respect to the incidence of nodes on edges and their directions is encoded in the *complete node-edge incidence matrix* [69BB, p. 73, Sec. 2.2] of the graph, which is denoted by the $\kappa \times \mu$ matrix $A_a = [a_{ij}]$ (the subscript a represents *all*) whose entries are:

$$a_{ij} = \begin{cases} 1 & \text{if the } j\text{th edge is incident on and directed away from the } i\text{th node,} \\ -1 & \text{if the } j\text{th edge is incident on and directed toward the } i\text{th node,} \\ 0 & \text{if the } j\text{th edge is not incident on the } i\text{th node.} \end{cases} \tag{3.1}$$

A *loop* is any simple closed path in the graph G and we enumerate them denoting the ith loop by l_i, for $i = 1, \ldots, \delta$, where δ is the number of loops in G. Next, we give a direction (i.e., orientation) to each loop and encode the information about the incidence of the edges on loops and their orientations, in the *complete loop matrix* [69BB, p. 77, Sec. 2.2, Chap. 2], which is denoted by the $\delta \times \mu$ matrix $B_a = [b_{ij}]$, whose entries are:

$$b_{ij} = \begin{cases} 1 & \text{if the } j\text{th edge is in the } i\text{th loop and their orientations are the same,} \\ -1 & \text{if the } j\text{th edge is in the } i\text{th loop and their orientations are different,} \\ 0 & \text{if the } j\text{th edge is not in the } i\text{th loop.} \end{cases} \tag{3.2}$$

A *tree* of a connected graph G is a connected subgraph of G that contains all the nodes of G but has no loops [69BB, p. 71, Sec. 2.2, Chap. 2]. The edges of a chosen tree are called its *twigs* and the edges not part of that chosen tree but in that connected component of G are called its *links* (i.e., "chords") [69BB, pp. 71, Sec. 2.2, Chap. 2]. In general (as G need not be a connected graph), the graph G has a finite number ς of maximally connected subgraphs, i.e., its connected components, and in this case a *forest* is a subgraph of G formed by selecting one tree from each of its connected components [60PB, p. 225], [69BB, p. 72, Chap. 2]. Note that according to this convention, for a connected graph G ($\varsigma = 1$), any forest in it contains only one tree; for a graph G that is not connected ($\varsigma > 1$), any forest in it contains ς trees (one tree for each of the ς connected components of G).

The *fundamental loops* (f-loops) in G with respect to a chosen forest are a special class of loops from G. There is a one-to-one correspondence between the links of each tree in the forest and the f-loops which can be described as follows: Choose any tree in the forest and choose any link of that tree. Then its incident nodes are in the tree. When we replace the link at those nodes in the tree a loop will form which is in G (i.e., it is the loop l_i for some $i \in \{1, \ldots, \delta\}$), this loop is the f-loop associated with that link [69BB, p. 79, Sec. 2.2, Chap. 2], [60PB, p. 219]. For a graph with ν f-loops, we enumerate them and denote by $(l_f)_i$ the ith f-loop for $i = 1, \ldots, \nu$. Then the loop matrix for these f-loops, called the *fundamental loop matrix* (f-loop matrix) [60PB, p. 219], [69BB, pp. 79, Sec. 2.2, Chap. 2], is denoted by the $\nu \times \mu$ matrix $B_f = [(b_f)_{ij}]$ whose entries are:

$$(b_f)_{ij} = \begin{cases} 1 & \text{if the } j\text{th edge is in the } i\text{th } f\text{-loop and their orientations are the same,} \\ -1 & \text{if the } j\text{th edge is in the } i\text{th } f\text{-loop and their orientations are different,} \\ 0 & \text{if the } j\text{th edge is not in the } i\text{th } f\text{-loop.} \end{cases} \tag{3.3}$$

The following well-known results related to the topology of the graph G (which need not be a connected graph) are fundamental in electrical network theory (see [60PB, pp. 216, 218, 219, 225], [69BB, pp. 74, 79-81], [61SR, Sec. 6, Chap. 5]):

(i) $A_a B_a^T = 0$ [i.e., $B_a A_a^T = 0$].

(ii) $\operatorname{rank} A_a = \kappa - \varsigma$.

(iii) the rows of B_f, which are rows of B_a, form a basis for the row space of B_a and

$$\operatorname{rank} B_a = \operatorname{rank} B_f = \nu = \mu - \kappa + \varsigma. \tag{3.4}$$

The following results are immediate consequences of (i), (ii), and (iii) by applying elementary results from linear algebra [02FIS]:

$$\operatorname{Ker} A_a = \operatorname{Ran} B_a^T = \operatorname{Ran} B_f^T, \;\; \operatorname{Ran} A_a^T = \operatorname{Ker} B_a = \operatorname{Ker} B_f. \tag{3.5}$$

We need to pointed out here two conventions in electrical network theory. The first convention is that typically the rows of B_f are ordered so that they are a submatrix of B_a, i.e., B_f is formed from B_a by just crossing out the rows of B_a corresponding to the indexed loops $l_1, \cdots, l_\delta$ that are not the f-loops and the f-loops would then be indexed in terms of the remaining loops: $(l_f)_1 = l_{i_1}, \ldots, (l_f)_\nu = l_{i_\nu}$ with $1 \leq i_1 < \cdots < i_\nu \leq \delta$. Now though, in our original definition of an f-loop matrix B_f above that we use, it may be that B_a is not a submatrix of B_a and so in order to make is so you just need to correctly reorder (i.e., permute) the rows of B_f so that it is and in which case it will still be an f-loop matrix according to our original definition, but now it is also a submatrix B_a. Second, the standard convention in electrical network theory is to choose the orientation of the f-loops in the same orientation of each f-loop's corresponding link in the tree that defines it [60PB, p. 219], [69BB, p. 79]. But then this matrix of f-loops would no longer be a submatrix of B_a in general. In order to adhere to the standard convention and for B_f to be a submatrix of B_a (3.2) we simply multiply the f-loop rows in B_a by -1 whenever the f-loop orientation and that f-loop's link direction are different. This can be a bit confusing when first being introduced to electrical network theory, but hopefully by pointing this issue out we have shown how it can be resolved in a satisfying manner from a engineering and mathematical perspective. This concludes the discussion on the topological aspects of the electrical network Γ.

Now we will discuss the constraints on the current and voltage, namely, Kirchhoff's laws [69BB, pp. 58-59, Sec. 2.1, Chap. 2]. First, Kirchhoff's current law (KCL) states that sum of all currents leaving any node equals zero, i.e.,

$$A_a i = 0. \qquad \text{(KCL)} \tag{3.6}$$

Second, Kirchhoff's voltage law (KVL) states that the sum of voltages of all edges forming any loop equals zero, i.e.,

$$B_a v = 0. \qquad \text{(KVL)} \tag{3.7}$$

These Kirchhoff's laws can be reinterpretated in the presence of independent sources as follows. There are two types of independent sources [69BB, p. 47-48, Sec. 1.5, Chap. 1], the $\mu \times 1$ column vectors of voltage sources v_g and current sources i_g, respectively, and v and i denote the $\mu \times 1$ column vectors of edge voltages [69BB, p. 95, Chap. 2] and edge currents [69BB, p. 90, Chap. 2] (which are implicitly functions of time t), respectively, then Kirchhoff's laws can be written as

$$A_a(i - i_g) = 0, \qquad \text{(KCL)} \tag{3.8}$$

$$B_a(v - v_g) = 0. \qquad \text{(KVL)} \tag{3.9}$$

We can then take the Laplace transform, which we will denote by $\mathcal{L}$, with the following relations [69BB, Appendix 3]

$$\mathcal{L}\{v(t)\} = V(s),\ \ \mathcal{L}\{i(t)\} = I(s),\ \ \mathcal{L}\{v_g(t)\} = V_g(s),\ \ \mathcal{L}\{i_g(t)\} = I_g(s). \tag{3.10}$$

With this notation, the Kirchhoff's laws become

$$A_a(I - I_g) = 0, \qquad \text{(KCL)} \tag{3.11}$$
$$B_a(V - V_g) = 0. \qquad \text{(KVL)} \tag{3.12}$$

It is this form of Kirchhoff's law that we need in Sections 3.2 and 3.3 when we introduce the loop impedance matrix (3.25) and the Kirchhoff matrix (3.73). This concludes the discussion on the constraint aspects of the electrical network Γ, namely, Kirchhoff's laws.

Now we will discuss the constitutive relations of the network Γ, which includes resistance R, induction L, or capacitance C. The notation of each relation (e.g., R, L, C) of an edge between two arbitrary nodes p_a and p_b, with the edge directed from p_a toward p_b, is shown graphically in Figure 3.1. The voltage drop across an edge with a

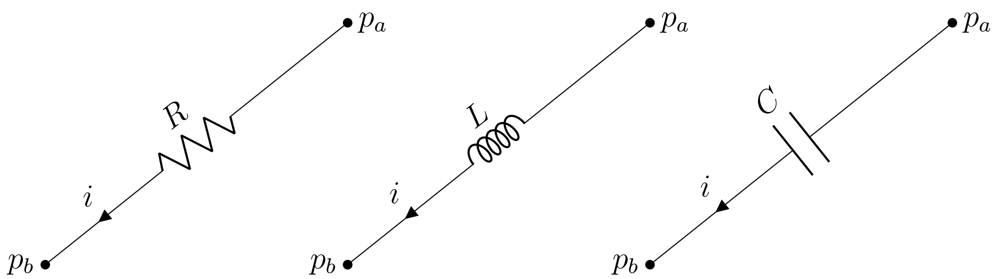


Figure 3.1: R, L, and C circuit elements.

resistor, inductor, and capacitor are, respectively,

$$v(t) = Ri(t),\ \ v(t) = L\frac{di}{dt},\ \text{and}\ v(t) = \frac{1}{C}\int_{-\infty}^{t} i(\tau)d\tau \tag{3.13}$$

[69BB, p. 40, Chap. 1]. To solve for the voltage drop of each equation in (3.13) we can take the Laplace transform $\mathcal{L}$ with the following relations [69BB, pp. 898-916, Appendix 3]

$$\mathcal{L}\left\{\frac{di}{dt}\right\} = sI(s),\ \text{and}\ \mathcal{L}\left\{\int_{-\infty}^{t} i(\tau)d\tau\right\} = \frac{1}{s}I(s). \tag{3.14}$$

Hence, after taking the Laplace transform of each constitutive relation in (3.13) we have

$$V = RI, \;\; V = sLI, \text{ and } V = \frac{1}{sC}I. \tag{3.15}$$

An *RLC network* Γ will be a network with only these types of network elements (cf. Remark 10). More specifically, it's constitutive relations in the Laplace domain can be put in the following matrix form:

$$V = \tilde{\mathcal{Z}} I, \tag{3.16}$$

[60PB, p. 221] where $\tilde{\mathcal{Z}}$ is called the *edge impedance matrix* [69BB, p. 101, Sec. .3, Chap. 2] and is an $\mu \times \mu$ diagonal matrix whose diagonal entries are:

$$\tilde{\mathcal{Z}}_{jj} = \begin{cases} R_j & \text{if the } j\text{th edge is resistive,} \\ sL_j & \text{if the } j\text{th edge is self-inductive,} \\ \frac{1}{sC_j} & \text{if the } j\text{th edge capacitive.} \end{cases} \tag{3.17}$$

More precisely, the constitutive relation on the jth edge between the jth edge voltage V_i and the jth edge current I_j is given by

$$V_j = \tilde{\mathcal{Z}}_{jj} I_j, \tag{3.18}$$

for each $j = 1, \dots, \mu$ [60PB, p. 221]. This concludes the discussion on the constitutive relations for an electrical network Γ.

In conclusion, this is the necessary background for this thesis on electrical networks which we will only use in the next two sections. In particular, in Section 3.2 we treat only these *RLC* networks and in Section 3.3 we treat resistor only networks (i.e., no inductors and capacitors).

Remark 10 *It should be remarked though that there are more general electrical networks with other constitutive relations such as those with mutual inductances which could be consider as RLC networks too and, more generally, RLCTG networks (T refers to transformers and G refers to gyrators). But for these the constitutive relations are beyond the scope of this introduction even though they are worth considering for the purposes of realizability and synthesis problems in passive electrical network theory.*

3.2 Impedance Matrices of an *RLC* Network

In this section we will introduce two other types of impedance matrices for an *RLC* network (namely, the loop and source impedance matrices denoted by $\mathcal{Z}$ and Z, respectively) derived from the edge impedance matrix $\tilde{\mathcal{Z}}$ [see (3.17)] introduced in Sec.

3.1. In order to do this, we also need to introduce the loop method (in this section) and "source loop method" (in Subsection 3.2.1).

Now we proceed to describe the loop method [60PB, p. 221]. Let Γ be an RLC network with $A_a, B_a, V_g, \tilde{\mathcal{Z}}$ given and

$$I_g = 0 \tag{3.19}$$

(i.e., independent voltage sources and no independent current sources). The fundamental network problem to solve now is for the edge voltages V and edge currents I in the network in terms of V_g. To do this we must solve the network equations:

$$A_a I = 0, \;\; B_a(V - V_g) = 0, \tag{3.20}$$

$$V = \tilde{\mathcal{Z}} I. \tag{3.21}$$

By considering only the f-loops, Kirchhoff's voltage law (KVL) in (3.20) reduces to the matrix equation

$$B_f(V - V_g) = 0 \tag{3.22}$$

[60PB, p. 220] and the edge currents I from Kirchhoff's currrent law (KCL) in (3.20) can be solved in terms of the $\nu \times 1$ column vector of f-loop currents $I_{\mathfrak{m}}$ by the following matrix equation

$$B_f^T I_{\mathfrak{m}} = I. \tag{3.23}$$

In particular, it should be noted that the equations (3.22) and (3.23) are equivalent to the equations (3.20) because of the equalities in (3.5).

The next proposition is well-known {see, for instance, [60PB, p. 221, Eq. (14)]}, where $\mathcal{Z}$ is the $\nu \times \nu$ *loop impedance matrix* [69BB, p. 105, Sec. 2.4, Chap. 2].

Proposition 11 (loop method) *Consider the network equations (3.20) and (3.21) for some RLC network Γ with complete node-edge matrix A_a (3.1), complete loop matrix B_a (3.2), independent source voltage V_g, no independent current source (i.e., $I_g = 0$), and diagonal edge impedance matrix $\tilde{\mathcal{Z}}$ with diagonal entries (3.17). Then the following equation (called the loop impedance equation) is satisfied:*

$$B_f V_g = \mathcal{Z} I_{\mathfrak{m}}, \tag{3.24}$$

where B_f is any f-loop matrix (3.3), $\mathcal{Z}$ is the loop impedance matrix defined by the formula

$$\mathcal{Z} = B_f \tilde{\mathcal{Z}} B_f^T, \tag{3.25}$$

and the f-loop currents $I_{\mathfrak{m}}$ are related to edge currents I by (3.23). Furthermore, if $\mathcal{Z}$ is invertible then the edge voltages V and the edge currents I are given by the following *formulas*

$$V = \tilde{\mathcal{Z}} I, \quad I = B_f^T I_{\mathfrak{m}} = B_f^T \mathcal{Z}^{-1} B_f V_g. \tag{3.26}$$

Moreover, if $\tilde{\mathcal{Z}}$ is a positive definite matrix (i.e., $\tilde{\mathcal{Z}} > 0$), then $\mathcal{Z}$ is invertible and, in particular, formulas (3.26) hold.

Proof. Let A_a, B_a, V_g, I_g, and $\tilde{\mathcal{Z}}$ be as in the hypotheses. Then equation (3.24) follows immediately from (3.22) and (3.23) since

$$B_f V_g = B_f V = B_f \tilde{\mathcal{Z}} I = B_f \tilde{\mathcal{Z}} B_f^T I_{\mathfrak{m}} = \mathcal{Z} I_{\mathfrak{m}}. \tag{3.27}$$

Now assume that $\mathcal{Z}$ is invertible. Then by (3.27) we can solve for $I_{\mathfrak{m}}$ with

$$I_{\mathfrak{m}} = \mathcal{Z}^{-1} B_f V_g. \tag{3.28}$$

Hence, using (3.23) we can solve for the edge current I with

$$I = B_f^T I_{\mathfrak{m}} = B_f^T \mathcal{Z}^{-1} B_f V_g. \tag{3.29}$$

Therefore, by this and the hypotheses that (3.21) is satisfied, the formulas (3.26) now follow.

Finally, assume that $\tilde{\mathcal{Z}}$ is a positive definite matrix, i.e., $\tilde{\mathcal{Z}} > 0$. Suppose $I_{\mathfrak{m}} \neq 0$ and $\langle I_{\mathfrak{m}}, \mathcal{Z} I_{\mathfrak{m}} \rangle = 0$. Then this implies

$$0 = \left\langle I_{\mathfrak{m}}, \mathcal{Z} I_{\mathfrak{m}} \right\rangle = \left\langle I_{\mathfrak{m}}, B_f \tilde{\mathcal{Z}} B_f^T I_{\mathfrak{m}} \right\rangle = \left\langle B_f^T I_{\mathfrak{m}}, \tilde{\mathcal{Z}} B_f^T I_{\mathfrak{m}} \right\rangle \tag{3.30}$$

and hence, since $\tilde{\mathcal{Z}} > 0$, this implies

$$\tilde{\mathcal{Z}} B_f^T I_{\mathfrak{m}} = 0. \tag{3.31}$$

But since $\tilde{\mathcal{Z}} > 0$ then $\tilde{\mathcal{Z}}$ must be invertible and thus,

$$B_f^T I_{\mathfrak{m}} = 0. \tag{3.32}$$

This shows that $\operatorname{Ker}(B_f^T) \neq \{0\}$ and hence $\operatorname{nullity}(B_f^T) \geq 1$. By (3.4), we know the rank of the $\nu \times \mu$ matrix B_f is ν and using "rank plus nullity theorem" [02FIS, p. 70, Theorem 2.3, Chap. 2], this implies

$$\nu = \operatorname{rank}(B_f^T) < \operatorname{rank}(B_f^T) + \operatorname{nullity}(B_f^T) = \dim(\mathbb{R}^\nu) = \nu, \tag{3.33}$$

a contradiction. This completes the proof. ∎

Remark 12 (Summary of loop method) *To summarize, we have provided sufficient conditions to answer the fundamental problem stated at the beginning of Section*

3.1 (i.e., to solve for the edge voltages V and edge currents I in a given network Γ). More specifically, given independent source voltages V_g and no independent source currents $I_g = 0$, the f-loop matrix B_f from (3.3), the edge impedance matrix $\tilde{\mathcal{Z}}$ (3.16), we can directly compute the edge voltages V and edge currents I in the network Γ (at least in the Laplace domain) with equation (3.26) provided $\mathcal{Z}$ is invertible.

Example 13 (Example of the loop method) *Consider the RLC network Γ as shown in Figure 3.1. We will use the loop method (see Proposition 11) to solve for the edge voltages V and edge currents I in terms of independent voltage sources V_g with no independent current sources ($I_g = 0$), where*

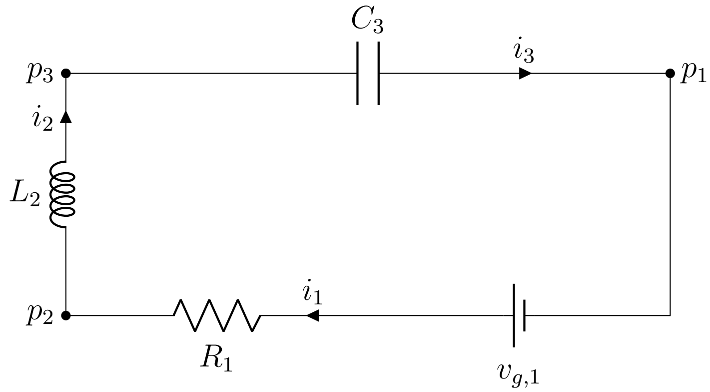


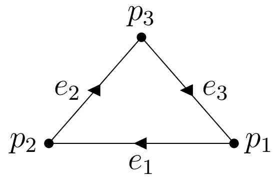
Figure 3.2: Example of an *RLC* network Γ with nodes p_1, p_2, p_3, and edges e_1, e_2, e_3 such that e_1 has resistor R_1 and a battery $v_{g,1}$, e_2 has an inductor L_2, and e_3 has a capacitor C_3.

$$V_g = \begin{bmatrix} V_{g,1} \\ 0 \\ 0 \end{bmatrix} \text{ and } I_g = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

First, we provide in Figure 3.3 the associated finite linear directed graph G of Γ, the complete node-edge incidence matrix A_a constructed from the topology of Γ [defined in (3.1)], and the associated edge impedance matrix $\tilde{\mathcal{Z}}$ [defined in (3.17)].

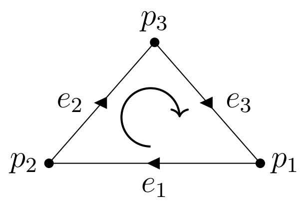
Next, in Figure 3.4 we give an orientation to each loop in G (there is only one loop and we choose a clockwise direction for its orientation) and construct the associated complete loop matrix B_a [defined in (3.2)].

Now we select a tree for each forest in G (in this example, there is only one forest since the graph G is connected). We illustrate in Figure 3.9 our chosen tree of G, the chosen tree's associated links (in this example there is only one link), and the corresponding f-loop matrix B_f [defined in (3.3)]. We are now ready to complete the



$$A_a = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad \tilde{\mathcal{Z}} = \begin{bmatrix} R_1 & 0 & 0 \\ 0 & sL_2 & 0 \\ 0 & 0 & \frac{1}{sC_3} \end{bmatrix}$$

Figure 3.3: Graph G, complete node-edge matrix A_a, and edge impedance matrix $\tilde{\mathcal{Z}}$ for Γ.



$$B_a = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

Figure 3.4: Clockwise orientation of the loop in G and the complete loop matrix B_a.

loop method. The reduced KVL (3.22) implies the following equality

$$B_f V = B_f V_g = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} V_{g,1} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} V_{g,1} \end{bmatrix}.$$

Furthermore, with the vector of loop currents denoted as

$$I_{\mathfrak{m}} = \begin{bmatrix} I_{\mathfrak{m},1} \end{bmatrix}, \tag{3.34}$$

we have by KCL (3.23) that the edge currents I are given in terms of the the f-loop currents by

$$I = B_f^T I_{\mathfrak{m}} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} I_{\mathfrak{m}} = \begin{bmatrix} I_{\mathfrak{m},1} \\ I_{\mathfrak{m},1} \\ I_{\mathfrak{m},1} \end{bmatrix}.$$

By Proposition 11, we can solve for our edge voltages V in the Laplace domain by first

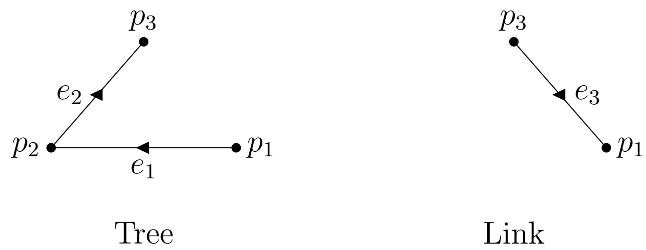


Figure 3.5: Tree, link, and f-loop matrix for the graph G of the network Γ.

calculating

$$\begin{aligned}
\begin{bmatrix} V_{g,1} \end{bmatrix} = B_f V_g = B_f \tilde{\mathcal{Z}} B_f^T I_{\mathfrak{m}} &= \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} R_1 & 0 & 0 \\ 0 & sL_2 & 0 \\ 0 & 0 & \frac{1}{sC_3} \end{bmatrix} \begin{bmatrix} I_{\mathfrak{m},1} \\ I_{\mathfrak{m},1} \\ I_{\mathfrak{m},1} \end{bmatrix} \\
&= \begin{bmatrix} R_1 & sL_2 & \frac{1}{sC_3} \end{bmatrix} \begin{bmatrix} I_{\mathfrak{m},1} \\ I_{\mathfrak{m},1} \\ I_{\mathfrak{m},1} \end{bmatrix} \\
&= \begin{bmatrix} R_1 I_{\mathfrak{m},1} + sL_2 I_{\mathfrak{m},1} + \frac{1}{sC_3} I_{\mathfrak{m},1} \end{bmatrix} \\
&= R_1 I_{\mathfrak{m}} + sL_2 I_{\mathfrak{m}} + \frac{1}{sC_3} I_{\mathfrak{m}}
\end{aligned}$$

and then solving for $I_{\mathfrak{m}}, I, V$ we get our desired solutions:

$$I_{\mathfrak{m}} = \left[\left(R_1 + sL_2 + \tfrac{1}{sC_3} \right)^{-1} V_{g,1} \right], \tag{3.35}$$

$$I = B_f^T I_{\mathfrak{m}} = \begin{bmatrix} \left(R_1 + sL_2 + \frac{1}{sC_3} \right)^{-1} V_{g,1} \\ \left(R_1 + sL_2 + \frac{1}{sC_3} \right)^{-1} V_{g,1} \\ \left(R_1 + sL_2 + \frac{1}{sC_3} \right)^{-1} V_{g,1} \end{bmatrix}, \tag{3.36}$$

$$V = \tilde{\mathcal{Z}} I = \begin{bmatrix} R_1 \left(R_1 + sL_2 + \frac{1}{sC_3} \right)^{-1} V_{g,1} \\ sL_2 \left(R_1 + sL_2 + \frac{1}{sC_3} \right)^{-1} V_{g,1} \\ \frac{1}{sC_3} \left(R_1 + sL_2 + \frac{1}{sC_3} \right)^{-1} V_{g,1} \end{bmatrix}. \tag{3.37}$$

3.2.1 Source Impedance Matrix

In this subsection, we begin to see exactly how the elementary electric network theory so far in this chapter is directly related to the results of this thesis.

The loop impedance matrix $\mathcal{Z}$ in (3.25) of an electrical network Γ is a multivariate rational $\nu \times \nu$ matrix function

$$\mathcal{Z} = \mathcal{Z}(z) = \mathcal{Z}(z_1, \ldots, z_\mu) \tag{3.38}$$

of μ complex variables $z_1, \ldots, z_\mu$ [$z = (z_1, \ldots, z_\mu)$], where μ is the number of edges in the network and the jth variable z_j is the jth diagonal entry of the diagonal matrix edge impedance matrix $\tilde{\mathcal{Z}}$ of the network Γ, i.e.,

$$z_j = \tilde{\mathcal{Z}}_{jj}, \;\; j = 1, \ldots, \mu, \tag{3.39}$$

$$\tilde{\mathcal{Z}} = z_1 E_{11} + \cdots + z_\mu E_{\mu\mu} = \begin{bmatrix} z_1 & & \\ & \ddots & \\ & & z_\mu \end{bmatrix}, \tag{3.40}$$

$$E_{jj} = e_j \otimes e_j, \;\; j = 1, \ldots, \mu \tag{3.41}$$

(where $\otimes$ is the Kronecker product and e_j, $j = 1, \ldots, \mu$ is the standard ordered basis of $\mu \times 1$ column vectors for $\mathbb{R}^\mu$).

Lemma 14 (loop impedance matrix is a linear pencil) *The loop impedance matrix $\mathcal{Z}$ of the network Γ with diagonal edge impedance matrix $\tilde{\mathcal{Z}}$, whose jth diagonal entry is $\tilde{\mathcal{Z}}_{jj} = z_j$, is a homogeneous real symmetric linear $\nu \times \nu$ matrix pencil in the complex variables z_j given by the formulas*

$$\mathcal{Z} = z_1 A_1 + \cdots + z_\mu A_\mu, \tag{3.42}$$

$$A_j = B_f (e_j \otimes e_j) B_f^T = (B_f e_j) \otimes (B_f e_j), \;\; j = 1, \ldots, \mu, \tag{3.43}$$

where $\otimes$ is the Kronecker product and e_j, $j = 1, \ldots, \mu$ is the standard ordered basis of $\mu \times 1$ column vectors for $\mathbb{R}^\mu$. Furthermore, A_j is a real symmetric and positive semidefinite $\nu \times \nu$ matrix with entries in $\{-1, 0, 1\}$ and rank

$$\operatorname{rank} A_j \in \{0, 1\} \tag{3.44}$$

whose range contains $B_f e_j$, the jth column of the f-loop matrix B_f, for each $j = 1, \ldots, \mu$. Moreover, $\operatorname{rank} A_j \neq 1$ *if and only if* $\operatorname{rank} A_j = 0$ *if and only if $A_j = 0$ if and only if $B_f e_j = 0$ if and only if $B_a e_j = 0$ if and only if the jth edge is not in any of the loops of the network Γ.*

Proof. The diagonal edge impedance matrix $\tilde{\mathcal{Z}}$ of the electrical network Γ can be written as (3.40) in terms of (3.39) and (3.41). It follows immediately from this that the loop impedance matrix $\mathcal{Z} = B_f \tilde{\mathcal{Z}} B_f^T$ in (3.25) of the network Γ is a multivariate rational $\nu \times \nu$ matrix function (3.38) of the μ complex variables $z_1, \ldots, z_\mu$ [$z = (z_1, \ldots, z_\mu)$] and is given by the formulas (3.42) and (3.43). The proof of the lemma now follows immediately from the formulas (3.43), the elementary properties of the Kronecker product $\otimes$, the definition of B_a as the complete loop matrix and B_f as an f-loop matrix of the network Γ, and the equalities (3.5). ■

Now let Γ be an electric network with graph G and $A_a, B_a, V_g, I_g, \tilde{\mathcal{Z}}$ given such that the diagonal edge impedance matrix $\tilde{\mathcal{Z}}$ has the form (3.39)-(3.41) and

$$I_g = 0 \tag{3.45}$$

(i.e., independent voltage sources and no independent current sources). Let B_f be an $\nu \times \nu$ submatrix of B_a which is an f-loop matrix for Γ as defined in (3.3). Then the jth row of $B_f V_g$ represents the algebraic sum of all the voltage sources contained in the jth f-loop [60PB, p. 221]. As such, it is possible for some of these rows to be zero and the rest being nonzero. This can be interpreted as follows:

$$(B_f V_g)_j = \begin{cases} \neq 0 & \text{the } j\text{th } f\text{-loop is a } \textit{boundary } f\textit{-loop}, \\ 0 & \text{the } j\text{th } f\text{-loop is an } \textit{internal } f\textit{-loop}. \end{cases} \tag{3.46}$$

We now reorder (i.e., permute) the rows of B_f to get a new f-loop matrix for Γ, which we will again denoted by B_f, such that

$$B_f V_g = \left[\begin{array}{c} \psi \\ \hdashline 0 \end{array}\right], \tag{3.47}$$

where the jth row of ψ represents the algebraic sum of all the voltage sources contained in the jth boundary f-loop, i.e., ψ is the vector of *boundary f-loop source voltages*. Next, partition the $\nu \times 1$ column vector of f-loop currents $I_{\mathfrak{m}}$ conformal to (3.47), i.e.,

$$I_{\mathfrak{m}} = \left[\begin{array}{c} I_{\mathfrak{m},\partial} \\ \hdashline I_{\mathfrak{m},\circ} \end{array}\right], \tag{3.48}$$

where $I_{\mathfrak{m},\partial}$ and $I_{\mathfrak{m},\circ}$ represent the vectors of *f-loop currents in the boundary and internal f-loops*, respectively. Then the loop impedance equation (3.24) becomes the following equation

$$\left[\begin{array}{c} \psi \\ \hdashline 0 \end{array}\right] = \mathcal{Z} \left[\begin{array}{c} I_{\mathfrak{m},\partial} \\ \hdashline I_{\mathfrak{m},\circ} \end{array}\right], \tag{3.49}$$

where $\mathcal{Z}$ is the loop impedance matrix, i.e.,

$$\mathcal{Z} = B_f \tilde{\mathcal{Z}} B_f^T. \tag{3.50}$$

Next, we partition $\mathcal{Z}$ conformal to (3.49) as a 2×2 block matrix

$$\mathcal{Z} = \mathcal{Z}(z) = \mathcal{Z}(z_1, \ldots, z_\mu) = [\mathcal{Z}_{ij}]_{i,j=1,2} = \left[\begin{array}{c:c} \mathcal{Z}_{11}(z) & \mathcal{Z}_{12}(z) \\ \hdashline \mathcal{Z}_{21}(z) & \mathcal{Z}_{22}(z) \end{array}\right]. \tag{3.51}$$

The next theorem introduces the *source loop impedance matrix* denoted by Z, which is the Schur complement of the loop impedance matrix $\mathcal{Z}$ of an electrical network Γ.

Theorem 15 (Source loop method) *Consider a electrical network* Γ *whose loop impedance equation has the partitioned form (3.49) with loop impedance matrix* $\mathcal{Z}$ *(3.50) in the conformal partitioned* 2×2 *block form (3.51). If* $\mathcal{Z}_{22}$ *is invertible (i.e.,* $\det \mathcal{Z}_{22} \neq 0$*) then the following equation (called the source impedance equation) is satisfied:*

$$\psi = Z I_{\mathfrak{m},\partial}. \tag{3.52}$$

where ψ *is the vector of boundary* f*-loop source voltages (3.47),* $I_{\mathfrak{m},\partial}$ *is the vector of boundary* f*-loop currents (3.48), and* Z *is the Schur complement of the loop impedance matrix, i.e.,*

$$Z = \mathcal{Z}/\mathcal{Z}_{22}. \tag{3.53}$$

Furthermore, if both $\mathcal{Z}_{22}$ *and* $\mathcal{Z}/\mathcal{Z}_{22}$ *are invertible then the edge voltages* V*, the edge currents* I*, and the* f*-loop currents* $I_{\mathfrak{m}}$ *of the network* Γ *are given by the formulas*

$$V = \tilde{\mathcal{Z}} I, \quad I = B_f^T I_{\mathfrak{m}}, \tag{3.54}$$

$$I_{\mathfrak{m}} = \begin{bmatrix} I_{\mathfrak{m},\partial} \\ \hline I_{\mathfrak{m},\circ} \end{bmatrix} = \begin{bmatrix} Z^{-1}\psi \\ \hline -\mathcal{Z}_{22}\mathcal{Z}_{21}Z^{-1}\psi \end{bmatrix}. \tag{3.55}$$

Moreover, if $\tilde{\mathcal{Z}}$ *is a positive definite matrix (i.e.,* $\tilde{\mathcal{Z}} > 0$*) then* $\mathcal{Z}_{22}$ *is invertible,* $\mathcal{Z}/\mathcal{Z}_{22}$ *is invertible and, in particular, formulas (3.54) and (3.55) hold.*

Proof. Let Γ be as in the hypothesis and $\mathcal{Z}_{22}$ be invertible. Then by Proposition 3 we have

$$I_{\mathfrak{m},\circ} = -\mathcal{Z}_{22}^{-1}\mathcal{Z}_{21} I_{\mathfrak{m},\partial}, \tag{3.56}$$

$$\psi = (\mathcal{Z}/\mathcal{Z}_{22}) I_{\mathfrak{m},\partial}. \tag{3.57}$$

Therefore, (3.52) holds. Also, if $\mathcal{Z}/\mathcal{Z}_{22}$ is invertible then together with Proposition 11, it follows immediately that formulas (3.54) and (3.55) hold. Now suppose that all we know is $\tilde{\mathcal{Z}}$ is a positive definite matrix, i.e., $\tilde{\mathcal{Z}} > 0$. Then by the proof of Proposition 11 it follows that $\mathcal{Z}$ is also invertible and hence by the Schur's determinant formula (i.e., Lemma 9), it follows that $Z = \mathcal{Z}/\mathcal{Z}_{22}$ is invertible. The proof now follows from this. ■

Remark 16 (Summary of source loop method) *To summarize, we have provided sufficient conditions to answer the fundamental problem stated at the beginning of Section 3.1 (i.e., to solve for the edge voltages* V *and edge currents* I *in a given network* Γ*). More specifically, given independent source voltages* V_g *and no independent source currents* $I_g = 0$*, the* f*-loop matrix* B_f *from (3.3) [with rows already reordered so that the block form (3.47) holds], and the loop impedance matrix* $\mathcal{Z}$ *(3.50) partitioned conformally as (3.51) with loop impedance equation (3.49), then we can compute the edge* *voltages V and edge currents I in the network Γ (at least in the Laplace domain) with equations (3.52), (3.54), and (3.55) provided both $\mathcal{Z}_{22}$ and $\mathcal{Z}/\mathcal{Z}_{22}$ are invertible, which is the case if the diagonal edge impedance matrix $\tilde{\mathcal{Z}}$ of the network Γ is a positive definite matrix (i.e., $\tilde{\mathcal{Z}} > 0$).*

An important observation relating the source impedance matrix Z and the loop impedance matrix $\mathcal{Z}$ to the Bessmertnyĭ realizability theorem of Chapter 5 is the following corollary.

Corollary 17 (Source impedance matrix has a Bessmertnyĭ realization) *The source impedance matrix Z (3.53) of an electric network Γ has a Bessmertnyĭ realization (see Def. 66) as the Schur complement of a homogeneous real symmetric matrix pencil, in particular, this linear matrix pencil can be chosen to be the loop impedance matrix $\mathcal{Z}$ (3.50) of the network Γ which has the pencil form given by the formulas (3.42) and (3.43).*

Proof. The proof follows immediately by Lemma 14 and Theorem 15. ■

In this next example, we show how to use the source loop method as described by Theorem 15 to solve for an electrical network's edge voltage V and edge currents I in terms of the boundary f-loop source voltages ψ.

Example 18 (Example of the source loop method) *Consider the electrical network Γ shown in Figure 3.6. The following example is a slightly modified electrical network from [69BB, p. 60, Chap. 2, Fig. 2]. We will use the source loop method (see Theorem 15) to solve for the edge voltage V and edge currents I of the network Γ in terms of the boundary f-loop source voltages ψ using the f-loop currents $I_{\mathfrak{m}}$ in terms of independent voltage source V_g with no independent current sources ($I_g = 0$).*

To begin, the source voltages, source currents, and the associated edge impedance matrix $\tilde{\mathcal{Z}}$ [defined in (3.17)] of Γ are

$$V_g = \begin{bmatrix} V_{g,1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \; I_g = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \; \text{and } \tilde{\mathcal{Z}} = \begin{bmatrix} R_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & R_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & R_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & R_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & R_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & R_6 \end{bmatrix}.$$

Next, we provide in Figure 3.7 the associated finite linear directed graph G of Γ and the complete node-edge incidence matrix A_a constructed from the topology of Γ [defined in (3.1)]. In Figure 3.8 we give a clockwise orientation to each loop of G ($\delta = 7$) and show the associated complete loop matrix B_a [defined in (3.2)]. Now we select a tree for each forest in G (in this example, there is only one forest since the graph G is connected). We illustrate in Figure 3.9 our chosen tree of G and the chosen tree's associated links (in this example e_1, e_2, e_3). Next, in Figure 3.10 we provide each f-loop of G ($\nu = 3$)

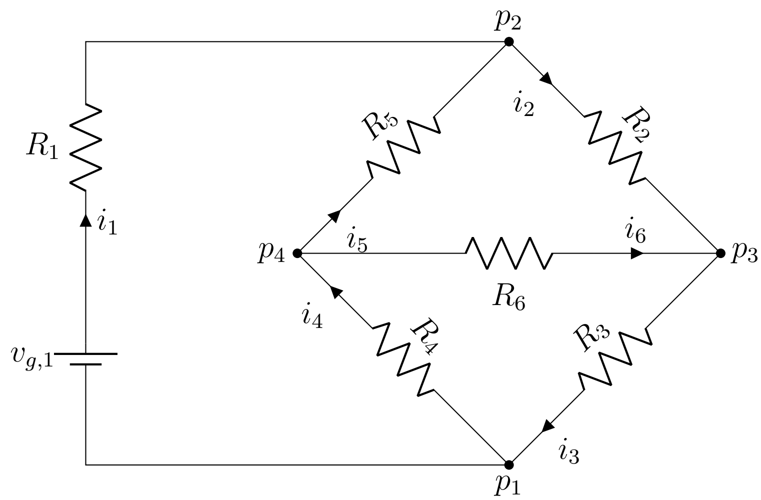
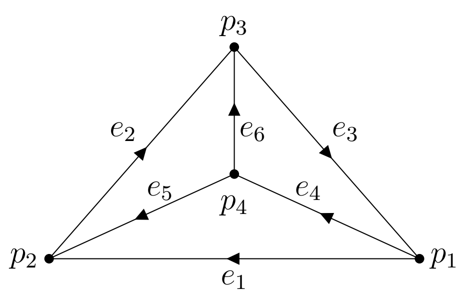


Figure 3.6: Example of a R-network Γ with a connected graph G ($\varsigma = 1$) with four nodes ($\kappa = 4$) denoted by p_1, p_2, p_3, p_4 and six edges ($\mu = 6$) denoted by e_1, e_2, e_3, e_4, e_5, e_6 such that e_1 has resistor R_1 and a battery $v_{g,1}$, e_2 has a resistor R_2, e_3 has resistor R_3, e_4 has a resistor R_4, e_5 has resistor R_5, and e_6 has resistor R_6.

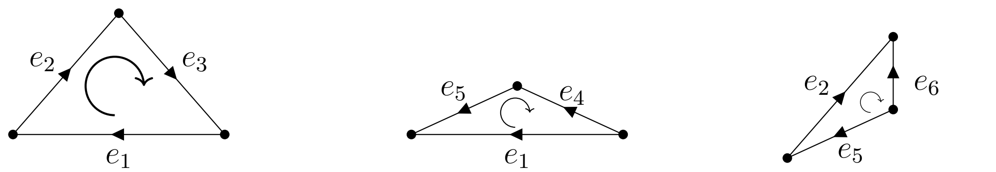


$$A_a = \begin{bmatrix} 1 & 0 & -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{bmatrix}$$

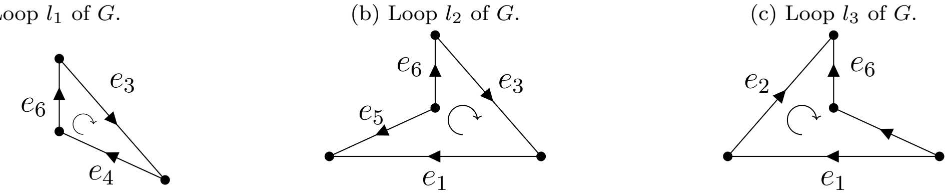
Figure 3.7: Graph G and complete node-edge matrix A_a for Γ.

and the the corresponding f-loop matrix B_f [defined in (3.3)]. Since the only voltage source is in loop l_2 we order the B_f matrix with the first row as the f-loop l_2. Using Proposition 11 (loop method), we can now compute the loop impedance matrix

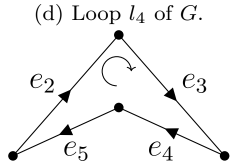
$$\mathcal{Z} = B_f \tilde{\mathcal{Z}} B_f^T = \begin{bmatrix} R_1 + R_4 + R_5 & -R_5 & -R_4 \\ -R_5 & R_2 + R_6 + R_5 & -R_6 \\ R_4 & -R_6 & R_3 + R_4 + R_6 \end{bmatrix}. \tag{3.58}$$



(a) Loop l_1 of G. (b) Loop l_2 of G. (c) Loop l_3 of G.



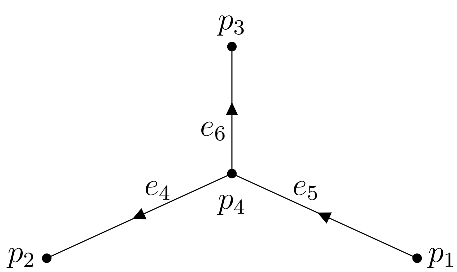
(d) Loop l_4 of G. (e) Loop l_5 of G. (f) Loop l_6 of G.



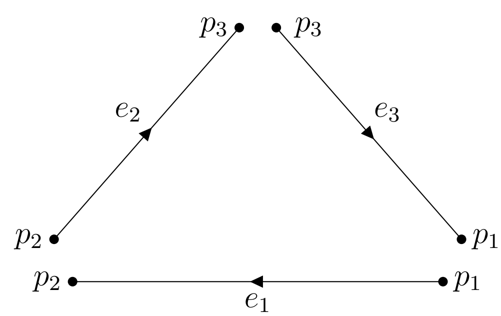
(g) Loop l_7 of G.

$$B_a = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & -1 & 1 \\ 1 & 1 & 0 & -1 & 0 & -1 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Figure 3.8: All loops of G ($\delta = 7$) with clockwise orientation and the complete loop matrix B_a.



Tree

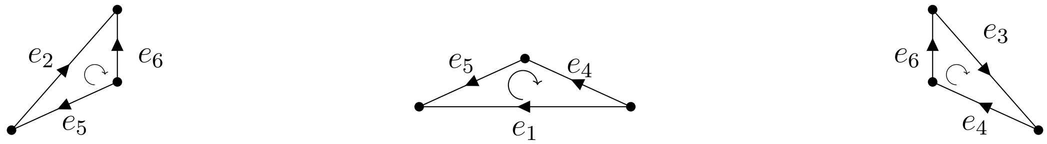


Links

Figure 3.9: Tree and its corresponding links for the graph G of the network Γ.

The reduced KVL (3.22) implies the following equality

$$\left[\begin{array}{c} \psi \\ \hdashline 0 \end{array}\right] = B_f V_g = B_f \begin{bmatrix} V_{g,1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \left[\begin{array}{c} V_{g,1} \\ \hdashline 0 \\ 0 \end{array}\right]. \tag{3.59}$$



(a) f-Loop l_3 of G. (b) f-Loop l_2 of G. (c) f-Loop l_4 of G.

$$B_f = \begin{bmatrix} 1 & 0 & 0 & -1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Figure 3.10: Corresponding f-loops ($\nu = 3$) to the chosen tree in G and the f-loop matrix B_f for Γ.

Furthermore, with the vector of loop currents partitioned confromal to (3.59) is

$$I_{\mathfrak{m}} = \begin{bmatrix} I_{\mathfrak{m},\partial} \\ \hdashline I_{\mathfrak{m},\circ} \end{bmatrix} = \begin{bmatrix} I_{\mathfrak{m},1} \\ \hdashline I_{\mathfrak{m},2} \\ I_{\mathfrak{m},3} \end{bmatrix}. \tag{3.60}$$

By Proposition 11, $\mathcal{Z}$ satisfies the following loop impedance equation (3.49)

$$\begin{bmatrix} V_{g,1} \\ \hdashline 0 \\ 0 \end{bmatrix} = \mathcal{Z} \begin{bmatrix} I_{\mathfrak{m},1} \\ \hdashline I_{\mathfrak{m},2} \\ I_{\mathfrak{m},3} \end{bmatrix}, \tag{3.61}$$

where

$$\mathcal{Z} = \left[\begin{array}{c:c} Z_{11} & Z_{12} \\ \hdashline Z_{21} & Z_{22} \end{array}\right] = \left[\begin{array}{c:cc} R_1 + R_4 + R_5 & -R_5 & -R_4 \\ \hdashline -R_5 & R_2 + R_5 + R_6 & -R_6 \\ -R_4 & -R_6 & R_3 + R_4 + R_6 \end{array}\right].$$

By Theorem 15, equation (3.56), we can solve for $I_{\mathfrak{m},\circ}$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \left[\begin{array}{c:cc} -R_5 & R_2 + R_5 + R_6 & -R_6 \\ -R_4 & -R_6 & R_3 + R_4 + R_6 \end{array}\right] \begin{bmatrix} I_{\mathfrak{m},1} \\ \hdashline I_{\mathfrak{m},2} \\ I_{\mathfrak{m},3} \end{bmatrix} \tag{3.62}$$

$$\begin{bmatrix} I_{\mathfrak{m},2} \\ I_{\mathfrak{m},3} \end{bmatrix} = - \begin{bmatrix} R_2 + R_5 + R_6 & -R_6 \\ -R_6 & R_3 + R_4 + R_6 \end{bmatrix}^{-1} \begin{bmatrix} -R_5 \\ -R_4 \end{bmatrix} [I_{\mathfrak{m},1}]. \tag{3.63}$$

By using back substitution,

$$
\begin{aligned}
\psi &= \begin{bmatrix} R_1 + R_4 + R_5 \end{bmatrix} [I_{\mathfrak{m},1}] + \begin{bmatrix} -R_5 & -R_4 \end{bmatrix} \begin{bmatrix} I_{\mathfrak{m},2} \\ I_{\mathfrak{m},3} \end{bmatrix} \\
&= \left(\begin{bmatrix} R_1 + R_4 + R_5 \end{bmatrix} - \begin{bmatrix} -R_5 & -R_4 \end{bmatrix} \begin{bmatrix} R_2 + R_5 + R_6 & -R_6 \\ -R_6 & R_3 + R_4 + R_6 \end{bmatrix}^{-1} \begin{bmatrix} -R_5 \\ -R_4 \end{bmatrix} \right) [I_{\mathfrak{m},1}] \\
&= \mathcal{Z}_{11} - \mathcal{Z}_{12}\mathcal{Z}_{22}^{-1}\mathcal{Z}_{21}[I_{\mathfrak{m},1}] \\
&= \mathcal{Z}/\mathcal{Z}_{22}[I_{\mathfrak{m},1}] \\
&= Z[I_{\mathfrak{m},1}].
\end{aligned}
$$

Therefore, solving for $I_{\mathfrak{m},\partial}$, $I_{\mathfrak{m},\circ}$, $I_{\mathfrak{m}}, I, V$ we get our desired solutions:

$$
\begin{bmatrix} I_{\mathfrak{m},\partial} \end{bmatrix} = [I_{\mathfrak{m},1}] = Z^{-1}\psi, \tag{3.64}
$$

$$
[I_{\mathfrak{m},\circ}] = - \begin{bmatrix} R_2 + R_5 + R_6 & -R_6 \\ -R_6 & R_3 + R_4 + R_6 \end{bmatrix}^{-1} \begin{bmatrix} -R_5 \\ -R_4 \end{bmatrix} Z^{-1}\psi, \tag{3.65}
$$

$$
I = B_f I_{\mathfrak{m}} = B_f \begin{bmatrix} Z^{-1}\psi \\ - \begin{bmatrix} R_2 + R_5 + R_6 & -R_6 \\ -R_6 & R_3 + R_4 + R_6 \end{bmatrix}^{-1} \begin{bmatrix} -R_5 \\ -R_4 \end{bmatrix} Z^{-1}\psi \end{bmatrix}, \tag{3.66}
$$

$$
V = \tilde{\mathcal{Z}} I. \tag{3.67}
$$

3.3 Dirichlet-to-Neumann Matrix

Following Section 3.1, we now introduce the Dirichlet-to-Neumann (DtN) map Λ of a resistor-only (i.e., R-only) electric network Γ whose graph G is connected with a boundary (which we will make precise below). This DtN map Λ takes as input an imposed voltage f on the boundary and outputs the currents ϕ into the boundary (i.e., $\Lambda f = \phi$).

To make this more precise, let Γ be an R-network whose graph G is connected ($\varsigma = 1$) with κ nodes $P_G = \{p_1, \ldots, p_\kappa\}$. The boundary and the interior of G is defined in terms of a partitioning of the nodes P_G into b boundary nodes, say, $P_{\partial G} = \{p_1, \ldots, p_b\}$ and interior nodes $P_{G^\circ} = P \setminus P_{\partial G}$, respectively. It is assumed below that $\emptyset \neq P_{\partial G} \neq P_G$ (i.e., $1 \leq b < \kappa$), in which case, we then say that Γ is a *connected R-only network with boundary.*

To compute the DtN map Λ we only need the complete node-edge matrix A_a of the network Γ as define by (3.1) along with the conductivity matrix γ and the Kirchhoff matrix K, which we define below. The KCL applied to all nodes becomes the following matrix equation

$$A_a(I - I_g) = 0 \tag{3.68}$$

[60PB, p. 220]. Let U be the $\kappa \times 1$ vector of voltage potentials throughout the network, then the vector of edge voltages V (with no independent voltage sources in the network Γ, i.e., $V_g = 0$) is related to U by the following matrix equation

$$A_a^T U = V. \tag{3.69}$$

We want to remark here that because of the equalities in (3.5), the relation (3.69) is equivalent to $B_a V = 0$. Furthermore, since we know from Section 3.1 and the definition of the edge impedance matrix (3.16) that $\tilde{\mathcal{Z}}$ is invertible, provided every resistor on every edge in the R-only network is nonzero, then the $\mu \times \mu$ *conductivity matrix* $\gamma = \tilde{\mathcal{Z}}^{-1}$ and Ohm's law becomes

$$I = \gamma V, \tag{3.70}$$

where, in terms of the resistors in the network, γ is a diagonal matrix whose jth diagonal entry γ_{jj}, which we will denote by γ_j, is

$$\gamma_j = \gamma_{jj} = \frac{1}{R_j}, \text{ for } j = 1, \ldots \mu. \tag{3.71}$$

Proposition 19 *Let γ be the diagonal $\mu \times \mu$ conductivity matrix of some R-network Γ. If $I = \gamma V$ then $A_a I_g = KU$, where K is a $\kappa \times \kappa$ matrix such that $K = A_a \gamma A_a^T$ and U is a $\kappa \times 1$ column vector of voltage potentials.*

Proof. The proof follows immediately from (3.68) a equation (3.69),

$$A_a I_g = A_a I = A_a \gamma V = A_a \gamma A_a^T U = KU. \tag{3.72}$$

Hence, K is the $\kappa \times \kappa$ *Kirchhoff matrix* [00CM, p. 34, Sec. 3.3, Chap. 3] function that satisfies the following voltage to current equation

$$A_a I_g = KU. \tag{3.73}$$

■

The Kirchhoff matrix, $K = K(\gamma_1, \ldots, \gamma_\mu)$, is a homogeneous real symmetric linear $\kappa \times \kappa$ matrix pencil of the μ complex variables $\gamma_1, \ldots, \gamma_\mu$. Furthermore, it is has the representation

$$K = A_a \gamma A_a^T, \tag{3.74}$$

where A_a is the complete node-edge matrix defined in (3.1) and

$$\gamma = \begin{bmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_\mu \end{bmatrix}. \tag{3.75}$$

As the jth row of $A_a I_g$ represents the algebraic sum of all the current sources entering the jth node (i.e., p_j) of the network Γ, then it is assumed that the current sources I_g are chosen so that there is no current entering the internal nodes P_{G°. This can be interpreted as follows:

$$A_a I_g = \begin{bmatrix} \phi \\ \hdashline 0 \end{bmatrix}, \tag{3.76}$$

where ϕ represents the algebraic sum of all the current sources entering the boundary nodes (i.e., $P_{\partial G}$) of the network Γ, i.e., ϕ is the vector of *boundary source currents.* Next, we partition the the vector of voltage potentials U conformal to (3.76), i.e.,

$$U = \begin{bmatrix} f \\ \hdashline g \end{bmatrix}, \tag{3.77}$$

where f and g are the *vector of voltage potentials at the boundary and interior nodes*, respectively. Then the voltage to current equation (3.73) becomes

$$\begin{bmatrix} \phi \\ \hdashline 0 \end{bmatrix} = K \begin{bmatrix} f \\ \hdashline g \end{bmatrix}, \tag{3.78}$$

where K is the Kirchhoff matrix (3.74) partitioned conformal to (3.78) as the 2×2 block matrix

$$K = K(\gamma) = K(\gamma_1, \ldots, \gamma_\mu) = [K_{ij}]_{i,j=1,2} = \left[\begin{array}{c:c} K_{11}(\gamma) & K_{12}(\gamma) \\ \hdashline K_{21}(\gamma) & K_{22}(\gamma) \end{array}\right]. \tag{3.79}$$

We now will introduce the *Dirichlet-to-Neumann matrix* (DtN matrix) Λ of the network Γ. The first part of the next theorem is well-known, for instance, see [00CM, p. 43, Theorem 3.9]. We provide an extension to the theorem to include the observation of symmetries that arise in the realization of Λ from the Kirchhoff matrix.

Theorem 20 (DtN matrix has a Bessmertnyĭ realization) *Let Γ be a connected R-network with boundary whose voltage to current equation in the partitioned form (3.78) with Kirchhoff matrix K (3.74) in the conformal partitioned 2×2 block form (3.79). If K_{22} is invertible (i.e., $\det K_{22} \neq 0$) then the following equation is satisfied:*

$$\phi = \Lambda f, \tag{3.80}$$

where f is the vector of voltage potentials at the boundary nodes, ϕ is the vector of boundary source currents, and the Dirichlet-to-Neumann matrix Λ is the Schur complement

$$\Lambda = K/K_{22}. \tag{3.81}$$

Furthermore, if γ is a positive definite matrix (i.e., $\gamma > 0$) then K_{22} is invertible and, in particular, formulas (3.80) and (3.81) hold. Moreover, the Kirchhoff matrix K with conductivity matrix γ is a homogeneous real symmetric matrix pencil in the complex variables γ_j and the Dirichlet-to-Neumann matrix Λ has a Bessmertnyĭ realization (3.81).

Proof. Assume the hypotheses with Γ a connected R-network with boundary. Suppose K_{22} is invertible. Then by Proposition 3 we have

$$\begin{aligned} g &= -K_{22}^{-1}K_{21}f, \\ \phi &= (K_{11} - K_{12}K_{22}^{-1}K_{21})f. \end{aligned}$$

Therefore, (3.80) and (3.81) hold. Suppose γ is a positive definite matrix. Then by [00CM, Lemma 3.8] the matrix K_{22} is positive definite hence invertible. The rest of the proof now follows immediately. ■

Example 21 *We will consider the connected R-network Γ shown in Figure 3.11 and with numerical values for $\gamma_1, \ldots, \gamma_4$ in Figure 3.12. Our set of boundary nodes and interior nodes are $P_{\partial G} = \{p_1, p_2, p_3\}$ and $P_{G^\circ} = \{p_4\}$, respectively. Hence, we can construct the complete node-edge matrix A_a from the topology of the network in Figure 3.11,*

$$A_a = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}.$$

Therefore, we can compute the Kirchhoff matrix by matrix multiplication $K = A_a\gamma A_a^T$, where

$$\gamma = \begin{bmatrix} \gamma_1 & 0 & 0 & 0 \\ 0 & \gamma_2 & 0 & 0 \\ 0 & 0 & \gamma_3 & 0 \\ 0 & 0 & 0 & \gamma_4 \end{bmatrix},$$

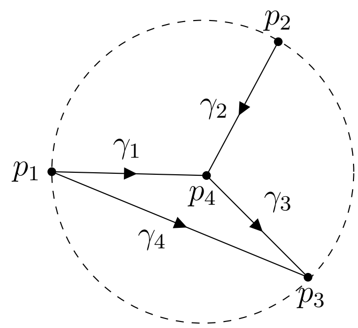


Figure 3.11: DtN map of R-network Γ with boundary.

$$K = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} \gamma_1 & 0 & 0 & 0 \\ 0 & \gamma_2 & 0 & 0 \\ 0 & 0 & \gamma_3 & 0 \\ 0 & 0 & 0 & \gamma_4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & -1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \gamma_1 + \gamma_4 & 0 & -\gamma_4 & -\gamma_1 \\ 0 & \gamma_2 & 0 & -\gamma_2 \\ -\gamma_4 & 0 & \gamma_3 + \gamma_4 & -\gamma_3 \\ -\gamma_1 & -\gamma_2 & -\gamma_3 & \gamma_1 + \gamma_2 + \gamma_3 \end{bmatrix}.$$

Suppose that the conductivity relationships γ_i are given the following values:

$$\gamma_1 = 2, \ \ \gamma_2 = 4, \ \ \gamma_3 = 1, \ \ \textit{and } \gamma_4 = 8.$$

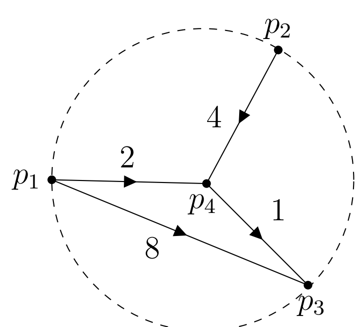


Figure 3.12: DtN map of R-network Γ with boundary and γ values.

Thus, we obtain the following Kirchhoff matrix

$$K = \begin{bmatrix} 10 & 0 & -8 & -2 \\ 0 & 4 & 0 & -4 \\ -8 & 0 & 9 & -1 \\ -2 & -4 & -1 & 7 \end{bmatrix}.$$

By Theorem 20, $\Lambda = K/K_{22}$ *where* $K_{22} = \begin{bmatrix} 7 \end{bmatrix}$ *and* $\det K_{22} \not\equiv 0$. *Hence, we can calculate the DtN matrix explicitly for this network,*

$$\begin{aligned}
K/K_{22} &= K_{11} - K_{12}K_{22}^{-1}K_{21} \\
&= \begin{bmatrix} 10 & 0 & -8 \\ 0 & 4 & 0 \\ -8 & 0 & 9 \end{bmatrix} - \begin{bmatrix} -2 \\ -4 \\ -1 \end{bmatrix} \begin{bmatrix} 7 \end{bmatrix}^{-1} \begin{bmatrix} -2 & -4 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 10 & 0 & -8 \\ 0 & 4 & 0 \\ -8 & 0 & 9 \end{bmatrix} - \begin{bmatrix} -2 \\ -4 \\ -1 \end{bmatrix} \begin{bmatrix} \frac{1}{7} \end{bmatrix} \begin{bmatrix} -2 & -4 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 10 & 0 & -8 \\ 0 & 4 & 0 \\ -8 & 0 & 9 \end{bmatrix} - \begin{bmatrix} \frac{99}{10} & \frac{33}{5} & \frac{6}{5} & \frac{3}{10} \\ \frac{33}{5} & \frac{22}{5} & \frac{4}{5} & \frac{1}{5} \\ \frac{6}{5} & \frac{4}{5} & \frac{8}{5} & \frac{2}{5} \\ \frac{3}{10} & \frac{1}{5} & \frac{2}{5} & \frac{1}{10} \end{bmatrix} \\
&= \begin{bmatrix} \frac{66}{7} & -\frac{8}{7} & -\frac{58}{7} \\ -\frac{8}{7} & \frac{12}{7} & -\frac{4}{7} \\ -\frac{58}{7} & -\frac{4}{7} & \frac{62}{7} \end{bmatrix} \\
&= \Lambda.
\end{aligned}$$

3.4 Abstract Theory of Composites

In this section we discuss relationship between effective operators in the abstract theory of composites and Bessmertnyĭ realizability theory. We conclude with a concrete example from the theory of composites of the periodic conductivity equation and the effective conductivity.

As we pointed out in Chapter 1, Chapter Summary 1.3, this section is the only instance in the thesis that linear operators are used instead of matrices. But we include this because the theory of composites is the most quintessential model for our methods, research, and motivation for this thesis. Furthermore, the overlay of our definitions, theorems, and techniques in our study on Schur complements with more difficult problems, as the one below, illustrates the potential of solving synthesis and inverse problems in realizability theory in a more operator theory and algorithmic approach. Moreover, the open problems discussed in Chapter 7 directly relate to this model (in particular Proposition 88) in the theory of composites.

3.4.1 Z-problems and effective operators

Let $\mathcal{H}$ be a Hilbert space with three mutually orthogonal subspaces $\mathcal{U}, \mathcal{E}, \mathcal{J} \subseteq \mathcal{H}$ such that

$$\mathcal{H} = \mathcal{U} \overset{\perp}{\oplus} \mathcal{E} \overset{\perp}{\oplus} \mathcal{J}, \tag{3.82}$$

where $\overset{\perp}{\oplus}$ denotes the orthogonal direct sum [16GM, p.56, Sec. 2.3, Chap. 2]. Also, suppose there exists a mutually orthogonal direct sum subspace decomposition of the Hilbert space $\mathcal{H}$ as

$$\mathcal{H} = \mathcal{P}_1 \overset{\perp}{\oplus} \mathcal{P}_2 \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} \mathcal{P}_n. \tag{3.83}$$

These collection of subspaces in (3.82) and (3.83) are called a *orthogonal $Z(n)$-subspace collection.*

Now, by hypotheses, any vector $H \in \mathcal{H}$ has a unique decomposition into component vectors

$$H = U + E + J = P_1 + P_2 + \cdots + P_n, \tag{3.84}$$

where $U \in \mathcal{U}$, $E \in \mathcal{E}$, $J \in \mathcal{J}$, and $P_i \in \mathcal{P}_i$, for each $i = 1, \ldots, n$. Next, let Γ_j for $j = 0, 1, 2$ be the orthogonal projections of $\mathcal{H}$ onto $\mathcal{U}, \mathcal{E}, \mathcal{J}$, respectively, and let Λ_i be the orthogonal projection of $\mathcal{H}$ onto $\mathcal{P}_i$, for each $i = 1, \ldots, n$. In particular,

$$\Gamma_0 H = U, \;\; \Gamma_1 H = E, \;\; \Gamma_2 H = J, \tag{3.85}$$

$$\Lambda_i H = P_i, \text{ for each } i = 1, \ldots, n. \tag{3.86}$$

Definition 22 (Z-problem) *The **Z-problem** [16GM, p. 189, Sec. 7.2, Chap. 7] is given $E_0 \in \mathcal{U}$, find $J_0 \in \mathcal{U}, E \in \mathcal{E}, J \in \mathcal{J}$ such that*

$$\sigma(E_0 + E) = J_0 + J, \tag{3.87}$$

where $\sigma = \sigma(z_1, \ldots, z_n)$ is the homogeneous self-adjoint linear operator pencil defined by

$$\sigma = \sum_{i=1}^{n} z_i \Lambda_i, \tag{3.88}$$

for $z_i \in \mathbb{C}$, in which each Λ_i is an orthogonal projection of $\mathcal{H}$ onto $\mathcal{P}_i$, for $i = 1, \ldots, n$.

Definition 23 (Effective operator) *For those $E_0 \in \mathcal{U}$ that a solution to the Z-problem exists, the **effective operator** is defined as any linear operator σ_* on $\mathcal{U}$ that satisfies*

$$\sigma_* E_0 = J_0.$$

Proposition 24 (Effective operator: Bessmertnyĭ realization) *If σ_{11} is invertible then the Z-problem has a unique solution for any $E_0 \in \mathcal{U}$ and there is only one effective operator σ_* and it is given by the Schur complement*

$$\sigma_* = \begin{bmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{10} & \sigma_{11} \end{bmatrix} / \sigma_{11} = \sigma_{00} - \sigma_{01}\sigma_{11}^{-1}\sigma_{10}. \tag{3.89}$$

Furthermore, σ_* is the Schur complement of the homogeneous self-adjoint linear operator pencil (i.e., an operator version of a Bessmertnyĭ realization of σ_*)

$$\begin{bmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{10} & \sigma_{11} \end{bmatrix} = \sigma_1 A_1 + \cdots + \sigma_n A_n, \tag{3.90}$$

$$A_i = (\Gamma_0 + \Gamma_1)\Lambda_i(\Gamma_0 + \Gamma_1)|_{\mathcal{U}\oplus\mathcal{E}}, \;\; i = 1, \ldots, n, \tag{3.91}$$

where $\Gamma_0+\Gamma_1$ is the orthogonal projection of $\mathcal{H}$ onto $\mathcal{U}\overset{\perp}{\oplus}\mathcal{E}$ and $|_{\mathcal{U}\oplus\mathcal{E}}$ denotes the operator restriction to that subspace.

Proof. We can write the Z-problem in 3×3 block matrix notation such that

$$\begin{bmatrix} \sigma_{00} & \sigma_{01} & \sigma_{02} \\ \sigma_{10} & \sigma_{11} & \sigma_{12} \\ \sigma_{20} & \sigma_{21} & \sigma_{22} \end{bmatrix} \begin{bmatrix} E_0 \\ E \\ 0 \end{bmatrix} = \begin{bmatrix} J_0 \\ 0 \\ J \end{bmatrix}, \tag{3.92}$$

which implies that

$$\begin{bmatrix} \sigma_{00}E_0 + \sigma_{01}E \\ \sigma_{10}E_0 + \sigma_{11}E \\ \sigma_{20}E_0 + \sigma_{21}E \end{bmatrix} = \begin{bmatrix} J_0 \\ 0 \\ J \end{bmatrix}. \tag{3.93}$$

Assuming σ_{11}^{-1} exists and using Proposition 3 we can find E, J, J_0 in terms of E_0 to solve for σ_*. And doing so we find that E is given by

$$E = -\sigma_{11}^{-1}\sigma_{10}E_0 \tag{3.94}$$

and from this J is given by

$$J = \sigma_{20}E_0 + \sigma_{21}E = (\sigma_{20} - \sigma_{21}\sigma_{11}^{-1}\sigma_{10})E_0. \tag{3.95}$$

Plugging (3.94) into the first equation of the system we find that

$$J_0 = \sigma_{00}E_0 - \sigma_{01}{\sigma_{11}}^{-1}\sigma_{10}E_0 = (\sigma_{00} - \sigma_{01}{\sigma_{11}}^{-1}\sigma_{10})E_0. \tag{3.96}$$

The proof of the proposition now follows immediately from this. ∎

3.4.2 Periodic conductivity equation and the effective conductivity

Example 25 (Effective Conductivity) *Consider the following conductivity equation*

$$\nabla \cdot \sigma \nabla u = 0, \tag{3.97}$$

of a periodic n-phase composite with unit cell

$$\Omega = [0, 2\pi]^d = \Omega_1 \cup \cdots \cup \Omega_n \tag{3.98}$$

with dimension d ($d = 2$ or $d = 3$), each Ω_j is Lebesgue measurable as a subset of $\mathbb{R}^d$, and each distinct pair Ω_i and Ω_j intersection on a set of Lebesgue measure zero. The Hilbert space $\mathcal{H}$ is

$$\mathcal{H} = \left[L^2_{per}(\Omega)\right]^d \tag{3.99}$$

with the inner product

$$\langle E, F\rangle = \frac{1}{|\Omega|}\int_\Omega \overline{E(x)}^T F(x)dx \tag{3.100}$$

and average

$$\langle F\rangle = \frac{1}{|\Omega|}\int_\Omega F(x)\,dx, \tag{3.101}$$

for all $E, F \in \mathcal{H}$. Then the $Z(n)$-subspace collection is the orthogonal direct sum

$$\mathcal{H} = \mathcal{U} \overset{\perp}{\oplus} \mathcal{E} \overset{\perp}{\oplus} \mathcal{J} = \chi_1\mathcal{H} \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} \chi_n\mathcal{H}, \tag{3.102}$$

where χ_j denotes the characteristic function of Ω_j and as a multiplication operator on $\mathcal{H}$ acts as an orthogonal projection, for each $j = 1, \ldots, n$, and

$$\mathcal{U} = \left\{U \in \mathcal{H} : U \equiv C, \text{ for some } C \in \mathbb{C}^d\right\}, \tag{3.103}$$

$$\mathcal{E} = \{E \in \mathcal{H} : \nabla \times E = 0 \text{ and } \langle E\rangle = 0\}, \tag{3.104}$$

$$\mathcal{J} = \{J \in \mathcal{H} : \nabla \cdot J = 0 \text{ and } \langle J\rangle = 0\}, \tag{3.105}$$

where $\nabla \times (\cdot) = 0$ and $\nabla \cdot (\cdot) = 0$ are meant to denote "curl free" and "divergence free" vector fields, respectively, in dimension d. The constitutive relation coming from Ohm's law for the n-phase conductivity is

$$J = \sigma E, \text{ where } E \in \mathcal{U} \overset{\perp}{\oplus} \mathcal{E},\ J \in \mathcal{U} \overset{\perp}{\oplus} \mathcal{J}, \tag{3.106}$$

and $\sigma = \sigma(\sigma_1, \ldots, \sigma_n)$ is the homogeneous linear matrix pencil

$$\sigma = \sigma_1\chi_1 + \cdots + \sigma_n\chi_n. \tag{3.107}$$

The effective conductivity σ_{eff} is then defined in terms of the constitutive relation and field averaging $\langle\cdot\rangle$ by the relationship

$$\sigma_{eff}\langle E\rangle = \langle J\rangle. \tag{3.108}$$

The relationship to Z-problem and the associated effective operator σ_* is as follows. First, the Z-problem in this setting is: Given $e \in \mathcal{U}$, find $(j, E, J) \in \mathcal{U} \times \mathcal{E} \times \mathcal{J}$ such that

$$j + J = \sigma\,(e + E)\,, \tag{3.109}$$

and the effective operator σ_* is defined by the relation

$$\sigma_* e = j. \tag{3.110}$$

Therefore, since $e = \langle e + E\rangle$, $j = \langle j + J\rangle$, we have the fundamental identity:

$$\sigma_{eff} = \sigma_*. \tag{3.111}$$

Chapter 4

Schur Complements: Algebra and Operations

4.1 Introduction

The Schur complement (as defined in Definition 2) naturally arises in linear algebra, for instance, in the solving of constrained linear systems (as shown in Proposition 3). As such, it is important to understanding the way the Schur complement interacts with the algebraic structure of matrices as was motivated by our research and its applications (see Chapters 3, 5, 6, and 7). Thus, the goal of this chapter is to show that elementary operations (whether algebraic like addition and products, functional like composition, or transformal like the principal pivot transform) when applied to Schur complements of block matrices will be equal to another Schur complement of a block matrix over the field of complex numbers $\mathbb{F} = \mathbb{C}$. This chapter accomplishes this goal thoroughly.

To give a flavor of this, consider the following example. Suppose we have two Schur complements, A/A_{22} and B/B_{22}, where their sum $A/A_{22} + B/B_{22}$ is well-defined, can we explicitly construct a matrix C such that $C/C_{22} = A/A_{22} + B/B_{22}$? We find that the answer is yes, see Proposition 30. But our statement contains more then just this, it also gives a explicit formula for C in such a manner that it inherits certain properties of the original matrices A and B. More precisely, if A and B are real, symmetric, Hermitian, or real and symmetric then the matrix C is real, symmetric, Hermitian, or real and symmetric. In addition, we follow the statement up with a concrete example, see Example 29.

More generally, we are able to do this with all of the following algebra and operations:

List 4.1

- Scalar multiplication of a Schur complements (Lemma 26, Example 27):

$$\lambda(A/A_{22});$$

- Sums of a Schur complement with a matrix (Lemma 28, Example 29):

$$A/A_{22} + B;$$

- Sums of a Schur complements (Proposition 30, Example 31):

$$A/A_{22} + B/B_{22};$$

- Shorted matrices are Schur complements (Lemma 32, Example 33):

$$C/C_{22} = A/A_{22} \oplus 0_l;$$

- Direct sum of Schur complements (Proposition 34, Example 35):

$$A/A_{22} \oplus B/B_{22};$$

- Matrix multiplication of two Schur complements (Proposition 37, Example 38):

$$(A/A_{22})(B/B_{22});$$

- Matrix product with a Schur complement (Proposition 39, Example 40):

$$B(A/A_{22})C;$$

- Inverse is a Schur complement (Lemma 41, Example 42):

$$A^{-1};$$

- Inverse of Schur complement (Proposition 43, Example 44):

$$(A/A_{22})^{-1};$$

- Kronecker product of a Schur complement with a matrix (Proposition 45, Example 46):

$$(A/A_{22}) \otimes B;$$

- Kronecker product of two Schur complements (Proposition 51, Example 53):

$$(A/A_{22}) \otimes (B/B_{22});$$

- Compositions of Schur complements (Proposition 57, Example 58):

$$(A/A_{22})/(A/A_{22})_{22};$$

- Principal pivot transform as a Schur complement (Proposition 62, Example 63):

$$C/C_{22} = \mathrm{ppt}_2(A);$$

- The other Principal pivot transform as a Schur complement (Corollary 64, Example 65):

$$D/D_{22} = \mathrm{ppt}_1(A).$$

Furthermore, we use the elementary operations as the most basic building blocks for producing more complicated ones. For example, using Proposition 30 and Lemma 32 to prove Proposition 34. Or using Lemma 41 together with Proposition 39 to prove Proposition 43. Another example of this is using Lemma 45 to prove Corollary 49 and then to use these together with Proposition 57 to prove Proposition 51.

This building block approach illustrates how one can attack problems by using our Schur complement algebra and operations in a more "natural," algorithmic, and potentially computational approach (see Chapter 7 for potential computational remarks).

4.2 Sums and scalar multiplication

This first lemma belongs to the set of results relating to linear combinations involving Schur complements. Although elementary, it should give the reader a feel for the style of statements and proofs that we give in the remaining part of this chapter which become progressively more difficult.

Lemma 26 (Scalar multiplication of a Schur complement) *If $A \in \mathbb{C}^{m\times m}$ is a 2×2 block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

and $\lambda \in \mathbb{C} \setminus \{0\}$, then

$$B/B_{22} = \lambda(A/A_{22}), \tag{4.1}$$

where $B \in \mathbb{C}^{m\times m}$ is the 2×2 block matrix

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \left[\begin{array}{c:c} \lambda A_{11} & \lambda A_{12} \\ \hdashline \lambda A_{21} & \lambda A_{22} \end{array}\right] = \lambda A, \tag{4.2}$$

and $B_{22} = \lambda A_{22}$ is invertible. Moreover, if λ is real and the matrix A is real, symmetric, Hermitian, or real and symmetric then the matrix B is real, symmetric, Hermitian, or real and symmetric, respectively.

Proof. From the block matrix equality

$$\begin{bmatrix} (\lambda A)_{11} & (\lambda A)_{12} \\ (\lambda A)_{21} & (\lambda A)_{22} \end{bmatrix} = \lambda A = \left[\begin{array}{c:c} \lambda(A_{11}) & \lambda(A_{12}) \\ \hdashline \lambda(A_{21}) & \lambda(A_{22}) \end{array}\right],$$

it follows that

$$\begin{aligned} (\lambda A)/(\lambda A)_{22} &= (\lambda A)_{11} - (\lambda A)_{12}[(\lambda A)_{22}]^{-1}(\lambda A)_{21} \\ &= \lambda(A_{11}) - \lambda(A_{12})[\lambda(A_{22})]^{-1}\lambda(A_{21}) \\ &= \lambda(A_{11} - A_{12}A_{22}^{-1}A_{21}) \\ &= \lambda(A/A_{22}). \end{aligned}$$

The remaining part of the proof is obvious. This proves the lemma. ■

Example 27 *Let*

$$A = \left[\begin{array}{c:c} 2 & 2 \\ \hdashline 3 & 2 \end{array}\right].$$

Then

$$\begin{aligned} A/A_{22} &= A_{11} - A_{12}A_{22}^{-1}A_{21} \\ &= [2] - [2][2]^{-1}[3] \\ &= [2] - [2]\left[\frac{1}{2}\right][3] \\ &= [2] - [3] \\ &= [-1]. \end{aligned}$$

Consider $4(A/A_{22}) = 4[-1] = [-4]$. *Suppose* $B = 4A$, *we have*

$$B = 4\begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix} = \left[\begin{array}{c:c} 8 & 8 \\ \hdashline 12 & 8 \end{array}\right],$$

which implies

$$\begin{aligned} B/B_{22} &= B_{11} - B_{12}B_{22}^{-1}B_{21} \\ &= [8] - [8][8]^{-1}[12] \\ &= [8] - [8]\left[\frac{1}{8}\right][12] \\ &= [8] - [12] \\ &= [-4]. \end{aligned}$$

Hence, $4B = 4(A/A_{22})$.

Lemma 28 (Sum of a Schur complement with a matrix) *If* $A \in \mathbb{C}^{m\times m}$ *is a* 2×2 *block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

such that A_{22} *is invertible and* $A/A_{22} \in \mathbb{C}^{k\times k}$ *then, for any matrix* $B \in \mathbb{C}^{k\times k}$,

$$C/C_{22} = A/A_{22} + B, \tag{4.3}$$

where $C \in \mathbb{C}^{m\times m}$ *is the* 2×2 *block matrix*

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \left[\begin{array}{c:c} A_{11}+B & A_{12} \\ \hdashline A_{21} & A_{22} \end{array}\right], \tag{4.4}$$

and $C_{22} = A_{22}$ *is invertible. Moreover, if both matrices* A *and* B *are real, symmetric, Hermitian, or real and symmetric then the matrix* C *is real, symmetric, Hermitian, or real and symmetric, respectively.*

Proof. The proof is a straightforward calculation using block matrix techniques, where from the definition of $C = [C_{ij}]_{i,j=1,2}$ in (4.4), we have

$$\begin{aligned} C/C_{22} &= C_{11} - C_{12}C_{22}^{-1}C_{21} \\ &= A_{11} + B - A_{12}A_{22}^{-1}A_{21} \\ &= A/A_{22} + B. \end{aligned}$$

The remaining part of the proof follows immediately now from the formula (4.4) in terms of the matrices A and B. This completes the proof. ■

Example 29 *Let*

$$A = \left[\begin{array}{cc:cc} 3 & 2 & 1 & 2 \\ 4 & 5 & 2 & 3 \\ \hdashline 1 & 2 & 2 & 1 \\ 2 & 3 & 1 & 1 \end{array}\right] \text{ and } B = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}.$$

Then

$$\begin{aligned} A/A_{22} &= \begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 2 \\ 4 & 5 \end{bmatrix} - \begin{bmatrix} 5 & 7 \\ 7 & 10 \end{bmatrix} \\ &= \begin{bmatrix} -2 & -5 \\ -3 & -5 \end{bmatrix}. \end{aligned}$$

Consider

$$A/A_{22} + B = \begin{bmatrix} -2 & -5 \\ -3 & -5 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} -1 & -3 \\ -1 & -2 \end{bmatrix}.$$

Suppose C is of the formula (4.4), then

$$C = \left[\begin{array}{cc:cc} 4 & 4 & 1 & 2 \\ 6 & 8 & 2 & 3 \\ \hdashline 1 & 2 & 2 & 1 \\ 2 & 3 & 1 & 1 \end{array}\right].$$

Thus,

$$C/C_{22} = \begin{bmatrix} -1 & -3 \\ -1 & -2 \end{bmatrix}.$$

Hence, we have shown $C/C_{22} = A/A_{22} + B$ by use of Lemma 28.

Proposition 30 (Sum of two Schur complements) *If $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ are 2×2 block matrices*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},\ B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

such that $A_{22} \in \mathbb{C}^{p\times p}$, $B_{22} \in \mathbb{C}^{q\times q}$ are invertible and A/A_{22}, $B/B_{22} \in \mathbb{C}^{k\times k}$ then

$$C/C_{22} = A/A_{22} + B/B_{22}, \tag{4.5}$$

where $C \in \mathbb{C}^{(k+p+q)\times(k+p+q)}$ is the 3×3 block matrix with the following block partitioned structure $C = [C_{ij}]_{i,j=1,2}$:

$$C = \left[\begin{array}{c:c} C_{11} & C_{12} \\ \hdashline C_{21} & C_{22} \end{array}\right] = \left[\begin{array}{c:cc} A_{11}+B_{11} & A_{12} & B_{12} \\ \hdashline A_{21} & A_{22} & 0 \\ B_{21} & 0 & B_{22} \end{array}\right], \tag{4.6}$$

and

$$C_{22} = \begin{bmatrix} A_{22} & 0 \\ 0 & B_{22} \end{bmatrix} \tag{4.7}$$

is invertible. Moreover, if both matrices A and B are real, symmetric, Hermitian, or real and symmetric then the matrix C is real, symmetric, Hermitian, or real and symmetric, respectively.

Proof. The proof is a straightforward calculation using block matrix techniques. First, since $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, $A_{22} \in \mathbb{C}^{p\times p}$, $B_{22} \in \mathbb{C}^{q\times q}$, A/A_{22}, $B/B_{22} \in \mathbb{C}^{k\times k}$ with A_{22} and B_{22} invertible then the 3×3 block matrix C defined in (4.6) belongs to $\mathbb{C}^{(k+p+q)\times(k+p+q)}$. Second, with its partitioned block structure $C = [C_{i,j}]_{i,j=1,2}$, its $(2,2)$-

block C_{22} in (4.7), belongs to $\mathbb{C}^{(p+q)\times(p+q)}$ and is invertible with the inverse

$$C_{22}^{-1} = \begin{bmatrix} A_{22}^{-1} & 0 \\ 0 & B_{22}^{-1} \end{bmatrix}.$$

Therefore, we have

$$\begin{aligned}
C/C_{22} &= C_{11} - C_{12}C_{22}^{-1}C_{21} \\
&= A_{11} + B_{11} - \begin{bmatrix} A_{12} & B_{12} \end{bmatrix} \begin{bmatrix} A_{22} & 0 \\ 0 & B_{22} \end{bmatrix}^{-1} \begin{bmatrix} A_{21} \\ B_{21} \end{bmatrix} \\
&= A_{11} + B_{11} - \begin{bmatrix} A_{12} & B_{12} \end{bmatrix} \begin{bmatrix} A_{22}^{-1} & 0 \\ 0 & B_{22}^{-1} \end{bmatrix} \begin{bmatrix} A_{21} \\ B_{21} \end{bmatrix} \\
&= A_{11} + B_{11} - \begin{bmatrix} A_{12}A_{22}^{-1} & B_{12}B_{22}^{-1} \end{bmatrix} \begin{bmatrix} A_{21} \\ B_{21} \end{bmatrix} \\
&= A_{11} + B_{11} - \left(A_{12}A_{22}^{-1}A_{21} + B_{12}B_{22}^{-1}B_{21}\right) \\
&= A_{11} - A_{12}A_{22}^{-1}A_{21} + B_{11} - B_{12}B_{22}^{-1}B_{21} \\
&= A/A_{22} + B/B_{22}.
\end{aligned}$$

The remaining part of the proof follows immediately now from the formula (4.6) in terms of the matrices A and B. This completes the proof. ■

Example 31 *Let*

$$A = \left[\begin{array}{c:cc} -9 & 1 & 2 \\ \hdashline 1 & 3 & 2 \\ 2 & 2 & 1 \end{array}\right] \quad \text{and } B = \left[\begin{array}{c:cc} 2 & 3 & 2 \\ \hdashline 1 & 2 & 1 \\ 6 & 3 & 2 \end{array}\right].$$

Then

$$\begin{aligned}
A/A_{22} &= A_{11} - A_{12}A_{22}^{-1}A_{21} \\
&= [-9] - \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\
&= [-9] - \begin{bmatrix} 3 & -4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\
&= [-9] - [-5] \\
&= [-4],
\end{aligned}$$

and

$$\begin{aligned} B/B_{22} &= B_{11} - B_{12}B_{22}^{-1}B_{21} \\ &= [2] - \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 6 \end{bmatrix} \\ &= [2] - \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 6 \end{bmatrix} \\ &= [2] - [6] \\ &= [-4]. \end{aligned}$$

This implies $A/A_{22} + B/B_{22} = [-8]$*. By Proposition 30,*

$$C = \begin{bmatrix} A_{11} + B_{11} & A_{12} & B_{12} \\ A_{21} & A_{22} & 0 \\ B_{21} & 0 & B_{22} \end{bmatrix} = \left[\begin{array}{c:cccc} -7 & 1 & 2 & 3 & 2 \\ \hdashline 1 & 3 & 2 & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 2 & 1 \\ 6 & 0 & 0 & 3 & 2 \end{array}\right].$$

Hence,

$$\begin{aligned} C/C_{22} &= C_{11} - C_{12}C_{22}^{-1}C_{21} \\ &= [-7] - \begin{bmatrix} 1 & 2 & 3 & 2 \end{bmatrix} \begin{bmatrix} -1 & 2 & 0 & 0 \\ 2 & -3 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & -3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \\ 6 \end{bmatrix} \\ &= [-7] - \begin{bmatrix} 3 & -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \\ 6 \end{bmatrix} \\ &= [-7] - [1] \\ &= [-8]. \end{aligned}$$

Therefore, $C/C_{22} = A/A_{22} + B/B_{22}$*.*

The next two results, namely Lemma 32 and Proposition 34, use the direct sum operation, denoted as $\oplus$, for the definition of a direct sum and elementary properties see Section 2.3. The next lemma is interesting in its own right due to the importance of shorted matrices and operators both in electrical network theory and operator theory, see [71WA, 74AT, 05TA, 15ACM, 47MK, 10MBM, 76NA, 14EP]. Furthermore, it is also an intermediate step in proving Proposition 34 using Proposition 30.

Lemma 32 (Shorted matrices are Schur complements) *If* $A \in \mathbb{C}^{m\times m}$ *and* $B \in$

$\mathbb{C}^{n\times n}$ *are* 2×2 *block matrices*

$$A=\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix},\ B=\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}$$

such that $A_{22}\in\mathbb{C}^{p\times p}$, $B_{22}\in\mathbb{C}^{q\times q}$ *are invertible and* $A/A_{22}\in\mathbb{C}^{k\times k}$, $B/B_{22}\in\mathbb{C}^{l\times l}$ *then the direct sums* $A/A_{22}\oplus 0_l, 0_k\oplus B/B_{22}\in\mathbb{C}^{(k+l)\times(k+l)}$ *are Schur complements*

$$C/C_{22}=A/A_{22}\oplus 0_l=\begin{bmatrix}A/A_{22} & 0\\ 0 & 0_l\end{bmatrix}, \tag{4.8}$$

$$D/D_{22}=0_k\oplus B/B_{22}=\begin{bmatrix}0_k & 0\\ 0 & B/B_{22}\end{bmatrix}, \tag{4.9}$$

where $C\in\mathbb{C}^{(k+l+p)\times(k+l+p)}, D\in\mathbb{C}^{(k+l+q)\times(k+l+q)}$ *are* 3×3 *block matrices with the following block partitioned structure* $C=[C_{ij}]_{i,j=1,2}, D=[D_{ij}]_{i,j=1,2}$*:*

$$C=\left[\begin{array}{c:c}C_{11} & C_{12}\\ \hdashline C_{21} & C_{22}\end{array}\right]=\left[\begin{array}{cc:c}A_{11} & 0 & A_{12}\\ 0 & 0_l & 0\\ \hdashline A_{21} & 0 & A_{22}\end{array}\right], \tag{4.10}$$

$$D=\left[\begin{array}{c:c}D_{11} & D_{12}\\ \hdashline D_{21} & D_{22}\end{array}\right]=\left[\begin{array}{cc:c}0_k & 0 & 0\\ 0 & B_{11} & B_{12}\\ \hdashline 0 & B_{21} & B_{22}\end{array}\right], \tag{4.11}$$

and $C_{22}=A_{22}$*,* $D_{22}=B_{22}$ *are invertible. Moreover, if the matrix* A *(the matrix* B*) is real, symmetric, Hermitian, or real and symmetric then the matrix* C *(the matrix* D*) is real, symmetric, Hermitian, or real and symmetric, respectively.*

Proof. The proof is again a straightforward calculation using block matrix techniques. First, we compute, from the definition of $C=[C_{ij}]_{i,j=1,2}$ in (4.10),

$$\begin{aligned}C/C_{22}&=C_{11}-C_{12}C_{22}^{-1}C_{21}\\ &=\begin{bmatrix}A_{11} & 0\\ 0 & 0\end{bmatrix}-\begin{bmatrix}A_{12}\\ 0\end{bmatrix}A_{22}^{-1}\begin{bmatrix}A_{21} & 0\end{bmatrix}\\ &=\begin{bmatrix}A_{11} & 0\\ 0 & 0\end{bmatrix}-\begin{bmatrix}A_{12}A_{22}^{-1}\\ 0\end{bmatrix}\begin{bmatrix}A_{21} & 0\end{bmatrix}\\ &=\begin{bmatrix}A_{11} & 0\\ 0 & 0\end{bmatrix}-\begin{bmatrix}A_{12}A_{22}^{-1}A_{21} & 0\\ 0 & 0\end{bmatrix}\\ &=\begin{bmatrix}A_{11}-A_{12}A_{22}^{-1}A_{21} & 0\\ 0 & 0\end{bmatrix}\\ &=\begin{bmatrix}A/A_{22} & 0\\ 0 & 0\end{bmatrix}.\end{aligned}$$

Similarly, we compute, from the definition of $D = [D_{ij}]_{i,j=1,2}$ in (4.11),

$$\begin{aligned} D/D_{22} &= D_{11} - D_{12}D_{22}^{-1}D_{21} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & B_{11} \end{bmatrix} - \begin{bmatrix} 0 \\ B_{12} \end{bmatrix} B_{22}^{-1} \begin{bmatrix} 0 & B_{21} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & B_{11} \end{bmatrix} - \begin{bmatrix} 0 \\ B_{12}B_{22}^{-1} \end{bmatrix} \begin{bmatrix} 0 & B_{21} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & B_{11} \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & B_{12}B_{22}^{-1}B_{11} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & B_{11} - B_{12}B_{22}^{-1}B_{11} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & B/B_{22} \end{bmatrix}. \end{aligned}$$

The remaining part of the proof follows immediately now from the formulas (4.10) and (4.11) in terms of the matrices A and B, respectively. This completes the proof. ■

Example 33 *Let*

$$A = \left[\begin{array}{c:c} 1 & 3 \\ \hdashline 2 & 3 \end{array}\right].$$

Then

$$\begin{aligned} A/A_{22} &= [1] - [3][3]^{-1}[2] \\ &= [1] - [3] \begin{bmatrix} \frac{1}{3} \end{bmatrix} [2] \\ &= [1] - [2] \\ &= [-1]. \end{aligned}$$

Consider

$$A/A_{22} \oplus 0_l = [-1] \oplus 0_l = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Suppose C is from the definition (4.10), we have

$$C = \left[\begin{array}{cc:c} 1 & 0 & 3 \\ 0 & 0 & 0 \\ \hdashline 2 & 0 & 3 \end{array}\right],$$

which implies

$$\begin{aligned} C/C_{22} &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 3 \\ 0 \end{bmatrix} [3]^{-1} \begin{bmatrix} 2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 3 \\ 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \end{bmatrix} \begin{bmatrix} 2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 3 \\ 0 \end{bmatrix} \begin{bmatrix} \frac{2}{3} & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Hence $C/C_{22} = A/A_{22} \oplus 0_l$*. Since an example of* D *in Lemma 32 is almost identical to the above example we leave it as an exercise.*

Proposition 34 (Direct sum of Schur complements) *If* $A \in \mathbb{C}^{m\times m}$ *and* $B \in \mathbb{C}^{n\times n}$ *are* 2×2 *block matrices*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \ B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

such that $A_{22} \in \mathbb{C}^{p\times p}$, $B_{22} \in \mathbb{C}^{q\times q}$ *are invertible and* $A/A_{22} \in \mathbb{C}^{k\times k}$, $B/B_{22} \in \mathbb{C}^{l\times l}$ *then*

$$C/C_{22} = A/A_{22} \oplus B/B_{22} = \begin{bmatrix} A/A_{22} & 0 \\ 0 & B/B_{22} \end{bmatrix} \tag{4.12}$$

where $C \in \mathbb{C}^{(k+l+p+q)\times(k+l+p+q)}$ *is the* 2×2 *block matrix*

$$C = \left[\begin{array}{c:c} C_{11} & C_{12} \\ \hdashline C_{21} & C_{22} \end{array}\right] = \left[\begin{array}{c:c} A_{11} \oplus B_{11} & A_{12} \oplus B_{12} \\ \hdashline A_{21} \oplus B_{21} & A_{22} \oplus B_{22} \end{array}\right] = \left[\begin{array}{cc:cc} A_{11} & 0 & A_{12} & 0 \\ 0 & B_{11} & 0 & B_{12} \\ \hdashline A_{21} & 0 & A_{22} & 0 \\ 0 & B_{21} & 0 & B_{22} \end{array}\right] \tag{4.13}$$

and $C_{22} = A_{22} \oplus B_{22}$ *is invertible. Moreover, if both matrices* A *and* B *are real, symmetric, Hermitian, or real and symmetric then the matrix* C *is real, symmetric, Hermitian, or real and symmetric, respectively.*

Proof. The proof, as we shall see, follows immediately from Lemma 32 using Proposition 30. From the formulas

$$A/A_{22} \oplus B/B_{22} = \begin{bmatrix} A/A_{22} & 0 \\ 0 & B/B_{22} \end{bmatrix} = A/A_{22} \oplus 0_l + 0_k \oplus B/B_{22},$$

and, by Lemma 32,

$$A/A_{22} \oplus 0_l = \left[\begin{array}{cc:c} A_{11} & 0 & A_{12} \\ 0 & 0_l & 0 \\ \hdashline A_{21} & 0 & A_{22} \end{array}\right] /A_{22},$$

$$0_k \oplus B/B_{22} = \left[\begin{array}{cc:c} 0_k & 0 & 0 \\ 0 & B_{11} & B_{12} \\ \hdashline 0 & B_{21} & B_{22} \end{array}\right] /B_{22},$$

it follows immediately from Proposition 30 that

$$C/C_{22} = A/A_{22} \oplus 0_l + 0_k \oplus B/B_{22},$$

where C is given by the formula (4.13). The remaining part of the proof follows immediately now from the formula (4.13) in terms of the matrices A and B. This completes the proof. ■

Example 35 *Let A be the same matrix from Example 31. Suppose*

$$D = \left[\begin{array}{cc:c} A_{11} & 0 & A_{12} \\ 0 & 0 & 0 \\ \hdashline A_{21} & 0 & A_{22} \end{array}\right] = \left[\begin{array}{cc:cc} -9 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ \hdashline 1 & 0 & 3 & 2 \\ 2 & 0 & 2 & 1 \end{array}\right],$$

and

$$E = \left[\begin{array}{c:cc} 0 & 0 & 0 \\ \hdashline 0 & B_{11} & B_{12} \\ 0 & B_{21} & B_{22} \end{array}\right] = \left[\begin{array}{c:ccc} 0 & 0 & 0 & 0 \\ \hdashline 0 & 2 & 3 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & 6 & 3 & 2 \end{array}\right].$$

We have

$$\begin{aligned} D/D_{22} &= D_{11} - D_{12}D_{22}^{-1}D_{21} \\ &= \begin{bmatrix} -9 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -9 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 3 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -9 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} -5 & 0 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -4 & 0 \\ 0 & 0 \end{bmatrix}, \end{aligned}$$

and

$$\begin{aligned}
E/E_{22} &= E_{11} - E_{12}E_{22}^{-1}E_{21} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 6 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 6 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 6 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & -4 \end{bmatrix}.
\end{aligned}$$

By Proposition 34,

$$F = \left[\begin{array}{c:cc} D_{11}+E_{11} & D_{12} & E_{12} \\ \hdashline D_{21} & D_{22} & 0 \\ E_{21} & 0 & E_{22} \end{array}\right] = \left[\begin{array}{cc:cccc} -9 & 0 & 1 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 & 3 & 2 \\ \hdashline 1 & 0 & 3 & 2 & 0 & 0 \\ 2 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 1 \\ 0 & 6 & 0 & 0 & 3 & 2 \end{array}\right].$$

Then

$$\begin{aligned}
F/F_{22} &= F_{11} - F_{12}F_{22}^{-1}F_{21} \\
&= \begin{bmatrix} -9 & 0 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 3 & 2 \end{bmatrix} \begin{bmatrix} -1 & 2 & 0 & 0 \\ 2 & -3 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & -3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \\ 0 & 6 \end{bmatrix} \\
&= \begin{bmatrix} -9 & 0 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 3 & -4 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \\ 0 & 6 \end{bmatrix} \\
&= \begin{bmatrix} -9 & 0 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} -5 & 0 \\ 0 & 6 \end{bmatrix} \\
&= \begin{bmatrix} -4 & 0 \\ 0 & -4 \end{bmatrix} \\
&= \begin{bmatrix} A/A_{22} & 0 \\ 0 & B/B_{22} \end{bmatrix} \\
&= D/D_{22} + E/E_{22} \\
&= A/A_{22} \oplus B/B_{22}.
\end{aligned}$$

Therefore, $F/F_{22} = A/A_{22} \oplus B/B_{22}$*.*

4.3 Matrix products and inverses

Remark 36 *Proposition 37 has a slight modification from [20SW], where now the resulting block matrix definition of* C *[i.e., (4.15)] can maintain the structural symmetry of both* A *and* B *if* $A = B^T$ *or* $A = B^*$*. Otherwise, without this relationship between* A *and* B*, the resulting matrix remains nonsymmetric after multiplication.*

Proposition 37 (Matrix multiplication of two Schur complements) *If* $A \in \mathbb{C}^{m\times m}$ *and* $B \in \mathbb{C}^{n\times n}$ *are* 2×2 *block matrices*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},\ B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

such that the matrices A_{22} *and* B_{22} *are invertible and* $A/A_{22}, B/B_{22} \in \mathbb{C}^{k\times k}$ *then*

$$C/C_{22} = (A/A_{22})(B/B_{22}), \tag{4.14}$$

where $C \in \mathbb{C}^{(m+n-k)\times(m+n-k)}$ *is the* 3×3 *block matrix with the following block partitioned structure* $C = [C_{ij}]_{i,j=1,2}$*:*

$$C = \left[\begin{array}{c:c} C_{11} & C_{12} \\ \hdashline C_{21} & C_{22} \end{array}\right] = \left[\begin{array}{c:cc} A_{11}B_{11} & A_{11}B_{12} & A_{12} \\ \hdashline A_{21}B_{11} & A_{21}B_{12} & A_{22} \\ B_{21} & B_{22} & 0 \end{array}\right], \tag{4.15}$$

where the matrix C_{22} *is invertible with*

$$C_{22}^{-1} = \begin{bmatrix} A_{21}B_{12} & A_{22} \\ B_{22} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & B_{22}^{-1} \\ A_{22}^{-1} & -A_{22}^{-1}A_{21}B_{12}B_{22}^{-1} \end{bmatrix}. \tag{4.16}$$

Moreover, if $A = B^T$ *and* A *is real, symmetric, Hermitian, or real and symmetric then the matrix* C *is real, symmetric, Hermitian, or real and symmetric, respectively.*

Proof. The proof is just a straightforward application of block matrix multiplication. First of all, by the hypotheses the matrix products in the statement of the proposition are well-defined and its easy to verify that $C \in \mathbb{C}^{(m+n-k)\times(m+n-k)}$ as well as the inverse formula (4.16) for C_{22}^{-1} is correct. Second, we compute, from the definition of $C =$

$[C_{ij}]_{i,j=1,2}$ in (4.15),

$$
\begin{aligned}
C/C_{22} &= C_{11} - C_{12}C_{22}^{-1}C_{21} \\
&= A_{11}B_{11} - \begin{bmatrix} A_{11}B_{12} & A_{12} \end{bmatrix} \begin{bmatrix} 0 & B_{22}^{-1} \\ A_{22}^{-1} & -A_{22}^{-1}A_{21}B_{12}B_{22}^{-1} \end{bmatrix} \begin{bmatrix} A_{21}B_{11} \\ B_{21} \end{bmatrix} \\
&= A_{11}B_{11} - \begin{bmatrix} A_{12}A_{22}^{-1} & A_{12}(-A_{22}^{-1}A_{21}B_{12}B_{22}^{-1}) + A_{11}B_{12}B_{22}^{-1} \end{bmatrix} \begin{bmatrix} A_{21}B_{11} \\ B_{21} \end{bmatrix} \\
&= A_{11}B_{11} - \{A_{12}A_{22}^{-1}A_{21}B_{11} + [A_{12}(-A_{22}^{-1}A_{21}B_{12}B_{22}^{-1}) + A_{11}B_{12}B_{22}^{-1}]B_{21}\} \\
&= A_{11}B_{11} - A_{12}A_{22}^{-1}A_{21}B_{11} + A_{12}A_{22}^{-1}A_{21}B_{12}B_{22}^{-1}B_{21} - A_{11}B_{12}B_{22}^{-1}B_{21} \\
&= (A_{11} - A_{12}A_{22}^{-1}A_{21})B_{11} - (A_{11} - A_{12}A_{22}^{-1}A_{21})B_{12}B_{22}^{-1}B_{21} \\
&= (A_{11} - A_{12}A_{22}^{-1}A_{21})(B_{11} - B_{12}B_{22}^{-1}B_{21}) \\
&= (A/A_{22})(B/B_{22}).
\end{aligned}
$$

The remaining part of the proof follows immediately now from the formula (4.15) in terms of the matrices A and B. This completes the proof. ∎

Example 38 *Let*

$$
A = B^T = \left[\begin{array}{c|cc} 3 & 1 & 0 \\ \hline 1 & 2 & 1 \\ 0 & 1 & 1 \end{array}\right].
$$

Then

$$
\begin{aligned}
A/A_{22} = B/B_{22} &= [3] - \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= [3] - \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= [3] - [1] \\
&= [2].
\end{aligned}
$$

Hence, $(A/A_{22})(B/B_{22}) = [2][2] = [4]$. *Let* C/C_{22} *be by definition (4.15) and* C^{-1} *be (4.16), then*

$$
C = \left[\begin{array}{c|cccc} 9 & 3 & 0 & 1 & 0 \\ \hline 3 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right]
$$

and

$$C_{22}^{-1} = \begin{bmatrix} 0 & 0 & 1 & -1 \\ 0 & 1 & -1 & 2 \\ 1 & -1 & -1 & 0 \\ -1 & 2 & 1 & 0 \end{bmatrix}.$$

Thus,

$$\begin{aligned} C/C_{22} &= C_{11} - C_{12}C_{22}^{-1}C_{21} \\ &= [9] - \begin{bmatrix} 3 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & -1 \\ 0 & 1 & -1 & 2 \\ 1 & -1 & -1 & 1 \\ -1 & 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 1 \\ 0 \end{bmatrix} \\ &= [9] - [5] \\ &= [4]. \end{aligned}$$

Therefore, we have shown $C/C_{22} = (A/A_{22})(B/B_{22})$ by the definitions of Proposition 37 and when $A = B^T$ and they are both real, C is also real and symmetric.

Proposition 39 (Matrix multiplication of a Schur complement) *If $A \in \mathbb{C}^{m\times m}$ is a 2×2 block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

such that $A_{22} \in \mathbb{C}^{p\times p}$ is invertible and $A/A_{22} \in \mathbb{C}^{k\times k}$ then, for any matrices $B \in \mathbb{C}^{l\times k}$ and $C \in \mathbb{C}^{k\times l}$,

$$D/D_{22} = B(A/A_{22})C, \tag{4.17}$$

where $D \in \mathbb{C}^{(l+p)\times(l+p)}$ is the 2×2 block matrix

$$D = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} = \left[\begin{array}{c|c} BA_{11}C & BA_{12} \\ \hline A_{21}C & A_{22} \end{array}\right] = \begin{bmatrix} B & 0 \\ 0 & I_p \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} C & 0 \\ 0 & I_p \end{bmatrix} \tag{4.18}$$

and $D_{22} = A_{22}$ is invertible. Moreover, the following statements are true:

(a) If A, B, C are real matrices then D is a real matrix.

(b) If A is a symmetric matrix and $C = B^T$ then D is a symmetric matrix.

(c) If the hypotheses of (a) and (b) are true then D is a real symmetric matrix.

(d) If A is a Hermitian matrix and $C = B^*$ then D is a Hermitian matrix.

Proof. By block multiplication the result follows immediately from the definition of $D = [D_{ij}]_{i,j=1,2}$ in (4.18) since

$$\begin{aligned} D/D_{22} &= D_{11} - D_{12}D_{22}^{-1}D_{21} \\ &= BA_{11}C - BA_{12}A_{22}^{-1}A_{21}C \\ &= B(A_{11} - A_{12}A_{22}^{-1}A_{21})C \\ &= B(A/A_{22})C. \end{aligned}$$

The remaining part of the proof follows immediately now from the formula (4.18) in terms of the matrices A, B and C. ■

Example 40 *Let*

$$A = \left[\begin{array}{c:c} 0 & 2 \\ \hdashline 2 & 1 \end{array}\right].$$

Then

$$\begin{aligned} A/A_{22} &= [0] - [2][1]^{-1}[2] \\ &= [0] - [2][2] \\ &= [0] - [4] \\ &= [-4]. \end{aligned}$$

Suppose

$$C = B^T = \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix}.$$

Consider

$$\begin{aligned} B(A/A_{22})C &= \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} [-4] \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -12 & -4 \\ -4 & 0 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -40 & -12 \\ -12 & -4 \end{bmatrix}. \end{aligned}$$

Suppose D is from the definition (4.18), we have

$$\begin{aligned}
D_{11} &= BA_{11}C \\
&= \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} [0] \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \\
D_{12} &= BA_{12} \\
&= \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} [2] \\
&= \begin{bmatrix} 6 & 2 \\ 2 & 0 \end{bmatrix}; \\
D_{21} &= A_{21}C \\
&= [2] \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 6 & 2 \\ 2 & 0 \end{bmatrix}; \\
D_{22} &= A_{22} \\
&= [1],
\end{aligned}$$

which implies

$$\begin{aligned}
D/D_{22} &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 6 & 2 \\ 2 & 0 \end{bmatrix} [1]^{-1} \begin{bmatrix} 6 & 2 \\ 2 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 6 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 6 & 2 \\ 2 & 0 \end{bmatrix} \\
&= - \begin{bmatrix} 40 & 12 \\ 12 & 4 \end{bmatrix}.
\end{aligned}$$

Hence $D/D_{22} = B(A/A_{22})C$. *Notice we also showed explicitly statement (c).*

Lemma 41 (Inverse is a Schur complement) *If* $A \in \mathbb{C}^{m\times m}$ *is a* 2×2 *block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{4.19}$$

such that $A_{22} \in \mathbb{C}^{p\times p}$ *is invertible then* A/A_{22} *is invertible and only if* A *is invertible, in which case*

$$A^{-1} = \begin{bmatrix} (A/A_{22})^{-1} & -(A/A_{22})^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}(A/A_{22})^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}(A/A_{22})^{-1}A_{12}A_{22}^{-1} \end{bmatrix} \tag{4.20}$$

and

$$(A/A_{22})^{-1} = (A^{-1})_{11}. \tag{4.21}$$

Furthermore,

$$B/B_{22} = A^{-1}, \tag{4.22}$$

where $B \in \mathbb{C}^{2m\times 2m}$ *is the* 2×2 *block matrix*

$$B = \left[\begin{array}{c:c} 0_m & I_m \\ \hdashline I_m & -A \end{array}\right] \tag{4.23}$$

and $B_{22} = -A$ *is invertible. In addition, if the matrix* A *is real, symmetric, Hermitian, or real and symmetric then the matrix* B *is real, symmetric, Hermitian, or real and symmetric, respectively.*

Proof. First, if $A \in \mathbb{C}^{m\times m}$ is a 2×2 block matrix of the form (4.19) and $A_{22} \in \mathbb{C}^{p\times p}$ is invertible, then by Lemma 4 and Corollary 5, A/A_{22} is invertible if and only if A is invertible, in which case we have the factorization

$$A^{-1} = \begin{bmatrix} I_{m-p} & 0 \\ -A_{22}^{-1}A_{21} & I_p \end{bmatrix} \begin{bmatrix} (A/A_{22})^{-1} & 0 \\ 0 & A_{22}^{-1} \end{bmatrix} \begin{bmatrix} I_{m-p} & -A_{12}A_{22}^{-1} \\ 0 & I_p \end{bmatrix}.$$

Hence, by block multiplication it follows immediately that the equality in (4.20) is true which implies the equality (4.21) is also true. Finally, for the matrix $B \in \mathbb{C}^{2m\times 2m}$ defined in (4.22) we have $B_{22} = -A$ is invertible so that

$$\begin{aligned} B/B_{22} &= B_{11} - B_{12}B_{22}^{-1}B_{21} \\ &= I_m A^{-1} I_m \\ &= A^{-1}. \end{aligned}$$

The remaining part of the proof follows immediately now from the formula (4.22) in terms of the matrix A. This completes the proof. ■

Example 42 *Recall* A *from Example 31. Let*

$$S = \begin{bmatrix} 0 & I \\ I & -A \end{bmatrix} = \left[\begin{array}{ccc:ccc} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hdashline 1 & 0 & 0 & 9 & -1 & -2 \\ 0 & 1 & 0 & -1 & -3 & -2 \\ 0 & 0 & 1 & -2 & -2 & -1 \end{array}\right].$$

Then

$$
\begin{aligned}
S/S_{22} &= S_{11} - S_{12}S_{22}^{-1}S_{21} \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & -\frac{3}{4} & 1 \\ -\frac{3}{4} & \frac{13}{4} & -5 \\ 1 & -5 & 7 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= - \begin{bmatrix} \frac{1}{4} & -\frac{3}{4} & 1 \\ -\frac{3}{4} & \frac{13}{4} & -5 \\ 1 & -5 & 7 \end{bmatrix} \\
&= \begin{bmatrix} -\frac{1}{4} & \frac{3}{4} & -1 \\ \frac{3}{4} & -\frac{13}{4} & 5 \\ -1 & 5 & -7 \end{bmatrix} \\
&= A^{-1}.
\end{aligned}
$$

Therefore, $S/S_{22} = A^{-1}$.

Proposition 43 (Inverse of a Schur complement) *If* $A \in \mathbb{C}^{m\times m}$ *is a* 2×2 *block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}, \end{bmatrix}$$

such that A_{22} *is invertible and* $A/A_{22} \in \mathbb{C}^{k\times k}$ *is invertible then*

$$C/C_{22} = (A/A_{22})^{-1} \tag{4.24}$$

where $C \in \mathbb{C}^{(k+m)\times(k+m)}$ *is the* 3×3 *block matrix with the following block partitioned structure* $C = [C_{ij}]_{i,j=1,2}$*:*

$$C = \left[\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right] = \left[\begin{array}{c|cc} 0_k & I_k & 0 \\ \hline I_k & -A_{11} & -A_{12} \\ 0 & -A_{21} & -A_{22} \end{array}\right] \tag{4.25}$$

with $C_{22} = -A$ *invertible. Moreover, if the matrix* A *is real, symmetric, Hermitian, or real and symmetric then the matrix* B *is real, symmetric, Hermitian, or real and symmetric, respectively.*

Proof. By Lemma 41 and, in particular, formulas (4.21)-(4.23) we have

$$(A/A_{22})^{-1} = (A^{-1})_{11} = \begin{bmatrix} I_k & 0 \end{bmatrix} A^{-1} \begin{bmatrix} I_k \\ 0 \end{bmatrix} = \begin{bmatrix} I_k & 0 \end{bmatrix} B/B_{22} \begin{bmatrix} I_k \\ 0 \end{bmatrix},$$

where the matrix $B \in \mathbb{C}^{2m\times 2m}$ is defined in terms of A in (4.22). Hence, by Proposition 39

$$C/C_{22} = \begin{bmatrix} I_k & 0 \end{bmatrix} B/B_{22} \begin{bmatrix} I_k \\ 0 \end{bmatrix},$$

where $C = [C_{ij}]_{i,j=1,2} \in \mathbb{C}^{(k+m)\times(k+m)}$ is given by

$$C_{11} = \begin{bmatrix} I_k & 0 \end{bmatrix} B_{11} \begin{bmatrix} I_k \\ 0 \end{bmatrix} = 0_k, \quad C_{12} = \begin{bmatrix} I_k & 0 \end{bmatrix} B_{12} = \begin{bmatrix} I_k & 0 \end{bmatrix}$$
$$= B_{21} \begin{bmatrix} I_k \\ 0 \end{bmatrix} = \begin{bmatrix} I_k \\ 0 \end{bmatrix}, \quad C_{22} = -A = \begin{bmatrix} -A_{11} & -A_{12} \\ -A_{21} & -A_{22} \end{bmatrix},$$

which yields the formula (4.25) for C. The remaining part of the proof follows immediately now from the formula (4.25) in terms of the matrix A. This completes the proof. ■

Example 44 *Continuing to use A from Example 31. Let*

$$G = \left[\begin{array}{c:cc} 0 & I & 0 \\ \hdashline I & -A_{11} & -A_{12} \\ 0 & -A_{21} & -A_{22} \end{array}\right] = \left[\begin{array}{c:ccc} 0 & 1 & 0 & 0 \\ \hdashline 1 & 9 & -1 & -2 \\ 0 & -1 & -3 & -2 \\ 0 & -2 & -2 & -1 \end{array}\right].$$

Then

$$\begin{aligned} G/G_{22} &= G_{11} - G_{12}G_{22}^{-1}G_{21} \\ &= [0] - \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & -\frac{3}{4} & 1 \\ -\frac{3}{4} & \frac{13}{4} & -5 \\ 1 & -5 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= -\begin{bmatrix} \frac{1}{4} & -\frac{3}{4} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{4} \end{bmatrix}. \end{aligned}$$

As we showed in Example 31, $A/A_{22} = -4$, which implies $(A/A_{22})^{-1} = \begin{bmatrix} -\frac{1}{4} \end{bmatrix}$. Therefore, $G/G_{22} = (A/A_{22})^{-1}$.

4.4 Kronecker products

The results in this section, use the Kronecker products [denoted as $\otimes$, for the definition of a Kronecker product and elementary properties see (2.13)] of matrices when one or more of the matrices is a Schur complement, are by far the most technical part of the chapter. The following are two major reasons for this.

First, the technique of finding Schur complement representation for $A/A_{22} \otimes B$ requires that B is invertible by Lemma 45. And from this simple result though we are able to "easily" build up on it to find Schur complement representations for $A \otimes B/B_{22}$

(Corollary 49), but requires A to be invertible, and $A/A_{22} \otimes B/B_{22}$ (Proposition 51), but it requires A/A_{22} and B/B_{22} to be invertible. Again, these invertibility requirements are just due to the invertibility hypothesis in Lemma 45. We will treat noninvertibility in general by using Lemma 54.

The second reason that this section is more technical is due to the fact that in the proof of Proposition 51, where we derive a Schur complement formula for the Kronecker product of two Schur complements, i.e., $A/A_{22} \otimes B/B_{22}$, we must use the result in Section 4.5 on composition of Schur complements (Proposition 57).

Lemma 45 (Kronecker product of a Schur complement with a matrix) *If $B \in \mathbb{C}^{n\times n}$ and $A \in \mathbb{C}^{m\times m}$ is a 2×2 block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

then the Kronecker product of A with B,

$$C = A \otimes B \in \mathbb{C}^{mn\times mn}, \tag{4.26}$$

has following 2×2 block matrix form $C = [C_{i,j}]_{i.j=1,2}$:

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \otimes B & A_{12} \otimes B \\ A_{21} \otimes B & A_{22} \otimes B \end{bmatrix}. \tag{4.27}$$

Furthermore, if A_{22} and B are invertible then $C_{22} = A_{22} \otimes B$ is invertible and

$$C/C_{22} = A/A_{22} \otimes B. \tag{4.28}$$

Moreover, if both matrices A and B are real, symmetric, Hermitian, or real and symmetric then the matrix C is real, symmetric, Hermitian, or real and symmetric, respectively.

Proof. The first part of the proof of this lemma, namely, that $C = A \otimes B \in \mathbb{C}^{mn\times mn}$ has the block form (4.26), follows immediately from the definition of the Kronecker product $A \otimes B = [a_{ij}B]_{i,j=1,\ldots,m}$ of the matrices A and B and the 2×2 block form of $A = [A_{ij}]_{i,j=1,2}$. Suppose now that A_{22} and B are invertible. Then it follows by elementary properties of Kronecker products that their Kronecker product $A_{22} \otimes B$ is invertible and by (2.16) that

$$\begin{aligned} C/C_{22} &= A_{11}\otimes B - (A_{12}\otimes B)(A_{22}\otimes B)^{-1}(A_{21}\otimes B) \\ &= A_{11}\otimes B - (A_{12}\otimes B)\left(A_{22}^{-1}\otimes B^{-1}\right)(A_{21}\otimes B) \\ &= A_{11}\otimes B - \left[(A_{12}A_{22}^{-1})\otimes (BB^{-1})\right](A_{21}\otimes B) \\ &= A_{11}\otimes B - \left[(A_{12}A_{22}^{-1}A_{21})\otimes B\right] \\ &= \left(A_{11} - A_{12}A_{22}^{-1}A_{21}\right)\otimes B \\ &= A/A_{22}\otimes B. \end{aligned}$$

The remaining part of the proof follows immediately now by elementary the properties of Kronecker products [(2.14), (2.15), and (2.17)]. This completes the proof. ■

Example 46 *Let*

$$A = \left[\begin{array}{c:c} A_{11} & A_{12} \\ \hdashline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{c:c} 0 & 2 \\ \hdashline 2 & 4 \end{array}\right], \quad B = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}.$$

Then

$$C = A\otimes B = \begin{bmatrix} 0B & 2B \\ 2B & 4B \end{bmatrix} = \begin{bmatrix} 0 & 0 & 4 & 6 \\ 0 & 0 & 6 & 10 \\ 4 & 6 & 8 & 12 \\ 6 & 10 & 12 & 20 \end{bmatrix}$$

and has the 2×2 *block matrix form*

$$C = \left[\begin{array}{c:c} C_{11} & C_{12} \\ \hdashline C_{21} & C_{22} \end{array}\right] = \left[\begin{array}{c:c} A_{11}\otimes B & A_{12}\otimes B \\ \hdashline A_{21}\otimes B & A_{22}\otimes B \end{array}\right] = \left[\begin{array}{cc:cc} 0 & 0 & 4 & 6 \\ 0 & 0 & 6 & 10 \\ \hdashline 4 & 6 & 8 & 12 \\ 6 & 10 & 12 & 20 \end{array}\right].$$

Now $A_{22} = [4]$ *and* B *are invertible which implies, by Lemma 45, that* $C_{22} = A_{22}\otimes B$ *is invertible with* $C_{22}^{-1} = A_{22}^{-1}\otimes B^{-1}$ *and* $C/C_{22} = A/A_{22}\otimes B$*, which we can show in this example by the following direct calculations:*

$$A_{22}^{-1} = \left[\tfrac{1}{4}\right], \quad B^{-1} = \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix}, \quad A_{22}^{-1}\otimes B^{-1} = \frac{1}{4}B^{-1} = \begin{bmatrix} \frac{5}{4} & -\frac{3}{4} \\ -\frac{3}{4} & \frac{2}{4} \end{bmatrix} = C_{22}^{-1},$$

$$A/A_{22} = [0] - [2][4]^{-1}[2] = \left[-1\right],$$

$$A/A_{22}\otimes B = (-1)B = \begin{bmatrix} -2 & -3 \\ -3 & -5 \end{bmatrix},$$

$$C/C_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 4 & 6 \\ 6 & 10 \end{bmatrix}\begin{bmatrix} 8 & 12 \\ 12 & 20 \end{bmatrix}^{-1}\begin{bmatrix} 4 & 6 \\ 6 & 10 \end{bmatrix} = \begin{bmatrix} -2 & -3 \\ -3 & -5 \end{bmatrix}.$$

The definition below comes from [91HJ, p. 259].

Definition 47 *For any positive integers m, n, the matrix $P(m,n) \in \mathbb{C}^{mn\times mn}$ is defined by*

$$P(m,n) = \sum_{i=1}^{m}\sum_{j=1}^{n} E_{ij}\otimes E_{ij}^T = [E_{ij}^T]_{i,j=1}^{m,n}, \tag{4.29}$$

where $\{E_{ij} : i = 1,\dots,m, j = 1,\dots,n\}$ is the standard basis for $\mathbb{C}^{m\times n}$, i.e., each $E_{ij} \in \mathbb{C}^{m\times n}$ has entry 1 in the ith row, jth column and all other entries are zero. The matrix $P(m,n)$ is called the commutation matrix (with respect to m and n).

The following lemma is proven in [91HJ, Corollary 4.3.10, p. 260].

Lemma 48 (Main properties of commutation matrices) *Let positive integers m, n, p, and q be given and let $P(p,m) \in \mathbb{C}^{pm\times pm}$ and $P(n,q) \in \mathbb{C}^{nq\times nq}$ denote the commutation matrices (as defined in Def. 47). Then $P(p,m)$ is a permutation matrix and $P(p,m) = \overline{P(p,m)} = P(m,p)^T = P(m,p)^{-1}$. Furthermore, for all $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{p\times q}$,*

$$B\otimes A = P(m,p)^T(A\otimes B)P(n,q). \tag{4.30}$$

Corollary 49 (Part 2 of Lemma 45) *If $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ is a 2×2 block matrix*

$$B = \begin{bmatrix} B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix},$$

such that A and B_{22} are invertible with $B/B_{22} \in \mathbb{C}^{l\times l}$ then

$$D/D_{22} = A\otimes B/B_{22}, \tag{4.31}$$

where

$$D = Q^T(B\otimes A)Q,\ Q \in \mathbb{C}^{mn\times mn} \tag{4.32}$$

have the following 2×2 block matrix forms $Q = [Q_{i,j}]_{i.j=1,2}$, $D = [D_{i,j}]_{i.j=1,2}$:

$$Q = \begin{bmatrix} Q_{11} & Q_{12}\\ Q_{21} & Q_{22}\end{bmatrix} = \begin{bmatrix} P(l,m) & 0\\ 0 & I_{mn-lm}\end{bmatrix}, \tag{4.33}$$

$$D = \begin{bmatrix} D_{11} & D_{12}\\ D_{21} & D_{22}\end{bmatrix} = \begin{bmatrix} A\otimes B_{11} & P(l,m)^T(B_{12}\otimes A)\\ (B_{21}\otimes A)P(l,m) & B_{22}\otimes A\end{bmatrix}, \tag{4.34}$$

in which $P(l,m) \in \mathbb{C}^{lm\times lm}$ is the commutation matrix (with respect to l and m as defined in Def. 47) and $D_{22} = B_{22}\otimes A$ is invertible. Furthermore, Q is a permutation matrix and $Q = \overline{Q}$, $Q^T = Q^{-1}$. Moreover, if both matrices A and B are real, symmetric, Hermitian, or real and symmetric then the matrix D is real, symmetric, Hermitian, or real and symmetric, respectively.

Proof. By the hypotheses, Lemma 48, and Lemma 45 we have

$$P(m,l)^T(A\otimes B/B_{22})P(m,l) = B/B_{22}\otimes A = C/C_{22},$$

where $C = B \otimes A \in \mathbb{C}^{nm \times nm}$ has the 2×2 block matrix form:

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} B_{11} \otimes A & B_{12} \otimes A \\ B_{21} \otimes A & B_{22} \otimes A \end{bmatrix},$$

in which $C_{22} = B_{22} \otimes A$ is invertible. By Lemma 48 and Proposition 39 it follows that

$$A \otimes B/B_{22} = [P(m,l)^T]^{-1} C/C_{22} P(m,l)^{-1} = P(l,m)^T C/C_{22} P(l,m) = D/D_{22},$$

where $D = [D_{ij}]_{i,j=1,2} \in \mathbb{C}^{mn \times mn}$ is the 2×2 block matrix defined by (4.34), where we have used the fact that

$$P(l,m)^T (B_{11} \otimes A) P(l,m) = A \otimes B_{11}, \tag{4.35}$$

which follows from Lemma 48, from which it follows immediately from block multiplication that $D = Q^T (B \otimes A) Q$, where $Q \in \mathbb{C}^{mn \times mn}$ is the matrix defined in (4.33). The fact that Q is a permutation matrix satisfying $Q = \overline{Q}$, $Q^T = Q^{-1}$ follows from its definition (4.33) and the corresponding properties of $P(l,m)$ in Lemma 48. The remaining part of the proof follows immediately now formula for D in (4.34) and the properties of Q. This completes the proof. ■

Example 50 *Let*

$$A = \begin{bmatrix} 0 & 2 \\ 2 & 4 \end{bmatrix}, \quad B = \left[\begin{array}{c:c} B_{11} & B_{12} \\ \hdashline B_{21} & B_{22} \end{array}\right] = \left[\begin{array}{c:c} 2 & 3 \\ \hdashline 3 & 5 \end{array}\right].$$

Then, for this example in the notation of Corollary 49, we have $m = n = 2, l = 1$ and

$$P(1,2) = \left[E_{ij}^T\right]_{i,j=1}^{1,2} = \begin{bmatrix} E_{11}^T & E_{12}^T \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix}^T & \begin{bmatrix} 0 & 1 \end{bmatrix}^T \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2,$$

$$Q = \begin{bmatrix} P(1,2) & 0 \\ 0 & I_2 \end{bmatrix} = \begin{bmatrix} I_2 & 0 \\ 0 & I_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = I_4,$$

$$D = Q^T (B \otimes A) Q = B \otimes A = \begin{bmatrix} 2A & 3A \\ 3A & 5A \end{bmatrix} = \begin{bmatrix} 0 & 4 & 0 & 6 \\ 4 & 8 & 6 & 12 \\ 0 & 6 & 0 & 10 \\ 6 & 12 & 10 & 20 \end{bmatrix}$$

and D has the 2×2 block matrix form

$$D = \left[\begin{array}{c:c} D_{11} & D_{12} \\ \hdashline D_{21} & D_{22} \end{array}\right] = \left[\begin{array}{c:c} A \otimes B_{11} & P(1,2)^T(B_{12} \otimes A) \\ \hdashline (B_{21} \otimes A)P(1,2) & B_{22} \otimes A \end{array}\right]$$

$$= \left[\begin{array}{c:c} A \otimes B_{11} & B_{12} \otimes A \\ \hdashline B_{21} \otimes A & B_{22} \otimes A \end{array}\right] = \left[\begin{array}{cc:cc} 0 & 4 & 0 & 6 \\ 4 & 8 & 6 & 12 \\ \hdashline 0 & 6 & 0 & 10 \\ 6 & 12 & 10 & 20 \end{array}\right].$$

Now A and $B_{22} = [5]$ are invertible which implies, by Corollary 49, that $D_{22} = B_{22} \otimes A$ is invertible with $D_{22}^{-1} = B_{22}^{-1} \otimes A^{-1}$ and $D/D_{22} = A \otimes B/B_{22}$, which we can show in this example by the following direct calculations:

$$B_{22}^{-1} = \left[\tfrac{1}{5}\right], \quad A^{-1} = \begin{bmatrix} -1 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \quad B_{22}^{-1} \otimes A^{-1} = \frac{1}{5}A^{-1} = \begin{bmatrix} -\frac{1}{5} & \frac{1}{10} \\ \frac{1}{10} & 0 \end{bmatrix} = D_{22}^{-1},$$

$$B/B_{22} = [2] - [3][5]^{-1}[3] = \left[\tfrac{1}{5}\right],$$

$$A \otimes B/B_{22} = \begin{bmatrix} 0\left[\frac{1}{5}\right] & 2\left[\frac{1}{5}\right] \\ 2\left[\frac{1}{5}\right] & 4\left[\frac{1}{5}\right] \end{bmatrix} = \begin{bmatrix} 0 & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix},$$

$$D/D_{22} = \begin{bmatrix} 0 & 4 \\ 4 & 8 \end{bmatrix} - \begin{bmatrix} 0 & 6 \\ 6 & 12 \end{bmatrix} \begin{bmatrix} 0 & 10 \\ 10 & 20 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 6 \\ 6 & 12 \end{bmatrix} = \begin{bmatrix} 0 & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix}.$$

Proposition 51 (Kronecker product of two Schur complements) *If $A \in \mathbb{C}^{m \times m}$ and $B \in \mathbb{C}^{n \times n}$ are 2×2 block matrices*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \; B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

such that the matrices A_{22}, B_{22}, A/A_{22}, and B/B_{22} are invertible with $B/B_{22} \in \mathbb{C}^{l \times l}$ then

$$M/M_{22} = A/A_{22} \otimes B/B_{22}, \tag{4.36}$$

where $M \in \mathbb{C}^{mn \times mn}$ is the invertible matrix

$$M = Q^T(B \otimes A)Q, \tag{4.37}$$

$Q \in \mathbb{C}^{mn \times mn}$ is the permutation matrix defined by (4.33), and $M = [M_{ij}]_{i,j=1,2}$ is the 2×2 matrix with the block partitioned structure:

$$M = \left[\begin{array}{c:c} M_{11} & M_{12} \\ \hdashline M_{21} & M_{22} \end{array}\right] = \left[\begin{array}{c:cc} D_{33} & D_{34} & D_{32} \\ \hdashline D_{43} & D_{44} & D_{42} \\ D_{23} & D_{24} & D_{22} \end{array}\right], \tag{4.38}$$

where

$$M_{22} = \begin{bmatrix} D_{44} & D_{42} \\ D_{24} & D_{22} \end{bmatrix}, \tag{4.39}$$

is invertible, and

$$\left[\begin{array}{c:c} D_{33} & D_{34} \\ \hdashline D_{43} & D_{44} \end{array}\right] = \left[\begin{array}{c:c} A_{11} \otimes B_{11} & A_{12} \otimes B_{11} \\ \hdashline A_{21} \otimes B_{11} & A_{22} \otimes B_{11} \end{array}\right] = A \otimes B_{11}, \tag{4.40}$$

$$D_{22} = B_{22} \otimes A, \tag{4.41}$$

$$\begin{bmatrix} D_{32} \\ D_{42} \end{bmatrix} = P(l,m)^T (B_{12} \otimes A), \tag{4.42}$$

$$\begin{bmatrix} D_{23} & D_{24} \end{bmatrix} = (B_{21} \otimes A) P(l,m), \tag{4.43}$$

in which $P(l,m) \in \mathbb{C}^{lm \times lm}$ is the commutation matrix (with respect to l and m as defined in Def. 47). Moreover, if both matrices A and B are real, symmetric, Hermitian, or real and symmetric then the matrix M is a real, symmetric, Hermitian, or real and symmetric, respectively.

Proof. As we shall see, the proof of this proposition will follow immediately from Lemma 45 and Corollary 49 on the Kronecker product of a Schur complement with a matrix, and Proposition 57 on the composition of Schur complements.

Let $A = [A_{ij}]_{i,j=1,2} \in \mathbb{C}^{m \times m}$ and $B = [B_{ij}]_{i,j=1,2} \in \mathbb{C}^{n \times n}$ be 2×2 block matrices such that the matrices A_{22}, B_{22}, A/A_{22}, and B/B_{22} are invertible with $B/B_{22} \in \mathbb{C}^{l \times l}$. Then by Lemma 45 we have

$$C/C_{22} = A/A_{22} \otimes B/B_{22}, \tag{4.44}$$

where $C = A \otimes (B/B_{22}) \in \mathbb{C}^{ml \times ml}$ with the 2×2 block matrix form

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \otimes (B/B_{22}) & A_{12} \otimes (B/B_{22}) \\ A_{21} \otimes (B/B_{22}) & A_{22} \otimes (B/B_{22}) \end{bmatrix} = A \otimes (B/B_{22}) \tag{4.45}$$

with $C_{22} = A_{22} \otimes (B/B_{22})$ invertible. By Corollary 49 we have

$$D/D_{22} = A \otimes (B/B_{22}) = C, \tag{4.46}$$

where $D = Q^T (B \otimes A) Q,\ Q \in \mathbb{C}^{mn \times mn}$ have the 2×2 block matrix forms (4.34) and (4.33), respectively, and $D_{22} = B_{22} \otimes A$ is invertible. We now write this all in another way in order to make it clear how we will use Proposition 57. First, we write

$$D/D_{22} = \begin{bmatrix} (D/D_{22})_{33} & (D/D_{22})_{34} \\ (D/D_{22})_{43} & (D/D_{22})_{44} \end{bmatrix} = \begin{bmatrix} A_{11} \otimes (B/B_{22}) & A_{12} \otimes (B/B_{22}) \\ A_{21} \otimes (B/B_{22}) & A_{22} \otimes (B/B_{22}) \end{bmatrix} \tag{4.47}$$

in which $(D/D_{22})_{44} = A_{22} \otimes (B/B_{22})$ is invertible. Now consider the 2×2 block matrix form $D = [D_{ij}]_{i,j=1,2}$ in (4.34). Then it follows that

$$D_{11} = A \otimes B_{11} = \begin{bmatrix} A_{11} \otimes B_{11} & A_{12} \otimes B_{11} \\ A_{21} \otimes B_{11} & A_{22} \otimes B_{11} \end{bmatrix} \tag{4.48}$$

(where the last equality follows from the first part of Lemma 45) which is conformal to the block structure of D/D_{22} in (4.47). Let us write now this 2×2 block form of D_{11} in (4.48) as

$$D_{11} = \begin{bmatrix} D_{33} & D_{34} \\ D_{43} & D_{44} \end{bmatrix}. \tag{4.49}$$

This yields a subpartitioning of the 2×2 block matrix D from (4.34) into a 3×3 block matrix as

$$D = \left[\begin{array}{c:c} D_{11} & D_{12} \\ \hdashline D_{21} & D_{22} \end{array} \right] = \left[\begin{array}{cc:c} D_{33} & D_{34} & D_{32} \\ D_{43} & D_{44} & D_{42} \\ \hdashline D_{23} & D_{24} & D_{22} \end{array} \right]. \tag{4.50}$$

On the other hand, we can repartition the matrix D in the following 2×2 block partitioned structure $D = M = [M_{ij}]_{i,j=1,2}$:

$$M = \left[\begin{array}{c:c} M_{11} & M_{12} \\ \hdashline M_{21} & M_{22} \end{array} \right] = \left[\begin{array}{c:cc} D_{33} & D_{34} & D_{32} \\ \hdashline D_{43} & D_{44} & D_{42} \\ D_{23} & D_{24} & D_{22} \end{array} \right], \tag{4.51}$$

where, in particular,

$$M_{22} = \begin{bmatrix} D_{44} & D_{42} \\ D_{24} & D_{22} \end{bmatrix}. \tag{4.52}$$

Therefore, Proposition 57 applies here since D_{22} is invertible as is $(D/D_{22})_{44}$, which together with the above implies

$$M/M_{22} = (D/D_{22})/(D/D_{22})_{44} = C/C_{22} = A/A_{22} \otimes B/B_{22}, \tag{4.53}$$

as desired. The remaining part of the proof follows immediately now formula for M in (4.37) and the properties of Q. ■

Remark 52 *In Corollary 49 and Proposition 51, we can write D and M in terms of $A \otimes B$ instead of $B \otimes A$ using the formula:*

$$Q^T (B \otimes A) Q = P^T (A \otimes B) P, \tag{4.54}$$

where P is the permutation matrix with $P = \overline{P}$, $P^T = P^{-1}$ is given by

$$P = P(m,n)Q, \tag{4.55}$$

which follows from Lemma 48 with the commutation matrix $P(m,n)$ since

$$B \otimes A = P(m,n)^T (A \otimes B) P(m,n). \tag{4.56}$$

Example 53 *To illustrate Proposition 51 and the proof, consider the following example which builds off the previous examples 46 and 50 (just as Proposition 51 and its proof builds off both Lemma 45, Corollary 49, and their proofs). Let*

$$A = \left[\begin{array}{c:c} A_{11} & A_{12} \\ \hdashline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{c:c} 0 & 2 \\ \hdashline 2 & 4 \end{array}\right], \quad B = \left[\begin{array}{c:c} B_{11} & B_{12} \\ \hdashline B_{21} & B_{22} \end{array}\right] = \left[\begin{array}{c:c} 2 & 3 \\ \hdashline 3 & 5 \end{array}\right].$$

Then, as calculated in Example 50, the matrix $D = Q^T(B \otimes A)Q$ has the 2×2 block matrix form

$$D = \left[\begin{array}{c:c} D_{11} & D_{12} \\ \hdashline D_{21} & D_{22} \end{array}\right] = \left[\begin{array}{c:c} A \otimes B_{11} & P(1,2)^T(B_{12} \otimes A) \\ \hdashline (B_{21} \otimes A)P(1,2) & B_{22} \otimes A \end{array}\right]$$
$$= \left[\begin{array}{cc:cc} 0 & 4 & 0 & 6 \\ 4 & 8 & 6 & 12 \\ \hdashline 0 & 6 & 0 & 10 \\ 6 & 12 & 10 & 20 \end{array}\right].$$

Now $A_{22} = [4]$, $B_{22} = [5]$, $A/A_{22} = [-1]$, and $B/B_{22} = \left[\frac{1}{5}\right]$ are invertible, which implies by Proposition 51 that $M = D$ is an invertible matrix and it has a 2×2 block form $M = [M_{ij}]_{i,j=1,2}$ such that M_{22} is invertible and $M/M_{22} = A/A_{22} \otimes B/B_{22}$. According to this Proposition 51 and its proof, we form this block structure $M = [M_{ij}]_{i,j=1,2}$ in the following manner: We first partition the matrix $D_{11} = A \otimes B_{11}$ into a 2×2 block matrix $D_{11} = [D_{ij}]_{i,j=3,4}$ as

$$D_{11} = \left[\begin{array}{c:c} D_{33} & D_{34} \\ \hdashline D_{44} & D_{44} \end{array}\right] = \left[\begin{array}{c:c} A_{11} \otimes B_{11} & A_{12} \otimes B_{11} \\ \hdashline A_{21} \otimes B_{11} & A_{22} \otimes B_{11} \end{array}\right] = A \otimes B_{11}$$
$$= \left[\begin{array}{c:c} 0 & 4 \\ \hdashline 4 & 8 \end{array}\right].$$

This yields a subpartitioning of the 2×2 block matrix D into a 3×3 block matrix as

$$D = \left[\begin{array}{c:c} D_{11} & D_{12} \\ \hdashline D_{21} & D_{22} \end{array}\right] = \left[\begin{array}{cc:c} D_{33} & D_{34} & D_{32} \\ D_{43} & D_{44} & D_{42} \\ \hdashline D_{23} & D_{24} & D_{22} \end{array}\right]$$

$$= \left[\begin{array}{cc:c} \begin{bmatrix}0\end{bmatrix} & \begin{bmatrix}4\end{bmatrix} & \begin{bmatrix}0 & 6\end{bmatrix} \\ \begin{bmatrix}4\end{bmatrix} & \begin{bmatrix}8\end{bmatrix} & \begin{bmatrix}6 & 12\end{bmatrix} \\ \hdashline \begin{bmatrix}0 \\ 6\end{bmatrix} & \begin{bmatrix}6 \\ 12\end{bmatrix} & \begin{bmatrix}0 & 10 \\ 10 & 20\end{bmatrix} \end{array}\right]$$

Next, we repartition $D = M = [M_{ij}]_{i,j=1,2}$ as

$$M = \left[\begin{array}{c:c} M_{11} & M_{12} \\ \hdashline M_{21} & M_{22} \end{array}\right] = \left[\begin{array}{c:cc} D_{33} & D_{34} & D_{32} \\ \hdashline D_{43} & D_{44} & D_{42} \\ D_{23} & D_{24} & D_{22} \end{array}\right] = \left[\begin{array}{c:cc} \begin{bmatrix}0\end{bmatrix} & \begin{bmatrix}4\end{bmatrix} & \begin{bmatrix}0 & 6\end{bmatrix} \\ \hdashline \begin{bmatrix}4\end{bmatrix} & \begin{bmatrix}8\end{bmatrix} & \begin{bmatrix}6 & 12\end{bmatrix} \\ \begin{bmatrix}0 \\ 6\end{bmatrix} & \begin{bmatrix}6 \\ 12\end{bmatrix} & \begin{bmatrix}0 & 10 \\ 10 & 12\end{bmatrix} \end{array}\right]$$

$$= \left[\begin{array}{c:ccc} 0 & 4 & 0 & 6 \\ \hdashline 4 & 8 & 6 & 12 \\ 0 & 6 & 0 & 10 \\ 6 & 12 & 10 & 20 \end{array}\right] = D.$$

Then by direct calculation we find that

$$M/M_{22} = M_{11} - M_{12}M_{22}^{-1}M_{21} = \begin{bmatrix}0\end{bmatrix} - \begin{bmatrix}4 & 0 & 6\end{bmatrix} \begin{bmatrix}8 & 6 & 12 \\ 6 & 0 & 10 \\ 12 & 10 & 20\end{bmatrix}^{-1} \begin{bmatrix}4 \\ 0 \\ 6\end{bmatrix} = \begin{bmatrix}-\frac{1}{5}\end{bmatrix},$$

$$A/A_{22} \otimes B/B_{22} = \begin{bmatrix}-1\end{bmatrix} \otimes \begin{bmatrix}\frac{1}{5}\end{bmatrix} = \begin{bmatrix}-\frac{1}{5}\end{bmatrix}.$$

We will now conclude this example by considering Remark 52. From this remark we know that

$$D = M = Q^T(B \otimes A)Q = P^T(A \otimes B)P,$$

where in this example we have $m = n = 2$, $\{E_{ij} : i = 1, 2, j = 1, 2\}$ is the standard basis for $C^{2\times 2}$ (as defined in Def. 47), and

$$P = P(m,n)Q = P(2,2)I_4 = P(2,2) = [E_{ij}^T]_{i,j=1}^{2,2} = \left[\begin{array}{c:c} E_{11}^T & E_{12}^T \\ \hdashline E_{21}^T & E_{22}^T \end{array}\right]$$

$$= \left[\begin{array}{cc:cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hdashline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

By a direct calculation we find that

$$P^T(A\otimes B)P = \begin{bmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}0&0&4&6\\0&0&6&10\\4&6&8&12\\6&10&12&20\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{bmatrix}$$
$$= \begin{bmatrix}0&4&0&6\\4&8&6&12\\0&6&0&10\\6&12&10&20\end{bmatrix} = Q^T(B\otimes A)Q = D = M.$$

.

The following lemma is well-known (see, for instance, [72PP, p. 123, Theorem 3] as proved in [66KP, p. 57, Theorem 1.1′], cf. [66KP, pp. 48-49, Theorem 1.1] and its proof).

Lemma 54 (Rank factorization) *If $\mathbb{C}$ is a field and $A\in\mathbb{C}^{m\times m}$ is a symmetric matrix with rank $r\geq 1$ then there exists an invertible matrix $Y\in\mathbb{C}^{m\times m}$ and a invertible symmetric matrix $B\in\mathbb{C}^{r\times r}$ such that*

$$B = Y^T\begin{bmatrix}F&0\\0&0_{m-r}\end{bmatrix}Y, \tag{4.57}$$

where the zero matrices bordering F are absent if $r=m$.

Proposition 55 (Scalar product of a Schur complement) *If $B\in\mathbb{C}^{n\times n}$ and $A\in\mathbb{C}^{m\times m}$ is a 2×2 block matrix*

$$A = \begin{bmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{bmatrix},$$

such that A_{22} is invertible and $A/A_{22}\in\mathbb{C}^{1\times 1}$ then there exists a 2×2 block matrix $C=[C_{ij}]_{i,j=1,2}$ with C_{22} invertible such that

$$C/C_{22} = A/A_{22}\otimes B = \det(A/A_{22})B. \tag{4.58}$$

Moreover, if both matrices A and B are real, symmetric, Hermitian, or real and symmetric then the matrix C is a real, symmetric, Hermitian, or real and symmetric, respectively. In addition, if $r=\operatorname{rank}(B)$ and using the factorization of B in Lemma 54, i.e., $B=Y^TFY$, where $Y\in\mathbb{C}^{n\times n}$ is invertible, $F=F_{11}\oplus 0_{n-r}$, and $F_{11}\in\mathbb{C}^{r\times r}$ is *invertible, then we can take the matrix* $C \in \mathbb{C}^{(mr+n-r)\times(mr+n-r)}$ *to be*

$$C = \left[\begin{array}{c:c} C_{11} & C_{12} \\ \hdashline C_{21} & C_{22} \end{array}\right] \tag{4.59}$$

$$= \left[\begin{array}{c:c} Y & 0 \\ \hdashline 0 & I_{(m-1)r} \end{array}\right] \left[\begin{array}{cc:c} A_{11}\otimes D_{11} & 0 & A_{12}\otimes D_{11} \\ 0 & 0_{n-r} & 0 \\ \hdashline A_{21}\otimes D_{11} & 0 & A_{22}\otimes D_{11} \end{array}\right] \left[\begin{array}{c:c} Y & 0 \\ \hdashline 0 & I_{(m-1)r} \end{array}\right]. \tag{4.60}$$

Proof. Suppose $B \in \mathbb{C}^{n\times n}$ and $A \in \mathbb{C}^{m\times m}$ is a 2×2 block matrix $A = [A_{ij}]_{i,j=1,2}$ such that A_{22} is invertible and $A/A_{22} \in \mathbb{C}^{1\times 1}$. Then, $\det(A/A_{22})B$ [i.e., the scalar multiplication of the scalar $\det(A/A_{22})$ with the matrix B] is equal to the Kronecker product of the 1×1 matrix A/A_{22} with the matrix B, that is,

$$A/A_{22} = \left[\det(A/A_{22})\right], \quad A/A_{22}\otimes B = \det(A/A_{22})B. \tag{4.61}$$

Now, we would like to apply Lemma 45 to this Kronecker product $A/A_{22}\otimes B$, but the hypothesis that B is invertible need not be satisfied. On the other hand, by Lemma 54 we have the factorization $B = Y^TFY$, where $Y \in \mathbb{C}^{n\times n}$ is invertible and $F = F_{11}\oplus 0_{n-r}$ is the matrix direct sum of an invertible matrix F_{11} and the $(n-r)\times(n-r)$ zero matrix 0_{n-r} (with no zero matrix present if B is invertible) in which statements (a)-(d) in that lemma are true. Thus, we have

$$\begin{aligned} A/A_{22}\otimes B = \det(A/A_{22})B = Y^T\det(A/A_{22})FY &= Y^T[(\det(A/A_{22})F_{11})\oplus 0_{n-r}]Y \\ &= Y^T[(A/A_{22}\otimes F_{11})\oplus 0_{n-r}]Y. \end{aligned}$$

By Lemma 45,

$$G/G_{22} = A/A_{22}\otimes F_{11}, \tag{4.62}$$

where

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \begin{bmatrix} A_{11}\otimes F_{11} & A_{12}\otimes F_{11} \\ A_{21}\otimes F_{11} & A_{22}\otimes F_{11} \end{bmatrix}. \tag{4.63}$$

By Lemma 32,

$$H/H_{22} = G/G_{22}\oplus 0_{n-r}, \tag{4.64}$$

where

$$H = \left[\begin{array}{c:c} H_{11} & H_{12} \\ \hdashline H_{21} & H_{22} \end{array}\right] = \left[\begin{array}{cc:c} G_{11} & 0 & G_{12} \\ 0 & 0_{n-r} & 0 \\ \hdashline G_{21} & 0 & G_{22} \end{array}\right]. \tag{4.65}$$

By Proposition 39,

$$C/C_{22} = Y^T(H/H_{22})Y \tag{4.66}$$

where

$$C = \begin{bmatrix} Y^T H_{11} Y & Y^T H_{12} \\ H_{21} Y & H_{22} \end{bmatrix} = \begin{bmatrix} Y & 0 \\ 0 & I_{(m-1)r} \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} Y & 0 \\ 0 & I_{(m-1)r} \end{bmatrix}. \tag{4.67}$$

Therefore, putting this all together we have proven

$$\begin{aligned} \det(A/A_{22})B &= A/A_{22} \otimes B \\ &= Y^T[(A/A_{22} \otimes F_{11}) \oplus 0_{n-r}]Y \\ &= Y^T(G/G_{22} \oplus 0_{n-r})Y \\ &= Y^T(H/H_{22})Y = C/C_{22}, \end{aligned}$$

and C has the desired properties. This completes the proof. ■

4.5 Compositions

This section is on compositions of Schur complements. For us, our result on compositions (Proposition 57) represents a fundamental result in this chapter which allow for producing from basic building blocks more complicated ones.

In order to understand the notion of compositions of Schur complements and the proposition that follows, we introduce first some notation. We also have provided a concrete example below (Example 58) that uses the notation and applies Proposition 57.

Definition 56 (The Schur complement function) *Suppose m, k are positive integers such that $m > k$. Then the Schur complement function with respect to the pair (m, k) is the function $f = f_{m,k} : D_{m,k} \to \mathbb{C}^{k\times k}$ defined by*

$$f(A) = f_{m,k}(A) = A/A_{22},$$

whose domain $D_{m,k}$ consists of all matrices $A \in \mathbb{C}^{m\times m}$ with a 2×2 block matrix form

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

such that $A_{22} \in \mathbb{C}^{(m-k)\times(m-k)}$ is invertible.

In this section we are interested in the composition of Schur complement functions, that is, using the definition above, the composition $h = g \circ f$ of the function $g = f_{k,l}$ with $f = f_{m,k}$. More precisely, let l be any positive integer such that $k > l$, then $g = f_{k,l} : D_{k,l} \to \mathbb{C}^{l\times l}$ is the function

$$g(B) = f_{k,l}(B) = B/B_{44},$$

whose domain $D_{k,l}$ consists of all matrices $B \in \mathbb{C}^{k\times k}$ with a 2×2 block matrix form

$$B = \begin{bmatrix} B_{33} & B_{34} \\ B_{43} & B_{44} \end{bmatrix},$$

such that $B_{44} \in \mathbb{C}^{(k-l)\times(k-l)}$ is invertible. Therefore, the composition function $h = g \circ f : \{A \in D_{m,k} : f(A) \in D_{k,l}\} \to \mathbb{C}^{l\times l}$ is defined by

$$h(A) = g(B),\ B = f(A),$$

that is,

$$h(A) = g(f(A)) = (A/A_{22})/(A/A_{22})_{44}.$$

The main question we address in this section is whether or not $h(A)$ is a Schur complement, i.e., for each $A \in D_{m,k}$ with $f(A) \in D_{k,l}$, does there exist a 2×2 block matrix $C = [C_{ij}]_{i,j=1,2}$ with C_{22} invertible such that $h(A) = C/C_{22}$? The next proposition tells us that the answer is yes and gives a formula for this matrix C in terms of A.

In order to state the next proposition and give a proof, we need to give some notation first. Begin by partitioning the $k\times k$ matrix $A_{11} = [A_{ij}]_{i,j=3,4}$ conformal to the block structure of the $k\times k$ matrix $A/A_{22} = B = [B_{ij}]_{i,j=3,4}$ so that

$$A_{11} = \begin{bmatrix} A_{33} & A_{34} \\ A_{43} & A_{44} \end{bmatrix}. \tag{4.68}$$

This yields a subpartitioning of the matrix A into a 3×3 block matrix as

$$A = \left[\begin{array}{c:c} A_{11} & A_{12} \\ \hdashline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{cc:c} A_{33} & A_{34} & A_{32} \\ A_{43} & A_{44} & A_{42} \\ \hdashline A_{23} & A_{24} & A_{22} \end{array}\right]. \tag{4.69}$$

On the other hand, we can repartition the matrix A in the following 2×2 block partitioned structure $A = C = [C_{ij}]_{i,j=1,2}$:

$$A = C = \left[\begin{array}{c:c} C_{11} & C_{12} \\ \hdashline C_{21} & C_{22} \end{array}\right] = \left[\begin{array}{c:cc} A_{33} & A_{34} & A_{32} \\ \hdashline A_{43} & A_{44} & A_{42} \\ A_{23} & A_{24} & A_{22} \end{array}\right], \tag{4.70}$$

where, in particular,

$$C_{22} = \begin{bmatrix} A_{44} & A_{42} \\ A_{24} & A_{22} \end{bmatrix}. \tag{4.71}$$

Our question is then answered with the following proposition since it tells us that C_{22} is invertible and

$$C/C_{22} = h(A) = g(f(A)).$$

Proposition 57 (Composition of Schur complements) *If $A \in \mathbb{C}^{m\times m}$ is a 2×2 block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

such that A_{22} is invertible and suppose that $A/A_{22} \in \mathbb{C}^{k\times k}$ is also a 2×2 block matrix

$$A/A_{22} = \begin{bmatrix} (A/A_{22})_{33} & (A/A_{22})_{34} \\ (A/A_{22})_{43} & (A/A_{22})_{44} \end{bmatrix},$$

such that $(A/A_{22})_{44} \in \mathbb{C}^{l\times l}$ is invertible then

$$C/C_{22} = (A/A_{22})/(A/A_{22})_{44},$$

where $A = C = [C_{ij}]_{i,j=1,2}$ is the 2×2 block matrix in (4.70) [defined in terms of the subpartitioning of A_{11} and A in (4.68) and (4.69)] with C_{22} [in (4.71)] invertible and

$$(A/A_{22})_{44} = C_{22}/A_{22}.$$

Proof. This proposition is essentially just the well-known Crabtree-Haynsworth quotient formula for Schur complements [05FZ, Theorem 1.2 (Quotient Formula), p. 25]. First, using the 2×2 block matrix form (4.68) for A_{11} and the 3×3 block matrix form for A in (4.69) we find that

$$\begin{aligned} A/A_{22} &= A_{11} - A_{12}A_{22}^{-1}A_{21} \\ &= \begin{bmatrix} A_{33} & A_{34} \\ A_{43} & A_{44} \end{bmatrix} - \begin{bmatrix} A_{32} \\ A_{42} \end{bmatrix} A_{22}^{-1} \begin{bmatrix} A_{23} & A_{24} \end{bmatrix} \\ &= \begin{bmatrix} A_{33} - A_{32}A_{22}^{-1}A_{23} & A_{34} - A_{32}A_{22}^{-1}A_{24} \\ A_{43} - A_{42}A_{22}^{-1}A_{23} & A_{44} - A_{42}A_{22}^{-1}A_{24} \end{bmatrix}. \end{aligned} \tag{4.72}$$

From this, it follows that

$$(A/A_{22})_{44} = A_{44} - A_{42}A_{22}^{-1}A_{24} = C_{22}/A_{22},$$

where C_{22} is the 2×2 block matrix in (4.71). Now by hypotheses, both A_{22} and $(A/A_{22})_{44}$ are invertible, and we just proved $C_{22}/A_{22} = (A/A_{22})_{44}$ so that by Lemma 41 it follows that C_{22} is also invertible and using inverse formula in (4.20) for C_{22}^{-1} we have

$$\begin{aligned} C_{22}^{-1} &= \begin{bmatrix} A_{44} & A_{42} \\ A_{24} & A_{22} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} ((A/A_{22})_{44})^{-1} & -((A/A_{22})_{44})^{-1}A_{42}A_{22}^{-1} \\ -A_{22}^{-1}A_{24}((A/A_{22})_{44})^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{24}((A/A_{22})_{44})^{-1}A_{42}A_{22}^{-1} \end{bmatrix}. \end{aligned} \tag{4.73}$$

Thus, by block multiplication, it follows from the formulas (4.72) and (4.73) that

$$C/C_{22} = A_{33} - \begin{bmatrix} A_{34} & A_{32} \end{bmatrix} \begin{bmatrix} A_{44} & A_{42} \\ A_{24} & A_{22} \end{bmatrix}^{-1} \begin{bmatrix} A_{43} \\ A_{23} \end{bmatrix}$$

$$= A_{33} - \begin{bmatrix} A_{34} & A_{32} \end{bmatrix} \begin{bmatrix} ((A/A_{22})_{44})^{-1} & -((A/A_{22})_{44})^{-1}A_{42}A_{22}^{-1} \\ -A_{22}^{-1}A_{24}((A/A_{22})_{44})^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{24}((A/A_{22})_{44})^{-1}A_{42}A_{22}^{-1} \end{bmatrix} \begin{bmatrix} A_{43} \\ A_{23} \end{bmatrix}$$

$$= A_{33} - [A_{34}((A/A_{22})_{44})^{-1} - A_{32}A_{22}^{-1}A_{24}((A/A_{22})_{44})^{-1}]A_{43}$$

$$-[-A_{34}((A/A_{22})_{44})^{-1}A_{42}A_{22}^{-1} + A_{32}(A_{22}^{-1} + A_{22}^{-1}A_{24}((A/A_{22})_{44})^{-1}A_{42}A_{22}^{-1})]A_{23}$$

$$= A_{33} - A_{34}((A/A_{22})_{44})^{-1}A_{43} + A_{32}A_{22}^{-1}A_{24}((A/A_{22})_{44})^{-1}A_{43}$$

$$+A_{34}((A/A_{22})_{44})^{-1}A_{42}A_{22}^{-1}A_{23} - A_{32}A_{22}^{-1}A_{23} - A_{32}A_{22}^{-1}A_{24}((A/A_{22})_{44})^{-1}A_{42}A_{22}^{-1}A_{23}$$

$$= A_{33} - A_{32}A_{22}^{-1}A_{23} - (A_{34} - A_{32}A_{22}^{-1}A_{24})((A/A_{22})_{44})^{-1}A_{43}$$

$$+(A_{34} - A_{32}A_{22}^{-1}A_{24})((A/A_{22})_{44})^{-1}A_{42}A_{22}^{-1}A_{23}$$

$$= (A_{33} - A_{32}A_{22}^{-1}A_{23}) - (A_{34} - A_{32}A_{22}^{-1}A_{24})((A/A_{22})_{44})^{-1}(A_{43} - A_{42}A_{22}^{-1}A_{23})$$

$$= (A/A_{22})_{33} - (A/A_{22})_{34}(A/A_{22})_{44}^{-1}(A/A_{22})_{43}$$

$$= (A/A_{22})/(A/A_{22})_{44}.$$

This completes the proof. ■

Example 58 *We will now work out a concrete example to to demonstrate the notation and Proposition 57 above. Consider the following* 2×2 *block matrix* $A = [A_{ij}]_{i,j=1,2}$ *and its Schur complement* A/A_{22}*,*

$$A = \left[\begin{array}{c:c} A_{11} & A_{12} \\ \hdashline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{cc:cc} 4 & 3 & 1 & 1 \\ 4 & 2 & 2 & 1 \\ \hdashline 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 \end{array}\right], \quad A/A_{22} = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}.$$

Suppose now we interested in taking the Schur complement of A/A_{22} *with respect to its lower left corner block (i.e., the invertible matrix* $[1] \in \mathbb{C}^{1\times 1}$*). Then we block partition* A/A_{22} *into the* 2×2 *block matrix* $A/A_{22} = [(A/A_{22})_{ij}]_{i,j=3,4}$ *and compute the desired Schur complement* $(A/A_{22})/(A/A_{22})_{44}$ *as*

$$A/A_{22} = \left[\begin{array}{c:c} (A/A_{22})_{33} & (A/A_{22})_{34} \\ \hdashline (A/A_{22})_{43} & (A/A_{22})_{44} \end{array}\right] = \left[\begin{array}{c:c} 3 & 2 \\ \hdashline 2 & 1 \end{array}\right],$$

$$(A/A_{22})/(A/A_{22})_{44} = [3] - [2][1]^{-1}[2] = [-1].$$

According to Proposition 57,

$$C/C_{22} = (A/A_{22})/(A/A_{22})_{44},$$

where $A = C = [C_{ij}]_{i,j=1,2}]$ *is the* 2×2 *block matrix in (4.70) [defined in terms of the subpartitioning of* A_{11} *and* A *in (4.68) and (4.69)] with* C_{22} *[in (4.71)] invertible* *and* $(A/A_{22})_{44} = C_{22}/A_{22}$. *Lets work this all out explicitly now for this example. We begin by partitioning* $A_{11} = [A_{ij}]_{i,j=3,4}$ *conformal to the block structure of* $A/A_{22} = [(A/A_{22})_{ij}]_{i,j=3,4}$ *so that*

$$A_{11} = \left[\begin{array}{c:c} A_{33} & A_{34} \\ \hdashline A_{43} & A_{44} \end{array}\right] = \left[\begin{array}{c:c} 4 & 3 \\ \hdashline 4 & 2 \end{array}\right]$$

This yields a subpartitioning of the matrix A *into the* 3×3 *block matrix as*

$$A = \left[\begin{array}{c:c} A_{11} & A_{12} \\ \hdashline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{cc:c} A_{33} & A_{34} & A_{32} \\ A_{43} & A_{44} & A_{42} \\ \hdashline A_{23} & A_{24} & A_{22} \end{array}\right] = \left[\begin{array}{cc:cc} 4 & 3 & 1 & 1 \\ 4 & 2 & 2 & 1 \\ \hdashline 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 \end{array}\right]$$

and from this we repartition A *to get the* 2×2 *block matrix*

$$A = C = \left[\begin{array}{c:c} C_{11} & C_{12} \\ \hdashline C_{21} & C_{22} \end{array}\right] = \left[\begin{array}{c:cc} A_{33} & A_{34} & A_{32} \\ \hdashline A_{43} & A_{44} & A_{42} \\ A_{23} & A_{24} & A_{22} \end{array}\right] = \left[\begin{array}{c:ccc} 4 & 3 & 1 & 1 \\ \hdashline 4 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 \end{array}\right].$$

We now verify that

$$\begin{aligned} C/C_{22} &= [4] - \begin{bmatrix} 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} = [4] - [5] \\ &= [-1] = (A/A_{22})/(A/A_{22})_{44}, \\ C_{22}/A_{22} &= \left[\begin{array}{c:cc} 2 & 2 & 1 \\ \hdashline 1 & 1 & 1 \\ 1 & 2 & 1 \end{array}\right] \Big/ \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} = [2] - \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = [2] - [1] \\ &= [1] = (A/A_{22})_{44}. \end{aligned}$$

4.6 Transforms

In this section we discuss transformations of matrices associated with the Schur complement. Our main focus will be on the *principal pivot transform* (PPT), see Definition 60 below, which can be considered as a matrix partial inverse. We begin by introducing the reader to PPT in the context of network synthesis problems. After this we will give our results.

The PPT was introduced by A. W. Tucker in [60AT] (see also [00MT]) as "an attempt to study for a general field, rather than an ordered field, the linear algebraic structure underlying the 'simplex method' of G. B. Dantzig, so remarkably effective in Linear Programming."

Later, R. J. Duffin, D. Hazony, and N. Morrison in [65DHM, 66DHM] studied the PPT for the purposes of solving certain network synthesis problems, although in the latter its called the *gyration* and they denote it by Γ. More generally, if you compare our Definition 60 of $\mathrm{ppt}_1(A)$ below to the definition of the r-fold gyration $\Gamma_{1,\ldots,r}(A)$ in [65DHM, Sec. 3.2, pp. 54-55, especially (11)] of a 2×2 block matrix $A=[A_{ij}]_{i,j=1,2}\in\mathbb{C}^{n\times n}$ with $A_{11}\in\mathbb{C}^{r\times r}$ invertible, you will see that $\mathrm{ppt}_1(A)=\Gamma_{1,\ldots,r}(A)$. As quoted in [66DHM, Sec. 1.1, p. 1], if $\Gamma(A)=B$ for two matrices A and B that "This relationship is sufficient to make the matrices A and B combinatorially equivalent," a term they say was coined by A. W. Tucker in [60AT] and "The impedance, admittance, chain, and hybrid matrices of network theory are all combinatorially equivalent. The work of A. W. Tucker emerged from the linear programming field and is applied here to network theory." They elaborate on this in [66DHM, Sec. 1.3, p. 394] by saying, "We wish to show in this paper that combinatorial equivalence has application in the entirely different field of network synthesis. It is worth noting that ideas similar to combinatorial equivalence have been applied to network algebra problems by Bott and Duffin..." and they cite [53BD, 59RD] (see also [78DM]). They further elaborate on their synthesis procedure in [66DHM, Sec. 2.3, p. 402] saying, "In what follows, we shall use the Γ operator to give a new extension of the Brune synthesis to n-port..." and that R. J. Duffin in [55RD] had showed how such a network Brune-type synthesis could be viewed as a purely algebraic process. The key point though that [66DHM] makes is that "Our extension of the Brune method differs from the above in that it is not necessary at any state to invert a matrix." In this regard, it is not surprising that R. J. Duffin and his colleagues would be interest in such network synthesis problems that don't require certain algebraic operations given his famous result with R. Bott in [49BD] on synthesis of network impedance functions of one-variable without the use of transformers (for R. Bott's perspective on this, see his interview [01AJ]).

The PPT and variations of its form also appear in other contexts such as in the work of M. G. Krein and I. E. Ovcharenko [94KO] on inverse problems for canonical differential equations or in the study of the analytic properties of the Dirichlet-to-Neumann (DtN) map in electromagnetism [16CWM] for layered media, for instance.

The above serves to give perspective and motivate our consideration of the PPT below in connection to electric circuit theory of Chapter 3 (and later in Chapter 5 for the realization problem of this using the Schur complement and the relationship with the PPT).

4.7 Principal pivot transform

There are two forms of the ppt of a 2×2 block matrix $A=[A_{ij}]_{i,j=1,2}\in\mathbb{C}^{m\times m}$ that we will discuss, denoted by ppt_1 and ppt_2, which can be written in terms of the two Schur complements either A/A_{11} or A/A_{22} of A with respect to the $(1,1)$-block A_{11} (if

A_{11} is invertible) or the $(2,2)$-block A_{22} (if A_{22} is invertible), where by definition

$$A/A_{11} = A_{22} - A_{21}A_{11}^{-1}A_{12}, \tag{4.74}$$

$$A/A_{22} = A_{11} - A_{12}A_{22}^{-1}A_{21}. \tag{4.75}$$

In fact, the results of this thesis could have been posed in terms of the first Schur complement A/A_{11} instead of second one A/A_{22}. The relationship between these two versions of the Schur complement is described in the next lemma. This lemma, together with Proposition 39, gives a means to easily transform our results, which are stated in terms of second form of the Schur complements, into similar statements in terms of the first form (or vice versa). As an example of this, compare Proposition 62 to its corollary (Corollary 64) by considering the proof of the latter.

Lemma 59 (Relationship between the two Schur complements) *If $A \in \mathbb{C}^{m\times m}$ is a 2×2 block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

such that $A_{11} \in \mathbb{C}^{k\times k}$ is invertible then

$$B/B_{22} = A/A_{11}, \tag{4.76}$$

where $B = [B_{ij}]_{i,j=1,2} \in \mathbb{C}^{m\times m}$ is the 2×2 block matrix

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{22} & A_{21} \\ A_{12} & A_{11} \end{bmatrix} = U^T AU, \tag{4.77}$$

such that $B_{22} = A_{11}$ is invertible and $U \in \mathbb{C}^{m\times m}$ is the 2×2 block matrix

$$U = \begin{bmatrix} 0 & I_k \\ I_{m-k} & 0 \end{bmatrix}. \tag{4.78}$$

Moreover, if A is real, symmetric, Hermitian, or real and symmetric then the matrix B is real, symmetric, Hermitian, or real and symmetric, respectively.

Proof. The proof is obvious from the definitions of A/A_{11}, B, and U. ■

Definition 60 *The principal pivot transform (PPT) of a matrix $A \in \mathbb{C}^{m\times m}$ in 2×2 block matrix form*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{4.79}$$

with respect to an invertible A_{22} is defined to be the matrix $\operatorname{ppt}_2(A) \in \mathbb{C}^{m\times m}$ with the 2×2 block matrix form

$$\operatorname{ppt}_2(A) = \begin{bmatrix} A/A_{22} & A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21} & A_{22}^{-1} \end{bmatrix}. \tag{4.80}$$

Similarly, the PPT of A with respect to an invertible A_{11} is defined to be the matrix $\operatorname{ppt}_1(A) \in \mathbb{C}^{m\times m}$ *with the 2×2 block matrix form*

$$\operatorname{ppt}_1(A) = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12} \\ A_{21}A_{11}^{-1} & A/A_{11} \end{bmatrix}. \tag{4.81}$$

The relationship between these two versions of the PPT is described in the next lemma.

Lemma 61 (Relationship between the two PPTs) *If $A \in \mathbb{C}^{m\times m}$ is a 2×2 block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{4.82}$$

such that $A_{11} \in \mathbb{C}^{k\times k}$ is invertible then

$$\operatorname{ppt}_1(A) = U \operatorname{ppt}_2(B) U^T, \tag{4.83}$$

where $B, U \in \mathbb{C}^{m\times m}$ are the 2×2 block matrices given in terms of A by (4.77) and (4.78), respectively.

Proof. By the definitions of the two PPTs (i.e., ppt_1 and ppt_2), the definitions of the 2×2 block matrices B, U in (4.77) and (4.78), respectively, in terms of A and the formula (4.76) from Lemma 59 we have

$$\begin{aligned} U \operatorname{ppt}_2(B) U^T &= \begin{bmatrix} 0 & I_k \\ I_{m-k} & 0 \end{bmatrix} \begin{bmatrix} B/B_{22} & B_{12}B_{22}^{-1} \\ -B_{22}^{-1}B_{21} & B_{22}^{-1} \end{bmatrix} \begin{bmatrix} 0 & I_{m-k} \\ I_k & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & I_k \\ I_{m-k} & 0 \end{bmatrix} \begin{bmatrix} A/A_{11} & A_{21}A_{11}^{-1} \\ -A_{11}^{-1}A_{12} & A_{11}^{-1} \end{bmatrix} \begin{bmatrix} 0 & I_{m-k} \\ I_k & 0 \end{bmatrix} \\ &= \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12} \\ A_{21}A_{11}^{-1} & A/A_{11} \end{bmatrix} \\ &= \operatorname{ppt}_1(A), \end{aligned}$$

which proves the lemma. ■

Proposition 62 (Principal pivot transform as a Schur complement) *If $A \in \mathbb{C}^{m\times m}$ is a 2×2 block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{4.84}$$

such that $A_{22} \in \mathbb{C}^{p\times p}$ is invertible and $A_{11} \in \mathbb{C}^{k\times k}$ then

$$C/C_{22} = \operatorname{ppt}_2(A), \tag{4.85}$$

where $C \in \mathbb{C}^{(k+p+p)\times(k+p+p)}$ is the 3×3 block matrix with the following block partititioned *structure* $C = [C_{ij}]_{i,j=1,2}$*:*

$$C = \left[\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right] = \left[\begin{array}{cc|c} A_{11} & 0 & A_{12} \\ 0 & 0_p & I_p \\ \hline A_{21} & -I_p & A_{22} \end{array}\right] \tag{4.86}$$

with $C_{22} = A_{22}$ *invertible. Furthermore,*

$$(JC)/(JC)_{22} = J_{11}\,\mathrm{ppt}_2(A) = \begin{bmatrix} A/A_{22} & A_{12}A_{22}^{-1} \\ A_{22}^{-1}A_{21} & -A_{22}^{-1} \end{bmatrix}, \tag{4.87}$$

$$JC = \left[\begin{array}{c|c} (JC)_{11} & (JC)_{12} \\ \hline (JC)_{21} & (JC)_{22} \end{array}\right] = \left[\begin{array}{c|c} J_{11}C_{11} & J_{11}C_{12} \\ \hline C_{21} & C_{22} \end{array}\right] = \left[\begin{array}{cc|c} A_{11} & 0 & A_{12} \\ 0 & 0_p & -I_p \\ \hline A_{21} & -I_p & A_{22} \end{array}\right], \tag{4.88}$$

where $J_{11} \in \mathbb{C}^{m\times m}$ *and* $J \in \mathbb{C}^{(k+p+p)\times(k+p+p)}$ *are the* 2×2 *block matrices*

$$J_{11} = \begin{bmatrix} I_k & 0 \\ 0 & -I_p \end{bmatrix}, \quad J = \begin{bmatrix} J_{11} & 0 \\ 0 & I_p \end{bmatrix}. \tag{4.89}$$

Moreover, if A *is real, symmetric, Hermitian, or real and symmetric then* $J_{11}\,\mathrm{ppt}_2(A)$ *and* JC *are both real, symmetric, Hermitian, or real and symmetric, respectively.*

Proof. The proof is straightforward via block multiplication. First,

$$\begin{aligned}
C/C_{22} &= C_{11} - C_{12}C_{22}^{-1}C_{21} \\
&= \begin{bmatrix} A_{11} & 0 \\ 0 & 0_p \end{bmatrix} - \begin{bmatrix} A_{12} \\ I_p \end{bmatrix} A_{22}^{-1} \begin{bmatrix} A_{21} & -I_p \end{bmatrix} \\
&= \begin{bmatrix} A_{11} & 0 \\ 0 & 0_p \end{bmatrix} - \begin{bmatrix} A_{12}A_{22}^{-1} \\ A_{22}^{-1} \end{bmatrix} \begin{bmatrix} A_{21} & -I_p \end{bmatrix} \\
&= \begin{bmatrix} A_{11} & 0 \\ 0 & 0_p \end{bmatrix} - \begin{bmatrix} A_{12}A_{22}^{-1}A_{21} & -A_{12}A_{22}^{-1} \\ A_{22}^{-1}A_{21} & -A_{22}^{-1} \end{bmatrix} \\
&= \begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21} & A_{22}^{-1} \end{bmatrix} \\
&= \begin{bmatrix} A/A_{22} & A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21} & A_{22}^{-1} \end{bmatrix} \\
&= \mathrm{ppt}_2(A)
\end{aligned}$$

and

$$J_{11}(C/C_{22}) = J_{11}\,\mathrm{ppt}_2(A) = \begin{bmatrix} A/A_{22} & A_{12}A_{22}^{-1} \\ A_{22}^{-1}A_{21} & -A_{22}^{-1} \end{bmatrix}.$$

By Proposition 39, it follows that

$$M/M_{22} = J_{11}(C/C_{22}),$$

where

$$M = \left[\begin{array}{c:c} M_{11} & M_{12} \\ \hdashline M_{21} & M_{22} \end{array}\right] = \left[\begin{array}{c:c} J_{11}C_{11} & J_{11}C_{12} \\ \hdashline C_{21} & C_{22} \end{array}\right].$$

Finally, by block multiplication, we verify that

$$\left[\begin{array}{c:c} J_{11}C_{11} & J_{11}C_{12} \\ \hdashline C_{21} & C_{22} \end{array}\right] = \left[\begin{array}{cc:c} A_{11} & 0 & A_{12} \\ 0 & 0_p & -I_p \\ \hdashline A_{21} & -I_p & A_{22} \end{array}\right] = JC.$$

The remaining part of the proof follows immediately now formulas (4.87) and (4.88). This completes the proof. ■

Example 63 *We will now work out a concrete example to to demonstrate the notation and Proposition 62 above. Consider the following* 2×2 *block matrix* $A = [A_{ij}]_{i,j=1,2}$*,*

$$A = \left[\begin{array}{c:c} A_{11} & A_{12} \\ \hdashline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{c:c} 1 & 2 \\ \hdashline 3 & 3 \end{array}\right].$$

Then,

$$\begin{aligned} A/A_{22} &= [-1] \\ A_{12}A_{22}^{-1} &= \left[\frac{2}{3}\right] \\ -A_{22}^{-1}A_{21} &= [-1] \\ A_{22}^{-1} &= \left[\frac{1}{3}\right], \end{aligned}$$

which implies by Definition 60

$$\operatorname{ppt}_2(A) = \begin{bmatrix} A/A_{22} & A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21} & A_{22}^{-1} \end{bmatrix} = \begin{bmatrix} -1 & \frac{2}{3} \\ -1 & \frac{1}{3} \end{bmatrix}.$$

Suppose C *is defined by (4.86). Then,*

$$C = \left[\begin{array}{cc:c} 1 & 0 & 2 \\ 0 & 0 & 1 \\ \hdashline 3 & -1 & 3 \end{array}\right],$$

and

$$
\begin{aligned}
C/C_{22} &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} 3 \end{bmatrix}^{-1} \begin{bmatrix} 3 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 2 & -\frac{2}{3} \\ 1 & -\frac{1}{3} \end{bmatrix} \\
&= \begin{bmatrix} -1 & \frac{2}{3} \\ -1 & \frac{1}{3} \end{bmatrix} \\
&= \operatorname{ppt}_2(A).
\end{aligned}
$$

Furthermore, by J_{11} and J defined in (4.89) we have

$$
J_{11} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
$$

Hence,

$$
J_{11} \operatorname{ppt}_2(A) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -1 & \frac{2}{3} \\ -1 & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} -1 & \frac{2}{3} \\ 1 & -\frac{1}{3} \end{bmatrix}
$$

and

$$
JC = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & 1 \\ 3 & -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & -1 \\ 3 & -1 & 3 \end{bmatrix}.
$$

Therefore,

$$
\begin{aligned}
(JC)/(JC)_{22} &= \left[\begin{array}{cc:c} 1 & 0 & 2 \\ 0 & 0 & -1 \\ \hdashline 3 & -1 & 3 \end{array}\right] \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 2 \\ -1 \end{bmatrix} \begin{bmatrix} 3 \end{bmatrix}^{-1} \begin{bmatrix} 3 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 2 \\ -1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 2 & -\frac{2}{3} \\ -1 & \frac{1}{3} \end{bmatrix} \\
&= \begin{bmatrix} -1 & \frac{2}{3} \\ 1 & -\frac{1}{3} \end{bmatrix} \\
&= J_{11}\,\mathrm{ppt}_2(A).
\end{aligned}
$$

Corollary 64 (The other PPT as a Schur complement) *If $A \in \mathbb{C}^{m\times m}$ is a 2×2 block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{4.90}$$

such that $A_{11} \in \mathbb{C}^{k\times k}$ is invertible and $A_{22} \in \mathbb{C}^{p\times p}$ then

$$D/D_{22} = \mathrm{ppt}_1(A), \tag{4.91}$$

where $D \in \mathbb{C}^{(k+p+k)\times(k+p+k)}$ is the 3×3 block matrix with the following block partitioned structure $D = [D_{ij}]_{i,j=1,2}$:

$$D = \left[\begin{array}{c:c} D_{11} & D_{12} \\ \hdashline D_{21} & D_{22} \end{array}\right] = \left[\begin{array}{cc:c} 0_k & 0 & I_k \\ 0 & A_{22} & A_{21} \\ \hdashline -I_k & A_{12} & A_{11} \end{array}\right] \tag{4.92}$$

with $D_{22} = A_{11}$ invertible. Furthermore,

$$(KD)/(KD)_{22} = K_{11}\,\mathrm{ppt}_1(A) = \begin{bmatrix} -A_{11}^{-1} & A_{11}^{-1}A_{12} \\ A_{21}A_{11}^{-1} & A/A_{11} \end{bmatrix}, \tag{4.93}$$

$$KD = \left[\begin{array}{c:c} (KD)_{11} & (KD)_{12} \\ \hdashline (KD)_{21} & (KD)_{22} \end{array}\right] = \left[\begin{array}{c:c} K_{11}D_{11} & K_{11}D_{12} \\ \hdashline D_{21} & D_{22} \end{array}\right] = \left[\begin{array}{cc:c} 0_k & 0 & -I_k \\ 0 & A_{22} & A_{21} \\ \hdashline -I_k & A_{12} & A_{11} \end{array}\right], \tag{4.94}$$

where $K_{11} \in \mathbb{C}^{m\times m}$ *and* $K \in \mathbb{C}^{(k+p+k)\times(k+p+k)}$ *are the* 2×2 *block matrices*

$$K_{11} = \begin{bmatrix} -I_k & 0 \\ 0 & I_p \end{bmatrix}, \quad K = \begin{bmatrix} K_{11} & 0 \\ 0 & I_k \end{bmatrix}. \tag{4.95}$$

Moreover, if A *is real, symmetric, Hermitian, or real and symmetric then* $K_{11}\,\mathrm{ppt}_1(A)$ *and* KD *are both real, symmetric, Hermitian, or real and symmetric, respectively.*

Proof. Although the proof of this corollary could be proved directly by verifying via block matrix methods the statements, our goal here is to give a proof based on the discussion in the introduction of Section 4.7, namely, to prove the corollary using Proposition 39, Lemma 61, and Proposition 62. First, by Lemma 61 we have

$$B = U^T A U, \quad \mathrm{ppt}_1(A) = U\,\mathrm{ppt}_2(B) U^T,$$

where $B, U \in \mathbb{C}^{m\times m}$ have the 2×2 block matrix forms in (4.77) and (4.78), respectively. Next, by Proposition 62 we know that

$$C/C_{22} = \mathrm{ppt}_2(B),$$

where, by the definition of B and the hypotheses that $A_{11} \in \mathbb{C}^{k\times k}$ and $A_{22} \in \mathbb{C}^{p\times p}$, the matrix $C \in \mathbb{C}^{(p+k+k)\times(p+k+k)}$ is the 3×3 block matrix with the following block partitioned structure $C = [C_{ij}]_{i,j=1,2}$:

$$C = \left[\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right] = \left[\begin{array}{cc|c} B_{11} & 0 & B_{12} \\ 0 & 0_k & I_k \\ \hline B_{21} & -I_k & B_{22} \end{array}\right] = \left[\begin{array}{cc|c} A_{22} & 0 & A_{21} \\ 0 & 0_k & I_k \\ \hline A_{12} & -I_k & A_{11} \end{array}\right]$$

with $C_{22} = B_{22} = A_{11}$ invertible. Thus, it follows by this and Proposition 39 that

$$D/D_{22} = UC/C_{22}U^T,$$

where $D \in \mathbb{C}^{(k+p+k)\times(k+p+k)}$ (with $m = k+p$) is the 2×2 block matrix

$$D = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} = \left[\begin{array}{c|c} UC_{11}U^T & UC_{12} \\ \hline C_{21}U^T & C_{22} \end{array}\right] = \left[\begin{array}{cc|c} 0_k & 0 & I_k \\ 0 & A_{22} & A_{21} \\ \hline -I_k & A_{12} & A_{11} \end{array}\right]$$
$$= \left[\begin{array}{c|c} U & 0 \\ \hline 0 & I_k \end{array}\right] \left[\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right] \left[\begin{array}{c|c} U & 0 \\ \hline 0 & I_k \end{array}\right]^T.$$

Furthermore, by Proposition 62 and the relation of B to A, we have

$$\begin{aligned}
(JC)/(JC)_{22} = J_{11}\,\mathrm{ppt}_2(B) &= \begin{bmatrix} B/B_{22} & B_{12}B_{22}^{-1} \\ B_{22}^{-1}B_{21} & -B_{22}^{-1} \end{bmatrix} \\
&= \begin{bmatrix} A/A_{11} & A_{21}A_{11}^{-1} \\ A_{11}^{-1}A_{12} & -A_{11}^{-1} \end{bmatrix}, \\
JC = \left[\begin{array}{c:c} (JC)_{11} & (JC)_{12} \\ \hdashline (JC)_{21} & (JC)_{22} \end{array}\right] &= \left[\begin{array}{c:c} J_{11}C_{11} & J_{11}C_{12} \\ \hdashline C_{21} & C_{22} \end{array}\right] \\
= \left[\begin{array}{cc:c} B_{11} & 0 & B_{12} \\ 0 & 0_k & -I_k \\ \hdashline B_{21} & -I_k & B_{22} \end{array}\right] &= \left[\begin{array}{cc:c} A_{22} & 0 & A_{21} \\ 0 & 0_k & -I_k \\ \hdashline A_{12} & -I_k & A_{11} \end{array}\right],
\end{aligned}$$

where $J_{11} \in \mathbb{C}^{m\times m}$ and $J \in \mathbb{C}^{(p+k+k)\times(p+k+k)}$ are the 2×2 block matrices

$$J_{11} = \begin{bmatrix} I_p & 0 \\ 0 & -I_k \end{bmatrix}, \quad J = \begin{bmatrix} J_{11} & 0 \\ 0 & I_k \end{bmatrix}.$$

Now, using block multiplication, it follows that

$$UJ_{11}U^T = K_{11}, \quad \left[\begin{array}{c:c} U & 0 \\ \hdashline 0 & I_k \end{array}\right] J \left[\begin{array}{c:c} U & 0 \\ \hdashline 0 & I_k \end{array}\right]^T = \begin{bmatrix} K_{11} & 0 \\ 0 & I_k \end{bmatrix} = K,$$

which implies

$$\begin{aligned}
KD &= \left[\begin{array}{c:c} U & 0 \\ \hdashline 0 & I_k \end{array}\right] J \left[\begin{array}{c:c} U & 0 \\ \hdashline 0 & I_k \end{array}\right]^T \left[\begin{array}{c:c} U & 0 \\ \hdashline 0 & I_k \end{array}\right] \left[\begin{array}{c:c} C_{11} & C_{12} \\ \hdashline C_{21} & C_{22} \end{array}\right] \left[\begin{array}{c:c} U & 0 \\ \hdashline 0 & I_k \end{array}\right]^T \\
&= \left[\begin{array}{c:c} U & 0 \\ \hdashline 0 & I_k \end{array}\right] JC \left[\begin{array}{c:c} U & 0 \\ \hdashline 0 & I_k \end{array}\right]^T = \left[\begin{array}{c:c} U(JC)_{11}U^T & U(JC)_{12} \\ \hdashline (JC)_{21}U^T & (JC)_{22} \end{array}\right].
\end{aligned}$$

Also, using block multiplication and the block forms for D and K, it follows that

$$KD = \left[\begin{array}{c:c} K_{11}D_{11} & K_{11}D_{12} \\ \hdashline D_{21} & D_{22} \end{array}\right] = \left[\begin{array}{cc:c} 0_k & 0 & -I_k \\ 0 & A_{22} & A_{21} \\ \hdashline -I_k & A_{12} & A_{11} \end{array}\right].$$

Now by Proposition 39 we know that

$$U(JC)/(JC)_{22}U^T = (KD)/(KD)_{22},$$

where $KD = [(KD)_{ij}]_{i,j=1,2}$ is given the 2×2 block matrix form

$$KD = \left[\begin{array}{c:c} (KD)_{11} & (KD)_{12} \\ \hdashline (KD)_{21} & (KD)_{22} \end{array}\right] = \left[\begin{array}{c:c} U(JC)_{11}U^T & U(JC)_{12} \\ \hdashline (JC)_{21}U^T & (JC)_{22} \end{array}\right].$$

From these facts we conclude that

$$\begin{bmatrix} -A_{11}^{-1} & A_{11}^{-1}A_{12} \\ A_{21}A_{11}^{-1} & A/A_{11} \end{bmatrix} = K_{11}\,\mathrm{ppt}_1(A) = K_{11}U\,\mathrm{ppt}_2(B)U^T$$
$$= UJ_{11}\,\mathrm{ppt}_2(B)U^T = U^T(JC)/(JC)_{22}U = (KD)/(KD)_{22}.$$

Finally, if A is real, symmetric, Hermitian, or real and symmetric then $B = U^TAU$ is real, symmetric, Hermitian, or real and symmetric, respectively, so by Proposition 62 it follows that $J_{11}\,\mathrm{ppt}_2(B)$ and JC are both real, symmetric, Hermitian, or real and symmetric, respectively, which implies $K_{11}\,\mathrm{ppt}_1(A) = UJ_{11}\,\mathrm{ppt}_2(B)U^T$ and $KD = \left[\begin{array}{c:c} U & 0 \\ \hdashline 0 & I_k \end{array}\right] JC \left[\begin{array}{c:c} U & 0 \\ \hdashline 0 & I_k \end{array}\right]^T$ are both real, symmetric, Hermitian, or real and symmetric, respectively. This completes the proof. ■

Example 65 *We will now work out a concrete example to demonstrate the notation and Proposition 62 above by direct calculation since the proof displays an alternative method using Section 4.7. Consider the following* 2×2 *block matrix* $A = [A_{ij}]_{i,j=1,2}$ *from the previous example,*

$$A = \left[\begin{array}{c:c} A_{11} & A_{12} \\ \hdashline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{c:c} 1 & 2 \\ \hdashline 3 & 3 \end{array}\right].$$

Then,

$$\begin{aligned} A/A_{11} &= [-3] \\ -A_{11}^{-1}A_{12} &= [-2] \\ A_{21}A_{11}^{-1} &= [3] \\ A_{11}^{-1} &= [1], \end{aligned}$$

which implies by Definition 60

$$\mathrm{ppt}_1(A) = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12} \\ A_{21}A_{11}^{-1} & A/A_{11} \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 3 & -3 \end{bmatrix}.$$

Suppose D *is defined by (4.92). Then,*

$$D = \left[\begin{array}{cc:c} 0 & 0 & 1 \\ 0 & 3 & 3 \\ \hdashline -1 & 2 & 1 \end{array}\right],$$

and

$$\begin{aligned} D/D_{22} &= \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix}^{-1} \begin{bmatrix} -1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} -1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} -1 & 2 \\ -3 & 6 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -2 \\ 3 & -3 \end{bmatrix} \\ &= \mathrm{ppt}_1(A). \end{aligned}$$

Furthermore, by K_{11} and K defined in (4.95) we have

$$K_{11} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad k = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Hence,

$$K_{11}\,\mathrm{ppt}_1(A) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 3 & -3 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 3 & -3 \end{bmatrix}$$

and

$$KD = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 3 & 3 \\ -1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 3 & 3 \\ -1 & 2 & 1 \end{bmatrix}.$$

Therefore,

$$
\begin{aligned}
(KD)/(KD)_{22} &= \left[\begin{array}{cc:c} 0 & 0 & -1 \\ 0 & 3 & 3 \\ \hdashline -1 & 2 & 1 \end{array}\right] \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} -1 \\ 3 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix}^{-1} \begin{bmatrix} -1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} -1 \\ 3 \end{bmatrix} \begin{bmatrix} -1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} 1 & -2 \\ -3 & 6 \end{bmatrix} \\
&= \begin{bmatrix} -1 & 2 \\ 3 & -3 \end{bmatrix} \\
&= K_{11}\,\mathrm{ppt}_1(A).
\end{aligned}
$$

Chapter 5

Extension of the Bessmertnyĭ Realization Theorem for Rational Functions of Several Complex Variables

In this chapter, we prove a realization theorem (see Theorem 77) for rational functions of several complex variables which extends an important theorem originally due to M. Bessmertnyĭ (see [02MB, Theorem 1.1]) that was proved in his 1982 Ph.D. thesis [82MB] in Russian. In contrast to Bessmertnyĭ's approach of solving large systems of linear equations (essentially based on Proposition 3 from Chapter 1), we use an operator theoretical approach based on the theory of Schur complements. This leads to a simpler and more "natural" construction to solving the realization problem as we need only apply elementary algebraic operations to Schur complements such as sums, products, inverses, and compositions from Chapter 4. As such, our approach leads to a solution of the realization problem that not only has a greater potential for further extensions and applications, but is also more natural within multidimensional systems theory, for instance, especially for those linear models associated with electric circuits, networks, and composites (some examples of this are discussed in Chapter 3).

5.1 Introduction

In this chapter, we are interested in giving an alternative construction (in comparison to the approach of M. Bessmertnyĭ [02MB] based on Proposition 3 or that which can be achieved using the methods in D. Alpay and C. Dubi [03AD] based on Gleason's problem) that solves the following version of the *Bessmertnyĭ realization problem*: Given a rational $\mathbb{C}^{k\times k}$-valued matrix function $f(z) = f(z_1, \ldots, z_n)$ of n complex variables $z_1, \ldots, z_n$ $[z = (z_1, \ldots, z_n)]$, find a linear matrix pencil $A(z) = A_0 + z_1A_1 +$

$\cdots + z_nA_n$ (see Definition 1) such that $f(z)$ is representable as the Schur complement $f(z) = A(z)/A_{22}(z)$ [for this notation recall (2.4)] of a 2×2 block matrix form $A(z) = [A_{ij}(z)]_{i,j=1,2}$ with respect to its $(2, 2)$-block $A_{22}(z)$. If this is possible, we say that $f(z)$ is *realizable* (or can be *realized*) and has a *realization* (see Definition 66).

To be more precise, we state the following definition:

Definition 66 (Bessmertnyĭ Realizable) *A rational $\mathbb{C}^{k\times k}$-valued matrix function $f(z) = f(z_1, \ldots, z_n)$ of n complex variables $z_1, \ldots, z_n$ [$z = (z_1, \ldots, z_n)$] is* ***Bessmertnyĭ realizable*** *if there exists a linear matrix pencil $A(z)$ and a partitioning of the matrix $A(z) = [A_{ij}(z)]_{i,j=1.2}$ into a 2×2 block matrix form such that $A_{22}(z)$ invertible as an element of $\mathbb{C}(z)^{(m-k)\times(m-k)}$ and $f(z)$ is the Schur complement of $A(z)$ with respect to $A_{22}(z)$, that is,*

$$f(z) = A(z)/A_{22}(z). \tag{5.1}$$

If, in addition, $A(z)$ is a linear symmetric matrix pencil then $f(z)$ is said to be ***symmetric Bessmertnyĭ realizable****.*

The main theorem of M. Bessmertnyĭ in [02MB, Theorem 1.1] solves this realization problem over the field $\mathbb{C}$, and his construction of the linear matrix pencil $A(z) = A_0 + z_1A_1 + \cdots + z_nA_n$ from $f(z)$ involves solving large systems of constrained linear equations in such a way that $A(z)$ inherits certain real, symmetric, or homogeneity properties from $f(z)$.

The main theorem of this chapter, namely, Theorem 77, also solves this realization problem, but our construction of the linear matrix pencil $A(z) = A_0 + z_1A_1 + \cdots + z_nA_n$ from $f(z)$ uses the theory of Schur complements that we develop in Chapter 4. In that chapter, we showed that for each elementary algebraic-functional operation (which we need for solving the realization problem) when applied to Schur complements (such as linear combinations, matrix products, inversion, and composition) is equal to another Schur complement of a block matrix and we give explicit formulas to compute it. The relationship between this chapter and Chapter 4 is quite evident, for instance, we can use Lemma 26 and Proposition 30 to prove Lemma 68 from Lemma 67 on realizing squares. In doing so, we now approach the realizability problem from a linear systems theory perspective, which is more "natural" and has the potential to allow further extensions of Bessmertnyĭ's Theorem [02MB, Theorem 1.1] [and other applications, for instance, see [21SW] and the contents of Chapter 6]. This "natural" approach is especially evident in our proof of Proposition 74 which uses Lemma 69, Lemma 71, and other basic results above (see Figure 5.1).

Our main theorem (i.e., Theorem 77) also provides an extension of M. Bessmertnyĭ's theorem [02MB, Theorem 1.1], namely, our extension is part (d) of Theorem 77 that allows us to construct a Hermitian matrix pencil $A(z) = A_0 + z_1A_1 + \ldots + z_nA_n$ (i.e., all the matrices A_j are Hermitian matrices) if $f(z)$ is "Hermitian" as well, [i.e., has the functional property $f(\overline{z})^* = f(z)$].

For simplicity sake, in the statement of Bessmertnyĭ Realization Theorem (Theorem 77) below, we will abuse this notation (just as M. Bessmertnyĭ does, see [02MB, p. 170, last para. in Sec. 1]) and treat a linear matrix pencil $A(z) = A_{11}(z) = A_0 + z_1 A_1 + \cdots + z_n A_n$ as a degenerate case of Schur complement $A(z) = A(z)/A_{22}(z)$, i.e., the matrix $A_{22}(z)$ is a 0×0 matrix, in which case we can ignore the statement $\det A_{22}(z) \not\equiv 0$. But throughout the rest of the this chapter, we will not abuse this notation of the Schur complement in order to avoid confusion and because the resulting statements we want to prove often are quite different in their proofs in the degenerate case vs the non-degenerate case (as can be seen in Chapter 4).

5.2 Elementary Bessmertnyĭ Realizations

The fundamental algebra and operations of the Schur complement which we will need for realizations have already been mostly developed in Chapter 4. Essentially all we need to do now is develop the realization theory for two of the most elementary types of realizations: a square and a simple product. Both are needed to prove Proposition 74 (namely, Realization of the Kronecker product of realizations) which is the underlying key operation in our construction of the proof to the Bessmertnyĭ Realization Theorem (Theorem 77).

Lemma 67 (Realization of squares) *The $\mathbb{C}^{1\times 1}$-valued function $f(z_1) = \begin{bmatrix} z_1^2 \end{bmatrix}$ of an independent variable z_1 has a symmetric Bessmertnyĭ realization, i.e.,*

$$f(z_1) = [z_1^2] = A(z_1)/A_{22}(z_1), \tag{5.2}$$

with linear matrix pencil

$$A(z_1) = A_0 + z_1 A_1 = \begin{bmatrix} A_{11}(z_1) & A_{12}(z_1) \\ A_{21}(z_1) & A_{22}(z_1) \end{bmatrix}, \tag{5.3}$$

where

$$A_0 = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{5.4}$$

and

$$\left[\begin{array}{c:c} A_{11}(z_1) & A_{12}(z_1) \\ \hdashline A_{21}(z_1) & A_{22}(z_1) \end{array}\right] = \left[\begin{array}{c:c} 0 & z_1 \\ \hdashline z_1 & -1 \end{array}\right] \tag{5.5}$$

such that

$$\det A_{22}(z_1) = -1 \not\equiv 0. \tag{5.6}$$

Moreover, the matrices A_j are real and symmetric, i.e., $A_j = \overline{A_j} = A_j^T$ for $j = 0, 1$.

Proof. The proof follows immediately from the definition of $A(z_1), A_0, A_1$ and the calculation

$$A(z_1)/A_{22}(z_1) = [0] - [z_1][-1]^{-1}[z_1] = -[z_1][-1][z_1] = -[-z_1^2] = [z_1^2].$$

■

Lemma 68 (Realization of the product of two independent variables) *The* $\mathbb{C}^{1\times 1}$*-valued function,* $f(z_1, z_2) = \begin{bmatrix} z_1 z_2 \end{bmatrix}$ *of two independent variables* z_1 *and* z_2 *has a symmetric Bessmertnyi realization, i.e.,*

$$f(z_1, z_2) = [z_1 z_2] = A(z_1, z_2)/A_{22}(z_1, z_2), \tag{5.7}$$

with linear matrix pencil

$$A(z_1, z_2) = A_0 + z_1 A_1 + z_2 A_2 = \begin{bmatrix} A_{11}(z_1, z_2) & A_{12}(z_1, z_2) \\ A_{21}(z_1, z_2) & A_{22}(z_1, z_2) \end{bmatrix}, \tag{5.8}$$

where

$$A_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ A_1 = \begin{bmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 0 & 0 \end{bmatrix},\ A_2 = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix}, \tag{5.9}$$

and

$$\left[\begin{array}{c:c} A_{11}(z_1, z_2) & A_{12}(z_1, z_2) \\ \hdashline A_{21}(z_1, z_2) & A_{22}(z_1, z_2) \end{array}\right] = \left[\begin{array}{c:cc} 0 & \frac{1}{2}(z_1 + z_2) & -\frac{1}{2}(z_1 - z_2) \\ \hdashline \frac{1}{2}(z_1 + z_2) & -1 & 0 \\ -\frac{1}{2}(z_1 - z_2) & 0 & 1 \end{array}\right] \tag{5.10}$$

such that

$$\det A_{22}(z_1, z_2) = -1 \not\equiv 0. \tag{5.11}$$

Moreover, the matrices A_j *are real and symmetric, i.e.,* $A_j = \overline{A_j} = A_j^T$ *for* $j = 0, 1, 2$.

Proof. The statement follows by applying Lemma 67, Lemma 26, and Proposition 30 in succession:

$$\begin{aligned}
[z_1 z_2] &= \left[\left(\frac{1}{2}(z_1 + z_2) \right)^2 + (-1) \left(\frac{1}{2}(z_1 - z_2) \right)^2 \right] \\
&= \left[\begin{array}{c:c} 0 & \frac{1}{2}(z_1 + z_2) \\ \hdashline \frac{1}{2}(z_1 + z_2) & -1 \end{array}\right] \Big/ [-1] + \left[\begin{array}{c:c} 0 & -\frac{1}{2}(z_1 - z_2) \\ \hdashline -\frac{1}{2}(z_1 - z_2) & 1 \end{array}\right] \Big/ [1] \\
&= \left[\begin{array}{c:cc} 0 & \frac{1}{2}(z_1 + z_2) & -\frac{1}{2}(z_1 - z_2) \\ \hdashline \frac{1}{2}(z_1 + z_2) & -1 & 0 \\ -\frac{1}{2}(z_1 - z_2) & 0 & 1 \end{array}\right] \Bigg/ \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.
\end{aligned}$$

The proof follows immediately from this representation. ■

5.2.1 Kronecker products of linear matrix pencils

This section contains the main technical portion of the statements (and their proofs) needed in the proof of the Bessmertnyĭ Realization Theorem (Theorem 77) that pertain to Kronecker products [denoted as $\otimes$, for the definition of a Kronecker product and elementary properties see (2.13)] of linear matrix pencils and realizations.

The main result in this section is Proposition 74. As the proof of it is rather technical (as it builds on many other basic building blocks in this section and the previous sections), we provide a flow diagram for the proof in Figure 5.1.

Lemma 69 (Realization of Kronecker products: Part I) *If $A(z)$ and $B(w)$ are two linear matrix pencils*

$$A(z) = A_0 + \sum_{i=1}^{s} z_i A_i, \quad B(w) = B_0 + \sum_{j=1}^{t} w_j B_j \tag{5.12}$$

where $A_i \in \mathbb{C}^{m\times m}$ (for $i = 0, \ldots, s$) and $B_j \in \mathbb{C}^{n\times n}$ (for $j = 0, \ldots, t$) then there exists a linear matrix pencil $C(z, w)$ in 2×2 block form

$$C(z,w) = C_0 + \sum_{i=1}^{s} z_i C_i + \sum_{j=1}^{t} w_j C_{s+j} = \begin{bmatrix} C_{11}(z,w) & C_{12}(z,w) \\ C_{21}(z,w) & C_{22}(z,w) \end{bmatrix}, \tag{5.13}$$

with $\det C_{22}(z,w) \not\equiv 0$, such that

$$C(z,w)/C_{22}(z,w) = A(z) \otimes B(w). \tag{5.14}$$

Moreover, the following statements are true:

(a) If all the matrices A_i and B_j (for $i = 0, \ldots, s$ and $j = 0, \ldots, t$) are real then one can choose all the matrices C_k (for $k = 0, \ldots, s+t$) to be real.

(b) If all the matrices A_i and B_j (for $i = 0, \ldots, s$ and $j = 0, \ldots, t$) are symmetric then one can choose all the matrices C_k (for $k = 0, \ldots, s+t$) to be symmetric.

(c) If all the matrices A_i and B_j (for $i = 0, \ldots, s$ and $j = 0, \ldots, t$) are Hermitian then one can choose all the matrices C_k (for $k = 0, \ldots, s+t$) to be Hermitian.

(d) If any combination of the (a)-(d) hypotheses are true then the matrices C_k can be chosen to satisfy the same combination of conclusions.

Proof. It follows from Lemma 45 and linearity properties of the Kronecker product $\otimes$

that

$$M(z,w) = A(z) \otimes B(w) \tag{5.15}$$

$$= A_0 \otimes B_0 + \sum_{i=1}^{s} z_i (A_i \otimes B_0) + \sum_{j=1}^{t} w_j (A_0 \otimes B_j) + \sum_{j=1}^{t}\sum_{i=1}^{s} (z_i w_j)(A_i \otimes B_j). \tag{5.16}$$

The first part of the sum for $M(z,w)$, i.e., $A_0 \otimes B_0 + \sum_{i=1}^{s} z_i(A_i \otimes B_0) + \sum_{j=1}^{t} w_j(A_0 \otimes B_j)$ is a linear matrix pencil and is already realized. The second part of the sum, i.e., $\sum_{j=1}^{t}\sum_{i=1}^{s}(z_i w_j)(A_i \otimes B_j)$ is realizable by Proposition 30, Proposition 55, and Lemma 68. Hence, the sum of these two parts, which is $M(z,w)$, is realizable by Lemma 28. This proves that $C(z,w)/C_{22}(z,w) = M(z,w) = A(z) \otimes B(w)$ for some linear matrix pencil $C(z,w)$ in the form (5.13) with $\det C_{22}(z,w) \not\equiv 0$. This completes the first part of the proof. The rest of the proof of statements (a)-(d) follow immediately from these results and the elementary properties of the Kronecker product [(2.14), (2.15), and (2.17)]. This completes the proof. ∎

Example 70 *Let $A(z) = 3 + z$ and $B(w) = 4w$. Then by Lemma 69 calculating $A(z) \otimes B(w)$ we have*

$$\begin{aligned} A_0 \otimes B_0 &= [3] \otimes [0] = [0]; \\ z(A_1 \otimes B_0) &= z([1] \otimes [0]) = z[0] = [0]; \\ w(A_0 \otimes B_1) &= w([3] \otimes [4]) = w[12]; \\ w(A_1 \otimes B_1) &= zw([1] \otimes [4]) = zw[4], \end{aligned}$$

which implies $M(z,w) = A(z) \otimes B(w) = [12w + 4zw] = w[12] + zw[4]$. Suppose we want to find $C(z,w)$ so that $C(z,w)/C_{22}(z,w) = M(z,w) = A(z) \otimes B(w)$. Since $w[12]$ is already a linear pencil, we need to only realize $zw[4]$ and by Lemma 68, Lemma 26, and Lemma 28

$$\begin{aligned} w[12] + zw[4] &= w[12] + 4\left(\left[\begin{array}{c:cc} 0 & \frac{1}{2}(z_1+z_2) & -\frac{1}{2}(z_1-z_2) \\ \hdashline \frac{1}{2}(z_1+z_2) & -1 & 0 \\ -\frac{1}{2}(z_1-z_2) & 0 & 1 \end{array}\right] \Big/ \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.\right) \\ &= w[12] + \left[\begin{array}{c:cc} 0 & 2(z_1+z_2) & -2(z_1-z_2) \\ \hdashline 2(z_1+z_2) & -4 & 0 \\ -2(z_1-z_2) & 0 & 4 \end{array}\right] \Big/ \begin{bmatrix} -4 & 0 \\ 0 & 4 \end{bmatrix} \\ &= \left[\begin{array}{c:cc} 12w & 2(z_1+z_2) & -2(z_1-z_2) \\ \hdashline 2(z_1+z_2) & -4 & 0 \\ -2(z_1-z_2) & 0 & 4 \end{array}\right] \Big/ \begin{bmatrix} -4 & 0 \\ 0 & 4 \end{bmatrix} \\ &= C(z,w), \end{aligned}$$

where $\det C_{22}(z,w) = -16 \not\equiv 0$. *Hence,*

$$\begin{aligned} C(z,w)/C_{22}(z,w) &= 12w - \begin{bmatrix} 2(z_1+z_2) & -2(z_1-z_2) \end{bmatrix} \begin{bmatrix} -4 & 0 \\ 0 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 2(z_1+z_2) \\ -2(z_1-z_2) \end{bmatrix} \\ &= 12w - \begin{bmatrix} 2(z_1+z_2) & -2(z_1-z_2) \end{bmatrix} \begin{bmatrix} -\frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 2(z_1+z_2) \\ -2(z_1-z_2) \end{bmatrix} \\ &= 12w - \begin{bmatrix} -\frac{1}{2}(z_1+z_2) & -\frac{1}{2}(z_1-z_2) \end{bmatrix} \begin{bmatrix} 2(z_1+z_2) \\ -2(z_1-z_2) \end{bmatrix} \\ &= 12w - \left(\left[-(z_1+z_2)^2 + (z_1-z_2)^2 \right] \right) \\ &= [12w - (-4zw)] \\ &= [12w + 4zw] \\ &= w[12] + zw[4] \\ &= M(z,w) \\ &= A(z) \otimes B(w). \end{aligned}$$

Therefore, we have shown that we can construct $C(z,w)$ *using (5.13), such that* $C(z,w)/C_{22}(z,w) = M(z,w) = A(z) \otimes B(w)$ *by Lemma 69.*

Lemma 71 (Realization of Kronecker products: Part II) *If* $B(w)$ *is a linear matrix pencil*

$$B(w) = B_0 + \sum_{j=1}^{t} w_j B_j \tag{5.17}$$

and $A(z)$ *is a linear matrix pencil in* 2×2 *block form*

$$A(z) = A_0 + \sum_{i=1}^{s} z_i A_i = \begin{bmatrix} A_{11}(z) & A_{12}(z) \\ A_{21}(z) & A_{22}(z) \end{bmatrix} \tag{5.18}$$

where $A_i \in \mathbb{F}^{m \times m}$ *(for* $i = 0, \ldots, s$*) and* $B_j \in \mathbb{F}^{n \times n}$ *(for* $j = 0, \ldots, t$*) such that* $\det A_{22}(z) \not\equiv 0$ *and* $\det B(w) \not\equiv 0$*, then there exists a linear matrix pencil* $D(z,w)$ *in* 2×2 *block form*

$$D(z,w) = D_0 + \sum_{i=1}^{s} z_i D_i + \sum_{j=1}^{t} w_j D_{s+j} = \begin{bmatrix} D_{11}(z) & D_{12}(z) \\ D_{21}(z) & D_{22}(z) \end{bmatrix}, \tag{5.19}$$

with $\det D_{22}(z,w) \not\equiv 0$*, such that*

$$D(z,w)/D_{22}(z,w) = A(z)/A_{22}(z) \otimes B(w). \tag{5.20}$$

Moreover, the following statements are true:

(a) *If all the matrices* A_i *and* B_j *(for* $i = 0, \ldots, s$ *and* $j = 0, \ldots, t$*) are real then one can choose all the matrices* D_k *(for* $k = 0, \ldots, s+t$*) to be real.*

(b) *If all the matrices* A_i *and* B_j *(for* $i = 0, \ldots, s$ *and* $j = 0, \ldots, t$*) are symmetric then one can choose all the matrices* D_k *(for* $k = 0, \ldots, s+t$*) to be symmetric.*

(c) *If all the matrices* A_i *and* B_j *(for* $i = 0, \ldots, s$ *and* $j = 0, \ldots, t$*) are Hermitian then one can choose all the matrices* D_k *(for* $k = 0, \ldots, s+t$*) to be Hermitian.*

(d) *If any combination of the (a)-(d) hypotheses are true then the matrices* D_k *can be chosen to satisfy the same combination of conclusions.*

Proof. By Lemma 45 we know that

$$M(z,w)/M_{22}(z,w) = A(z)/A_{22}(z) \otimes B(w), \tag{5.21}$$

where

$$M(z,w) = \begin{bmatrix} M_{11}(z,w) & M_{12}(z,w) \\ M_{21}(z,w) & M_{22}(z,w) \end{bmatrix} \tag{5.22}$$

$$= \begin{bmatrix} A_{11}(z) \otimes B(w) & A_{12}(z) \otimes B(w) \\ A_{21}(z) \otimes B(w) & A_{22}(z) \otimes B(w) \end{bmatrix} = A(z) \otimes B(w) \tag{5.23}$$

and $\det M_{22}(z,w) \not\equiv 0$. By Lemma 69 we know that there exists a linear matrix pencil $C(z,w)$ in the 2×2 block form (5.13) with $\det C_{22}(z,w) \not\equiv 0$ such that

$$C(z,w)/C_{22}(z,w) = A(z) \otimes B(w) = M(z,w). \tag{5.24}$$

It now follows from this that

$$[C(z,w)/C_{22}(z,w)]/[C(z,w)/C_{22}(z,w)]_{22} = M(z,w)/M_{22}(z,w) \tag{5.25}$$

$$= A(z)/A_{22}(z) \otimes B(w), \tag{5.26}$$

and thus by Proposition 57 the statement is proven since by this proposition we can take $D(z,w) = C(z,w)$ [although with a possibly different 2×2 block form described in that proposition in which $\det D_{22}(z,w) \not\equiv 0$]. This proves the first part of the lemma and statements (a)-(d) of this lemma follow immediately from this representation of $D(z,w)$ and Lemma 69. ■

Example 72 *Consider the rational* $\mathbb{C}^{1\times 1}$*-valued function of the two independent variables* z_1, w_1 *defined by*

$$f(z_1, w_1) = \left[\frac{1}{3+3z_1}(9+55w_1)\right].$$

Then this can be written in terms of a Kronecker product as

$$\left[\frac{1}{3+3z_1}\,(9+55w_1)\right] = [3+3z_1]^{-1}\otimes[9+55w_1] = A(z)/A_{22}(z)\otimes B(w),$$

where

$$\begin{aligned}
A(z) = A_0 + z_1 A_1 &= \begin{bmatrix} A_{11}(z) & A_{12}(z) \\ A_{21}(z) & A_{22}(z)\end{bmatrix} = \left[\begin{array}{c:c} 0 & 1 \\ \hdashline 1 & -(3+3z_1)\end{array}\right],\\
A_0 &= \begin{bmatrix} 0 & 1 \\ 1 & -3\end{bmatrix},\ \ A_1 = \begin{bmatrix} 0 & 0 \\ 0 & -3\end{bmatrix},\\
B(w) &= B_0 + w_1 B_1 = [9+55w_1],\ \ B_0 = [9],\ \ B_1 = [55].
\end{aligned}$$

By Lemma 45 we know that

$$M(z,w)/M_{22}(z,w) = A(z)/A_{22}(z)\otimes B(w),$$

where

$$\begin{aligned}
M(z,w) &= \begin{bmatrix} M_{11}(z,w) & M_{12}(z,w) \\ M_{21}(z,w) & M_{22}(z,w)\end{bmatrix} = \begin{bmatrix} A_{11}(z)\otimes B(w) & A_{12}(z)\otimes B(w) \\ A_{21}(z)\otimes B(w) & A_{22}(z)\otimes B(w)\end{bmatrix}\\
&= A(z)\otimes B(w),\\
M_{11}(z,w) &= A_{11}(z)\otimes B(w) = [0],\\
M_{12}(z,w) &= A_{12}(z)\otimes B(w) = [9+55w_1] = A_{21}(z)\otimes B(w) = M_{21}(z,w),\\
M_{22}(z,w) &= A_{22}(z)\otimes B(w) = [-(3+3z_1)(9+55w_1)]
\end{aligned}$$

and

$$\det M_{22}(z,w) = \det[A_{22}(z)\otimes B(w)] = -(3+3z_1)(9+55w_1) \not\equiv 0.$$

By Lemma 69 we know that there exists a linear matrix pencil $C(z,w)$ in the 2×2 block form (5.13) with $\det C_{22}(z,w)\not\equiv 0$ such that

$$C(z,w)/C_{22}(z,w) = A(z)\otimes B(w) = M(z,w). \tag{5.27}$$

Let us now calculate this $C(z,w)$ using the method described in the proof of Lemma 69.

First, we have

$$\begin{aligned}
M(z,w) &= A(z)\otimes B(w)\\
&= A_0\otimes B_0 + z_1(A_1\otimes B_0) + w_1(A_0\otimes B_1) + (z_1w_1)\,(A_1\otimes B_1)\,,\\
A_0\otimes B_0 &= \begin{bmatrix}0 & 1\\ 1 & -3\end{bmatrix}\otimes[9] = \begin{bmatrix}0 & 9\\ 9 & -27\end{bmatrix},\\
A_1\otimes B_0 &= \begin{bmatrix}0 & 0\\ 0 & -3\end{bmatrix}\otimes[9] = \begin{bmatrix}0 & 0\\ 0 & -27\end{bmatrix},\\
A_0\otimes B_1 &= \begin{bmatrix}0 & 1\\ 1 & -3\end{bmatrix}\otimes[55] = \begin{bmatrix}0 & 55\\ 55 & -165\end{bmatrix},\\
A_1\otimes B_1 &= \begin{bmatrix}0 & 0\\ 0 & -3\end{bmatrix}\otimes[55] = \begin{bmatrix}0 & 0\\ 0 & -165\end{bmatrix}.
\end{aligned}$$

The first part of the sum for $M(z,w)$, i.e.,

$$\begin{aligned}
&A_0\otimes B_0 + z_1(A_1\otimes B_0) + w_1(A_0\otimes B_1)\\
&= \begin{bmatrix}0 & 9\\ 9 & -27\end{bmatrix} + z_1\begin{bmatrix}0 & 0\\ 0 & -27\end{bmatrix} + w_1\begin{bmatrix}0 & 55\\ 55 & -165\end{bmatrix}\\
&= \begin{bmatrix}0 & 9+55w_1\\ 9+55w_1 & -27-27z_1-165w_1\end{bmatrix},
\end{aligned}$$

is a linear matrix pencil and is already realized. The second part of the sum, i.e., $(z_1w_1)\,(A_1\otimes B_1)$, is realizable by Lemma 68 and Proposition 55 (in fact, for this example Lemma 32 could be used instead of the latter proposition to speed up the calculation), which we can calculate as

$$(z_1w_1)\,(A_1\otimes B_1) = (z_1w_1)\,(A_1\otimes B_1) = (z_1w_1)\begin{bmatrix}0 & 0\\ 0 & -165\end{bmatrix}$$

$$\left[\begin{array}{c:cc}
0 & \frac{1}{2}(z_1+z_2) & -\frac{1}{2}(z_1-z_2)\\ \hdashline
\frac{1}{2}(z_1+z_2) & -1 & 0\\
-\frac{1}{2}(z_1-z_2) & 0 & 1
\end{array}\right] \Big/ \begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\otimes\begin{bmatrix}0 & 0\\ 0 & -165\end{bmatrix}$$

$$= \left[\begin{array}{cc:cc}
0 & 0 & 0 & 0\\
0 & 0 & \frac{-165}{2}(z_1+w_1) & \frac{165}{2}(z_1-w_1)\\ \hdashline
0 & \frac{-165}{2}(z_1+w_1) & 165 & 0\\
0 & \frac{165}{2}(z_1-w_1) & 0 & -165
\end{array}\right] \Big/ \begin{bmatrix}165 & 0\\ 0 & -165\end{bmatrix}.$$

Hence, the sum of these two parts, which is $M(z,w)$, is realizable by Lemma 28, which *we can calculate as*

$$C(z,w)/C_{22}(z,w) = A(z) \otimes B(w) = M(z,w),$$
$$C(z,w) = \begin{bmatrix} C_{11}(z,w) & C_{12}(z,w) \\ C_{21}(z,w) & C_{22}(z,w) \end{bmatrix}$$
$$= \left[\begin{array}{cc:cc} 0 & 9+55w_1 & 0 & 0 \\ 9+55w_1 & -27-27z_1-165w_1 & \frac{-165}{2}(z_1+w_1) & \frac{165}{2}(z_1-w_1) \\ \hdashline 0 & \frac{-165}{2}(z_1+w_1) & 165 & 0 \\ 0 & \frac{165}{2}(z_1-w_1) & 0 & -165 \end{array}\right] \Bigg/ \begin{bmatrix} 165 & 0 \\ 0 & -165 \end{bmatrix},$$
$$\det C_{22}(z,w) = (165)(-165) \not\equiv 0.$$

It now follows that we have

$$[C(z,w)/C_{22}(z,w)]/[C(z,w)/C_{22}(z,w)]_{22} = A(z)/A_{22}(z) \otimes B(w)$$

and, by Proposition 57,

$$D(z,w)/D_{22}(z,w) = [C(z,w)/C_{22}(z,w)]/[C(z,w)/C_{22}(z,w)]_{22},$$

where

$$D(z,w) = C(z,w),$$

with the 2×2 *block form*

$$D(z,w) = \left[\begin{array}{c:c} D_{11}(z_1,w_1) & D_{12}(z_1,w_1) \\ \hdashline D_{21}(z_1,w_1) & D_{22}(z_1,w_1) \end{array}\right]$$
$$= \left[\begin{array}{c:ccc} 0 & 9+55w_1 & 0 & 0 \\ \hdashline 9+55w_1 & -27-27z_1-165w_1 & \frac{-165}{2}(z_1+w_1) & \frac{165}{2}(z_1-w_1) \\ 0 & \frac{-165}{2}(z_1+w_1) & 165 & 0 \\ 0 & \frac{165}{2}(z_1-w_1) & 0 & -165 \end{array}\right]$$

and

$$\det D_{22}(z_1,w_1) = -(165)^2(3+3z_1)(9+55w_1) \not\equiv 0.$$

Therefore, $f(z,w) = \left[\frac{1}{3+3z_1}(9+55w_1)\right] = D(z,w)/D_{22}(z,w)$ *has the desired Bess-*

mertnyĭ realization with the linear matrix pencil $D(z,w)$ given by

$$D(z,w) = D_0 + z_1 D_1 + w_1 D_2,$$

$$D_0 = \begin{bmatrix} 0 & 9 & 0 & 0 \\ 9 & -27 & 0 & 0 \\ 0 & 0 & 165 & 0 \\ 0 & 0 & 0 & -165 \end{bmatrix},$$

$$D_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -27 & \frac{-165}{2} & \frac{165}{2} \\ 0 & \frac{-165}{2} & 0 & 0 \\ 0 & \frac{165}{2} & 0 & 0 \end{bmatrix},$$

$$D_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -165 & \frac{-165}{2} & \frac{-165}{2} \\ 0 & \frac{-165}{2} & 0 & 0 \\ 0 & \frac{-165}{2} & 0 & 0 \end{bmatrix},$$

in which D_0, D_1, D_2 are all real and symmetric matrices.

Remark 73 *Lemma 69 can also be extended to the realization $A(z) \otimes B(w)/B_{22}(w)$ in a similar fashion with a slight nuance. First we have $A(z) \otimes B(w)/B_{22}(w) = P^T(B(w)/B_{22}(w) \otimes A(z))P$, with permutation matrix P independent of the variables z and w. Then the desired result follows from the lemma above together with Proposition 39. This is similar to Corollary 49 with Remark 52.*

Proposition 74 (Realization of the Kronecker product of realizations) *If $A(z)$ and $B(w)$ are two linear matrix pencils in 2×2 block form*

$$A(z) = A_0 + \sum_{i=1}^{s} z_i A_i = \begin{bmatrix} A_{11}(z) & A_{12}(z) \\ A_{21}(z) & A_{22}(z) \end{bmatrix},$$

$$B(w) = B_0 + \sum_{j=1}^{t} w_j B_j = \begin{bmatrix} B_{11}(z) & B_{12}(z) \\ B_{21}(z) & B_{22}(z) \end{bmatrix},$$

where $A_i \in \mathbb{F}^{m\times m}$ (for $i = 0,\ldots,s$), $B_j \in \mathbb{F}^{n\times n}$ (for $j = 0,\ldots,t$) such that $\det A_{22}(z) \not\equiv 0$, $\det B_{22}(w) \not\equiv 0$, $\det A(z)/A_{22}(z) \not\equiv 0$, and $\det B(w)/B_{22}(w) \not\equiv 0$ then there exists a linear matrix pencil

$$D(z,w) = D_0 + \sum_{i=1}^{s} z_i D_i + \sum_{j=1}^{t} w_j D_{s+j} = \begin{bmatrix} D_{11}(z) & D_{12}(z) \\ D_{21}(z) & D_{22}(z) \end{bmatrix}, \tag{5.28}$$

with $\det D_{22}(z,w) \not\equiv 0$, such that

$$D(z,w)/D_{22}(z,w) = A(z)/A_{22}(z) \otimes B(w)/B_{22}(w). \tag{5.29}$$

Moreover, the following statements are true:

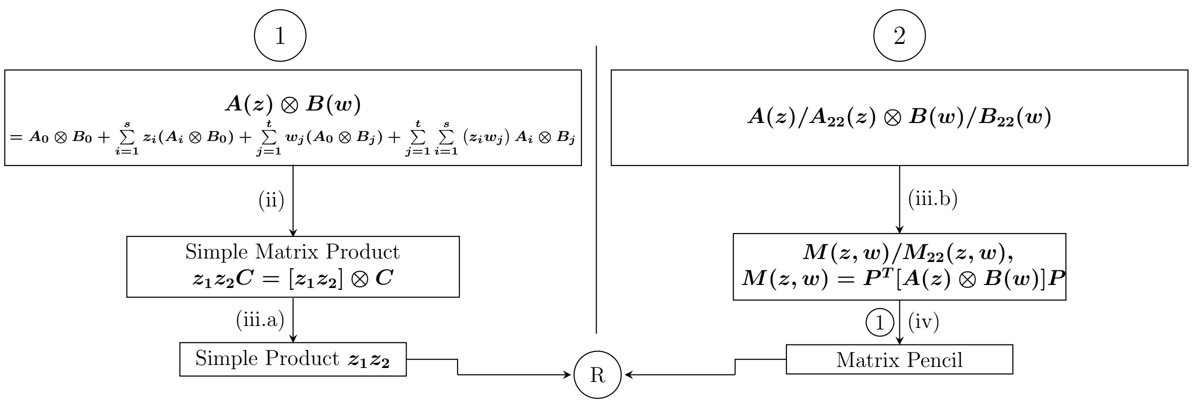
(a) If all the matrices A_i and B_j (for $i = 0, \ldots, s$ and $j = 0, \ldots, t$) are real then one can choose all the matrices D_k (for $k = 0, \ldots, s+t$) to be real.

(b) If all the matrices A_i and B_j (for $i = 0, \ldots, s$ and $j = 0, \ldots, t$) are symmetric then one can choose all the matrices D_k (for $k = 0, \ldots, s+t$) to be symmetric.

(c) If all the matrices A_i and B_j (for $i = 0, \ldots, s$ and $j = 0, \ldots, t$) are Hermitian then one can choose all the matrices D_k (for $k = 0, \ldots, s+t$) to be Hermitian.

(d) If any combination of the (a)-(d) hypotheses are true then the matrices D_k can be chosen to satisfy the same combination of conclusions.

Figure 5.1: Flow diagram for the proof of the realization of the Kronecker product of realizations (see Proposition 74)



(ii) Linear Combinations

(iii.a) Scalar Multiplication (scalar is a Schur complement)

(iii.b) Kronecker Product of Two Schur Complements

(iv) Composition of a Schur Complements

Proof. The proof of this statement is very similar to the proof of Lemma 71, but is slightly more technical because of Proposition 51. By the hypotheses, Proposition 51, and the linearity properties of the Kronecker product $\otimes$, it follows that

$$M(z,w)/M_{22}(z,w) = A(z)/A_{22}(z) \otimes B(w)/B_{22}(w), \tag{5.30}$$

where

$$M(z,w) = Q^T[B(w) \otimes A(z)]Q = P^T[A(z) \otimes B(w)]P \tag{5.31}$$

in which $M = M(z,w) = [M_{ij}(z,w)]_{i,j=1,2}$ is the 2×2 block matrix in terms of $A = A(z) = [A_{ij}(z)]_{i,j=1,2}$ and $B = B(w) = [B_{ij}(w)]_{i,j=1,2}$ in Proposition 51, $Q \in \mathbb{C}^{mn \times mn}$ is the constant (independent of z, w) permutation matrix given by the formula (4.33) satisfying $Q = \overline{Q}, Q^T = Q^{-1}$ (similarly for the permutation matrix P defined in terms of Q in (4.55) discussed in Remark 52), and

$$\det M_{22}(z,w) \not\equiv 0. \tag{5.32}$$

The proof of this proposition now follows immediately from this (in a similar manner as the proof of Lemma 71) by Lemma 69, Proposition 39, and Proposition 57. This completes the proof. ■

We now extend our results to realize an arbitrary monomial z^α.

Proposition 75 (Realizability of Monomial z^α) *An arbitrary monomial $z^\alpha = z_1^{\alpha(1)} \cdots z_n^{\alpha(n)}$, where $\alpha = (\alpha(1), \ldots, \alpha(n)) \in [\mathbb{N} \cup \{0\}]^n$, $n \in \mathbb{N}$, is Bessmertnyĭ realizable by a linear matrix pencil $A(z) = A_0 + z_1 A_1 + \cdots + z_n A_n$ with matrices*

$$A_j = A_j^T = \overline{A_j},\ j = 0, \ldots, n.$$

Proof. From Lemma 68, the result is true for the product $w_1 w_2$ of two independent variables w_1 and w_2. By Proposition 74, the result is true for the product $w_1 w_2 \cdots w_{2m-1} w_{2m}$ for any $m \in \mathbb{N}$. Hence by taking m large enough, changing variables to the z_j's, and possibly setting some of the w_l equal to 1, it follow that the monomial $z^\alpha = z_1^{\alpha(1)} \cdots z_n^{\alpha(n)}$ has the desired Bessmertnyĭ realization. This completes the proof. ■

Combining all of Chapter 4 with the results thus far of this chapter, we now have all the "tools" needed to prove Theorem 77, our extension of the main theorem of M. Bessmertnyĭ [02MB, Theorem 1.1].

5.3 Extension of the Bessmertnyĭ Realization Theorem

The following is our extension of the main theorem of M. Bessmertnyĭ [02MB, Theorem 1.1]. The proof of statements (a)-(c) are originally due to M. Bessmertnyĭ [82MB, 02MB]. The statements (d) and (e) are new and extends his results (our modest contributions). Our constructive proof of the theorem below is also new and is based on a completely new approach using the results in Chapter 4 (see Fig. 5.2) which may be of independent interest in the areas of linear algebra, operator theory, and multidimensional systems theory. To illustrate our approach of this theorem, we give two examples of realizations in Section 5.4. Before we state our main theorem though, we want to make a remark regarding statement (a).

Remark 76 *It appears that there is an error or at least some confusion that needs to be cleared up in regards to one of the statements in the main theorem of M. Bessmertnyĭ*

[02MB, Theorem 1.1.c)] which seems to have propagated in the literature (see [04KV, p. 256]). Namely, if a rational $\mathbb{C}^{k\times k}$-valued matrix function $f(z_1,\ldots,z_n)$ of n-variables satisfies $f(\lambda z) = \lambda f(z)$, i.e., is a homogeneous degree one function, which can be represented as a Schur complement $f(z) = A(z)/A_{22}(z)$ of a linear matrix pencil $A(z) = [A_{ij}(z)]_{i,j=1,2} = A_0 + z_1A_1 + \cdots + z_nA_n$, where $\det A_{22}(z) \not\equiv 0$, then it need not be the case that $A_0 = 0$. To see this consider the following simple example:

$$f(z) = [z_1] = \left[\begin{array}{c:c} z_1 & 0 \\ \hdashline 0 & 1 \end{array}\right] \Big/ [1] = A(z)/A_{22}(z), \tag{5.33}$$

where

$$A(z) = A_0 + z_1A_1 = \left[\begin{array}{c:c} A_{11}(z) & A_{12}(z) \\ \hdashline A_{21}(z) & A_{22}(z) \end{array}\right] = \left[\begin{array}{c:c} z_1 & 0 \\ \hdashline 0 & 1 \end{array}\right], \tag{5.34}$$

$$A_0 = \left[\begin{array}{c:c} 0 & 0 \\ \hdashline 0 & 1 \end{array}\right], \quad A_1 = \left[\begin{array}{c:c} 1 & 0 \\ \hdashline 0 & 0 \end{array}\right]. \tag{5.35}$$

This example and our statement below in Theorem 77.(a) below should now help to clear up any confusion regarding [02MB, Theorem 1.1.c)].

Theorem 77 (Bessmertnyĭ Realization Theorem) *Every rational $\mathbb{C}^{k\times k}$-valued matrix function $f(z_1,\ldots,z_n)$ of n-variables can be represented as a Schur complement $f(z) = A(z)/A_{22}(z)$ of a linear matrix pencil $A(z) = [A_{ij}(z)]_{i,j=1,2} = A_0 + z_1A_1 + \cdots + z_nA_n$, where $\det A_{22}(z) \not\equiv 0$. Moreover, the following functional relations are true:*

(a) $f(\lambda z) = \lambda f(z)$ (i.e., f is a homogeneous degree-one function) if and only if one can choose $A_0 = 0$.

(b) $f(z) = \overline{f(\overline{z})}$ if and only if one can choose $A_j = \overline{A_j}$, for all $j = 0,\ldots n$.

(c) $f(z) = f(z)^T$ if and only if one can choose $A_j = A_j^T$, for all $j = 0,\ldots n$.

(d) $f(z) = f(\overline{z})^*$ if and only if one can choose $A_j = A_j^*$, for all $j = 0,\ldots n$.

(e) f satisfies any combination of the (a)-(d) if and only if one can choose the A_j to have simultaneously the associated properties.

Proof. The theorem will be proved in a series of steps that reduce the complexity of the problem into simpler realization problems, as is illustrated in the flow diagram of Figure 5.2. Let $f(z) = f(z_1,\ldots,z_n)$ be a rational $\mathbb{C}^{k\times k}$-valued matrix function of n complex variables $z_1,\ldots,z_n$. Then there exists a nonzero scalar polynomial $q(z)$ and a polynomial $\mathbb{C}^{k\times k}$-valued matrix function $P(z)$ of these n-variables such that

$$f(z) = \frac{1}{q(z)} P(z).$$

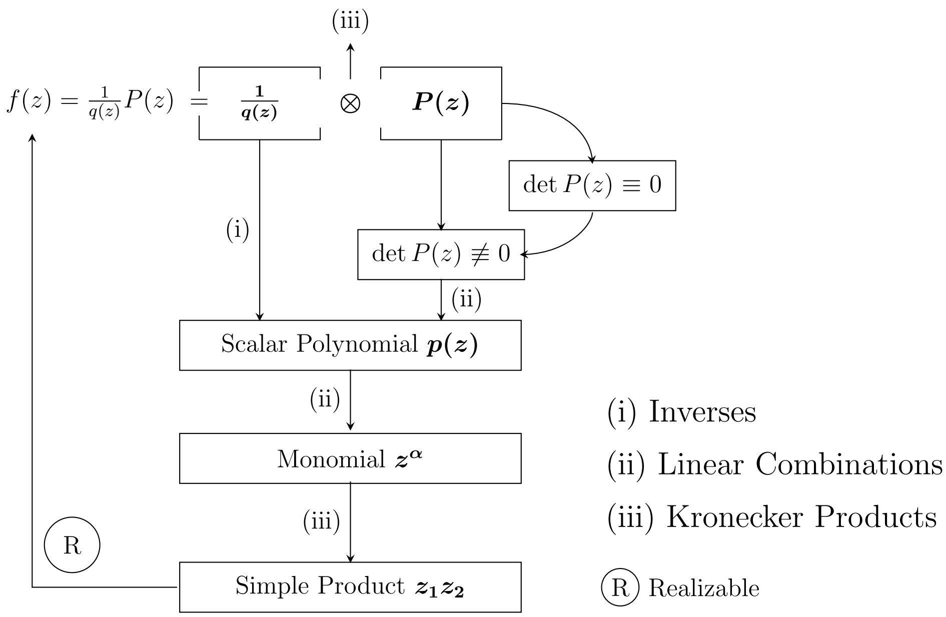


Figure 5.2: Flow diagram for the proof of the Bessmertnyĭ Realization Theorem

Using the Kronecker product ⊗ [and treating $q(z)$ as a polynomial $\mathbb{C}^{1\times 1}$-valued matrix function], we can rewrite this as

$$f(z) = \frac{1}{q(z)} \otimes P(z).$$

Before we can proceed further, there are two cases we must consider corresponding to whether $\det P(z) \not\equiv 0$ or not.

First, consider the case that $\det P(z) \not\equiv 0$. By Proposition 74 (as well as Lemmas 69 and 71), we can realize $\frac{1}{q(z)} \otimes P(z)$ if both $\frac{1}{q(z)}$ and $P(z)$ are realizable. By Proposition 43, we can realize $\frac{1}{q(z)}$ if $q(z)$ is realizable.

Second, consider the case $\det P(z) \equiv 0$. As the theorem is obviously true if $P(z) \equiv 0$, we may assume $P(z) \not\equiv 0$. Then there exists $z_0 \in \mathbb{C}^n$ such that $P(z_0) \neq 0$. Fix a nonzero real number λ_0 that is not an eigenvalue of $P(z_0)$ and consider the two matrix polynomials $P_1(z) = P(z) + \lambda_0 I_k - P(z_0)$ and $P_2(z) = P(z_0) - \lambda_0 I_k$. They both satisfy $\det P_j(z) \not\equiv 0$, for $j = 1, 2$ and $f(z) = \frac{1}{q(z)}P(z) = \frac{1}{q(z)}P_1(z) + \frac{1}{q(z)}P_2(z)$. Hence, by Lemma 28 and Proposition 30, $f(z)$ is realizable if both $\frac{1}{q(z)}P_1(z)$ and $\frac{1}{q(z)}P_2(z)$ are realizable. Thus, we are back to the first case again.

From considering both of the cases above, it now becomes clear that we just need to be able to realize any arbitrary matrix polynomial $P(z)$ and scalar polynomial $q(z)$. We will begin by investigating the realizability of the former and show it reduces to the realizability of the latter.

Suppose $P(z)$ is a polynomial $\mathbb{C}^{k\times k}$-valued matrix function of the n complex variables $z_1, \ldots, z_n$. Then, there exists scalar polynomial functions $P_{ij}(z)$, for $i,\ j = 1, \ldots, k$ such that

$$P(z) = [P_{ij}(z)]_{i,j=1}^{k} = \textstyle\sum_{i=1}^{k}\sum_{j=1}^{k} P_{ij}(z)\, E_{ij},$$

where E_{ij}, for $i,\ j = 1, \ldots, k$, are the standard basis vectors for $\mathbb{C}^{k\times k}$ (i.e., E_{ij} is the $k \times k$ matrix whose entry in ith row, jth column is 1 and the remaining entries are all 0).

Therefore, by Lemma 28, Proposition 30, and Proposition 55, $P(z)$ is realizable if each scalar polynomial functions $P_{ij}(z)$, for $i,\ j = 1, \ldots, k$ are realizable. Thus, we have reduced our problem to realizing an arbitrary scalar polynomial $p(z)$.

Suppose that $p(z)$ is an arbitrary scalar polynomial of the n complex variables $z_1, \ldots, z_n$ [e.g., $q(z)$ or one of the $P_{ij}(z)$ above]. Then $p(z)$ can be written uniquely as a linear combination of monomials,

$$p(z) = \textstyle\sum_{j=0}^{m} a_j z^{\alpha_j},$$

where a_j are scalars and z^{α_j} are monomials, for $j = 0, \ldots, m$. Hence, by Lemma 26, Lemma 28, and Proposition 30, it follows that $p(z)$ is realizable if each monomial z^{α_j} is realizable. Thus, we have reduced our problem to realizing an arbitrary monomial z^{α}.

Suppose that z^{α} is a monomial. Then it is realizable by Proposition 75. Let us explain the reason why. The monomial can be written uniquely as products of powers of the independent variables, $z^{\alpha} = \prod_{j=1}^{n} z_j^{\alpha_j}$. As this can be written as the Kronecker product $\prod_{j=1}^{n} z_j^{\alpha_j} = z_1^{\alpha_1} \otimes \cdots \otimes z_n^{\alpha_n}$, then by Proposition 74 we can realize the monomial z^{α} if we can realize the product $w_1 w_2$ of two independent complex variables w_1 and w_2, which we can by Lemma 68.

Therefore, we have proven that the rational $\mathbb{C}^{k\times k}$-valued matrix function $f(z) = f(z_1, \ldots, z_n)$ of the n complex variables $z_1, \ldots, z_n$ is realizable.

In the second part of this theorem, we will prove statements a)-e) are true for the rational function $f(z)$. We will achieve this by modifying the proof of the first part of the theorem above, when and where necessary, for each statements (a)-(e).

First, we will prove statement (a). Suppose that $f(\lambda z) = \lambda f(z)$, i.e., $f(z)$ is also a homogeneous function of degree one. Then

$$f(z) = z_1 f\left(\frac{z}{z_1}\right) = z_1 f\left(1, \frac{z_2}{z_1}, \ldots, \frac{z_n}{z_1}\right).$$

Hence,

$$g(w) = f(1, w_2, \ldots, w_n), \quad w = (w_2, \ldots, w_n),$$

satisfies the hypotheses of the first part of the theorem, and has a realization

$$g(w) = B(w)/B_{22}(w), \quad B(w) = A_1 + w_2 A_2 + \cdots + w_n A_n,$$

implying $f(z)$ has the realization

$$f(z) = A(z)/A_{22}(z), \quad A(z) = z_1A_1 + z_2A_2 + \cdots + z_nA_n,$$

since

$$A(z) = z_1 A\left(\frac{z}{z_1}\right) = z_1 B\left(\frac{z_2}{z_1}, \ldots, \frac{z_n}{z_1}\right)$$

and, by property (2.7),

$$\begin{aligned} f(z) &= z_1 f\left(\frac{z}{z_1}\right) = z_1 g\left(\frac{z_2}{z_1}, \ldots, \frac{z_n}{z_1}\right) \\ &= z_1\left[B\left(\frac{z_2}{z_1}, \ldots, \frac{z_n}{z_1}\right) / B_{22}\left(\frac{z_2}{z_1}, \ldots, \frac{z_n}{z_1}\right)\right] \\ &= \left[z_1 B\left(\frac{z_2}{z_1}, \ldots, \frac{z_n}{z_1}\right)\right] / \left[z_1 B_{22}\left(\frac{z_2}{z_1}, \ldots, \frac{z_n}{z_1}\right)\right] \\ &= \left[z_1 A\left(\frac{z}{z_1}\right)\right] / \left[z_1 A_{22}\left(\frac{z}{z_1}\right)\right] \\ &= A(z)/A_{22}(z). \end{aligned}$$

Conversely, suppose that $f(z)$ has a realization

$$f(z) = A(z)/A_{22}(z), \quad A(z) = z_1A_1 + z_2A_2 + \cdots + z_nA_n.$$

Then, since

$$A(\lambda z) = \lambda A(z),$$

it follows by property (2.7) that

$$f(\lambda z) = A(\lambda z)/A_{22}(\lambda z) = [\lambda A(z)] / [\lambda A_{22}(z)] = \lambda [A(z)/A_{22}(z)] = \lambda f(z).$$

Therefore, statement (a) is true.

Next, we will prove statement (b). Suppose that $f(z) = \overline{f(\overline{z})}$, [i.e., $f(z)$ is a $k \times k$ matrix whose entries are real rational scalar functions of z]. Then in the proof above in which we constructed a realization for $f(z)$ from the factorization $f(z) = \frac{1}{q(z)}P(z)$, we may assume that the nonzero scalar polynomial $q(z)$ is a real polynomial and the polynomial $\mathbb{C}^{k\times k}$-valued matrix function $P(z)$ is a real matrix polynomial [i.e., $P(z)$ is a $k \times k$ matrix whose entries are real polynomial scalar functions]. In this case, the proof of the realization of such an input $f(z) = \frac{1}{q(z)}P(z)$, as shown by the flow diagram in Fig. 5.2, would output a realization of $f(z)$ with real matrices (i.e., a Bessmertnyi realization in which the matrices in the linear matrix pencil are all real matrices). The converse of statement (b) is obviously true by property (2.8). Therefore, we have proven statement (b).

Next, we will prove statement (c). Suppose that $f(z) = f(z)^T$. Then in the proof above, in which we constructed a realization for $f(z)$ from the factorization $f(z) = \frac{1}{q(z)}P(z)$, we may assume that the polynomial $\mathbb{C}^{k\times k}$-valued matrix function $P(z)$ satisfies $P(z) = P(z)^T$. In this case, the proof of the realization of such an input $f(z) = \frac{1}{q(z)}P(z)$, as shown by the flow diagram in Fig. 5.2, would output a symmetric realization of $f(z)$ (i.e., a Bessmertnyi realization in which the matrices in the linear matrix pencil are all symmetric matrices) provided we can prove that $P(z)$ has a symmetric realization. To prove this, we need only make one slight modification of our proof using the fact that since $P(z) = P(z)^T$ then $P_{ji}(z) = P_{ij}(z)$ for all $i, j = 1, \ldots, k$ and $P_{ii}(z)E_{ii}$ and $P_{ij}(z)(E_{ij}+E_{ji})$ are symmetric for all $i, j = 1, \ldots, k$ so that by Lemma 26, Lemma 28, Proposition 30, Proposition 55, and Proposition 75 they have symmetric realizations which implies by Lemma 28 and Proposition 30 that their sum

$$\begin{aligned}\textstyle\sum_{i=1}^{k} P_{ii}(z)\,E_{ii} + \sum_{1\leq i<j\leq k} P_{ij}(z)\,(E_{ij}+E_{ji}) &= \textstyle\sum_{i=1}^{k}\sum_{j=1}^{k} P_{ij}(z)\,E_{ij} \\ &= P(z),\end{aligned}$$

has a symmetric realization. The converse of statement (c) is obviously true by property (2.9). Therefore, we have proven statement (c).

Next, we will prove statement (d). Suppose that $f(z) = f(\overline{z})^*$. Then in the proof above, in which we constructed a realization for $f(z)$ from the factorization $f(z) = \frac{1}{q(z)}P(z)$, we may assume that the nonzero scalar polynomial $q(z)$ is a real polynomial [i.e., $q(z) = \overline{q(\overline{z})}$] and the polynomial $\mathbb{C}^{k\times k}$-valued matrix function $P(z)$ satisfies $P(z) = P(\overline{z})^*$. In this case, the proof of the realization of such an input $f(z) = \frac{1}{q(z)}P(z)$, as shown by the flow diagram in Fig. 5.2, would output a Hermitian realization of $f(z)$ (i.e., a Bessmertnyi realization in which the matrices in the linear matrix pencil are all Hermitian matrices) provided we can prove that $P(z)$ has a Hermitian realization. To prove this, we need only make one slight modification to our proof of part (c). We separate $P(z)$ into its symmetric $Q_s(z)$ and skew-symmetric $Q_a(z)$ parts, i.e.,

$$\begin{aligned}P(z) &= \textstyle\sum_{i=1}^{k}\sum_{j=1}^{k} P_{ij}(z)\,E_{ij} \\ &= Q_s(z) + Q_a(z),\end{aligned}$$

where

$$\begin{aligned}Q_s(z) &= \textstyle\sum_{i=1}^{k} P_{ii}(z)\,E_{ii} + \sum_{1\leq i<j\leq k}\left[\dfrac{P_{ij}(z)+\overline{P_{ij}(\overline{z})}}{2}\right](E_{ij}+E_{ji}), \\ Q_a(z) &= \textstyle\sum_{1\leq i<j\leq k}\left[\dfrac{P_{ij}(z)-\overline{P_{ij}(\overline{z})}}{2}\right](E_{ij}-E_{ji}) \\ &= \textstyle\sum_{1\leq i<j\leq k}\left[\dfrac{P_{ij}(z)-\overline{P_{ij}(\overline{z})}}{2i}\right][i\,(E_{ij}-E_{ji})].\end{aligned}$$

Notice that for all $i, j = 1, \dots, k$, the scalar polynomials

$$\frac{P_{ij}(z) + \overline{P_{ij}(\overline{z})}}{2}, \quad \frac{P_{ij}(z) - \overline{P_{ij}(\overline{z})}}{2i}$$

are all real polynomials, the matrices

$$E_{ij} + E_{ji}$$

are all real and symmetric (hence Hermitian), and the matrices

$$i\left(E_{ij} - E_{ji}\right)$$

are all Hermitian. Thus, it follows by Proposition 55 that for any real scalar polynomial $p(z)$, if B is real and symmetric then $p(z)B$ has a real symmetric realization (i.e., a Bessmertnyi realization in which each matrix in the linear matrix pencil is a real and symmetric matrix) and, if instead B is a Hermitian matrix then $p(z)B$ has a Hermitian realization. From these facts and Lemma 28 and Proposition 30 on realizations of sums, it follows that $Q_s(z)$ has a real symmetric realization (which is a Hermitian realization) and $Q_a(z)$ has a Hermitian realization, and thus, Lemma 28 and Proposition 30 implies their sum $Q_s(z)+Q_a(z) = P(z)$ has a Hermitian realization. The converse of statement (d) is obviously true by property (2.10). This proves statement (d).

Finally, we will prove statement (e). Suppose $f(z)$ has any combination of two of the functional properties in (b), (c), or (d). Then $f(z)$ must satisfy $f(z) = f(z)^T$ and $f(z) = f(\overline{z})^*$ and hence we can proceed as in the proof of (d), in which case this we can assume that the nonzero scalar polynomial $q(z)$ is a real polynomial and $P(\overline{z})^* = P(z)^T = P(z) = Q_s(z) + Q_a(z)$ implying $Q_a(z)$ is the zero matrix and hence $P(z) = Q_s(z)$ has a real symmetric realization from which we conclude that in the proof of the realization of such an input $f(z) = \frac{1}{q(z)}P(z)$, as shown by the flow diagram in Fig. 5.2, would output a real symmetric realization of $f(z)$ which is automatically also a Hermitian realization. Now suppose that $f(z)$ has any combination of functional properties in (a)-(d). To complete the proof of statement (e), we need only prove the statement now in the case one of these functional properties is (a) [which we do by slightly modifying the proof of statement (a)]. By our proof of (a), it follows that the function $g(w) = f(1, w_2, \dots, w_n)$ inherits the same combination of functional properties (b)-(d) that $f(z)$ has. From our proof of statements (b)-(d) and the first part of our proof of (e) above, it follows that $g(w)$ has a real realization if (b) is true, a symmetric realization if (c) is true, a Hermitian realization if (d) is true, and a real symmetric realization if it has any combination of two of the functional properties in (b), (c), or (d). From this and the proof of statement (a) using such a realization for $g(w)$ as the choice of the linear matrix pencil $B(w) = A_1 + w_2A_1 + \cdots + w_nA_n$ in the proof of (a), it follows that $f(z)$ can be realized with the linear matrix pencil $A(z) = z_1A_1 + z_2A_1 + \cdots + z_nA_n$ which has the desired properties. The converse of statement (e) is obviously true by the elementary properties (2.7)-(2.10) of Schur complements. This proves statement

(e) and completes the proof of the theorem. ■

5.4 Examples

Example 78 *To illustrate our approach of the Bessmertnyĭ Realization Theorem in the case in which the hypotheses of statements (b), (c), and (e) apply, we will work out the realization of the following rational* $\mathbb{C}^{1\times 1}$*-valued function of* 2*-variables*

$$f(z) = \begin{bmatrix} \frac{z_2}{z_1} \end{bmatrix}.$$

As a first step, we write this in the form of a Kronecker product of matrices

$$f(z) = \frac{1}{q(z)}P(z) = \frac{1}{q(z)} \otimes P(z) = [q(z)]^{-1} \otimes P(z),$$

where

$$q(z) = z_1, \quad P(z) = \begin{bmatrix} z_2 \end{bmatrix}.$$

Next, we have $\det P(z) = z_2 \not\equiv 0$ *and* $P(z)$ *is already in the desired realized form. The next step is to realize* $q(z)$*, but in this case its already in the desired realized form, so we can realize its inverse,*

$$\frac{1}{q\left(z\right)} = \left[q\left(z\right)\right]^{-1} = \left[z_1\right]^{-1} = \left[\begin{array}{c:c} 0 & 1 \\ \hdashline 1 & -z_1 \end{array}\right] \Big/ \left[-z_1\right].$$

Finally, we complete this part of the example by realizing the Kronecker product of realizations

$$\begin{aligned} f(z) = [q(z)]^{-1} \otimes P(z) &= \left(\left[\begin{array}{c:c} 0 & 1 \\ \hdashline 1 & -z_1 \end{array}\right] \Big/ \left[-z_1\right] \right) \otimes [z_2] \\ &= A(z)/A_{22}(z), \end{aligned}$$

in which

$$A(z) = A_0 + z_1 A_1 + z_2 A_2$$

is a linear matrix pencil such that the matrices $A_j \in \mathbb{C}^{m\times m}$*, for some positive integer* m *(in this example we will have* $m = 4$*), are real and symmetric for* $j = 0, 1, 2$*. To compute this pencil, we follow Lemma 71 and its proof. First, by Lemma 45,*

$$\left(\left[\begin{array}{c:c} 0 & 1 \\ \hdashline 1 & -z_1 \end{array}\right] \Big/ \left[-z_1\right] \right) \otimes [z_2] = \left[\begin{array}{c:c} [0] \otimes [z_2] & [1] \otimes [z_2] \\ \hdashline [1] \otimes [z_2] & [-z_1] \otimes [z_2] \end{array}\right] \Big/ \left([-z_1] \otimes [z_2]\right).$$

Second, we compute

$$\left[\begin{array}{c:c} [0]\otimes[z_2] & [1]\otimes[z_2] \\ \hdashline [1]\otimes[z_2] & [-z_1]\otimes[z_2] \end{array}\right] = \left[\begin{array}{c:c} 0 & z_2 \\ \hdashline z_2 & -z_1z_2 \end{array}\right]$$
$$= z_2 \left[\begin{array}{c:c} 0 & 1 \\ \hdashline 1 & 0 \end{array}\right] + (z_1z_2)\left[\begin{array}{c:c} 0 & 0 \\ \hdashline 0 & -1 \end{array}\right].$$

Third, by Lemma 32 and Lemma 68,

$$(z_1z_2)\left[\begin{array}{c:c} 0 & 0 \\ \hdashline 0 & -1 \end{array}\right] = \left[\begin{array}{c:c} [0] & [0] \\ \hdashline [0] & [-z_1z_2] \end{array}\right]$$
$$= \left[\begin{array}{cc:cc} 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2}(z_1+z_2) & \frac{1}{2}(z_1-z_2) \\ \hdashline 0 & -\frac{1}{2}(z_1+z_2) & 1 & 0 \\ 0 & \frac{1}{2}(z_1-z_2) & 0 & -1 \end{array}\right] \Big/ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Fourth, by Lemma 28,

$$z_2 \left[\begin{array}{c:c} 0 & 1 \\ \hdashline 1 & 0 \end{array}\right] + (z_1z_2)\left[\begin{array}{c:c} 0 & 0 \\ \hdashline 0 & -1 \end{array}\right]$$
$$= \left[\begin{array}{cc:cc} 0 & z_2 & 0 & 0 \\ z_2 & 0 & -\frac{1}{2}(z_1+z_2) & \frac{1}{2}(z_1-z_2) \\ \hdashline 0 & -\frac{1}{2}(z_1+z_2) & 1 & 0 \\ 0 & \frac{1}{2}(z_1-z_2) & 0 & -1 \end{array}\right] \Big/ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Finally, we compute

$$A(z) = A_0 + z_1A_1 + z_2A_2 = \left[\begin{array}{c:ccc} 0 & z_2 & 0 & 0 \\ \hdashline z_2 & 0 & -\frac{1}{2}(z_1+z_2) & \frac{1}{2}(z_1-z_2) \\ 0 & -\frac{1}{2}(z_1+z_2) & 1 & 0 \\ 0 & \frac{1}{2}(z_1-z_2) & 0 & -1 \end{array}\right],$$
$$A_0 = \left[\begin{array}{c:ccc} 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{array}\right], \quad A_1 = \left[\begin{array}{c:ccc} 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \end{array}\right],$$
$$A_2 = \left[\begin{array}{c:ccc} 0 & 1 & 0 & 0 \\ \hdashline 1 & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \end{array}\right].$$

Example 79 *To illustrate our approach to the Bessmertnyĭ Realization Theorem in the case in which the hypotheses of statements (a)-(c) and (e) apply, we will work out* *the realization of the following rational $\mathbb{C}^{1\times 1}$-valued function of 3-variables*

$$f(z) = \left[\tfrac{z_2 z_3}{z_1}\right].$$

As the function $f(z)$ is homogeneous degree one [i.e., $f(\lambda z) = \lambda f(z)$] then following the proof of part (a) we start by realizing the function:

$$g(w) = f(1, w_2, w_3) = \left[w_2 w_3\right].$$

This has the realization

$$g(w) = B(w)/B_{22}(w),$$
$$B(w) = A_1 + w_2 A_2 + w_3 A_3 = \left[\begin{array}{c|cc} 0 & \frac{1}{2}(w_2+w_3) & -\frac{1}{2}(w_2-w_3) \\ \hline \frac{1}{2}(w_2+w_3) & -1 & 0 \\ -\frac{1}{2}(w_2-w_3) & 0 & 1 \end{array}\right],$$
$$A_1 = \left[\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & -1 & 0 \\ 0 & 0 & 1 \end{array}\right], \quad A_2 = \left[\begin{array}{c|cc} 0 & \frac{1}{2} & -\frac{1}{2} \\ \hline \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 0 & 0 \end{array}\right] \quad A_3 = \left[\begin{array}{c|cc} 0 & \frac{1}{2} & \frac{1}{2} \\ \hline \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{array}\right].$$

Finally, since $f(z) = z_1 g(\frac{z_2}{z_1}, \frac{z_3}{z_1})$, we get the realization of $f(z)$ as

$$f(z) = \left[\tfrac{z_2 z_3}{z_1}\right] = A(z)/A_{22}(z),$$
$$A(z) = z_1 A_1 + z_2 A_2 + z_3 A_3 = \left[\begin{array}{c|cc} 0 & \frac{1}{2}(z_2+z_3) & -\frac{1}{2}(z_2-z_3) \\ \hline \frac{1}{2}(z_2+z_3) & -z_1 & 0 \\ -\frac{1}{2}(z_2-z_3) & 0 & z_1 \end{array}\right].$$

Chapter 6

Symmetric Determinantal Representations of Polynomials

In this chapter we will provide a short and direct proof, using the results in this thesis, that every real multivariate polynomial has a symmetric determinantal representation. We then provide an example using our approach and extend our results from the real field $\mathbb{R}$ to an arbitrary field $\mathbb{F}$ different from characteristic 2. Although this result was first proved over the field $\mathbb{R}$ in 2006 by J. W. Helton, S. A. McCullough, and V. Vinnikov in [06HMV] using advanced results from multidimensional system theory, our new approach is only based on elementary results from the theory of determinants (see prelims Definition 2.3 and Lemma 9), the theory of Schur complements from Chapters 4 and 5, and basic properties of polynomials.

6.1 Introduction

As was first proven in 2006 by J. W. Helton, S. McCullough, and V. Vinnikov [06HMV], any real polynomial $p(z) \in \mathbb{R}[z]$ in n variables $[z = (z_1, \ldots, z_n)]$ has a *symmetric determinant representation*, i.e., there exists a linear matrix pencil $A_0 + \sum_{i=1}^{n} z_i A_i$ with symmetric matrices $A_0, \ldots, A_n \in \mathbb{R}^{m \times m}$ i.e., a linear real and symmetric matrix pencil (see Definition 1) such that

$$p(z) = \det \left(A_0 + \sum_{i=1}^{n} z_i A_i \right). \tag{6.1}$$

We give a new proof of this theorem, which we will refer to as the *HMV Theorem*.

One of the merits of our proof that makes it short and elementary (see Figure 6.1) is that it requires no prior knowledge of multidimensional systems theory (as compared to [06HMV] and [12RQ]) or advanced representation theory for multivariate polynomials (as compared to [11BG]). For instance, the proof of the HMV theorem in both [06HMV] and [12RQ], it is necessary to know or first prove that an arbitrary real polynomial has a "linear description." Thus, in [06HMV, p. 106] they say that "Our determinantal representation theorem is a bi-product of the theory of systems realizations of noncommutative rational functions..." (cf. Theorem 14.1 and its proof in [06HMV]), whereas R. Quarez in [12RQ] gives a more elementary proof of the HMV Theorem, yet has to first derive a result on symmetrizable unipotent linear descriptions for homogeneous polynomials (see Proposition 4.2 and the proof of Theorem 4.4 in [12RQ]).

An alternative proof of the HMV Theorem was given by B. Grenet et al. in [11BG] which is based on symmetrizing the algebraic complexity theoretic construction by L. Valiant in [79LV]. More precisely, it is proved in [79LV] that every polynomial has a nonsymmetric determinantal representation which is proved using a weighted digraph construction and then in [11BG] they use a similar yet symmetrized construction to get a symmetric determinantal representation to prove the HMV Theorem. As such, the proof in [11BG] requires a bit of effort to derive the HMV Theorem as there is quite a bit of prior knowledge needed to do this construction, especially if you are not familiar with algebraic complexity theory.

In this chapter we give a short proof of the HMV Theorem (see Theorem 84 and its proof) using some elementary results from the theory of determinants (see Lemmas 8 and 9), the theory of Schur complements from Chapter 4 (see 30 and 68), and basic properties of polynomials (see Lemma 81). We then provide an example that not only demonstrates our approach, but also illustrates the known problem of the HMV theorem not being true in a field of characteristic 2 (e.g., $\mathbb{F}_2$). As such, we provide a discussion on this and relate it to Chapter 5 (specifically Lemma 68), then extend our results of HMV theorem to arbitrary fields different from characteristic 2.

6.2 Polynomial Realization

The next lemma is the key to the short proof of the HMV Theorem in this chapter and its proof is somewhat obvious and likely known (in fact, the statement and proof is valid over any field $\mathbb{F}$ not just for $\mathbb{R}$). Before we state it, we need the following definition.

Definition 80 (Simple product substitution) *For any real polynomial $q = q(z) \in \mathbb{R}[z]$ in n variables $z = (z_1, \ldots, z_n)$, any $k \in \{1, \ldots, n\}$, and any pair of variables u and v, the real polynomial $p = q(z)|_{z_k=uv}$ obtained from $q(z)$ by making the substitution $z_k = uv$ is called a **simple product substitution** on q.*

Lemma 81 (Polynomial realization) *Each real polynomial p can be constructed from an linear real polynomial q by applying a finite number of simple product substitutions to q.*

Proof. First, Lemma 81 is obviously true for any linear real polynomial, i.e., for any real polynomial of degree less than or equal to one. We will now prove it is true for any real polynomial of degree 2. Let $n \in \mathbb{N}$, $a_0, \ldots, a_n, b_1, \ldots, b_{\frac{n(n+1)}{2}}$ be real scalars, and $z_1, \ldots, z_n, w_1, \ldots, w_{\frac{n(n+1)}{2}}$ be independent variables. Consider the linear real polynomial

$$q(z, w) = a_0 + \sum_{l=1}^{n} a_l z_l + \sum_{k=1}^{\frac{n(n+1)}{2}} b_k w_k.$$

Let $S_k = |_{w_k = z_i z_j}$ denote the operation of simple product substitution of $w_k = z_i z_j$, for all integer pairs i, j with $1 \leq i \leq j \leq n$ with integers $k = 1, \ldots, \frac{n(n+1)}{2}$ ordered by the lexicographical order on the pairs (i, j). Applying the operations consecutively of $S_1, \ldots, S_{\frac{n(n+1)}{2}}$ to q we get the polynomial

$$S_{\frac{n(n+1)}{2}} \cdots S_1 q = a_0 + \sum_{l=1}^{n} a_l z_l + \sum_{k=1}^{\frac{n(n+1)}{2}} b_k z_i z_j.$$

This proves the lemma for any real polynomial of degree 2. Suppose the lemma is true for all real polynomials of degree less than or equal to d for some natural number $d \geq 2$. We will now prove the lemma is true for any real polynomial of degree $d+1$. Let $n \in \mathbb{N}$, $b_1, \ldots, b_{M_{n,d+1}}$ be real scalars, and $z_1, \ldots, z_n, w_1, \ldots, w_{M_{n,d+1}}$ be independent variables, where $M_{n,d+1}$ is the number of monomials of degree $d+1$ in n independent variables. To be explicit, it is a well-known result that this number is given by the following formula

$$M_{n,d+1} = \binom{n+d}{d+1} = \frac{(n+d)!}{(d+1)!(n-1)!}.$$

Let $q_d(z)$ be a real polynomial of degree less than or equal to d in the n independent variables $z = (z_1, \ldots, z_n)$. Consider the real polynomial $q(z, w)$ of degree less than or equal to d in the $n + M_{n,d+1}$ independent variables $z_1, \ldots, z_n, w_1, \ldots, w_{M_{n,d+1}}$ defined by

$$r(z, w) = q_d(z) + \sum_{k=1}^{M_{n,d+1}} b_k z^{\alpha_k} w_k,$$

where the set $\{z^{\alpha_k} : k = 1, \ldots, M_{n,d+1}\}$ equals the set of all monomials in the n variables $z_1, \ldots, z_n$ of degree $d-1$ such that $z^{\alpha_k} z_i z_j$, $k = 1, \ldots, M_{n,d+1}, 1 \leq i \leq j \leq n$ is a list, with no repeats, of all the monomials in the n variables $z_1, \ldots, z_n$ of degree $d+1$. By the induction hypothesis, there exists a linear real polynomial q and a finite number T_l, $l = 1, \ldots M$ of simple product substitution operations such that

$$r(z, w) = T_M \cdots T_1 q.$$

Let $S_k = |_{w_k=z_iz_j}$ denote the operation of simple product substitution of $w_k = z_iz_j$, for all integer pairs i, j with $1 \leq i \leq j \leq n$ with integers $k = 1, \ldots, \frac{n(n+1)}{2}$ ordered by the lexicographical order on the pairs (i, j). Applying the operations consecutively of $S_1, \ldots, S_{\frac{n(n+1)}{2}}$ to $r(z, w)$ we get the polynomial

$$S_{\frac{n(n+1)}{2}} \cdots S_1 r = S_{\frac{n(n+1)}{2}} \cdots S_1 T_M \cdots T_1 q = q_d(z) + \sum_{k=1}^{M_{n,d+1}} b_k z^{\alpha_k} z_i z_j = p(z).$$

This proves the lemma for any real polynomial of degree $d+1$. Therefore, by induction the lemma is true for any real polynomial of any degree. This proves the lemma. ■

6.3 Symmetric Determinantal Representation

In this section we will prove the HMV Theorem over the field $\mathbb{R}$ (i.e., Theorem 84). To do this we will need the following two lemmas (the first lemma and its proof are valid over any field $\mathbb{F}$ not just $\mathbb{R}$; the second lemma and its proof are valid over any field $\mathbb{F}$ different from characteristic 2).

Lemma 82 (Scalar multiplication) *If c is a real number and $p(z) \in \mathbb{R}[z]$ has a symmetric determinantal representation then $cp(z)$ also has a symmetric determinantal representation.*

Proof. Suppose $c \in \mathbb{R}$ and $p(z) = \det[A(z)]$, where $A(z) = A_0 + \sum_{i=1}^n z_i A_i$ is an linear matrix pencil with symmetric matrices $A_0, \ldots, A_n \in \mathbb{R}^{m\times m}$. Then it follows from Lemma 8 that

$$B(z) = A(z) \oplus [c] = A_0 \oplus [c] + \sum_{i=1}^n z_i(A_i \oplus [c])$$

is an linear matrix pencil with symmetric matrices $A_i \oplus [c] \in \mathbb{R}^{(m+1)\times(m+1)}$ such that

$$cp(z) = \det[B(z)].$$

This proves that $cp(z)$ has a symmetric determinantal representation. ■

Lemma 83 (Substitution) *If $q = q(z) \in \mathbb{R}[z]$ has a symmetric determinant representation and p is obtained from q by a simple product substitution then p also has a symmetric determinant representation.*

Proof. Suppose $q = q(z)$ is a real polynomial in n variables $z = (z_1, \ldots, z_n)$ with a symmetric determinant representation

$$q(z) = \det[A(z)],$$

where $A(z) = A_0 + \sum_{i=1}^{n} z_i A_i$ is an linear matrix pencil with symmetric matrices $A_0, \ldots, A_n \in \mathbb{R}^{m\times m}$. Let p be a polynomial obtained from q by a simple product substitution. Then, by definition, there is a $k \in \{1, \ldots, n\}$ and a pair of variables u and v such that p is the real polynomial $p = q(z)|_{z_k=uv}$ in the variables z, u, v. Hence,

$$p = q(z)|_{z_k=uv} = \det\left[A(z)\right]|_{z_k=uv} = \det\left[A(z)|_{z_k=uv}\right].$$

It follows immediately from Lemma 30 and Lemma 68 that

$$A(z)|_{z_k=uv} = uvA_k + \left(A_0 + \sum_{i=1, i\neq k}^{n} z_i A_i\right) = B(z,u,v)/B_{22}(z,u,v),$$

where $B(z, u, v)$ is an linear matrix pencil of the form

$$B(z,u,v) = \left(B_0 + \sum_{i=1, i\neq k}^{n} z_i B_i\right) + uB_n + vB_{n+1} = \begin{bmatrix} B_{11}(z,u,v) & B_{12}(z,u,v) \\ B_{21}(z,u,v) & B_{22}(z,u,v) \end{bmatrix}, \quad (6.2)$$

such that, for some $N \in \mathbb{N}$, the matrices $B_j \in \mathbb{R}^{N\times N}$ for each $j = 0, \ldots, n+1$ are symmetric and $B_{22}(z, u, v)$ is a constant invertible matrix. In particular,

$$B_{22}(z,u,v) \equiv \det B_{22}(0,0,0) = d \in \mathbb{R}\setminus\{0\}. \quad (6.3)$$

Therefore, it follows from this, Lemma 8, and Lemma 9 that

$$p = \det\left[B(z,u,v)/B_{22}(z,u,v)\right] = \frac{1}{d}\det\left[B(z,u,v)\right] = \det\left\{B(z,u,v) \oplus [d^{-1}]\right\},$$

which is a symmetric determinantal representation for p. This completes the proof. ■

We are now ready to prove the HMV Theorem as an immediate corollary of Lemma 81 and Lemma 83 (the statement and its proof is valid for any field $\mathbb{F}$ different from characteristic 2; for more details on this see Sec. 6.5).

Theorem 84 (Symmetric determinantal representation) *Any real polynomial $p(z) \in \mathbb{R}[z]$ in n variables $z = (z_1, \ldots, z_n)$ has a symmetric determinant representation, i.e., there exists an linear matrix pencil $A_0 + \sum_{i=1}^{n} z_i A_i$ with real symmetric matrices $A_0, \ldots, A_n \in \mathbb{R}^{m\times m}$ such that*

$$p(z) = \det\left(A_0 + \sum_{i=1}^{n} z_i A_i\right). \quad (6.4)$$

Proof. Let $p = p(z) \in \mathbb{R}[z]$ be a real polynomial in n variables $z = (z_1, \ldots, z_n)$. If p is an linear real polynomial then $p(z) = \det[p(z)]$ is symmetric determinantal representation. Suppose that p is not an linear real polynomial. Then by Lemma 81, there exists an linear real polynomial q such that $p = S_l \cdots S_1 q$, where S_k is an operation of simple product substitution for each $k = 1, \ldots, l$ for some $l \in \mathbb{N}$. By Lemma 83, $S_1 q$ has a symmetric determinantal representation. If $l = 1$ then we are done. Thus, assume $l \geq 2$. If $S_k \cdots S_1 q$ has a symmetric determinantal representation for some integer $1 \leq k < l$ then by Lemma 83, $S_{k+1} S_k \cdots S_1 q = S_{k+1}(S_k \cdots S_1 q)$ has a symmetric determinantal representation. Therefore, by induction $p = S_l \cdots S_1 q$ has a symmetric determinantal representation. This proves the theorem. ■

We provide below a proof diagram to illustrate its algorithmic and elementary structure.

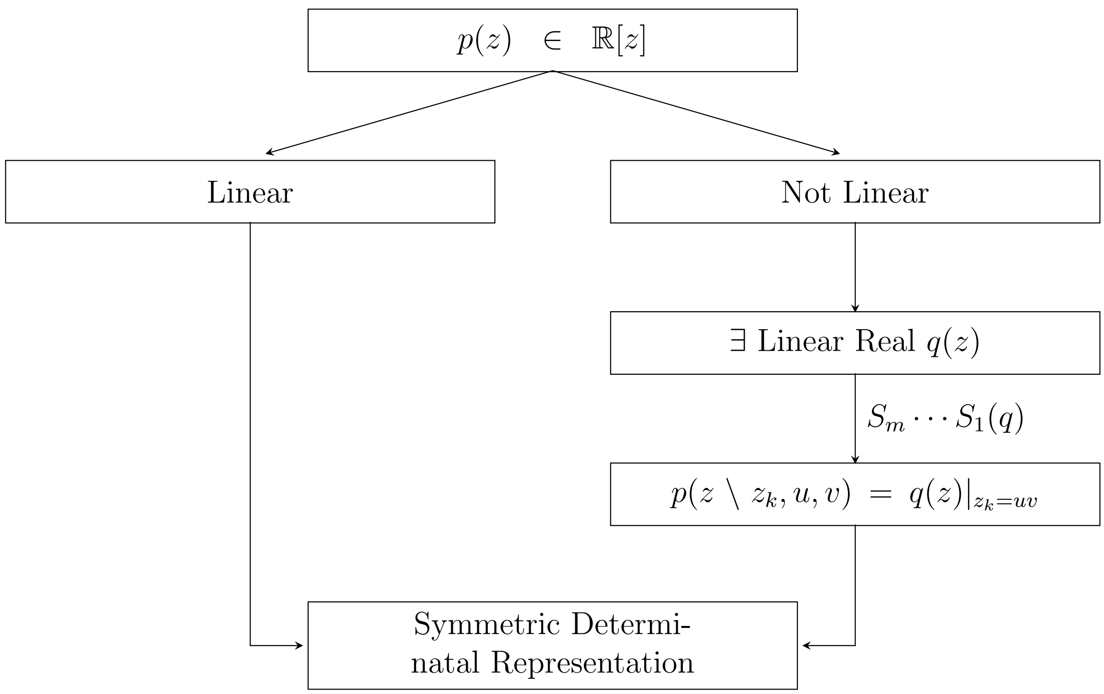


Figure 6.1: Flow diagram for the proof of HMV Theorem.

6.4 Example of HMV Theorem using our method

Example 85 *Here we will use the results in this chapter to show how to produce a symmetric determinantal representation for the real polynomial*

$$p(z_1, z_2, z_3) = z_1 + z_2 z_3. \tag{6.5}$$

First, (85) can be obtained from the linear real polynomial

$$q(z_1, w_1) = z_1 + w_1$$

by making the simple product substitution $w_1 = z_2z_3$ *since*

$$p(z_1, z_2, z_3) = z_1 + z_2z_3 = q(z_1, w_1)|_{w_1=z_2z_3}.$$

Next, $q(z_1, w_1)$ *has the symmetric determinantal representation*

$$q(z_1, w_1) = \det\left[q(z_1, w_1)\right] = \det(z_1[1] + w_1[1])$$

and hence,

$$p(z_1, z_2, z_3) = z_1 + z_2z_3 = \det\left[q_0(z_1, w_1)\right]|_{w_1=z_2z_3} = \det(z_1[1] + z_2z_3[1]).$$

Next, by Lemma 9 and Lemma 30, and following the proofs of Lemma 68 and Lemma 82 we have

$$\begin{aligned}
&\det(z_1[1] + z_2z_3[1]) \\
&= \det\left\{[z_1] + \left[\begin{array}{c:cc} 0 & \frac{1}{2}(z_2+z_3) & -\frac{1}{2}(z_2-z_3) \\ \hdashline \frac{1}{2}(z_2+z_3) & -1 & 0 \\ -\frac{1}{2}(z_2-z_3) & 0 & 1 \end{array}\right] \Big/ \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\right\} \\
&= \det\left\{\left[\begin{array}{c:cc} z_1 & \frac{1}{2}(z_2+z_3) & -\frac{1}{2}(z_2-z_3) \\ \hdashline \frac{1}{2}(z_2+z_3) & -1 & 0 \\ -\frac{1}{2}(z_2-z_3) & 0 & 1 \end{array}\right] \Big/ \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\right\} \\
&= \left(\det\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\right)^{-1} \det\begin{bmatrix} z_1 & \frac{1}{2}(z_2+z_3) & -\frac{1}{2}(z_2-z_3) \\ \frac{1}{2}(z_2+z_3) & -1 & 0 \\ -\frac{1}{2}(z_2-z_3) & 0 & 1 \end{bmatrix} \\
&= (-1)\det\begin{bmatrix} z_1 & \frac{1}{2}(z_2+z_3) & -\frac{1}{2}(z_2-z_3) \\ \frac{1}{2}(z_2+z_3) & -1 & 0 \\ -\frac{1}{2}(z_2-z_3) & 0 & 1 \end{bmatrix} \\
&= \det\left\{\begin{bmatrix} z_1 & \frac{1}{2}(z_2+z_3) & -\frac{1}{2}(z_2-z_3) \\ \frac{1}{2}(z_2+z_3) & -1 & 0 \\ -\frac{1}{2}(z_2-z_3) & 0 & 1 \end{bmatrix} \oplus [-1]\right\} \\
&= \det\begin{bmatrix} z_1 & \frac{1}{2}(z_2+z_3) & -\frac{1}{2}(z_2-z_3) & 0 \\ \frac{1}{2}(z_2+z_3) & -1 & 0 & 0 \\ -\frac{1}{2}(z_2-z_3) & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.
\end{aligned}$$

Therefore, the real polynomial $p(z_1, z_2, z_3) = z_1 + z_2z_3$ *has the symmetric determinantal representation*

$$p(z_1, z_2, z_3) = z_1 + z_2z_3 = \det\left(A_0 + z_1A_1 + z_2A_2 + z_3A_3\right), \tag{6.6}$$

with the linear matrix pencil

$$A_0 + z_1A_1 + z_2A_2 + z_3A_3 = \begin{bmatrix} z_1 & \frac{1}{2}(z_2+z_3) & -\frac{1}{2}(z_2-z_3) & 0 \\ \frac{1}{2}(z_2+z_3) & -1 & 0 & 0 \\ -\frac{1}{2}(z_2-z_3) & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{6.7}$$

and symmetric matrices $A_0, A_1, A_2, A_3 \in \mathbb{R}^{4\times 4}$ *given by*

$$A_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix},\ A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{6.8}$$

$$A_2 = \begin{bmatrix} 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},\ A_3 = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.9}$$

6.5 Extensions to arbitrary fields not characteristic two

In this section we will discuss how to extended our proof of the HMV Theorem (i.e., Theorem 84) on symmetric determinantal representations of real polynomials to polynomials with coefficients in an arbitrary field $\mathbb{F}$ different from characteristic 2.

We will begin in Subsection 6.5.1 with a discussion of some previous results in this regard and compare it to ours. Furthermore, we include a brief discussion on how we modify our proof of the HMV Theorem from the field $\mathbb{R}$ to an arbitrary field $\mathbb{F}$ different from characteristic 2. Moreover, we explain why the proof of our theorem would fail for fields of characteristic 2 by pointing out which step in our proof causes the problem and illustrate the issue in the example from Section 6.4. Finally, in Subsection 6.5.2 we give our short proof of the extension of the HMV Theorem to such fields.

6.5.1 Comparison to previous results

In [12RQ, Sec. 5.1], R. Quarez extends the HMV Theorem to polynomials with coefficients in an arbitrary ring R of characteristic different from 2 (see [12RQ, Theorem 5.1]). To do this, he has to adjust the steps in his construction to work not only over $\mathbb{R}$, but also over the ring R. The main issue he has to resolve is adjusting his proof (mainly of [12RQ, Theorem 4.4]) to avoid issues with inversion and using the diagonalization theorem for real symmetric matrices. However, these matrices could involve, for instance, square roots and hence not allowable over a general ring (cf. [12RQ, Theorem 5.1] and its proof).

In comparison with our methods, although we just treat a field $\mathbb{F}$ different from characteristic 2 instead of a ring R (for the sake of simplicity), we only have to adjust one step in our proof of the HMV Theorem. The adjustment we have to make is in the proof of Lemma 68, namely, for the case when the symmetric matrix $B \in \mathbb{F}^{m\times m}$ is not invertible. This is due to the issue that it is not always possible to find a scalar $\lambda_0 \in \mathbb{F} \setminus \{0\}$ such that $B - \lambda_0 I_m$ is invertible. This issue would not occur if $\mathbb{F}$ is an infinite field (e.g., a field of characteristic 0) as B then can only have a finite number of eigenvalues [02FIS, p. 249, Theorem 5.3], but it is an issue when $\mathbb{F}$ is a finite field (hence it must be a field of characteristic p, for some prime number p). To see this, let $\mathbb{F}$ be a finite field. Then, it has p^k elements in it for some prime number p (so that $\mathbb{F}$ has characteristic p) and some integer $k \geq 1$. Let $0, \lambda_1, \ldots, \lambda_{p^k-1}$ be all the distinct elements of $\mathbb{F}$ and consider the diagonal matrix $B = \operatorname{diag}\{0, \lambda_1, \ldots, \lambda_{p^k-1}\} \in \mathbb{F}^{p^k\times p^k}$. This is a symmetric matrix in $\mathbb{F}^{p^k\times p^k}$ that is not invertible such that $B - \lambda_0 I_{p^k}$ is not invertible for every scalar $\lambda_0 \in \mathbb{F} \setminus \{0\}$.

In [11BG], B. Grenet et al. gives a proof of the HMV Theorem using a construction that, from the very beginning, was meant to be valid for any field of characteristic different from 2. As they point out (see [11BG, Sec. 5]), their constructions are not valid for fields of characteristic 2 because they need to use the scalar $1/2$ (i.e., the multiplicative inverse of 2 which, of course, does not exist in a field of characteristic 2).

Comparing our methods, the reason that our construction will fail as well, for a field $\mathbb{F}$ of characteristic 2, is for a similar reason as in [11BG]. More precisely, the only issue in our proof of the HMV Theorem is in the proof of Lemma 68, where in order to realize the simple product uvB for a symmetric matrix $B \in \mathbb{F}^{m\times m}$, we need to use the scalar $1/2$.

In [13GMT], B. Grenet, T. Monteil, and S. Thomassé study the problem of the symmetric determinantal representations of polynomials with coefficients in fields with characteristic 2, i.e., on the extension of the HMV Theorem to such fields. One of the motivations for their paper was to prove a conjecture from [11BG] that such representations do not always exist for such fields. This conjecture is proven in [13GMT, Sec. 4] and, in particular, [13GMT, Sec. 4.2] gives as an example the polynomial $xy + z$ in the three variables x, y, z from the field $\mathbb{F}_2$, the field with two elements, and prove it has no symmetric determinantal representation. More specifically, they prove by their result [13GMT, Theorem 4.2], that the polynomial $p(x, y, z) = xy + z$ can *not* be represented as the determinant of a symmetric matrix with entries in $\mathbb{F}_2 \cup \{x, y, z\}$.

Therefore, if you compare this to our example from Section 6.4, since we need to use the scalar $1/2$ as well, we face the same issue if our field was indeed characteristic 2.

6.5.2 Extension of the HMV Theorem

In this subsection we will prove an extension of the HMV Theorem from the field $\mathbb{R}$ (i.e., Theorem 84) to an arbitrary field $\mathbb{F}$ different from characteristic 2 (i.e., Theorem 86). We will need the following elementary lemma (which are true over any field $\mathbb{F}$ with any characteristic).

Theorem 86 (Symmetric determinantal representation) *Let $\mathbb{F}$ be a field of characteristic different from 2. Then any polynomial $p(z) \in \mathbb{F}[z]$ in n variables $z = (z_1, \ldots, z_n)$ has a symmetric determinant representation, i.e., there exists an linear matrix pencil $A_0 + \sum_{i=1}^{n} z_i A_i$ with symmetric matrices $A_0, \ldots, A_n \in \mathbb{F}^{m \times m}$ such that*

$$p(z) = \det\left(A_0 + \sum_{i=1}^{n} z_i A_i\right). \tag{6.10}$$

Proof. The proof of this theorem is the same as the proof of the case when the field is $\mathbb{R}$ (i.e., the same proof as we gave for Theorem 84, but just replace the field $\mathbb{R}$ in that proof with the field $\mathbb{F}$ of characteristic different from 2) except we have to make one change to the proof of Lemma 68 in Chapter 5, namely, for the case when the symmetric matrix $B \in \mathbb{F}^{m \times m}$ is not invertible we need to prove Lemma 68 for the matrix polynomial uvB. If $B = 0$ then the proof is immediate. Thus, assume $B \neq 0$. In this case, we prove Lemma 68 by first applying Lemma 54 to the matrix B to get the factorization

$$uvB = Y^T \begin{bmatrix} uvB_1 & 0 \\ 0 & 0_{m-r} \end{bmatrix} Y, \tag{6.11}$$

where $Y \in \mathbb{F}^{m \times m}$ and $B_1 \in \mathbb{F}^{r \times r}$ are invertible matrices and B_1 is symmetric. Now we can appeal to the first part of the proof of Lemma 68 applied to the matrix polynomial uvB_1 and then use Lemma 32 followed by Lemma 39 to prove Lemma 68 for the matrix polynomial uvB. This proves the theorem. ■

Chapter 7

Open Problems and Future Directions

7.1 Open Problems

The two open problems we discuss are associated to the class of functions $f(z)$ that are rational positive-real functions of n-variables and are also homogeneous degree-one functions. For such functions, the main question we are interested in answering is whether or not there exists a linear matrix pencil $A(z) = z_1A_1 + \cdots + z_nA_n$ that also has these functional properties and gives a realization of $f(z)$ (the class of such functions that have such a realization are known as the rational *Bessmertnyĭ class of functions*, see [11JB, 04KV]). The Bessmertnyĭ class is as follows.

Proposition 87 (Bessmertnyĭ class) *Let* $f(z_1, \ldots, z_n)$ *is an* n*-variable,* $\mathbb{C}^{k\times k}$*-valued matrix function representable as a Schur complement* $f(z) = A(z)/A_{22}(z)$ *of a matrix pencil* $A(z) = [A_{ij}(z)]_{i,j=1,2} = z_1A_1 + \cdots + z_nA_n$*, where* $\det A_{22}(z) \not\equiv 0$*, with* $A_j \geq 0$ *(*$j = 1, \ldots, n$*) then it has the properties:*

(i) f *is a rational function,*

(ii) f *is homogeneous degree* 1*, and*

(iii) $\operatorname{Re} f(z) = \frac{1}{2}[f(z) + f(z)^*] \geq 0$ *if* $\operatorname{Re} z_j > 0$ *(*$j = 1, \ldots, n$*).*

Functions with the hypothesis are called the rational Bessmerntyĭ class of functions. The converse of Proposition 87 for $n = 2$ has been proven in the real symmetric case is due to M. F. Bessmertnyĭ [82MB] and the Hermitian case is due to D. Kalyuzhnyĭ-Verbovetzkiĭ [04KV]. On March 3rd, 2021, M. F. Bessmertnyĭ uploaded a possible proof to ArXiv for $n \geq 3$ [21MB], this is a very exciting result and although we hope it is a correct proof, it still needs to be confirmed. Related to this open problem is the associated *Milton class of functions* which is of particular importance in the theory of composites (see [16GM10, 87GM1, 87GM2], [02GM, Chap. 29] and [16GM7]). The associated functions $f(z)$ have the property that, in addition to being in the rational Bessmertnyĭ class of functions, also satisfy the normalization condition $f(1,\ldots,1)=I_k$ (I_k is the $k\times k$ identity matrix). Then the goal is find a Bessmertnyĭ realization with a positive linear matrix pencil $A(z)=z_1A_1+\ldots+z_nA_n$ also satisfying the normalization condition, i.e., $A_1+\cdots+A_n=I_m$.

The *Milton class of functions* is as follows.

Proposition 88 (Milton class) *Let $f(z_1,\ldots,z_n)$ is an n-variable, $\mathbb{C}^{k\times k}$-valued matrix function representable as an effective operator $f(z)=\sigma_*(z)$ of a finite-dimensional orthogonal $Z(n)$-subspace collection ($\mathcal{H}=\mathcal{U}\overset{\perp}{\oplus}\mathcal{E}\overset{\perp}{\oplus}\mathcal{J}=\mathcal{P}_1\overset{\perp}{\oplus}\cdots\overset{\perp}{\oplus}\mathcal{P}_n$, $\dim\mathcal{H}<\infty$) then it has the properties:*

(i) f is in the rational Bessmerntyĭ class and

(ii) $f(1,\ldots,1)=I_k$.

Functions with the hypothesis are called the rational Milton class of functions. The converse of Proposition 88 for $n=2$ real symmetric case is due to G. W. Milton [87GM1], [87GM2] using continued fraction expansion of $f(z_1,z_2)$. The Hermitian case for $n=2$ was also solved by Milton [16GM] using the integral representations of Herglotz functions. The converse of Proposition 88 for $n\geq 3$ is completely open.

7.2 Future Directions

This thesis is only the starting point to our research. Using the Schur complement algebra and operations perspective we can extend the Bessmertnyĭ Realization (Theorem 77) to include additional symmetries that arise in applications. For instance, incorporate nonreciprocal systems (e.g., lossless circuits with gyrators break time-reversal symmetry) by using our extensions [namely, e) and d)] of Hermitian matrices and any combination with the other symmetries to the Bessmertnyĭ Realization Theorem. Other symmetries to investigate are realizability for multivariate reactance functions. Further developments in characterising what rotational symmetries (which naturally occur in physics and applied models) pose on our functions. As we know, Kronecker product is "natural" for realizations, but how might one use matrix products in applied models? The condition of matrix products of two Schur complements (i.e., Proposition 37) requiring $A=B^T$ or $A=B^*$ in order to have a symmetric matrix C, could be useful in specific applications.

The more algorithmic structure that Schur complement algebra and operations as seen our proofs of both Bessmertnyĭ Realization Theorem and HMV Theorem motivate the development of computational methods to implement the Schur complement algebra and operations to compute realizations with symmetries when they exist and test conjectures. The program has already developed a tool kit of our Schur complement algebra and operations, the next step is to start incorporating this tool kit in symbolic computations for realizations. We seek to investigate minimization and bounds on the size of the matrices from realizations.

As seen in Chapter 6, Section 6.5.2 we extended HMV to arbitrary fields not characteristic two, we are able to identify the exact source of failure to the the Bessmertnyĭ realizability theorem in fields characteristic two (realizations of simple products), thus are finalizing an extension to the Bessmertnyĭ Realizability Theorem over any arbitrary field using our methods. Finally, to study the recent developments on the Bessmertnyĭ class of functions by M. Bessmertnyĭ [21MB] can not only give insights on how we can approach using our techniques to solve the Bessmertnyĭ class open problem (Proposition 87) but also approach solving the Milton class (Proposition 88) of functions. The Milton class motivates its applications in the theory of composites (e.g., effective operators and their bounds, limitations, as well as realizability problems), in particular, the effective operators σ_* in the abstract theory of composites and connection to the Schur-Agler class of multivariate functions.

References

[03AD] D. Alpay and C. Dubi. A realization theorem for rational functions of several complex variables. Systems & Control Letters 49, 225-229 (2003). doi:10.1016/S0167-6911(02)00326-2.

[71WA] W. N. Anderson Jr. Shorted operators. SIAM J. Appl. Math. 20, 520-525 (1971). doi:10.1137/0120053.

[74AT] W. N. Anderson Jr. and G. E. Trapp. Shorted operators. II. SIAM J. Appl. Math. 28, 60-71 (1975). doi:10.1137/0128007.

[05TA] T. Ando. Schur complements and matrix inequalities: Operator-Theoretical Approach. In: Zhang, F. (ed.) The Schur Complement and its Applications, pp. 137-162. Springer (2005).

[15ACM] M. Arias, G. Corach, and A. Maestripieri. Range additivity, shorted operator and the Sherman-Morrison-Woodbury formula. Linear Algebra and its Applications 467, 86-99 (2015). doi:10.1016/j.laa.2014.11.001.

[69BB] N. Balabanian and T. Bickart. Electrical network theory. With Contributions from the work of the late Sundaram Seshu. Wiley, 1969.

[11JB] J. A. Ball. Multidimensional circuit synthesis and multivariable dilation theory. Multidim. Syst. Sign. Process. 22, 27–44 (2011). doi:10.1007/s11045-010-0123-2.

[82MB] M. F. Bessmertnyĭ. Functions of Several Variables in the Theory of Finite Linear Structures. Ph. D. Thesis, Kharkov University, Kharkov (1982).

[02MB] M. F. Bessmertnyĭ. On realizations of rational matrix functions of several complex variables. In: Alpay, D., Gohberg, I., Vinnikov, V. (eds.) Interpolation Theory, Systems Theory and Related Topics (Oper. Theory: Adv. Appl., Vol. 134), pp. 157-185. Birkhäuser Verlag, Basel (2002). doi:10.1007/978-3-0348-8215-6_1.

[21MB] M. F. Bessmertnyĭ. Representation of rational positive real functions of several variables by means of positive long resolvent, arXiv:2103.02541 (2021).

[49BD] R. Bott and R. J. Duffin. Impedance synthesis without use of transformers. Journal of Applied Physics 20, 816 (1949). doi:10.1063/1.1698532.

[53BD] R. Bott and R. J. Duffin. The algebra of networks. Trans. Amer. Math. Soc. 74, 99-109 (1953). doi:10.1090/S0002-9947-1953-0056573-X.

[16CWM] M. Cassier, A. Welters, and G. W. Milton. Analyticity of the Dirichlet-to-Neumann map for the time-harmonic Maxwell's equations. In: Milton, G. W. (ed.) Extending the Theory of Composites to Other Areas of Science, pp. 95-121, Milton-Patton Publishers, Salt Lake City, Utah (2016). doi:10.1115/1.4035525.

[16GM10] M. Cassier, A. Welters, and G. W. Milton. A rigorous approach to the field recursion method for two-component composites with isotropic phases. In: Milton, G. W. (ed.) Extending the Theory of Composites to Other Areas of Science, pp. 287-307, Milton-Patton Publishers, Salt Lake City, Utah (2016). doi:10.1115/1.4035525.

[00CM] E. B. Curtis and J. Morrow. Inverse problems in electric circuits. (2000). doi:10.1142/4306.

[59RD] R. J. Duffin. An analysis of the Wang algebra of networks. Trans. Amer. Math. Soc. 93, 114-131 (1959). doi:10.1090/S0002-9947-1959-0109161-6.

[78DM] R. J. Duffin and T. D. Morley. Wang algebra and matroids. IEEE Transactions on Circuits and Systems 25, 755-762 (1978). doi:10.1109/TCS.1978.1084531.

[55RD] R. J. Duffin. Elementary operations which generate network matrices. Proc. Amer. Math. Soc. 6, 335-339 (1955). doi:10.1090/S0002-9939-1955-0072725-4.

[65DHM] R. J. Duffin, D. Hazony, and N. Morrison. The gyration operator in network theory. Scientific Report No. 7, AF 19 (628) 1699, CRST I Sills Bld 5285 Port Royal Road, Springfield, Virginia (1965). doi:10.21236/ad0616329.

[66DHM] R. J. Duffin, D. Hazony, and N. Morrison. Network synthesis through hybrid matrices. SIAM Journal on Applied Mathematics 14, 390–413 (1966). doi:10.1137/0114032.

[02FIS] S. Friedberg, A. Insel, and L. Spence. Linear algebra, 4th ed., Pearson, 2002.

[11BG] B. Grenet, E. Kaltofen, P. Koiran, and N. Portier. Symmetric determinantal representation of formulas and weakly skew circuits, Contemporary Mathematics Volume 566, (2011). doi:10.1090/conm/556.

[13GMT] B. Grenet, T. Monteil, and S. Thomassé. Symmetric determinantal representations in characteristic 2, Linear Algebra and its Applications 439 (2013), 1364–1381. doi:10.1016/j.laa.2013.04.022.

[06HMV] J. W. Helton, S. A. McCullough, and V. Vinnikov. Noncommutative convexity arises from linear matrix inequalities, J. Funct. Anal. 240 (2006), 105-191. doi:10.1016/j.jfa.2006.03.018.

[91HJ] R. Horn and C. Johnson. Topics in matrix analysis, Cambridge University Press, 1991. doi:10.1017/CBO9780511840371.

[01AJ] A. Jackson. Interview with Raoul Bott. Notices of the American Mathematical Society 48, 374–382 (2001).

[04KV] D. S. Kalyuzhnyĭ-Verbovetzkiĭ. On the Bessmertnyĭ Class of Homogeneous Positive Holomorphic Functions of Several Variables. In: Ball, J. A., Helton, J. W., Klaus, M., Rodman, L. (eds.) Current Trends in Operator Theory and its Applications (Oper. Theory: Adv. Appl., Vol. 149), pp. 255-289. Birkhäuser-Verlag, Basel (2004). doi:10.1007/978-3-0348-7881-4_12.

[66KP] I. J. Katz and M. H. Pearl. On EP_r and normal EP_r matrices, J. Res. Nat. Bur. Stand., Series B, 70B (1966), 47-75. doi:10.6028/jres.070B.004.

[47MK] M. G. Krein. Theory of selfadjoint extensions of semi-bounded Hermitian operators (in Russian). Rec. Math. [Mat. Sbornik] N.S., Vol. 20(62), No. 3, 431–495 (1947).

[94KO] M. G. Krein and I. E. Ovcharenko. On the theory of inverse problems for the canonical differential equation. In: Gohberg, I., Sakhnovich, L. A. (eds.) Matrix and Operator Valued Functions (Oper. Theory: Adv. Appl., Vol. 72), pp. 162-170. Birkhäuser-Verlag, Basel (1994). doi:10.1007/978-3-0348-8532-4_7.

[05LPR] A. S. Lewis, P. A. Parrilo, and M. V. Ramana. The Lax conjecture is true, American Mathematical Society 9, 133, (2005), 2495–2499.

[87GM1] G. W. Milton. Multicomponent composites, electrical networks and new types of continued fractions I. Commun. Math. Phys. 111, 281-327 (1987). doi:10.1007/BF01217763.

[87GM2] G. W. Milton. Multicomponent composites, electrical networks and new types of continued fractions II., Commun. Math. Phys. 111, 329-372, (1987). doi:10.1007/BF01238903

[02GM] G. W. Milton. The theory of composites, Cambridge University Press, 2002. doi:10.1017/CBO9780511613357.

[16GM] G. W. Milton. (ed.). Extending the Theory of Composites to Other Areas of Science. Milton-Patton Publishers, Salt Lake City, Utah (2016). doi:10.1115/1.4035525.

[16GM7] G. W. Milton. Superfunctions and the algebra of subspace collections and their association with rational functions of several complex variables. In: Milton, G. W. (ed.) Extending the Theory of Composites to Other Areas of Science, pp. 179-234, Milton-Patton Publishers, Salt Lake City, Utah (2016). doi:10.1115/1.4035525.

[10MBM] S. Mitra, P. Bhimasankaram, and S. Malik. Matrix partial orders, Shorted Operators and Applications. World Sci. Publ. (2010). doi:10.1142/7170.

[76NA] K. Nishio and T. Ando. Characterizations of operations derived from network connections. J. Math. Anal. 53, 539-549 (1976). doi:10.1016/0022-247X(76)90090-1.

[14EP] E. L. Pekarev. A note on characterization of the shorted operation. Elec. J. Linear Algebra 27, 155-161 (2014). doi:10.13001/1081-3810.1609.

[12RQ] R. Quarez. Symmetric determinantal representation of polynomials, Linear Algebra and its Applications 436 (2012), 3642-3660. doi:10.1016/j.laa.2012.01.004.

[00MT] M. J. Tsatsomeros. Principal pivot transforms: properties and applications. Linear Algebra and its Applications 307, 151-165 (2000). doi:10.1016/S0024-3795(99)00281-5.

[60AT] A. W. Tucker. A combinatorial equivalence of matrices. In: Bellman, R. Hall, M., Jr. (eds.), Combinatorial Analysis, American Mathematical Society, Providence, pp. 129–140 (1960). doi:10.1090/psapm/010.

[60PB] P. R. Bryant. The algebra and topology of electrical networks. Proceedings of the IEE - Part C: Monographs, 108, (13), p. 215-229 (1960). doi:10.1049/pi-c.1961.0030.

[61SR] S. Seshu and M. Reed. Linear Graphs and Electrical Networks, Addison-Wesley, 1961.

[72PP] Martin H. Pearl and Alan I. Penn. Normal matrices with entries from an arbitrary field of characteristic $\neq 2$, J. Res. Nat. Bur. Stand., Series B, 76B (1972), 119-143. doi:10.6028/jres.076B.009.

[21SW] A. Stefan and A. Welters. A Short Proof of the Symmetric Determinantal Representation of Polynomials, arXiv:2101.03589 (2021).

[20SW] A. Stefan and A. Welters. Extension of the Bessmertnyĭ realization theorem for rational functions of several complex variables, arXiv:2010.08088 (2020).

[79LV] L. G. Valiant. Completeness classes in algebra. Proc. 11th STOC '79. ACM, New York, NY, USA, pp. 249-261, (1979). doi:10.1145/800135.804419.

[05FZ] F. Zhang, The Schur complement and its applications, Springer, 2005. doi:10.1007/b105056.