THE SCHUR COMPLEMENT AND ITS APPLICATIONS

Numerical Methods and Algorithms

VOLUME 4

Series Editor:

Claude Brezinski
Université des Sciences et Technologies de Lille, France

THE SCHUR COMPLEMENT AND ITS APPLICATIONS

Edited by

FUZHEN ZHANG
Nova Southeastern University, Fort Lauderdale, U.S.A.
Shenyang Normal University, Shenyang, China

Springer

Library of Congress Cataloging-in-Publication Data

A C.I.P. record for this book is available from the Library of Congress.

ISBN 0-387-24271-6 e-ISBN 0-387-24273-2 Printed on acid-free paper.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1 SPIN 11161356

springeronline.com

To our families, friends, and the matrix community

Issai Schur (1875-1941)

This portrait of Issai Schur was apparently made by the "Atelieir Hanni Schwarz, N. W. Dorotheenstraße 73" in Berlin, *c.* 1917, and appears in *Ausgewählte Arbeiten zu den Ursprüngen der Schur-Analysis: Gewidmet dem großen Mathematiker Issai Schur (1875–1941)* edited by Bernd Fritzsche & Bernd Kirstein, pub. B. G. Teubner Verlagsgesellschaft, Stuttgart, 1991.

Emilie Virginia Haynsworth (1916-1985)

This portrait of Emilie Virginia Haynsworth is on the Auburn University Web site www.auburn.edu/~fitzpjd/ben/images/Emilie.gif and in the book *The Education of a Mathematician* by Philip J. Davis, pub. A K Peters, Natick, Mass., 2000.

Contents

Preface

What's in a name? To paraphrase Shakespeare's Juliet, that which Emilie Haynsworth called the *Schur complement*, by any other name would be just as beautiful. Nevertheless, her 1968 naming decision in honor of Issai Schur (1875–1941) has gained lasting acceptance by the mathematical community. The Schur complement plays an important role in matrix analysis, statistics, numerical analysis, and many other areas of mathematics and its applications.

Our goal is to expose the Schur complement as a rich and basic tool in mathematical research and applications and to discuss many significant results that illustrate its power and fertility. Although our book was originally conceived as a research reference, it will also be useful for graduate and upper division undergraduate courses in mathematics, applied mathematics, and statistics. The contributing authors have developed an exposition that makes the material accessible to readers with a sound foundation in linear algebra.

The eight chapters of the book (Chapters 0–7) cover themes and variations on the Schur complement, including its historical development, basic properties, eigenvalue and singular value inequalities, matrix inequalities in both finite and infinite dimensional settings, closure properties, and applications in statistics, probability, and numerical analysis. The chapters need not be read in the order presented, and the reader should feel at leisure to browse freely through topics of interest.

It was a great pleasure for me, as editor, to work with a wonderful group of distinguished mathematicians who agreed to become chapter contributors: T. Ando (Hokkaido University, Japan), C. Brezinski (Université des Sciences et Technologies de Lille, France), R. A. Horn (University of Utah, Salt Lake City, USA), C. R. Johnson (College of William and Mary, Williamsburg, USA), J.-Z. Liu (Xiangtang University, China), S. Puntanen (University of Tampere, Finland), R. L. Smith (University of Tennessee, Chattanooga, USA), and G. P. H. Styan (McGill University, Canada).

I am particularly thankful to George Styan for his great enthusiasm in compiling the master bibliography for the book. We would also like to acknowledge the help we received from Gülhan Alpargu, Masoud Asgharian, M. I. Beg, Adi Ben-Israel, Abraham Berman, Torsten Bernhardt, Eva Brune, John S. Chipman, Ka Lok Chu, R. William Farebrother, Bernd Fritsche, Daniel Hershkowitz, Jarkko Isotalo, Bernd Kirstein, André Klein, Jarmo Niemelä, Geva Maimon Reid, Timo Mäkeläinen, Lindsey E. McQuade, Aliza K. Miller, Ingram Olkin, Emily E. Rochette, Vera Rosta, Eugénie Roudaia, Burkhard Schaffrin, Hans Schneider, Shayle R. Searle, Daniel N. Selan, Samara F. Strauber, Evelyn M. Styan, J. C. Szamosi, Garry J. Tee, Götz Trenkler, Frank Uhlig, and Jürgen Weiß. We are also very grateful to the librarians in the McGill University Interlibrary Loan and Document Delivery Department for their help in obtaining the source materials for many of our references. The research of George P. H. Styan was supported in part by the Natural Sciences and Engineering Research Council of Canada.

Finally, I thank my wife Cheng, my children Sunny, Andrew, and Alan, and my mother-in-law Yun-Jiao for their understanding, support, and love.

Fuzhen Zhang
September 1, 2004
Fort Lauderdale, Florida

Chapter 0

Historical Introduction: Issai Schur and the Early Development of the Schur Complement

0.0 Introduction and mise-en-scène

In this introductory chapter we comment on the history of the Schur complement from 1812 through 1968 when it was so named and given a notation. As Chandler & Magnus [113, p. 192] point out, "The coining of new technical terms is an absolute necessity for the evolution of mathematics." And so we begin in 1968 when the mathematician Emilie Virginia Haynsworth (1916–1985) introduced a name and a notation for the Schur complement of a square nonsingular (or invertible) submatrix in a partitioned (two-way block) matrix [210, 211].

We then go back fifty-one years and examine the seminal lemma by the famous mathematician Issai Schur (1875–1941) published in 1917 [404, pp. 215–216], in which the *Schur determinant formula* (0.3.2) was introduced. We also comment on earlier implicit manifestations of the Schur complement due to Pierre Simon Laplace, later Marquis de Laplace (1749–1827), first published in 1812, and to James Joseph Sylvester (1814–1897), first published in 1851.

Following some biographical remarks about Issai Schur, we present the *Banachiewicz inversion formula* for the inverse of a nonsingular partitioned matrix which was introduced in 1937 [29] by the astronomer Tadeusz Banachiewicz (1882–1954). We note, however, that closely related results were obtained earlier in 1933 by Ralf Lohan [290], following results in the book [66] published in 1923 by the geodesist Hans Boltz (1883–1947).

We continue with comments on material in the book *Elementary Matrices and Some Applications to Dynamics and Differential Equations* [171], a classic by the three aeronautical engineers Robert Alexander Frazer (1891–1959), William Jolly Duncan (1894–1960), and Arthur Roderick Collar (1908–1986), first published in 1938, and in the book *Determinants and Matrices* [4] by the mathematician and statistician Alexander Craig Aitken (1895–1967), another classic, and first published in 1939.

We introduce the *Duncan inversion formula* (0.8.3) for the sum of two matrices, and the very useful *Aitken block-diagonalization formula* (0.9.1), from which easily follow the *Guttman rank additivity formula* (0.9.2) due to the social scientist Louis Guttman (1916–1987) and the *Haynsworth inertia additivity formula* (0.10.1) due to Emilie Haynsworth.

We conclude this chapter with some biographical remarks on Emilie Haynsworth and note that her thesis adviser was Alfred Theodor Brauer (1894–1985), who completed his Ph.D. degree under Schur in 1928.

This chapter builds on the extensive surveys of the Schur complement published (in English) by Brezinski [73], Carlson [105], Cottle [128, 129], Ouellette [345], and Styan [432], and (in Turkish) by Alpargu [8]. In addition, the role of the Schur complement in matrix inversion has been surveyed by Zielke [472] and by Henderson & Searle [219], with special emphasis on inverting the sum of two matrices, and by Hager [200], with emphasis on the inverse of a matrix after a small-rank perturbation.

0.1 The Schur complement: the name and the notation

The term *Schur complement* for the matrix

$$S - RP^{-1}Q, \tag{0.1.1}$$

where the nonsingular matrix P is the leading submatrix of the complex partitioned matrix

$$M = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}, \tag{0.1.2}$$

was introduced in 1968 in two papers [210, 211] by Emilie Haynsworth published, respectively, in the *Basel Mathematical Notes* and in *Linear Algebra and its Applications.*

The notation

$$(M/P) = S - RP^{-1}Q \tag{0.1.3}$$

for the Schur complement of P in $M = \left(\begin{smallmatrix} P & Q \\ R & S \end{smallmatrix}\right)$ was apparently first used in 1968 by Haynsworth, in the *Basel Mathematical Notes* [210] but not in *Linear Algebra and its Applications* [211], where its first appearance seems to be in the 1970 paper by Haynsworth [212]. This notation does appear, however, in the 1969 paper [131] by Haynsworth with Douglas E. Crabtree in the *Proceedings of the American Mathematical Society* and is still in use today, see e.g., the papers by Brezinski & Redivo Zaglia [88] and N'Guessan [334] both published in 2003; the notation (0.1.3) is also used in the six surveys [8, 73, 128, 129, 345, 432].

The notation $(M|P)$, with a vertical line separator rather than a slash, was introduced in 1971 by Markham [295] and is used in the book by Prasolov [354, p. 17]; see also [296, 332, 343] published in 1972–1980. The notation $M|P$ without the parentheses was used in 1976 by Markham [297].

In this book we will use the original notation (0.1.3) but without the parentheses,

$$M/P = S - RP^{-1}Q, \tag{0.1.4}$$

for the Schur complement of the nonsingular matrix P in the partitioned matrix $M = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}$. This notation (0.1.4) without the parentheses was introduced in 1974 by Carlson, Haynsworth & Markham [106] and seems to be very popular today, see, e.g., the recent books by Ben-Israel & Greville [45, p. 30], Berman & Shaked-Monderer [48, p. 24], and by C. R. Rao & M. B. Rao [378, p. 139], and the recent papers [160, 287, 471].

0.2 Some implicit manifestations in the 1800s

According to David Carlson in his 1986 survey article [105] entitled "What are Schur complements, anyway?":

> The idea of the Schur complement matrix goes back to the 1851 paper [436] by James Joseph Sylvester. It is well known that the entry a_{ij} of [the Schur complement matrix] A, $i = 1, \ldots, m-k$, $j = 1, \ldots, n-k$, is the minor of [the partitioned matrix] M determined by rows $1, \ldots, k, k+i$ and columns $1, \ldots, k, k+j$, a property which was used by Sylvester as his definition. For a discussion of this and other appearances of the Schur complement matrix in the 1800s, see the paper by Brualdi & Schneider [99].

Farebrother [162, pp. 116–117] discusses work by Pierre Simon Laplace, later Marquis de Laplace, and observes that Laplace [273, livre II, §21 (1812); *Œuvres*, vol. 7, p. 334 (1886)] obtained a ratio that we now recognize as the ratio of two successive leading principal minors of a symmetric positive definite matrix. Then the ratio $\det(M)/\det(M_1)$ is the determinant of what we now know as the Schur complement of M_1 in M, see the

Schur determinant formula (0.3.2) below. Laplace [273, §3 (1816); *Œuvres*, vol. 7, pp. 512–513 (1886)] evaluates the ratio $\det(M)/\det(M_1)$ with $n = 3$.

0.3 The lemma and the Schur determinant formula

The adjectival noun "Schur" in "Schur complement" was chosen by Haynsworth because of the lemma (Hilfssatz) in the paper [404] by Issai Schur published in 1917 in the *Journal für die reine und angewandte Mathematik*, founded in Berlin by August Leopold Crelle (1780–1855) in 1826 and edited by him until his death. Often called Crelle's *Journal* this is apparently the oldest mathematics periodical still in existence today [103]; Frei [174] summarizes the long history of the *Journal* in volume 500 (1998).

The picture of Issai Schur facing the opening page of this chapter appeared in the 1991 book *Ausgewählte Arbeiten zu den Ursprüngen der Schur-Analysis: Gewidmet dem großen Mathematiker Issai Schur (1875–1941)* [177, p. 20]; on the facing page [177, p. 21] is a copy of the title page of volume 147 (1917) of the *Journal für die reine und angewandte Mathematik* in which the Schur determinant lemma [404] was published.

This paper [404] is concerned with conditions for power series to be bounded inside the unit circle; indeed a polynomial with roots within the unit disk in the complex plane is now known as a *Schur polynomial*, see e.g., Lakshmikantham & Trigiante [271, p. 49].

The lemma appears in [404, pp. 215–216], see also [71, pp. 148–149], [177, pp. 33–34]. Our English translation, see also [183, pp. 33–34], follows. The Schur complement $S - RP^{-1}Q$ is used in the proof but the lemma holds even if the square matrix P is singular. We refer to this lemma as the *Schur determinant lemma.*

LEMMA. *Let P, Q, R, S denote four $n \times n$ matrices and suppose that P and R commute. Then the determinant* $\det(M)$ *of the* $2n \times 2n$ *matrix*

$$M = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}$$

is equal to the determinant of the matrix $PS - RQ$.

Proof. We assume that the determinant of P is not zero. Then, with I denoting the $n \times n$ identity matrix,

$$\begin{pmatrix} P^{-1} & 0 \\ -RP^{-1} & I \end{pmatrix} \begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \begin{pmatrix} I & P^{-1}Q \\ 0 & S - RP^{-1}Q \end{pmatrix}$$

Taking determinants yields $\det(P^{-1}) \cdot \det(M) = \det(S - RP^{-1}Q)$ and so

$$\begin{aligned}\det(M) &= \det(P)\cdot\det(S - RP^{-1}Q) \qquad (0.3.1)\\ &= \det(PS - PRP^{-1}Q) = \det(PS - RQ).\end{aligned}$$

If, however, $\det(P) = 0$, we replace matrix M with the matrix

$$M_1 = \begin{pmatrix} P + xI & Q \\ R & S \end{pmatrix}.$$

The matrices R and $P + xI$ commute. For the absolute value $|x|$ sufficiently small (but not zero), the determinant of $P + xI$ is not equal to 0 and so $\det(M_1) = \det\big((P + xI)S - RQ\big)$. Letting x converge to 0 yields the desired result. ∎

We may write (0.3.1) as the *Schur determinant formula*

$$\det(M) = \det(P)\cdot\det(M/P) = \det(P)\cdot\det(S - RP^{-1}Q) \qquad (0.3.2)$$

and so determinant is multiplicative on the Schur complement, which suggests the notation M/P for the Schur complement of P in M.

Schur [404, pp. 215–216] used this lemma to show that the complex $2k \times 2k$ determinant

$$\delta_k = \det\begin{pmatrix} P_k & Q_k \\ Q_k^* & P_k^* \end{pmatrix} = \det(P_k P_k^* - Q_k^* Q_k), \quad k = 1, \dots, n, \qquad (0.3.3)$$

where

$$P_k = \begin{pmatrix} a_0 & 0 & \dots & 0 \\ a_1 & a_0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{k-1} & a_{k-2} & \dots & a_0 \end{pmatrix}, \quad Q_k = \begin{pmatrix} a_n & a_{n-1} & \dots & a_{n-k+1} \\ 0 & a_n & \dots & a_{n-k+2} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix},$$

and so $P_k Q_k^* = Q_k^* P_k$, $k = 1, \dots, n$. What are now known as *Schur conditions*,

$$\delta_k > 0, \; k = 1, \dots, n,$$

are necessary and sufficient for the roots of the polynomial

$$f(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n = 0 \qquad (0.3.4)$$

to lie within the unit circle of the complex plane, see e.g., Chipman [116, p. 371 (1950)].

Schur's paper [404] and its sequel [405] were selected by Fritzsche & Kirstein in the *Ausgewählte Arbeiten* [177] as two of the six influential papers considered as "fundamental for Schur analysis"; the book [177] is dedicated to the "great mathematician Issai Schur". The four other papers in [177] are by Gustav Herglotz (1881–1953), Rolf Nevanlinna (1895–1980), Georg Pick (1859–1942), and Hermann Weyl (1885–1955).

0.4 Issai Schur (1875–1941)

Issai Schur was born on 10 January 1875, the son of Golde Schur (née Landau) and the *Kaufmann* Moses Schur, according to Schur's *Biographische Mitteilungen* [406]. In a recent biography of Issai Schur, Vogt [449] notes that Schur used the first name "Schaia" rather than "Issai" until his mid-20s and that his father was a *Großkaufmann.*

Writing in German in [406], Schur gives his place of birth as Mohilew am Dnjepr (Russland)—in English: Mogilev on the Dnieper, Russia. Founded in the 13th century, Mogilev changed hands frequently among Lithuania, Poland, Sweden, and Russia, and was finally annexed to Russia in 1772 in the first partition of Poland [31, p. 155]. By the late 19th century, almost half of the population of Mogilev was Jewish [262]. About 200 km east of Minsk, Mogilev is in the eastern part of the country now known as Belarus (Belorussia, White Russia) and called Mahilyow in Belarusian [306].

In 1888 when he was 13, Schaia Schur, as he was then known [449], went to live with his older sister and brother-in-law in Libau (Kurland), about 640 km northwest of Mogilev. Also founded in the 13th century, Libau (Liepāja in Latvian or Lettish) is on the Baltic coast of what is now Latvia in the region of Courland (Kurland in German, Kurzeme in Latvian), which from 1562–1795 was a semi-independent duchy linked to Poland but with a prevailing German influence [60, 423]. Indeed the German way of life was dominant in Courland in 1888, with mostly German (not Yiddish) being the spoken language of the Jewish community until 1939 [39]. In the late 19th century there were many synagogues in Libau, the Great Synagogue in Babylonian style with three cupolas being a landmark [60].

Schur attended the German-language Nicolai Gymnasium in Libau from 1888–1894 and received the highest mark on his final examination and a gold medal [449]. It was here that he became fluent in German (we believe that his first language was probably Yiddish). In Germany the Gymnasium is a "state-maintained secondary school that prepares pupils for higher academic education" [158]. We do not know why the adjectival noun Nicolai is used here but in Leipzig the *Nikolaischule* was so named because of the adjacent *Nikolaikirche*, which was founded *c.* 1165 and named after Saint Nicholas of Bari [207, 224], the saint who is widely associated with Christmas and after whom Santa Claus in named [248, ch. 7].

In October 1894, Schur enrolled in the University of Berlin, studying mathematics and physics; on 27 November 1901 he passed his doctoral examination *summa cum laude* with the thesis entitled "Über eine Klasse von Matrizen, die sich einer gegebenen Matrix zuordnen lassen" [402]: his thesis adviser was Ferdinand Georg Frobenius (1849–1917). According to Vogt [449], in this thesis Schur used his first name "Issai" for the first time.

Feeling that he "had no chance whatsoever of sustaining himself as a mathematician in czarist Russia" [113, p. 197] and since he now wrote and spoke German so perfectly that one would guess that German was his native language, Schur stayed on in Germany. According to [406], he was *Privatdozent* at the University in Berlin from 1903 till 1913 and *außerordentlicher Professor* (associate professor) at the University of Bonn from 21 April 1913 till 1 April 1916 [425, p. 8], as successor to Felix Hausdorff (1868–1942); see also [276, 425]. In 1916 Schur returned to Berlin where in 1919 he was appointed full professor; in 1922 he was elected a member of the Prussian Academy of Sciences to fill the vacancy caused by the death of Frobenius in 1917. We believe that our portrait of Issai Schur in the front of this book was made in Berlin, *c.* 1917; for other photographs see [362].

Schur lived in Berlin as a highly respected member of the academic community and was a quiet unassuming scholar who took no part in the fierce struggles that preceded the downfall of the Weimar Republic. "A leading mathematician and an outstanding and highly successful teacher, [Schur] occupied for 16 years the very prestigious chair at the University of Berlin" [113, p. 197]. Until 1933 Schur's algebraic school at the University of Berlin was, without any doubt, the single most coherent and influential group of mathematicians in Berlin and among the most important in all of Germany. With Schur as its charismatic leader, the school centered around his research on group representations, which was extended by his students in various directions (soluble groups, combinatorics, matrix theory) [100, p. 25]. "Schur made fundamental contributions to algebra and group theory which, according to Hermann Weyl, were comparable in scope and depth to those of Emmy Amalie Noether (1882–1935)" [353, p. 178].

When Schur's lectures were canceled (in 1933) there was an outcry among the students and professors, for he was respected and very well liked [100, p. 27]. Thanks to his colleague Erhard Schmidt (1876–1959), Schur was able to continue his lectures till the end of September 1935 [353, p. 178], Schur being the last Jewish professor to lose his job at the Universität Berlin at that time [425, p. 8]. Schur's "lectures on number theory, algebra, group theory and the theory of invariants attracted large audiences. On 10 January 1935 some of the senior postgraduates congratulated [Schur] in the lecture theatre on his sixtieth birthday. Replying in mathematical language, Schur hoped that the good relationship between himself and his student audience would remain invariant under all the transformations to come" [353, p. 179].

Indeed Schur was a superb lecturer. His lectures were meticulously prepared and were exceedingly popular. Walter Ledermann (b. 1911) remembers attending Schur's algebra course which was held in a lecture theatre filled with about 400 students [276]: "Sometimes, when I had to be content with a seat at the back of the lecture theatre, I used a pair of opera glasses to get a glimpse of the speaker." In 1938 Schur was pressed to resign from the Prussian Academy of Sciences and on 7 April 1938 he resigned "voluntarily" from the Commissions of the Academy. Half a year later, he had to resign from the Academy altogether [100, p. 27].

The names of the 22 persons who completed their dissertations from 1917–1936 under Schur, together with the date in which the Ph.D. degree was awarded and the dissertation title, are listed in the *Issai Schur Gesammelte Abhandlungen* [71, *Band III*, pp. 479–480]; see also [100, p. 23], [249, p. xviii]. One of these 22 persons is Alfred Theodor Brauer (1894–1985), who completed his Ph.D. dissertation under Schur on 19 December 1928 and with Hans Rohrbach edited the *Issai Schur Gesammelte Abhandlungen* [71]. Alfred Brauer was a faculty member in the Dept. of Mathematics at The University of North Carolina at Chapel Hill for 24 years and directed 21 Ph.D. dissertations, including that of Emilie Haynsworth, who in 1968 introduced the term "Schur complement" (see §0.1 above).

A remark by Alfred Brauer [70, p. xiii], see also [100, p. 28], sheds light on Schur's situation after he finally left Germany in 1939: "When Schur could not sleep at night, he read the *Jahrbuch über die Fortschritte der Mathematik* (now *Zentralblatt MATH*). When he came to Tel Aviv (then British Mandate of Palestine, now Israel) and for financial reasons offered his library for sale to the Institute for Advanced Study in Princeton, he finally excluded the *Jahrbuch* in a telegram only weeks before his death."

Issai Schur died of a heart attack in Tel Aviv on his 66th birthday, 10 January 1941. Schur is buried in Tel Aviv in the Old Cemetery on Trumpeldor Street, which was "reserved for the Founders' families and persons of special note. Sadly this was the only tribute the struggling Jewish Home could bestow upon Schur" [249, p. clxxxvi]; see also [331, 362].

Schur was survived by his wife, medical doctor Regina (née Frumkin, 1881–1965), their son Georg (born 1907 and named after Frobenius), and daughter Hilde (born 1911, later Hilda Abelin-Schur), who in "A story about father" [1] in *Studies in Memory of Issai Schur* [249] writes

> One day when our family was having tea with some friends, [my father] was enthusiastically talking about his work. He said: "I feel like I am somehow moving through outer space. A particular idea leads me to a nearby star on which I decide to land. Upon my arrival I realize that somebody already lives there. Am I disappointed? Of course not. The inhabitant and I are cordially welcoming each other, and we are happy about our common discovery." This was typical of my father; he was never envious.

0.5 Schur's contributions in mathematics

Many of Issai Schur's contributions to linear algebra and matrix theory are reviewed in [152] by Dym & Katsnelson in *Studies in Memory of Issai Schur* [249]. Among the topics covered in [249] are estimates for matrix and integral operators and bilinear forms, the Schur (or Hadamard) product of matrices, Schur multipliers, Schur convexity, inequalities between eigenvalues and singular values of a linear operator, and triangular representations of matrices. Schur is considered as a "pioneer in representation theory" [136], and Haubrich [208] surveys Schur's contributions in linear substitutions, locations of roots of algebraic equations, pure group theory, integral equations, and number theory.

Soifer [425] discusses the origins of certain combinatorial problems nowadays seen as part of Ramsey theory, with special reference to a lemma, now known as Schur's theorem, embedded in a paper on number theory. Included in *Studies in Memory of Issai Schur* [249] are over 60 pages of biographical and related material (including letters and documents in German, with translations in English) on Issai Schur, as well as reminiscences by his former students Bernhard Hermann Neumann (1909–2002) and Walter Ledermann, and by his daughter Hilda Abelin-Schur [1] and his granddaughter Susan Abelin.

In the edited book [183] entitled *I. Schur Methods in Operator Theory and Signal Processing*, Thomas Kailath [252] briefly reviews some of the "many significant and technologically highly relevant applications in linear algebra and operator theory" arising from Schur's seminal papers [404, 405]. For some comments by Paul Erdős (1913–1996) on the occasion of the 120th anniversary of Schur's birthday in 1995, see [159].

0.6 Publication under J. Schur

Issai Schur published under "I. Schur" and under "J. Schur". As is pointed out by Ledermann in his biographical article [276] on Schur, this has caused some confusion: "For example I have a scholarly work on analysis which lists amongst the authors cited both J. Schur and I. Schur, and an author on number theory attributes one of the key results to I. J. Schur."

We have identified 81 publications by Issai Schur which were published before he died in 1941; several further publications by Schur were, however, published posthumously including the book [408] published in 1968. On the title page of the (original versions of the) articles [404, 405], the author is given as "J. Schur"; indeed for all but one of the other 11 papers by Issai Schur that we found published in the *Journal für die reine und angewandte Mathematik* the author is given as "J. Schur". For the lecture notes [407] published in Zürich in 1936, the author is given as J. Schur on the title page and so cited in the preface. For all other publications by Issai Schur that we have found, however, the author is given as "I. Schur", and posthumously as "Issai Schur"; moreover Schur edited the *Mathematische Zeitschrift* from 1917–1938 and he is listed there on the journal title pages as I. Schur.

The confusion here between "I" and "J" probably stems from there being two major styles of writing German: *Fraktur script*, also known as *black letter script* or *Gothic script*, in use since the ninth century and prevailing until 1941 [130, p. 26], and *Roman* or *Latin*, which is common today [237]. According to Mashey [302, p. 28], "it is a defect of most styles of German type that the same character $\mathfrak{J}$ is used for the capitals I (i) and J (j)"; when followed by a vowel it is the consonant "J" and when followed by a consonant, it is "I", see also [46, pp. 4–5], [220, pp. 166–167], [444, p. 397].

The way Schur wrote and signed his name, as in his *Biographische Mitteilungen* [406], his first name could easily be interpreted as "Jssai" rather than "Issai"; see also the signature at the bottom of the photograph in the front of this book and at the bottom of the photograph in the *Issai Schur Gesammelte Abhandlungen* [71, *Band I*, facing page v (1973)]. The official letter, reprinted in Soifer [425, p. 9], dated 28 September 1935 and signed by Kunisch [270], relieving Issai Schur of his duties at the University of Berlin, is addressed to "Jssai Schur"; the second paragraph starts with "Jch übersende Jhnen ... " which would now be written as "Ich übersende Ihnen ... "; see also [249, p. lxxiv (2003)]. Included in the article by Ledermann & Neumann [277, (2003)] are copies of many documents associated with Issai Schur. These are presented in chronological order, with a transcription first, followed by a translation. It is noted there [277, p. lx] that "Schur used Roman script" but "sometimes, particularly in typed official letters after 1933, initial letters I are rendered as J."

0.7 Boltz 1923, Lohan 1933, Aitken 1937, and the Banachiewicz inversion formula 1937

In 1937 the astronomer and mathematician Tadeusz Banachiewicz (1882–1954) established in [29, p. 50] the Schur determinant formula (0.3.2) with P nonsingular,

$$\det(M) = \det\begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \det(P)\cdot\det(S - RP^{-1}Q)\,. \tag{0.7.1}$$

Also in 1937, the mathematician and statistician Alexander Craig Aitken (1895–1967) gave [3, p. 172] "a uniform working process for computing" the triple matrix product $RP^{-1}Q$, and noted explicitly that when the matrix

R is a row vector $-r'$, say, and Q is a column vector q, say, then

$$\det\begin{pmatrix} P & q \\ -r' & 0 \end{pmatrix} / \det(P) = r'P^{-1}q.$$

From (0.7.1), it follows at once that the square matrix M is nonsingular if and only if the Schur complement $M/P = S - RP^{-1}Q$ is nonsingular. We then obtain the *Banachiewicz inversion formula* for the inverse of a partitioned matrix

$$M^{-1} = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}^{-1} = \begin{pmatrix} P^{-1} + P^{-1}Q(M/P)^{-1}RP^{-1} & -P^{-1}Q(M/P)^{-1} \\ -(M/P)^{-1}RP^{-1} & (M/P)^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} P^{-1} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} -P^{-1}Q \\ I \end{pmatrix} (M/P)^{-1} \begin{pmatrix} -RP^{-1} & I \end{pmatrix}. \qquad (0.7.2)$$

Banachiewicz [29, p. 54] appears to have been the first to obtain (0.7.2); his proof used Cracovians, a special kind of matrix algebra in which columns multiply columns, and which is used, for example, in spherical astronomy (polygonometry), geodesy, celestial mechanics, and in the calculation of orbits; see e.g., Bujakiewicz-Korońska & Koroński [101], Ouellette [345, pp. 290–291],

Fourteen years earlier in 1923, the geodesist Hans Boltz (1883–1947) implicitly used partitioning to invert a matrix (in scalar notation), see [66, 181, 225, 240]. According to the review by Forsythe [170] of the book *Die Inversion geodätischer Matrizen* by Ewald Konrad Bodewig [63], Boltz's interest concerned the "inverse of a *geodetic matrix* G in which a large submatrix A is mostly zeros and depends only on the topology of the geodetic network of stations and observed directions. When the directions are given equal weights, A has 6 on the main diagonal and ± 2 in a few positions off the diagonal. Boltz proposed first obtaining A^{-1} (which can be done before the survey), and then using it to obtain G^{-1} by partitioning G; see also Wolf [460]. Bodewig [62] refers to the "method of Boltz and Banachiewicz". Nistor [335] used the "method of Boltz" applied to partitioning in the solution of *normal equations* in statistics; see also Householder [234].

The Banachiewicz inversion formula (0.7.2) appears in the original version of the book *Matrix Calculus* by Bodewig published in 1956 [64, Part IIIA, §2, pp. 188–192] entitled "Frobenius' Relation" and in the second edition, published in 1959 [64, Part IIIA, ch. 2, pp. 217–222] entitled "Frobenius–Schur's Relation". In [65, p. 20], Bodewig notes that it was Aitken who referred him to Frobenius. No specific reference to Frobenius is given in [64, 65]. Lokki [291, p. 22] refers to the "Frobenius–Schur–Boltz–Banachiewicz method for partitioned matrix inversion".

In 1933 Ralf Lohan, in a short note [290] "extending the results of Boltz [66]", solves the system of equations

$$\begin{pmatrix} P & Q \\ R & S \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} v \\ w \end{pmatrix} \tag{0.7.3}$$

for the vectors x and y and explicitly gives the solution as

$$\begin{aligned} x &= \left(P^{-1} + P^{-1}Q(M/P)^{-1}RP^{-1}\right)v - P^{-1}Q(M/P)^{-1}w, \\ y &= -(M/P)^{-1}RP^{-1}v + (M/P)^{-1}w. \end{aligned} \tag{0.7.4}$$

While Lohan [290] does not explicitly present the inversion formula (0.7.2), he does use it to compute the inverse (presented explicitly, correct to 4 decimal places) of a specific real symmetric indefinite 5×5 matrix A with positive and negative integer elements in the range $[-17, +36]$. Letting A_j denote the top left $j \times j$ principal leading submatrix of A with $j = 3, 4$, Lohan [290] first computes A_3^{-1}, and then using A_3^{-1} and the scalar Schur complement A_4/A_3 he obtains A_4^{-1}. His inversion of A is then completed using A_4^{-1} and the scalar Schur complement A/A_4. A similar method was given in 1940 by Jossa [250]; see also Forsythe [170].

Following up on the results of Banachiewicz (1937), the well-known mathematician and statistician Bartel Leendert van der Waerden (1903–1996) gives the formula

$$\begin{pmatrix} P & Q \\ R & S \end{pmatrix}^{-1} = \begin{pmatrix} I & -P^{-1}Q(M/P)^{-1} \\ 0 & (M/P)^{-1} \end{pmatrix} \begin{pmatrix} P^{-1} & 0 \\ -RP^{-1} & I \end{pmatrix} \tag{0.7.5}$$

in a short note [446] in the "Notizen" section of the *Jahresbericht der Deutschen Mathematiker Vereinigung* in 1938. The formula (0.7.5) follows at once from (0.7.2) and from the Schur determinant formula (0.3.2).

0.8 Frazer, Duncan & Collar 1938, Aitken 1939, and Duncan 1944

The three aeronautical engineers Robert Alexander Frazer (1891–1959), William Jolly Duncan (1894–1960) and Arthur Roderick Collar (1908–1986) established the Banachiewicz inversion formula (0.7.2) in their classic book entitled *Elementary Matrices and Some Applications to Dynamics and Differential Equations* [171, p. 113] first published in 1938, just one year after Banachiewicz (1937). The appearance in [171] of the Banachiewicz inversion formula is almost surely its first appearance in a book; the Schur determinant formula also appears here for the special case when the Schur complement is a scalar. We find no mention in [171], however, of Banachiewicz, Boltz or Schur.

Let us consider again the nonsingular partitioned matrix $M = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}$ as above, but now with S nonsingular and where the Schur complement $M/S = P - QS^{-1}R$. Then, in parallel to the Banachiewicz inversion formula (0.7.2) above, we have

$$M^{-1} = \begin{pmatrix} (M/S)^{-1} & -(M/S)^{-1}QS^{-1} \\ -S^{-1}R(M/S)^{-1} & S^{-1} + S^{-1}R(M/S)^{-1}QS^{-1} \end{pmatrix} \tag{0.8.1}$$

with P not necessarily nonsingular (but square so that M is square). When, however, both S and P are nonsingular, then (0.7.2) also holds, i.e.,

$$M^{-1} = \begin{pmatrix} P^{-1} + P^{-1}Q(M/P)^{-1}RP^{-1} & -P^{-1}Q(M/P)^{-1} \\ -(M/P)^{-1}RP^{-1} & (M/P)^{-1} \end{pmatrix}. \tag{0.8.2}$$

Equating the top left-hand corners in (0.8.1) and (0.8.2) yields

$$(M/S)^{-1} = P^{-1} + P^{-1}Q(M/P)^{-1}RP^{-1},$$

or explicitly

$$(P - QS^{-1}R)^{-1} = P^{-1} + P^{-1}Q(S - RP^{-1}Q)^{-1}RP^{-1}, \tag{0.8.3}$$

which we refer to as the *Duncan inversion formula.* We believe that (0.8.3) was first explicitly established by William Jolly Duncan in 1944, see [151, equation (4.10), p. 666]. See also the 1946 paper by Guttman [197]. Piegorsch & Casella [351] call (0.8.3) the *Duncan–Guttman inverse* while Grewal & Andrews [189, p. 366] call (0.8.3) the *Hemes inversion formula* with reference to Bodewig [64, p. 218 (1959)], who notes that (0.8.3) "has, with another proof, been communicated to the author by H. Hemes."

The survey paper by Hager [200] focuses on the special case of (0.8.3) when $S = I$

$$(P - QR)^{-1} = P^{-1} + P^{-1}Q(I - RP^{-1}Q)^{-1}RP^{-1}, \tag{0.8.4}$$

which he calls the *inverse matrix modification formula* and observes that the matrix $I - RP^{-1}Q$ is often called the *capacitance matrix*, see also [356]. Hager [200] notes that (0.8.4) is frequently called the *Woodbury formula* and the special case of (0.8.4) when Q and R are vectors the *Sherman–Morrison formula*, following results by Sherman & Morrison [416, 417, 418] and Woodbury [325, 461] in 1949–1950; see also Bartlett [36] and our Chapter 6 on Schur complements in statistics and probability.

When P, Q, R and S are all $n \times n$ as in the Schur determinant lemma in §0.3 above, and if P, Q, R and S are all nonsingular, then Aitken [4, Example #27, p. 148] also obtained the additional formula involving four Schur complements:

$$M^{-1} = \begin{pmatrix} (M/S)^{-1} & (M/Q)^{-1} \\ (M/R)^{-1} & (M/P)^{-1} \end{pmatrix}, \tag{0.8.5}$$

where $M/Q = R - SQ^{-1}P$ and $M/R = Q - PR^{-1}S$. The formula (0.8.5) was obtained by Aitken in his classic book *Determinants and Matrices* [4] first published in 1939, just one year after Frazer, Duncan & Collar [171] was first published; the formula (0.8.5) appears in Example #27 in the section entitled "Additional Examples" in [4, p. 148].

Duncan [151, equation (3.3), p. 664] also gives the Banachiewicz inversion formula explicitly and notes there that it "has been given by A. C. Aitken in lectures to his students, together with some alternative equivalent forms which are now included in this paper", see also [65, p. 20].

0.9 The Aitken block-diagonalization formula 1939 and the Guttman rank additivity formula 1946

With P nonsingular, the useful *Aitken block-diagonalization formula*

$$\begin{pmatrix} I & 0 \\ -RP^{-1} & I \end{pmatrix} \begin{pmatrix} P & Q \\ R & S \end{pmatrix} \begin{pmatrix} I & -P^{-1}Q \\ 0 & I \end{pmatrix} = \begin{pmatrix} P & 0 \\ 0 & M/P \end{pmatrix} \tag{0.9.1}$$

was apparently first established explicitly by Aitken and first published in 1939, see [4, ch. III, §29]. In (0.9.1), neither M nor S need be square.

While the Aitken formula (0.9.1) holds even if neither M nor S is square, when both M and S are square, (0.9.1) immediately yields the Schur determinant formula (0.3.2), and when M is square and nonsingular, (0.9.1) immediately yields the Banachiewicz inversion formula (0.7.2).

From the Aitken formula (0.9.1) we obtain at once the *Guttman rank additivity formula*

$$\operatorname{rank}(M) = \operatorname{rank}(P) + \operatorname{rank}(M/P),$$

or equivalently

$$\operatorname{rank}\begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \operatorname{rank}(P) + \operatorname{rank}(S - QP^{-1}R), \tag{0.9.2}$$

which we believe was first established in 1946 by the social scientist and statistician Louis Guttman (1916–1987) in [197, p. 339].

0.10 Emilie Virginia Haynsworth (1916–1985) and the Haynsworth inertia additivity formula

Emilie Haynsworth, in addition to introducing the term Schur complement in [210, 211], also showed there that inertia is "additive on the Schur complement". The *inertia* or *inertia triple* of the partitioned Hermitian matrix

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}$$

is defined to be the ordered integer triple

$$\text{In}(H) = \{\pi, \nu, \delta\},$$

where the nonnegative integers $\pi = \pi(H)$, $\nu = \nu(H)$, and $\delta = \delta(H)$ give the numbers, respectively, of positive, negative and zero eigenvalues of H. Here H_{11} is nonsingular and H_{12}^* is the conjugate transpose of H_{12}. This leads to the *Haynsworth inertia additivity formula*

$$\text{In}(H) = \text{In}(H_{11}) + \text{In}(H/H_{11}),$$

or equivalently

$$\text{In}\begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix} = \text{In}(H_{11}) + \text{In}(H_{22} - H_{12}^* H_{11}^{-1} H_{12}), \tag{0.10.1}$$

proved in 1968, apparently for the first time, by Haynsworth [210, 211]. From (0.10.1), it follows at once that rank is additive on the Schur complement in a Hermitian matrix. As Guttman showed, see (0.9.2) above, this rank additivity holds more generally: H need not even be square—we need only that H_{11} be square and nonsingular. As we will see in Chapter 6, however, such rank additivity also holds in a Hermitian matrix when H_{11} is rectangular or square and singular but with the generalized Schur complement $H_{22} - H_{12}^* H_{11}^- H_{12}$, where H_{11}^- is a generalized inverse of H_{11}; moreover inertia additivity then also holds provided H_{11} is square.

To prove the Haynsworth inertia additivity formula (0.10.1) we apply the Aitken factorization formula (0.9.1) to the Hermitian matrix H with H_{11} square and nonsingular, then we have

$$\begin{pmatrix} I & 0 \\ -H_{12}^* H_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix} \begin{pmatrix} I & -H_{11}^{-1} H_{12} \\ 0 & I \end{pmatrix} = \begin{pmatrix} H_{11} & 0 \\ 0 & H/H_{11} \end{pmatrix},$$

which immediately leads to (0.10.1) by Sylvester's Law of Inertia: The inertia $\text{In}(H) = \text{In}(THT^*)$ for any nonsingular matrix T, see also §1.3 of Chapter 1.

Emilie Virginia Haynsworth was born on 1 June 1916 and died on 4 May 1985, both at home in Sumter, South Carolina. As observed in the obituary article [108] by Carlson, Markham & Uhlig, "In her family there have been Virginia Emilies or Emilie Virginias for over 200 years. From childhood on, Emilie had a strong and independent mind, so that her intellectual pursuits soon gained her the respect and awe of all her relatives and friends".

Throughout her life Emilie Haynsworth was eager to discuss any issue whatsoever. From Carlson, Markham & Uhlig [108] we quote Philip J. Davis (b. 1923): "She was a strong mixture of the traditional and the unconventional and for years I could not tell beforehand on what side of the line she would locate a given action". In *The Education of a Mathematician* [144, p. 146], Davis observes that Emilie Haynsworth "had a fine sense of mathematical elegance—a quality not easily defined. Her research can be found in a number of books on advanced matrix theory under the topic: 'Schur complement'. Emilie taught me many things about matrix theory."

The portrait of Emilie Haynsworth reproduced on page ix in the frontal matter of this book is on the Auburn University Web site [214] and in the book *The Education of a Mathematician* by Philip J. Davis [144] We conjecture that the portrait was made *c.* 1968, the year in which the term Schur complement was introduced by Haynsworth [210, 211].

In 1952 Emilie Haynsworth received her Ph.D. degree in mathematics at The University of North Carolina at Chapel Hill with Alfred Brauer as her dissertation adviser. We note that Issai Schur was Alfred Brauer's Ph.D. dissertation adviser and that the topic of Haynsworth's dissertation was determinantal bounds for diagonally dominant matrices. From 1960 until retirement in 1983, Haynsworth taught at Auburn University (Auburn, Alabama) "with a dedication which honors the teaching profession" [108] and supervised 18 Ph.D. students.

The mathematician Alexander Markowich Ostrowski (1893–1986), with whom Haynsworth co-authored the paper [216] on the inertia formula for the apparently not-then-yet-publicly-named Schur complement, wrote the following upon her death:

> I lost a very good, life-long friend and mathematics [lost] an excellent scientist. I remember how on many occasions I had to admire the way in which she found a formulation of absolute originality.

Chapter 1

Basic Properties of the Schur Complement

1.0 Notation

Most of our notation is standard, and our matrices are complex or real (though greater algebraic generality is often possible). We designate the set of all $m \times n$ matrices over $\mathbb{C}$ (or $\mathbb{R}$) by $\mathbb{C}^{m\times n}$ (respectively $\mathbb{R}^{m\times n}$), and denote the conjugate transpose of a matrix A by $A^* = (\bar{A})^T$. A matrix A is *Hermitian* if $A^* = A$, and a Hermitian matrix is *positive semidefinite* (*positive definite*) if all its eigenvalues are nonnegative (positive). The *Löwner partial order* $A \geq B$ ($A > B$) on Hermitian matrices means that $A - B$ is positive semidefinite (positive definite). For $A \in \mathbb{C}^{m\times n}$, we denote the *matrix absolute value* by $|A| = (A^*A)^{1/2}$. A nonsingular square matrix has *polar decompositions* $A = U\,|A| = |A^*|\,U$ in which the positive definite factors $|A|$ and $|A^*|$, and the unitary factor $U = A\,|A|^{-1} = |A^*|^{-1} A$ are uniquely determined; if A is singular then the respective positive semidefinite factors $|A|$ and $|A^*|$ are uniquely determined and the left and right unitary factor U may be chosen to be the same, but U is not uniquely determined. Two matrices A and B of the same size are said to be **-congruent* if there is a nonsingular matrix S of the same size such that $A = SAS^*$; *-congruence is an equivalence relation. We denote the (multi-) set of eigenvalues of A (its *spectrum*) by $s(A) = \{\lambda_i(A)\}$ (including multiplicities).

1.1 Gaussian elimination and the Schur complement

One way to solve an $n \times n$ system of linear equations is by row reduction—Gaussian elimination that transforms the coefficient matrix into upper triangular form. For example, consider a homogeneous system of linear equations $Mz = 0$, where M is an $n \times n$ coefficient matrix with a nonzero $(1,1)$ entry. Write $M = \begin{pmatrix} a & b^T \\ c & D \end{pmatrix}$, where b and c are column vectors of size $n-1$, D is a square matrix of size $n-1$, and $a \neq 0$. The equations

$$Mz = 0 \quad \text{and} \quad \begin{pmatrix} a & b^T \\ 0 & D - ca^{-1}b^T \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

are equivalent, so the original problem reduces to solving a linear equation system of size $n-1$: $(D - ca^{-1}b)y = 0$.

This idea extends to a linear system $Mz = 0$ with a nonsingular leading principal submatrix. Partition M as

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{1.1.1}$$

suppose A is nonsingular, and partition $z = \begin{pmatrix} x \\ y \end{pmatrix}$ conformally with M. The linear system $Mz = 0$ is equivalent to the pair of linear systems

$$Ax + By = 0 \tag{1.1.2}$$

$$Cx + Dy = 0 \tag{1.1.3}$$

If we multiply (1.1.2) by $-CA^{-1}$ and add it to (1.1.3), the vector variable x is eliminated and we obtain the linear system of smaller size

$$(D - CA^{-1}B)y = 0.$$

We denote the matrix $D - CA^{-1}B$ by M/A and call it the *Schur complement of A in M*, or *the Schur complement of M relative to A*. In the same spirit, if D is nonsingular, the Schur complement of D in M is

$$M/D = A - BD^{-1}C.$$

For a non-homogeneous system of linear equations

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix},$$

we may use Schur complements to write the solution as (see Section 0.7)

$$x = (M/D)^{-1}(u - BD^{-1}v), \quad y = (M/A)^{-1}(v - CA^{-1}u).$$

The Schur complement is a basic tool in many areas of matrix analysis, and is a rich source of matrix inequalities. The idea of using the Schur complement technique to deal with linear systems and matrix problems is