

John von Neumann's Analysis of Gaussian Elimination and the Origins of Modern Numerical Analysis*

Joseph F. Grcar[†]

Abstract. Just when modern computers (digital, electronic, and programmable) were being invented, John von Neumann and Herman Goldstine wrote a paper to illustrate the mathematical analyses that they believed would be needed to use the new machines effectively and to guide the development of still faster computers. Their foresight and the congruence of historical events made their work the first modern paper in numerical analysis. Von Neumann once remarked that to found a mathematical theory one had to prove the first theorem, which he and Goldstine did for the accuracy of mechanized Gaussian elimination—but their paper was about more than that. Von Neumann and Goldstine described what they surmized would be the significant questions once computers became available for computational science, and they suggested enduring ways to answer them.

Key words. backward error, CFL condition, computer architecture, computer arithmetic, computer programming, condition number, decomposition paradigm, Gaussian elimination, history, matrix norms, numerical analysis, rounding error analysis, stability, stochastic linear algebra, von Neumann

AMS subject classifications. 01-08, 65-03, 65F05, 65F35, 65G50, 65M12, 68-03

DOI. 10.1137/080734716

I	The First Modern Paper in Numerical Analysis	608
1.1	Introduction	608
1.2	Overview	610
1.3	Heritage versus History	610
1.4	Heritage of Error Analysis	611
2	Von Neumann and Goldstine	612
2.1	Introduction to Biographical Material	612
2.2	Lives through World War II	612
2.3	Von Neumann's Constant	615
2.4	Open Source Computing	618
2.5	Rounding Error Brouhaha	620
2.6	A Dog Named Inverse	622
2.7	Stump Speech	625
2.8	Numerical Analysis as Mathematics	626
2.9	Good Start for Numerical Analysis	629

*Received by the editors September 8, 2008; accepted for publication (in revised form) July 6, 2010; published electronically November 7, 2011.
<http://www.siam.org/journals/sirev/53-4/73471.html>

[†]6059 Castlebrook Drive, Castro Valley, CA 94552-1645 (jfgrcar@comcast.net, na.grcar@na-net.ornl.gov).

2.10	Von Neumann and Turing	630
2.11	Final Years	633
3	Inversion Paper Highlights	634
3.1	Introduction to Research Summaries	634
3.2	Mathematical Stability	634
3.3	Machine Arithmetic	637
3.4	Matrix Norms	641
3.5	Decomposition Paradigm	642
3.5.1	Triangular Factoring	642
3.5.2	Small A Priori Error Bounds	645
3.5.3	A Posteriori Bounds	647
3.6	Normal Equations	650
3.7	Condition Numbers	651
3.8	Backward Errors	653
3.9	Accuracy Criterion	655
3.10	Stochastic Linear Algebra	656
3.11	Remembering the Inversion Paper Imperfectly	658
4	Synopsis of the Error Analysis	658
4.1	Algorithm	658
4.2	Plan of the Error Analysis	659
4.3	(Step 1) Triangular Factoring	659
4.3.1	(Step 1) Factoring Algorithm	660
4.3.2	(Step 1, Part 1) Proof Factoring Succeeds	661
4.3.3	(Step 1, Part 2) Error of Factoring	662
4.4	Programming with Scale Factors	663
4.5	(Step 2') Inverting the Triangular Matrix	664
4.6	(Step 3') Inverting the Diagonal	665
4.7	(Step 4') Forming the Scaled Inverse	665
4.8	(Step F) Bounding the Residual of the Inverse	666
4.9	Forward, Backward, or Residual Analysis	667
5	Conclusion	667
	Epilogue	668
	Notes on Biography, Computer History, and Historiography	669
	Acknowledgments	670
	References	670

I. The First Modern Paper in Numerical Analysis.

I.1. Introduction.

In the older days the objective of the applied mathematician was to reduce physical problems to mathematical form, and then to show how the solution of the mathematical problem could be expressed in terms of known functions—particularly in terms of a finite number of the elementary functions.

— George Stibitz¹ [263, p. 15]

¹George Stibitz 1904–1995 [173] built relay calculators at Bell Telephone Laboratories.

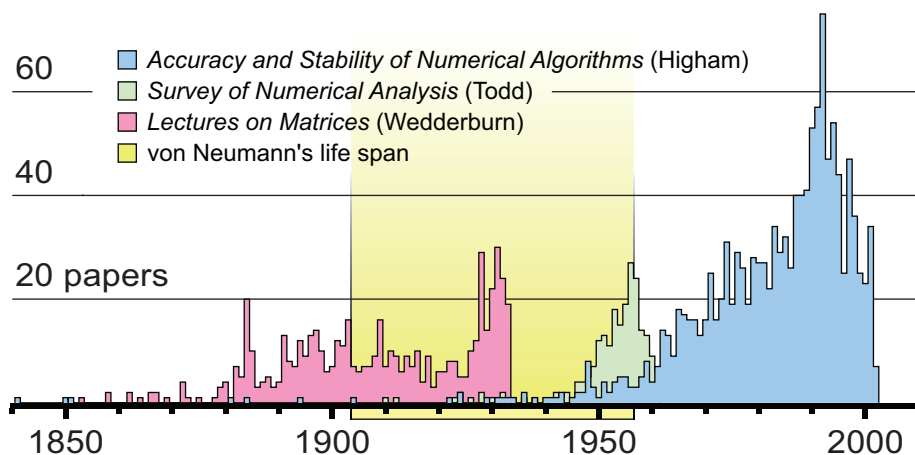


Fig. 1.1 Illustrating that “numerical mathematics stayed as Gauss left it until World War II” [115, p. 287], a histogram of Wedderburn’s bibliography of matrix theory [319] has little overlap with Todd’s and Higham’s bibliographies for numerical linear algebra [273, Chaps. 2, 6, 8, 10, 11], [140]. A bibliography similar to Todd’s was given by Forsythe [96]. Carl Friedrich Gauss died in 1855, and World War II began in 1939.

Before computers, numerical analysis consisted of stopgap measures for the physical problems that could not be analytically reduced. The resulting hand computations were increasingly aided by mechanical tools which are comparatively well documented, but little was written about numerical algorithms because computing was not considered an archival contribution.² “The state of numerical mathematics stayed pretty much the same as Gauss left it until World War II” [115, p. 287] (Figure 1.1). “Some astronomers and statisticians did computing as part of their research, but few other scientists were numerically oriented. Among mathematicians, numerical analysis had a poor reputation and attracted few specialists” [10, pp. 49–50]. “As a branch of mathematics, it probably ranked the lowest, even below statistics, in terms of what most university mathematicians found interesting” [141, p. 316].

In this environment John von Neumann and Herman Goldstine wrote the first modern paper on numerical analysis, “Numerical Inverting of Matrices of High Order” [314],³ and they audaciously published the paper in the journal of record for the American Mathematical Society (AMS). The inversion paper was part of von Neumann’s efforts to create a mathematical discipline around the new computing machines. Gaussian elimination was chosen to focus the paper, but matrices were not its only subject. The paper was the first to distinguish between the stability of a mathematical problem and of its numerical approximation, to explain the significance in this context of the “Courant criterium” (later CFL condition), to point out the advantages of computerized mixed precision arithmetic, to use a matrix decomposition to prove the accuracy of a calculation, to describe a “figure of merit” for calculations that became the matrix condition number, and to explain the concept of inverse, or

²Hand computers were discussed by Campbell-Kelly et al. [46] and Grier [132]. Computing devices were discussed by Aspray [9], Baxandall [25], Bush and Caldwell [43], Eckert [87], Martin [185], and Murray [200, 201].

³I will refer to [314] as the “inversion paper” and cite it without a reference number in braces, e.g., {p. 1021}.

backward, error. The inversion paper thus marked the first appearance in print of many basic concepts in numerical analysis.

The inversion paper may not be the source from which most people learn of von Neumann's ideas, because he disseminated his work on computing almost exclusively outside refereed journals. Such communication occurred in meetings with the many researchers who visited him at Princeton and with the staff of the numerous industrial and government laboratories whom he advised, in the extemporaneous lectures that he gave during his almost continual travels around the country, and through his many research reports which were widely circulated, although they remained unpublished.⁴ As von Neumann's only archival publication about computers, the inversion paper offers an integrated summary of his ideas about a rapidly developing field at a time when the field had no publication venues of its own.

The inversion paper was a seminal work whose ideas became so fully accepted that today they may appear to lack novelty or to have originated with later authors who elaborated on them more fully. It is possible to trace many provenances to the paper by noting the sequence of events, similarities of presentation, and the context of von Neumann's activities.

1.2. Overview. This review begins in section 2 by sketching the careers of von Neumann and Goldstine. The highlights of the inversion paper are then examined in almost a dozen vignettes in section 3. In section 4, a synopsis of the error analysis for matrix inversion explains the material that seems to have discouraged many readers. Sections 2, 3, and 4 are increasingly technical, but they can be read in any order and are best read separately because of their differing subject matter. The remainder of this introduction explains the historical narrative and the inversion paper's namesake analysis of matrix inversion.

1.3. Heritage versus History. There are two versions of the mathematical past: history recounts the development of ideas in the context of contemporary associations, while heritage is reinterpreted work that embodies the present state of mathematical knowledge [127]. Very little mathematics began in the form in which it is presently known; how it changed to become the heritage is the history. For example, person A may create concept X which person B later modifies to Y , which then appears in a textbook. Evidently, Y is the heritage, but the textbook might judge either A or B to be the inventor, and would be wrong in each case.

The heritage of "Gaussian elimination" is a collection of algorithms interpreted as decomposing square matrices into triangular factors for various uses. The most general such decomposition of a matrix A is

$$(1.1) \quad PAQ = LDU,$$

where L , U are, respectively, lower and upper triangular matrices with unit diagonal entries, D is a diagonal matrix, and P , Q are variously chosen permutation matrices that are often suppressed by assuming A is suitably ordered.

It may be a surprise to learn there was no "Gaussian elimination" before the mid-20th century [128].⁵ That is, the present heritage is very recent. Algebra textbooks

⁴Von Neumann's visitors, consulting, travels, and reports are, respectively, discussed in [115, p. 292], [10, p. 246], [181, pp. 366–367], and [181, pp. 307–308]. For example, his contributions to the Monte Carlo method [10, pp. 110–117] and dual linear programs [67, p. 24], [165, p. 85] were made entirely in letters and private meetings.

⁵I follow current usage by writing "Gaussian elimination," whereas the inversion paper wrote "elimination."

taught a method called “elimination” to solve simultaneous linear equations, while professional computers performed algorithms without algebraic notation to solve normal equations for the method of least squares. Human computers knew the algorithms by names such as Gauss’s procedure, Doolittle’s method, the square-root method, and many others.

The inversion paper helped create the present heritage. Von Neumann and Goldstine were among a few early authors who did not cite similar work when they used matrix algebra to describe the algorithms. Whereas the other authors began from the professional methods, instead, von Neumann and Goldstine showed how school-book elimination calculates $A = L(DU)$, which led them to suggest the more balanced $A = LDU$.

1.4. Heritage of Error Analysis. The mathematical heritage includes methods to study the errors of calculation. Von Neumann and Goldstine assembled many concepts that they thought might help to analyze algorithms. Consequently much of the history of analysis passes through either themselves or the inversion paper, hence this review. Readers interested in the inversion paper for specific results should be aware that the terms used in this section are anachronisms, in that they developed after von Neumann and Goldstine.

Many problems reduce to equations $r(d, s) = 0$, where r is a given function, d is given data, and s is the solution. If s' is an approximate solution, then the various errors are

$$\begin{aligned} \text{forward error, } f: \quad & s' = s + f, \\ \text{backward error, } b: \quad & r(d + b, s') = 0, \\ \text{residual error:} \quad & r(d, s') \neq 0. \end{aligned}$$

The forward error is the genuine error of the approximate solution. The backward error interprets the approximate solution as solving a perturbed problem.

Rounding error analysis finds bounds for one of the three kinds of errors when the approximate solution comes from machine calculation. Bounds are of the form

$$\begin{aligned} \text{a priori error bound:} \quad & B(d)u, \\ \text{a posteriori error bound:} \quad & B'(d, s')u, \end{aligned}$$

where B, B' are functions, and u is the magnitude of the worst rounding error in the fractional arithmetic of the computer. Error bounds that are formulas only of the data are called “a priori” because $B(d)u$ predicts whether a calculation *will be* accurate. Bounds that are formulas of the computed solution are called “a posteriori” because $B'(d, s')u$ can be evaluated only after the calculation to determine whether it *was* accurate. There are three types of error and two types of error bounds, giving in principle six types of rounding error analyses.

Following von Neumann and Goldstine, overbars mark machine-representable quantities. If \bar{A} is a symmetric, positive definite (SPD) matrix, then a calculation of the decomposition in (1.1) produces \bar{D} and \bar{L} , so that (suppressing permutations)

$$(1.2) \quad \bar{A} = \bar{L}\bar{D}\bar{L}^t + E.$$

An E always exists for any \bar{L} and \bar{D} . When \bar{L} and \bar{D} are calculated by a procedure close to that which Gauss actually used, then von Neumann and Goldstine showed E is the sum of the errors incurred on each elimination step. By this trick, they constructed an a priori error bound for E . Note that (1.2) can be interpreted as making E either

$$\begin{aligned} \text{backward error: } & (\bar{A} - E) - \bar{L}\bar{D}\bar{L} = 0 \quad \text{or} \\ \text{residual error: } & \bar{A} - \bar{L}\bar{D}\bar{L}^t = E. \end{aligned}$$

Thus, whether von Neumann and Goldstine performed an “a priori backward error analysis” or an “a priori residual error analysis” is a matter of later interpretation. Their proof is given in Theorem 4.1 and (4.7) at the end of this paper.

Von Neumann and Goldstine continued from (1.2) to study matrix inversion. Because the decomposition is the essence of the heritage, von Neumann and Goldstine also contributed to the history of analysis with respect to other uses of the decomposition, such as solving equations. Much of the present heritage was fully developed within about thirty years after their paper.

2. Von Neumann and Goldstine. This chapter explains the circumstances that led von Neumann and Goldstine to write the inversion paper.

2.1. Introduction to Biographical Material. Histories of technology have been classified as internalist, externalist, and contextualist.⁶ This chapter is a contextualist history of the genesis of numerical analysis as the mathematical part of computational science. Many people have noted that the mathematics of computing was reinvented in the 1940s: Goldstine [115] found that computing changed little until World War II; von Neumann [300] and Rees [229] believed new types of analysis were needed; Fox [102] and Householder [149] understood that “numerical analysis” was chosen to name the new subject, and Householder [149], Parlett [217], and Traub [276] remarked that von Neumann and Goldstine wrote the first paper. Their thesis is supported here by information culled from the secondary literature about the early years of computer science. Sifting that literature for pertinence to the inversion paper places a new emphasis on some aspects: that von Neumann used his computer project to encourage research in computers and computing, the curious period of the von Neumann constant, and the acquaintanceship of von Neumann and Alan Turing.

2.2. Lives through World War II. John von Neumann (Figure 2.1) was born into a prominent Budapest family during the Christmas holidays of 1903. He attended a secondary school that would become famous for its Nobel laureate alumni, and when his mathematical talents were recognized the school arranged for tutoring from Mihály Fekete and Gábor Szegő. Von Neumann’s university education took place in Berlin and Zurich in the high tech discipline of his time, chemical engineering, and he simultaneously enrolled in mathematics at Budapest, where he famously attended classes only during final exams [10, p. 7].⁷

⁶These terms were defined by Pannabecker [215] and Staudenmaier [255], and they were used to characterize computer history by Campbell-Kelly [44].

⁷Fekete 1886–1957 was an assistant professor and Szegő 1895–1985 was a student at the University of Budapest. They eventually joined the Hebrew University of Jerusalem and Stanford University, respectively. Von Neumann’s chemical engineering Masters thesis [290] appears to be lost. The Masters qualification was then the terminal engineering degree. Aftalion [3, pp. 102–107] describes the technical preeminence of the German chemical industry.



Fig. 2.1 *John von Neumann in March 1947. Courtesy of the Los Alamos National Laboratory Archives.*

“The last of the great mathematicians” [79], von Neumann began his career working for another, as David Hilbert’s assistant in 1926.⁸ Over the next dozen years von Neumann wrote upwards of ten papers annually (Figure 2.2). Many either originated or reoriented major areas of mathematics, among them set theory, game theory, Hilbert spaces, operator theory and operator algebras, foundations of quantum mechanics [167], harmonic analysis, Lie groups and Hilbert’s fifth problem, measure theory, statistical mechanics, and economic theory. During and after World War II came automata theory, computer science, linear programming, numerical analysis, and computational science in the form of hydrodynamics, especially for meteorology and shock physics. This work was done while von Neumann was an assistant or part-time faculty member at universities in Göttingen, Berlin, Hamburg, and Princeton, and from 1933 as a member of the Institute for Advanced Study (IAS).

Coincidentally, Herman Goldstine’s hometown of Chicago resembled Budapest as a fast growing hub for agricultural commerce. Goldstine was born in Chicago in 1913 and was educated entirely in the city, first in public schools and then at the University of Chicago. In the days when most innovative research was done in Europe [27, p. 31], Goldstine’s pedigree as a homegrown mathematician was impeccable, having studied the calculus of variations [116] under Lawrence Graves, who came from the

⁸Hilbert 1862–1943 [231] influenced many branches of mathematics either directly via his work or indirectly by his advocacy of abstraction and axiomatization.

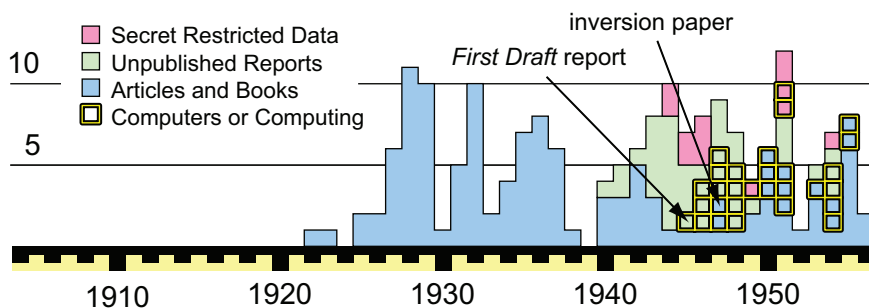


Fig. 2.2 Histogram of John von Neumann's publications, from the enumeration in [10], which is a superset of the collected works [266]. Restricted items are likely undercounted. The inversion paper was the earliest archival publication about computers by von Neumann or, likely, by anyone. Computing papers do not include automata theory. The gaps reflect that in 1930 von Neumann began working in both the U.S. and Europe, in 1933 he obtained a full position in the U.S., in 1939 World War II started, and in 1951–1952 he was appointed to four boards overseeing government research.

mathematical family tree of Gilbert Bliss.⁹ Goldstine declined a postdoctoral position at the IAS [181, p. 277] to begin his career in Chicago and Ann Arbor.

World War II abruptly changed the careers of both von Neumann and Goldstine. Von Neumann easily acquired expertise in explosives and shock physics from his background in chemical engineering, and he understood what today would be called operations research from his boyhood administrative training given by his father [317, pp. 23–25, 33], all of which recommended him to several military advisory panels. Von Neumann shared an ability for mental calculations with his maternal grandfather [181, p. 50], but his military work involved practical questions about ordnance that led to a professional interest in computation. By 1944 he was leading an IAS project in numerical mathematics sponsored by the wartime Applied Mathematics Panel [10, p. 27] and soliciting advice on computing technology for the Manhattan Project [298]. Because of his many advisory duties, von Neumann was one of the few scientists allowed to come and go from Los Alamos.

Meanwhile, Goldstine had already set applied mathematics on a new course. His introduction to the field was at the University of Chicago, where besides the calculus of variations he studied mathematical astronomy and taught Bliss's class on ballistics [117, p. 8]. After joining the U.S. Army in 1942, he oversaw production of firing tables at the University of Pennsylvania, where Aberdeen Proving Ground made ballistics calculations with a differential analyzer and a hundred human computers. He soon persuaded the Army to build John Mauchly and J. Presper Eckert's electronic calculator, ENIAC, for which Goldstine became the program manager.¹⁰

Von Neumann and Goldstine first met at the Aberdeen railroad station in 1944 [115, p. 182]. Thereafter, until von Neumann's death—barely a dozen years later—

⁹Bliss 1876–1951 [190] was highly regarded as an authority because he was entrusted to study the apportionment of seats in the U. S. House of Representatives (as was von Neumann) [14]. Bliss's student Graves 1896–1973 introduced functional analytic methods to the study of the calculus of variations at the University of Chicago.

¹⁰The ballistic calculations and the start of the ENIAC project are described in [27, pp. 24–37], [115, p. 149], [132, pp. 258–261], [188, pp. 52–61], and [220]. Eckert and Mauchly's proposal had been discarded by the university and needed to be reconstructed from a stenographer's notes once Goldstine took interest.

the two made common cause to promote computing. Their partnership began when Goldstine involved von Neumann in planning ENIAC's successor, resulting in the 1945 report about what came to be called *von Neumann machines* [299]. One of von Neumann's special talents was seeing through to the heart of problems, whether in pure mathematics or in advanced technology. Von Neumann borrowed ideas and terminology from discussions of cybernetics with Warren McCulloch and Walter Pitts [189] and Norbert Wiener [322] to make abstract the plans at the University of Pennsylvania.¹¹ He described in general terms the minimum that was needed technically to build a practical, universal machine.

The genius of von Neumann is that he came to the subject at the right time with a critical view and he selected out of a lot of possibilities what was really important.

— Konrad Zuse¹² [340, p. 19]

Johnny von Neumann went off to Los Alamos for the summer as he always did, and he wrote me a series of letters. Those letters were essentially the thing called the *First draft of a report on the EDVAC*. I bashed those letters together into a document without any footnote references to who was responsible for what. It formed a blueprint for the engineers and people to use. Then, it somehow took off in a peculiar way. People... began to talk about this new idea, and letters kept pouring into the Ordnance Office asking for permission to have a copy of this report. And pretty soon, we had distributed a few hundred.

— Herman Goldstine [27, p. 34]

Von Neumann's description of what did not yet exist, the architecture of a modern computer, remains intelligible and relevant to this day: separate facilities for arithmetic, control, and memory repeating a fetch-and-execute cycle of consecutive instructions alterable by conditional branching (modern terms).¹³ The 1945 report and another by Arthur Burks, Goldstine, and von Neumann [42] in 1946 were the direct antecedents of most computers built through 1960, and indirectly of nearly all the machines built thereafter.¹⁴ The *First Draft* report is arguably the most important document in computer science, but it was significant beyond its contents because, as the first publication of its kind, it ensured that the stored program would not be among the intellectual property that was subject to patent.

2.3. Von Neumann's Constant. The University of Pennsylvania's Moore School of Electrical Engineering stood at the apex of computer research with the commissioning of ENIAC in 1946 [47, p. xiv]. Howard Aiken, John Atanasoff and Clifford Berry, Charles Babbage and Ada Lovelace, Tommy Flowers and Maxwell Newman, George Stibitz, Konrad Zuse, and perhaps others who are now less prominent had previously

¹¹Aspray [8], [10, p. 183] describes von Neumann's participation in meetings about cybernetics with McCulloch 1898–1969 and Pitts 1923–1969, who together wrote the first paper on neural nets (modern terminology), and with Wiener 1894–1964, who pioneered what is now systems theory [156]. Goldstine describes von Neumann's participation in planning what was supposed to follow ENIAC at Pennsylvania [27, pp. 33–34], [115, p. 188].

¹²Zuse 1910–1995 built a relay computer with tape control in 1941 [238].

¹³Rather than the original name, “storage,” von Neumann introduced the cybernetically inspired, anthropomorphic name, “memory.” Remember von Neumann the next time you buy memory.

¹⁴Burks 1915–2008 was a logician and member of the ENIAC project. Campbell-Kelly [45, pp. 150–151] and Ceruzzi [52, p. 196] position the *First Draft* at the root of the computer family tree.

built or planned digital machines.¹⁵ Some could step through commands inscribed on mechanical media, but only ENIAC combined a capacity for general-purpose calculations with electronics. No way existed to issue orders at electronic speeds, so ENIAC had to be configured with patch cords and rotary switches. ENIAC itself had been secret and so inconspicuous that high-ranking government officials were not aware of it, but then newspaper reports so stimulated interest that even the Soviet Union asked to buy a copy [188, p. 107]. The reason for von Neumann's year-old report on machines with internally stored programs was suddenly clear.

In the same year, the Moore School sponsored the first workshop on electronic computers. Invited luminaries gave lectures, while the bulk of the coursework came from Mauchly and Eckert, and from Goldstine, who delivered an extemporaneous survey of numerical methods. Interest in the workshop was so great that institutions were limited to one participant each. Maurice Wilkes recalled the excitement of attending [324, p. 116] and credited the lectures for establishing the principles of modern computers [47, p. xxiii].¹⁶ All that remained was to build an electronic, programmable machine.

Several projects got underway during this period (Figure 2.3). Besides the continuing effort at the Moore School, there were projects at Cambridge, at Eckert and Mauchly's commercial venture, at Manchester, the Massachusetts Institute of Technology (MIT), the National Bureau of Standards (NBS), the National Physical Laboratory (NPL), at Engineering Research Associates (ERA), which was a contractor for a predecessor of the National Security Agency, and at Princeton (IAS). The United Kingdom had personnel experienced in building several Colossi, but these cryptanalysis machines were so secret that even years later scarcely anyone knew of them.¹⁷ The U.S. had the only publicly known, all-electronic calculator, but its builders had been forced to leave the project and become entrepreneurs. Mauchly and Eckert's firm suffered so many reversals that the NBS built two machines before a UNIVAC arrived [151]. In the end, Manchester demonstrated the first digital, electronic, programmable computer. It was not until roughly three years after the Moore School Lectures that any new machines were ready for work, and those were at Cambridge and Manchester in the United Kingdom. The wait was so long in the U.S. that by 1948 there had been time to retrofit ENIAC to run a form of stored program; this was done at von Neumann's suggestion, so that his earliest scientific computations could be done on that machine.¹⁸

The impatient wait for new machines found voice in von Neumann's humor. His childhood had honed his skills as a raconteur [317, pp. 38–43], which he used to enliven

¹⁵Aiken 1900–1973 developed a series of computers beginning with the electromechanical Harvard Mark I in 1944 [57]. Atanasoff 1903–1995 and his student Berry 1918–1963 built an electronic calculator to solve simultaneous linear equations at Iowa State University in 1943 [41]. Babbage 1791–1871 and his protégé Lovelace 1815–1852 planned a mechanical computer that could not be built with 19th century technology [152]. Regarding Lovelace, see especially [159]. Flowers 1905–1998 and Newman 1897–1984 [1] led development of the Colossus electronic machine for cryptanalysis [226]. For Stibitz and Zuse, see footnotes 1 and 12.

¹⁶Wilkes [324] built the Cambridge EDSAC. Early computer conferences are listed in [323, p. 299]. Grier [132, p. 288] describes the significance of the workshop. Some lectures and Goldstine's bibliography can be found in [47]. The antecedents of modern control units include the *sequencers* that operated linked IBM tabulating machines [51], which in later forms were called *selective* when different sequences of activity could be chosen, but the word *program* was first used as a verb with its modern meaning at the Moore School Lectures [131].

¹⁷Colossus became well known only 30 years after the machines were destroyed [226].

¹⁸The modified ENIAC is discussed in [194] and is mentioned in [10, pp. 238–239] and [27, pp. 27, 30].

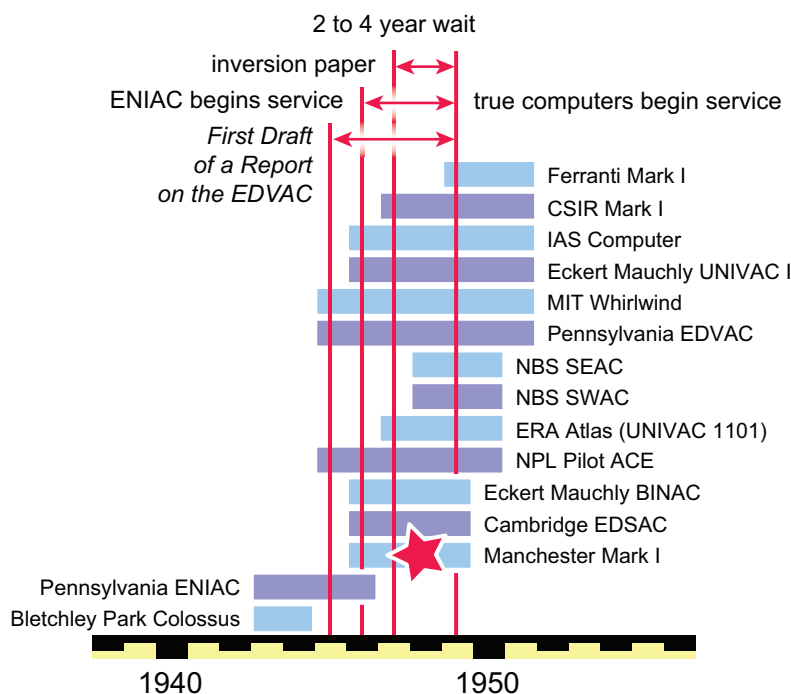


Fig. 2.3 The inversion paper was written about computer calculations before computers existed. Against the time frame of von Neumann's two seminal computer publications, timelines show the approximate development periods for most if not all of the modern computers (digital, electronic, programmable) that were completed by 1951. Colossus and ENIAC did not have internally stored programs. A "baby" version of the Manchester machine ran the first stored program of about 20 instructions in June 1948. In the words of Mina Rees [229, p. 834], there was a long wait for "the powerful machines that, we were always confident, would one day come."

his famous parties and to relax in lectures and meetings [181, p. 149]. His jokes could also be pointedly sarcastic [222, pp. 21–23], as is the von Neumann constant. The origin of the joke is lost, but the story lives on in many retellings:¹⁹

The long delays then common in computer creation... had led to von Neumann's 1948 remark, which [IBM executive Cuthbert] Hurd always remembered, "When any engineer is asked when his computer would be running, the reply is always the same, 'In six months.'" This invariant time period came to be known as "The von Neumann Constant." By 1951, a few of the some 20... computer building programs then under way worldwide had "von Neumann constants" of as long as three years.

— Cuthbert Hurd and Eric Weiss [320, p. 68]

So many humorous stories survive from the Princeton years that one suspects von Neumann encouraged them to soften his formidable reputation. Julian Bigelow told of visiting von Neumann at home to apply for the job of IAS chief computer engineer. When Johnny opened the door at Westcott Road, he let the great dane on the front lawn squeeze past them to play inside the house. At the end of the interview, von Neumann asked whether Bigelow always traveled with the dog.²⁰

¹⁹For other mention of the von Neumann constant see [279, p. 29] and [133, p. 120]. There are serious von Neumann constants for soap bubbles [246, 308] and for Banach spaces [56, 158].

²⁰The stray dog story appears in [181, p. 311] and [230, p. 110].

2.4. Open Source Computing. Goldstine demobilized and joined von Neumann when the IAS computer project commenced in 1946. Many accounts dwell on the academic politics of building a machine at the “Princetintute,” in Norbert Wiener’s phrase [321], but the early and generous financial support of the IAS enabled von Neumann and Goldstine to accomplish more than just building a computer.

The remarkably free exchange of computer technology in the decade following World War II occurred because von Neumann and Goldstine made computers a topic of academic interest, superseding government and proprietary desires for secrecy:

About this time, Paul Gillon and I decided that we ought to get some people over from abroad to get the idea of the computer out to the world—as much as we could. . . . So Paul Gillon got permission from the government to bring [Douglas Hartree] over, and he programmed a big problem. . . . He received, while he was in the States, an offer to be a professor at Cambridge University, which he accepted. He took documents from us like the *First Draft* report and got Maurice Wilkes interested. — Herman Goldstine describing ENIAC [27, pp. 34–35] in 1945²¹

Several machines working on the same principles. . . are now in operation in the United States and England. These principles derived from a report drafted by J. von Neumann in 1946. . . . — Maurice Wilkes [325, p. 3] in 1951

Many of us who are. . . making copies of the IAS machine have a tendency to emphasize our deviations and forget the tremendous debt that we owe Julian Bigelow and others at the Institute. — William Gunning recalling building JOHNNIAC [270] in 1953

The success of von Neumann and Goldstine’s precedent was not assured. The British government so well suppressed memory of the Colossus that subsequent developments had to be undertaken without public examination of its success. In the U.S. the exchange of information was threatened by the property rights that Mauchly and Eckert sought and by the military applications that paid for the machines.²²

Mauchly and Eckert’s idea for a computer venture dated from 1944, when the University of Pennsylvania ceded them its rights to file patents under the ENIAC contract. After some faculty questioned this arrangement, Dean Harold Pender insisted that Mauchly and Eckert give the university their rights to the stored-program successor, EDVAC. Patent waivers had not been a condition of their hiring, so they instead resigned in March 1946.²³ The same month, von Neumann notified Pentagon lawyers that he had contributed to the design of EDVAC and cited his report as evidence; he learned that the lawyers were already working with Mauchly and Eckert and would evaluate his claims in that light.²⁴ The Pentagon was expected to file all the patents on this military work, as it customarily did, on behalf of its contractors.

²¹Colonel Gillon was a senior officer at Aberdeen and the man who committed to building ENIAC. Hartree was a “numerical analyst” in the original sense of computational scientist; for examples of this usage see [69, p. 104].

²²The sharing of computer knowledge that is documented by Aspray [7], [11, pp. 187–188] was at cross-purposes with the national security interests that Pugh and Aspray [223] identify as the earliest customers for the machines. A thesis of this article is that von Neumann and Goldstine were the catalysts for the dissemination.

²³University authorities saw no value in ENIAC and declined patent rights when the project began. McCartney [188, pp. 129, 132] and Stern [257, pp. 48–52, 90–92] chronicle the award of the rights to Mauchly and Eckert and their humiliating forced departure from their own project.

²⁴Aspray [10, pp. 43–44] describes von Neumann’s contact with the lawyers.

The university's impromptu policies for intellectual property, the Pentagon's unwillingness to adjudicate conflicting claims, and Mauchly and Eckert's tardiness in filing any patents left unsettled the ownership of all the basic inventions at the Moore School. Other projects besides EDVAC got underway (Figure 2.3) while the Pentagon lawyers held meetings and wrote letters that demanded von Neumann's time. Finally, in April 1947, the lawyers astounded the principals by declaring that von Neumann's unclassified *First Draft* report was a public announcement, so the allotted year to file patents on the contents had lapsed.²⁵

In contrast, von Neumann used the freedom afforded by the IAS project to maximize the unencumbered dispersal of technology. The staff were told they could patent their contributions, but no action was taken and the option was eventually withdrawn.²⁶ Von Neumann personally forsook patents to avoid distractions like those that slowed the EDVAC project [10, pp. 45, 266]. His project began without government funds so the computer was IAS property, safe from government meddling [181, p. 306]. The several military sponsors, who ultimately paid the bulk of the costs at the IAS, were expected to build their own copies. Conversely, lest the IAS develop proprietary ambitions, it was obligated to disseminate the plans to at least the government sponsors. In practice the IAS circuit diagrams were simply given away to laboratories and universities around the world and over a dozen clones were built, including the first IBM mainframe, the 701 [320, p. 68] (see Figure 2.4). These early computers, with their proven circuitry and common architecture, cannot be overestimated for their impact in seeding computer science research.

Von Neumann, with his experience in computing for the Manhattan Project, and Goldstine, with his experience supervising the construction of ENIAC, shared most of the world's expertise in what then existed of computer science. The first tangible product of the IAS project was a series of important memoranda about computer architecture and "coding" (programming), written before anyone else could have done so, in the years 1946–1948, by Goldstine, von Neumann, and initially Burks. This subject matter was ambitious given that there were no modern computers at the time, but today the material is so well known that the reports could be mistaken for early and elementary textbooks. Many first-generation computer scientists learned the subject by reading the IAS plans and reports.²⁷ Von Neumann and Goldstine invented early staples of good programming practice such as the *subroutine* (original

²⁵Stern [258] discovered a transcript of the eventful meeting. The Pentagon lawyers evidently waited for von Neumann's rights to expire before informing him of the time limit. There still remained Mauchly and Eckert's ENIAC rights, for which the lawyers had prepared patent disclosures (thereby extending the filing period). Mauchly and Eckert passed the basic computer patent to their firm, which was acquired and reacquired by other corporations. U.S. District Court Judge Miles Lord eviscerated the patent in 1973, citing John Atanasoff's prior work, for the purpose of abrogating licensing agreements that gave a competitive advantage to IBM [188, pp. 198–202]. Mauchly and Eckert have perhaps the strongest claim to having invented the electronic (not to say, stored-program) computer, but no one with a plausible claim appears to have earned any great wealth from the invention.

²⁶Bigelow [193, pp. 25–26, 31–32] describes the change of policy in 1949 and the cavalier disposition of intellectual property at the IAS. "If we had an idea for something, and it looked as if it might work, we just ran a pencil sketch through [the ozalid machine] and sent it to everybody else It isn't that Von Neumann wanted us not to have patents, but that nothing could interest him less than patents."

²⁷The reports are [42, 120, 121, 122]. Mention of their importance is made by Aspray [7, p. 356], Knuth [163, p. 202], and Macrae [181, pp. 307–308]. For example, the Burks, Goldstine, and von Neumann report was the prototype for an early computer design manual prepared at ERA [264, p. 65].

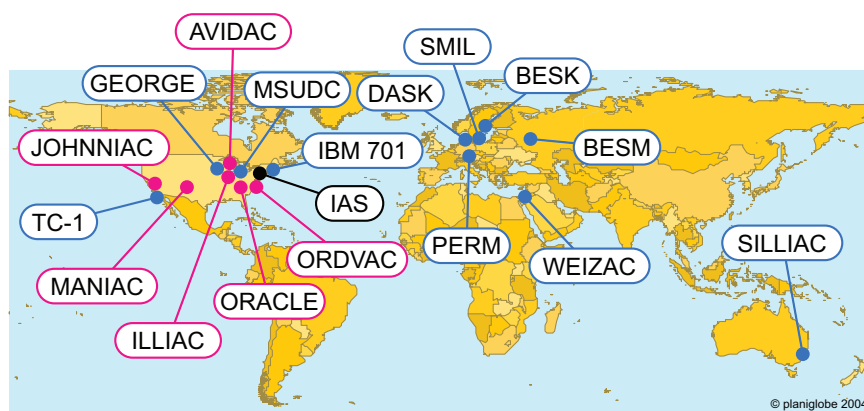


Fig. 2.4 Sometimes whimsical names were given to the contracted (red) and either derivative or unofficial clones (blue) of the IAS computer that were built from 1952 to 1957 [10, p. 91], [115, p. 307].

terminology) and the *flow chart* (originally *flow diagram*).²⁸ Indeed, the IAS reports served as programming manuals until the first actual book by Wilkes, Wheeler, and Gill [325, p. ix]. All the IAS reports were mimeographed for rapid distribution and, in any case, no journals then published this subject matter. Upwards of 175 initial copies were distributed around the world with legal notices that the contents were being placed in the public domain [181, p. 306]. The computer may be the only significant technology for which a principal inventor distributed his contributions without any expectation of remuneration, ownership, or (in view of the unarchived reproduction) even attribution. How differently computer science might have developed if John von Neumann had behaved like computer industry tycoons of past and present whose fortunes ultimately derive from his foresight.

2.5. Rounding Error Brouhaha. Within a few years after World War II, von Neumann and Goldstine had designed a computer architecture, formed an engineering group to build it, begun studying how programs could be written to solve scientific problems, and, not waiting for journals and publishers to address these issues, had communicated all their results in the form of mimeographed reports. They had not yet addressed the questions of approximation that would be relevant to the new calculations. Such mathematics at the time usually consisted of painstaking analysis to formulate problems in ways that would be tractable to hand computing.

Stibitz [263, p. 15] began the Moore School Lectures by asking, “Should machines be used to generate tables of mathematical functions, or should they be used to solve problems directly in numerical form? There is at present a considerable difference of opinion among workers in the field as to which procedure will be dominant.” The two sides, respectively, expected or feared that rounding errors might render machine calculations useless. Some took this problem so seriously that Mauchly [186] closed the Moore School Lectures by calling for “an intelligent approach to error buildup.”

²⁸Flow charts and subroutines are introduced in [120, pp. 84–104], [122, p. 217]. The originality of the contribution is attested to by Bashe et al. [23, p. 327], Hartree [137, p. 112], and Newman and Todd [205, p. 176–177].

The lack of contemporary journals about computing makes it difficult to assess opinions about rounding error. Hans Rademacher [225, p. 176] reported without references that “astronomers seem to have been the first to discuss rounding-off error” when preparing tables by integration (numerical solution of ordinary differential equations (ODEs) in modern terms), but “their arguments are somewhat qualitative in character.”²⁹

An activity sponsored by governments throughout the 19th and early 20th centuries, and for which there is some literature about computing, was the adjustment of geodetic survey triangulations by least squares methods [128]. Gauss [108] reduced the geodetic problems to simultaneous linear equations called normal equations (his terminology) that were solved by various forms of symmetric elimination. Human computers were advised to retain only the digits deemed significant [22, p. 24], [244, pp. 17–19], and they had experience recognizing normal equations, that were poorly formulated [338, pp. 204–206, par. 154]. In these circumstances elimination may well have acquired a reputation for sensitivity to rounding.

In the 20th century, normal equations were solved by inner product forms of elimination (modern terminology) which had developed because mechanical calculators could accumulate series. For other equations Prescott Crout [63] introduced similar methods that quickly became popular.³⁰ However, these inner product methods could give poor results because they omitted provisions to reorder the equations and variables [327, p. 49]; such a contingency is unnecessary for normal equations, but it can be important for others. This neglect of reordering may have contributed to the poor reputation of elimination.

The concerns about rounding error appear to have been magnified by prejudices about numerical calculations. For example, Rademacher drew positive conclusions from a probabilistic treatment of rounding errors in solving ODEs.³¹ He showed that smaller time steps had smaller truncation error but larger probable rounding error because more steps were taken [225, p. 185]. Nevertheless, he estimated and observed that the overall error from ENIAC could be acceptably small [224, p. 236]. Rademacher’s paper was rejected by *Mathematical Tables and Other Aids to Computation* (now *Mathematics of Computation*) because the referees refused to believe his findings [47, p. 222].

In contrast, Harold Hotelling [143] had been able to publish negative conclusions about the accuracy of elimination.³² He suggested that errors could quadruple with the elimination of each variable, so exponentially large numbers of digits might be required for accurate solutions. This prediction could not be easily tested because new computers were delayed by “von Neumann’s constant.” Hotelling’s paper is the only work ever cited to impeach the method used by Gauss, but he apparently gave voice to widespread concerns about mechanized computation. Those building computers were prudent to dispel this pessimism that increased the apparent risk of

²⁹Rademacher 1892–1969 [28] moved from Germany to the University of Pennsylvania in the 1930s, where he became acquainted with ENIAC.

³⁰Waugh and Dwyer [318, p. 259] comment on the popularity of “abbreviated” or “compact” methods. Crout 1907–1984 was a mathematician at MIT.

³¹Rademacher studied Heun’s method, which was used to prepare firing tables [73]. Aerial warfare had made it necessary to determine complete ballistic trajectories since World War I [132, Chap. 9].

³²Hotelling 1895–1973 [6] was a statistician and economist. As government bureaus increasingly gathered economic data after World War I, statisticians performing regression analyses gradually supplanted geodesists as the majority of computers who solved normal equations. The growth of statistical work is described by Fox [100], Grier [132, Chaps. 10–11], and Hotelling [142].



Fig. 2.5 *Mrs. von Neumann (Klara Dan), Inverse, and John von Neumann at home in Princeton circa 1954. Photo by Alan Richards courtesy of the Archives of the Institute for Advanced Study.*

rote calculation. The mathematically oriented builders at both the IAS and the NPL [141, p. 344] undertook research to that end.³³

2.6. A Dog Named Inverse. Once von Neumann took up matrix inversion in spite of his many commitments, a good part of his unscheduled time was apparently occupied by inverse matrices. He liked to work at home on the dining room table with friends who were involved in his favorite projects. Indicating just how much time von Neumann spent on the project, Mrs. von Neumann named the family dog “Inverse” [115, pp. 290–292] (see Figure 2.5).

Von Neumann and his collaborators undertook three investigations of matrix inversion. The first study was conducted for the IAS numerical analysis project. It resulted in a report by Valentine Bargmann, Deane Montgomery, and von Neumann

³³Hotelling and von Neumann appear to have been acquainted because they sparred over linear programming at the Econometric Society [67, p. 25]. Von Neumann’s interest in matrix inversion seems to have been in response to the apocryphal concerns that Hotelling fueled. Hotelling’s paper alone is cited to motivate studying Gaussian elimination in the inversion paper itself {p. 1022} and by Aspray [10, p. 98], Bargmann et al. [21, pp. 430–431], and Goldstine [115, pp. 289–290], [117, p. 10]. The IAS computer was not expected to use Gaussian elimination. Von Neumann’s interests were partial differential equations (PDEs) of hydrodynamics for the Manhattan Project and for meteorology—the latter was so important to him that it was the focus of an IAS subgroup. Von Neumann knew Gaussian elimination would not be used for such problems. As a consultant for the Standard Oil Company he wrote two reports assessing how best to solve certain flow problems [304, 305]. By 1960 it would become routine to apply iteration to tens of thousands of unknowns [288, p. 1], whereas Gaussian elimination would rarely be applied to equations with more than 100 variables [140, p. 185], [204, p. 474].

[21] that explored alternatives to Gaussian elimination and may have made the first comparison of algorithms by operation counts.³⁴ The report contrasted Hotelling's pessimistic assessment of the precision needed to invert matrices by Gaussian elimination with a modification of an iterative method for SPD matrices.³⁵ The analysis of the iterative method featured the ratio of extreme eigenvalues [21, pp. 461–477].

The second project stemmed from von Neumann and Goldstine's conviction that Gauss would not have used a flawed method, so Hotelling was likely wrong [117, p. 10]. The result was the inversion paper [314]. Von Neumann and Goldstine interpreted Gaussian elimination as an algorithm for factoring matrices, and they proved that the algorithm is necessarily accurate for factoring SPD matrices under mild restrictions. The inverse matrix can then be formed as the product of the inverse factors with an error that is bounded by a small constant times the ratio of the extreme singular values (modern terminology). In short, the error was not exponential. The inversion paper further showed that any matrix could be inverted by a formula related to the normal equations of least squares problems, but then the square of the singular value ratio figured in the error bound. The basic discoveries were made by von Neumann in January of 1947. He wrote to Goldstine from Bermuda that he was returning with proofs for the SPD case and the extension based on normal equations [302] (Figure 2.6).

Goldstine recalls that during the summer of 1947 he was corresponding with the peripatetic von Neumann about completing the inversion paper. They continually revised the error analysis to improve the error bounds. Von Neumann referred to the ratio of singular values as ℓ . Both his letter and an undated IAS manuscript [119, p. 14] had an extra ℓ in the error bound for inverting SPD matrices (visible in line -2 of Figure 2.6), which was expressed as $\mathcal{O}(\ell^2 n^2)$ times the maximum absolute size of an individual rounding error (n is the order of the matrix). Von Neumann did not say what matrix factorization he considered, but evidently it was a variation of $A = LDL^t$, where L is lower triangular with unit diagonal entries and D is diagonal. By July 19, Goldstine had used the grouping $(LD)L^t$ to reduce the bound to $\mathcal{O}(\tau^{3/2} n^2)$, writing τ for von Neumann's ℓ . "When I got up this morning," Goldstine [113, p. 5] thought to try $(LD^{1/2})(LD^{1/2})^t$ because "the estimates for the bounds of these matrices are more favorable." This approach had a smaller bound, $\mathcal{O}(\tau n^2)$, and it needed less storage, but there were square roots. On July 23, Goldstine [114] was "putting finishing touches on the inversion paper" and planning to include both of his bounds: the smaller one for the calculation with square roots and the larger one for the calculation without. The printed paper instead had the decomposition LDL^t that was ultimately found to combine the best aspects of the other two.

Von Neumann clearly participated in writing the paper. Several of Goldstine's suggestions were not adopted: a technical estimate for the norm of the computed inverse [114], double precision accumulation of inner products (modern terminology) [113], and presenting the two error bounds. The text alternates between Goldstine's unobtrusive style and von Neumann's unmistakable English, which Paul Halmos³⁶ [135, p. 387] described as having "involved" sentence structures with "exactly right"

³⁴Aspray [10, p. 98] remarks on the "first." The same approach was used after 1946 by Bodewig [32] and by Cassina [50]. Von Neumann's collaborator Bargmann [161] belonged to the mathematics faculty at Princeton University. Montgomery [37] subsequently became a member of the IAS.

³⁵The iterative method is also discussed by Goldstine and von Neumann [119, p. 19], by Hotelling [143, pp. 14–16], and more recently by Higham [140, pp. 183, 287].

³⁶Halmos 1916–2006 [92] was known for exposition and had been a postdoc under von Neumann at the IAS.

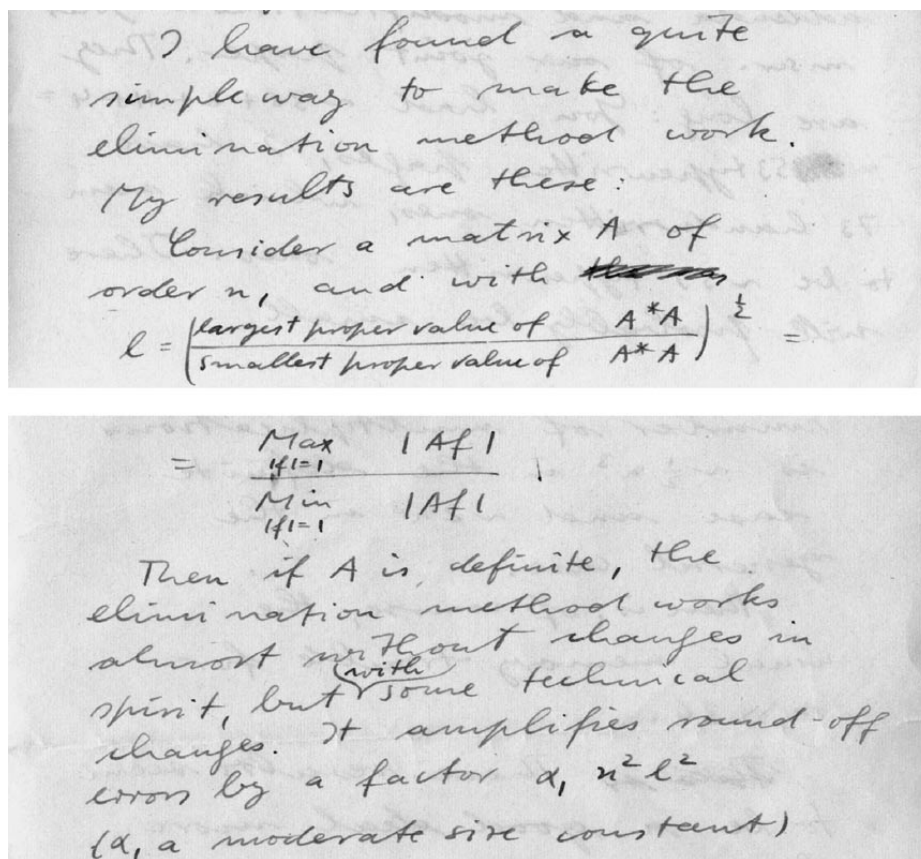


Fig. 2.6 “I have found a quite simple way . . .” Letter from von Neumann [302] to Goldstine dated January 11, 1947 announcing the discovery of an error bound for inverting SPD matrices by Gaussian elimination. The value l is the first appearance in a rounding error bound of what became the matrix condition number. Note the use of the matrix lower bound, $\min_{\|f\|=1} \|Af\|$ [129]. Von Neumann’s overestimate of the error bound (an extra factor of l in line -2) was corrected in the final paper. From the Herman Heine Goldstine Collection, American Philosophical Society.

vocabulary. Scattered throughout the paper are many passages that are identifiably von Neumann’s.

Whether Goldstine or von Neumann had the last word cannot be determined from the available documentation. The paper entered publication about two months after Goldstine’s last archived letter. It was presented at an AMS meeting on September 5, submitted to the *Bulletin of the AMS* on October 1, and published on November 25. Perhaps from haste in revising the error analysis, parts of Chapter 6 exhibit what appears to be an arduous version of von Neumann’s telegraphic style. Halmos [135, pp. 387–388] amusingly described this style and sarcastically explained its effect on lesser mathematicians. James Wilkinson³⁷ commented that the proof “is not exactly

³⁷Wilkinson 1919–1986 [101] finished building the NPL Pilot ACE. He became the leading authority on the numerical linear algebra subject matter of the inversion paper.

bedside reading" [332, p. 551] and left a reader "nursing the illusion that he has understood it" [333, p. 192].

The error bounds derived in modern textbooks by what now is called rounding error analysis are in essence the same as von Neumann and Goldstine's bounds and methods. It was of continuing disappointment that analyses of this kind considerably overestimate the error [332, p. 567]. When questioned on the sharpness of the bounds a few years later, von Neumann [307] privately explained that the purpose of the inversion paper was to see if it were possible to obtain strict error bounds, rather than estimates which may be closer to optimum.

The inversion paper contrasted its approach with probabilistic estimates {p. 1036, par. (e)}. A third project took the statistical approach and replaced the bounds by estimates that are smaller by a factor of n , the matrix order [115, p. 291]. This work resulted in part two of the inversion paper, which was completed only in 1949 and appeared still later [123].

2.7. Stump Speech. Von Neumann and Goldstine avoided speculative opinions in the inversion paper, so it is necessary to look elsewhere for their characterization of the results. They appear to have informally discussed the paper in their crusade to gain acceptance for computers. An IAS manuscript mixes lectures from as early as 1946 with references to the inversion work in 1947.³⁸ During the IAS project, von Neumann made many proselytizing speeches [181, pp. 307–308] about the "extent human reasoning in the sciences can be more efficiently replaced by mechanisms" [119, pp. 2–5]. Von Neumann meant the practice of reducing physical problems to forms permitting analytic estimation, which he believed was often impossible especially for nonlinear problems, citing shock waves and turbulence. The speeches gave von Neumann free rein to answer doubts about the new machines.

Von Neumann and Goldstine addressed three objections to machine calculations. The last objection was "that the time of coding and setting up problems is the dominant consideration" [119, p. 29]. They acknowledged this delay would be possible when the machines were new, but they pointed out that eventually instructions could be selected from a "library of previously prepared routines." Moreover, the computers were intended for longer calculations than were feasible by hand, so "solution times of a few seconds are quite unrealistic."

Second, "the most commonplace objection... is that, even if extreme speed were achievable, it would not be possible to introduce the data or to extract (and print) the results at a corresponding rate. Furthermore, even if the results could be printed at such a rate, nobody could or would read them and understand (or interpret) them in a reasonable time.... We now proceed to analyze this situation...." See [119, pp. 20–23] for the elaborate rebuttal.

The first and most serious objection was "the question of stability" [119, pp. 13–14]. The inversion paper used this word with reference to a strict definition, "the limits of the change of the result, caused by changes of the parameters (data) of the problem within given limits" {p. 1027}. Von Neumann's speeches introduced an informal meaning, "a danger of an amplification of errors." After explaining Hotelling's speculation that Gaussian elimination could magnify rounding errors by the factor 4^n [143, pp. 6–8] (n is the matrix order), von Neumann and Goldstine announced their own result:

³⁸A recently discovered transcript of remarks by von Neumann [313] contains passages that appear in an unpublished manuscript [119] that is known to have been composed from lectures given during 1946–1947.

We have worked out a rigorous theory. . . . [T]he actual [bound] of the loss of precision (i.e. of the amplification factor for errors, referred to above) depends not on n only, but also on the ratio ℓ of the upper and lower absolute bounds of the matrix. (ℓ is the ratio of the maximum and minimum vector length dilations caused by the linear transformation associated with the matrix. . . . It appears to be the “figure of merit” expressing the difficulties caused by inverting the matrix in question, or by solving simultaneous equations systems in which [the matrix] appears.)

— Goldstine and von Neumann [119, p. 14]

In modern notation $\ell = \|A\|_2 \|A^{-1}\|_2$ is the matrix condition number. Von Neumann discovered, essentially, that this quantity is needed to bound the rounding errors.³⁹

2.8. Numerical Analysis as Mathematics. Because von Neumann preferred questions related to the systematization of subjects both mathematical and scientific [138, pp. 126, 128–129], he is recognized as founding several fields of mathematics, particularly applied mathematics. His approach was formal, but unlike David Hilbert he was not drawn into debates about the philosophical ramifications of formal methodologies.⁴⁰ Instead, von Neumann [303, pp. 6, 9] preferred the empiricism of science, which he suspected was “probably not wholly shared by many other mathematicians.”

Von Neumann took the methodological view that all mathematical ideas, no matter how abstruse, “originate in empirics,” so “it is hardly possible to believe in an absolute, immutable concept of mathematical rigor, dissociated from all human experience.” A corollary was that to be viable, mathematics required research goals set by uses rather than aesthetics [79]. The inversion paper fit this pattern because it inaugurated the mathematical study of numerical algorithms for the purpose of making better use of computing machines, by showing that the errors committed in scientific computing could be studied by strict error bounds, a rigor that appeared relevant.

The theory of numerical computation developed in an unusual manner that can be understood by contrasting it with other mathematics. For game theory, for example, von Neumann’s 1926 paper [291] on parlor games contained the original minimax theorem which became a subject of research after a lapse of about ten years. Von Neumann [296] then returned to the topic and proved a more abstract minimax theorem approximately as it is known today.⁴¹ The second paper also made a connection to economic equilibrium which it treated by fixed point methods. Von Neumann’s last work on the subject was the 1944 book with Morgenstern [315] that developed economic theory in terms of multiperson games.⁴²

Mathematical research on games continued after the contributions of von Neumann with limited application, while his book was disparaged for its difficulty and low readership [222, pp. 41–42]. However, in the mid-1970s economics dissertations began

³⁹The name “condition number” was first used by Turing [285] in 1948 for different values. The name was applied to von Neumann’s figure of merit by Todd [272] in 1949 and reemphasized by him in 1958 [204].

⁴⁰Rowe [241] describes these disputes. Formalism is controversial in so far as it entails an epistemology and ontology for mathematics. Brown [40] and Shapiro [249] survey philosophies of mathematics.

⁴¹Kjeldsen [160] and Simons [252] trace the history of minimax theorems.

⁴²These contributions were so early that they receded from memory as the subject developed. It came to be said that Émile Borel had done as much to invent game theory and in prior papers. Von Neumann [309] responded that to found a mathematical theory one had to prove the first theorem, which Borel had not done. Borel actually conjectured against the possibility of a theory for optimal strategies.

to use game-theoretic methods [136], and since 1994 five economics Nobel laureates have been named for game theory.

By and large it is uniformly true in mathematics that there is a time lapse between a mathematical discovery and the moment when it is useful; and that this lapse of time can be anything from thirty to a hundred years. . . and that [in the meantime] the whole system seems to function without any direction, without any reference to usefulness, and without any desire to do things which are useful.

— von Neumann [310]

In marked contrast, numerical analysis had no leisurely gestation in mathematics before the uses for its subject matter arose. Elaborate calculations had always been organized independent of university mathematics faculties. The French, British, and U.S. governments sponsored observations and computations for annual ephemerides beginning in 1679, 1767, and 1855, respectively [207, p. 206]. Government bureaus made substantial hand calculations for ballistics, geodesy, or nautical and astronomical tables through World War II, and for economic forecasting after World War I. Among the most prominent calculators were extra-academic organizations such as Karl Pearson's privately funded Biometrics Laboratory, Mauro Picone's Istituto Nazionale per le Applicazioni del Calcolo, Leslie Comrie's commercial Scientific Computing Service, and Gertrude Blanch and Arnold Lowan's storied Mathematical Tables Project of the Works Progress Administration. Many universities started calculating groups between the world wars, including Benjamin Wood's Statistical (later Wallace Eckert's Astronomical) Computing Bureau at Columbia University and George Snedecor's Statistical Computing Service at Iowa State University (which introduced computer inventor John Atanasoff to calculating machines). The history of all these organizations is only now being written, but with the possible exception of Picone's institute, their purpose was to calculate rather than to study the mathematics of calculation.⁴³

In this environment von Neumann worked in two ways to build a mathematical research community. First, when he accepted administrative duties he used the positions to bring mathematicians into contact with scientists doing calculations. These responsibilities included the IAS numerical analysis project during the war, the advisory capacity in which he served at Los Alamos, the IAS meteorological project, and his membership in the Atomic Energy Commission (AEC). The numerical analysis project introduced von Neumann's ideas to IAS visitors, some of whom went on to lead computer science departments [115, p. 292]. He was wildly successful at bringing computers to meteorology, although not meteorology to mathematics, and at his death the IAS meteorological group led by Jule Charney found employment outside academic mathematics.⁴⁴ With the exception of Stanislaw Ulam, whom von Neumann recruited to Los Alamos, the senior people on the mesa were scientists [287, p. 161]. Some laboratory physicists such as Nicholas Metropolis and Robert Richtmyer visited the IAS to work with von Neumann [115, p. 292] and are now remembered as pioneering numerical analysts. Von Neumann belonged to the AEC only long enough to start financial support for independent research in computing [10, pp. 248–249]. He wanted to duplicate in civilian science the direct cooperation between mathematical and scientific researchers that he experienced in the Manhattan Project. Von

⁴³Calculations for mathematical tables are discussed in [46, 62, 132].

⁴⁴Charney is remembered for developing scientific weather prediction [219]. The history of the IAS weather project is explained by Aspray [10, pp. 121–154]. The ramifications of the project for mathematical modeling are discussed by Dalmedico [66].

Neumann brought John Pasta to Washington from Los Alamos to begin programs in computer science and computational mathematics which continue at the present Department of Energy.⁴⁵

Second, to build a research community, von Neumann undertook his own exemplary research in numerical analysis. His work for the IAS project included planning what today would be called the computer architecture and the computational science [10, p. 58]. He and Goldstine believed that new algorithms would be needed to make best use of the new machines, so they insisted that mathematical analysis must precede computation.⁴⁶ Privately, von Neumann wrote to his friend Maxwell Newman:

I am convinced that the methods of “approximation mathematics” will have to be changed very radically in order to use ... [a computer] sensibly and effectively—and to get into the position of being able to build and to use still faster ones.

— von Neumann [300, p. 1]

Mina Rees uncannily echoed von Neumann’s sentiment about the need for

numerical analysis that would affect the design of computers and be responsive to ... [the use] of the powerful machines that, we were always confident, would one day come.

— Mina Rees [229, p. 832]

Rees led the Mathematics Branch of the Office of Naval Research, which was the primary agency supporting university research, either directly or through the NBS, from 1946 until the creation of the National Science Foundation in the 1950s. The subjects von Neumann pioneered largely coincided with the highlights of this first U.S. research program in numerical analysis, as explained by Rees [229, p. 834]: “numerical stability, matrix inversion and diagonalization, finite difference methods for the solution of PDEs, the propagation, accumulation, and statistical distribution of errors in long computations, and convergence questions in connection with finite difference schemes approximating PDEs.” Von Neumann worked with Charney, Goldstine, Metropolis, Richtmyer, Ulam, and others to study these topics, often in the context of various applications, and to describe how to prepare the applications for the computer. Most of his results were delivered orally or recorded in reports often written by others with few of the titles easily obtained outside his collected works.

Von Neumann described his numerical work as “approximation mathematics,” for want of a better word. John Curtiss applied the name *numerical analysis* to the study of numerical algorithms after World War II.⁴⁷ From a prominent mathematical family, Curtiss evidently meant to position this new subject alongside other kinds of analysis

⁴⁵Pasta 1918–1981 [64] coauthored a famous early example of scientific discovery through computation [94] and subsequently led computer science programs at the University of Illinois and the National Science Foundation.

⁴⁶Their view that analyzing should come before computing is stated in [119, pp. 6, 15] and [120, p. 113].

⁴⁷Householder [149, p. 59] finds the first use of *numerical analysis* to name a branch of mathematics in the name of the Institute for Numerical Analysis, which was founded by Curtiss at the University of California at Los Angeles. The first usage is attributed specifically to Curtiss by Fox [102]. Curtiss 1909–1977 [275] was an official at the NBS. He obtained funds for basic research on numerical analysis from Rees and from closing the labor-intensive Mathematical Tables Project. The Institute closed after a few years when the Navy withdrew support. Both closures were causes célèbres: the first lost hundreds of people their livelihoods and the second is remembered in political conspiracy theories [132, Chap. 18]. For computation during the unsettled time of von Neumann’s constant, see Cohen [58], Hestenes and Todd [139], Rees [229], Stern [257], Terrall [269], and Tropp [278, 279].

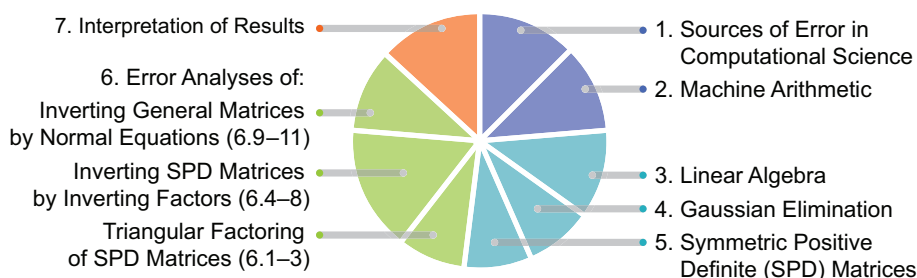


Fig. 2.7 Page budget for the inversion paper [314]. The rounding error analyses are in Chapter 6. The Bulletin of the AMS in 1947 had much smaller pages than a journal such as SIAM Review today, so the 79-page paper would occupy about 52 pages in a modern journal.

in mathematics. In the sciences, *numerical analysis* always had—and continues to have—the quite different meaning of analyzing data and models numerically.

2.9. Good Start for Numerical Analysis. Of all von Neumann’s writings on computers before 1950, only a short prospectus for meteorological computing and the inversion paper appeared in archival journals (Figure 2.2). Matrix inversion offered the perfect foil to illustrate in print the new numerical analysis that von Neumann and Goldstine (and Curtiss and Rees) intended. The subject was easily grasped and was of use in operations research [106] and statistics [143]. More importantly, matrix inversion was unencumbered by military and proprietary interests (see section 2.4). It was topical because of the prevailing concerns about the accuracy of automated computation (see sections 2.5 and 2.7). It was relevant because the delay in building computers meant that only analysis could address the question of stability (see section 2.3 and Figure 2.3). Finally, it seemed that a proof might be constructed, thereby justifying, at last, a publication about computers in a journal that was read by academic mathematicians (see section 2.8).

The inversion paper has a didactic tone because Goldstine actually intended to write the first modern paper in numerical analysis. He included lengthy preparatory material to “start numerical analysis off on the right foot” [115, p. 290]. As a result, many “firsts” and much of the text lie outside error analysis itself, so in a real sense the surrounding material is the heart of the inversion paper (see Figure 2.7).

1. A quarter of the paper examines the sources of error in scientific computing which future numerical analysts would have to confront. The first chapter enumerates the four primary sources of error and identifies rounding error as a phenomenon in need of study. For that purpose, Chapter 2 introduces an axiomatic model of machine arithmetic.
2. Another quarter of the paper supplies mathematical constructions that can be used to express and study numerical algorithms. Von Neumann and Goldstine successively introduced matrices and matrix norms for this purpose, interpreted Gaussian elimination in this manner by constructing what is now called the LDU factorization, and finally they derived properties of the factors of SPD matrices.
3. The preceding preparation reduces the analysis of matrix inversion to only a third of the paper. The three-part error analysis in Chapter 6 covers (i) what today is called the LDL^t factorization of an SPD matrix; (ii) its use to invert the matrix, $L^{-t}D^{-1}L^{-1}$; and (iii) what today is called a normal equations

formula to invert general matrices, $A^{-1} = A^t(AA^t)^{-1}$, where $(AA^t)^{-1}$ is formed by (i) and (ii).

4. The error analysis shows that Gaussian elimination can produce accurate results. This point having been made, the final chapter returns to the sources of error. It interprets the rounding errors as having produced the inverse of a perturbed matrix. The data perturbations (backward errors, in today's terminology) are likened to those caused by another basic source of error: data preparation, or measurement.

It was not necessary to follow the long proof in Chapter 6 to understand that Hotelling had been proved wrong by using mathematics to study numerical algorithms. Von Neumann and Goldstine thus both lent credence to computerized calculations and heightened interest in additional work on what became modern numerical analysis. Almost all the highlights of their paper can be viewed by reading the easily accessible chapters shown in Figure 2.8. Note that all these contributions were made in a paper that was written before modern computers existed.

2.10. Von Neumann and Turing. No less a figure than Alan Turing produced the first derivative work from the inversion paper. He and von Neumann appear to have had a synergistic relationship for many years.⁴⁸ They first met while von Neumann lectured at Cambridge University in the summer of 1935. Turing had improved one of von Neumann's papers [293, 281], and they had a mutual friend in Maxwell Newman. Von Neumann helped arrange a fellowship that enabled Turing to study at Princeton through 1938. During this time Turing again wrote a paper inspired by von Neumann's work [294, 283]. After Turing finished his American Ph.D. dissertation under Alonzo Church, von Neumann offered Turing an IAS postdoc position that would have been the next step to a career in the U.S.

The introverted Turing probably never realized the impression he made on the urbane von Neumann, who in 1939 remarked to Ulam about the "schemata of Turing" and his "brilliant ideas" [141, p. 145]. The paper on what became Turing machines [282] was required reading for the IAS computer project [10, p. 178], and it influenced von Neumann's formulation of automata theory during this time.⁴⁹ Reciprocally, Turing's proposal [284] to build a programmable machine at the NPL praised von Neumann's *First Draft* report on programmable computers.

Thus Turing and von Neumann were accustomed to exchanging ideas. After Turing attended the first Harvard symposium he spent two weeks with von Neumann and Goldstine in January 1947 during which they discussed computers and the inversion paper in progress.⁵⁰ Olga Taussky [269, p. 34] notes that von Neumann was very rarely at the Institute and was very busy when he was there, so it is remarkable that he spent so much time with Turing. The visit coincided with von Neumann's return from Bermuda, bringing new results for Gaussian elimination. By then the IAS researchers were adept at triangular factoring and they understood the ramifications of the symmetric and nonsymmetric cases. They had to explain it all to Turing in order to discuss their work, so they would have described the figure of merit ℓ , even as

⁴⁸The early contacts between von Neumann and Turing are described by Aspray [10, pp. 177–178], Goldstine [27, pp. 32–33], and Hodges [141, pp. 94–95, 131, 144].

⁴⁹Von Neumann's primary work on automata is the posthumous [312]. He mentioned automata and Turing together in letters such as [301]. Aspray [8], [10, pp. 189–194] traces the genesis of von Neumann's theory.

⁵⁰The symposium proceedings are given in [4]. Turing's visit to von Neumann is again described by Aspray [10, p. 100], Goldstine [27, p. 32] [115, pp. 191, 291], and Hodges [141, p. 355].

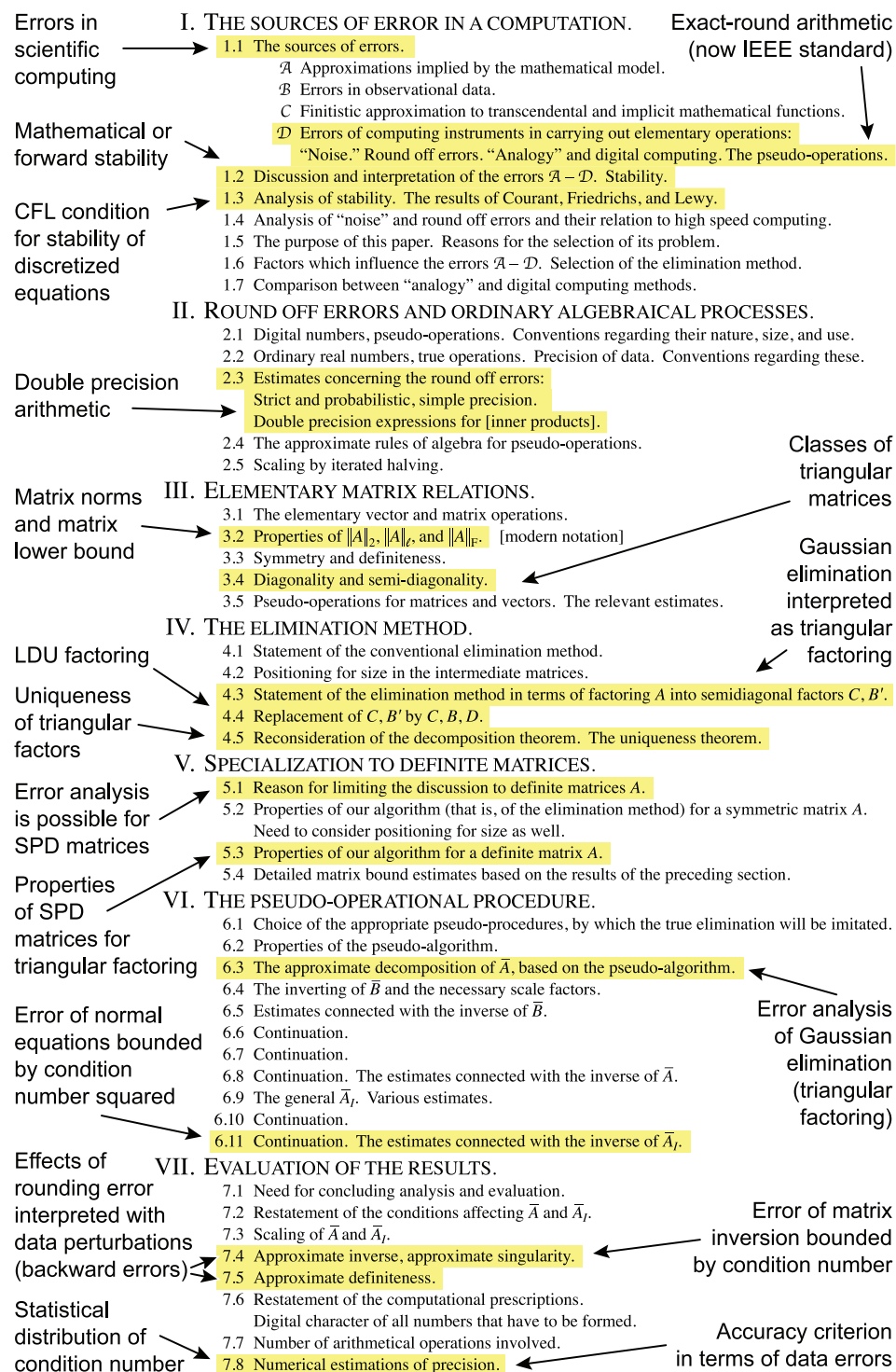


Fig. 2.8 *Firsts in the first modern paper on numerical analysis, by location in its own table of contents.*

von Neumann had just done in a letter to Goldstine and would proceed to do in the IAS transcript of his speeches.⁵¹

[Turing] was a very stubborn man. Von Neumann and I had the idea that Gauss's method of elimination was probably the correct [one] to use on a computer. We had a mathematical proof, and Turing came over for a few weeks and worked with us. He had a somewhat different idea, and we never were able to get him to agree with what we said.

— Herman Goldstine [27, p. 32]

The disagreement is plain if Goldstine's words are understood to mean Gauss's procedure for SPD matrices rather than all of Gaussian elimination. The inversion paper focused on the SPD case, whereas Turing's own paper of 1948 considered general matrices. The first half of his paper described the *LDU* decomposition and interpreted algorithms as constructing decompositions [285, pp. 289–296]. Turing's "main concern" was the difficulties that rounding errors were reputed to cause for these "matrix processes" [285, pp. 287, 297]. The second half of his paper considered solving equations and inverting matrices.

For equations, Turing supposed that $Ax = b$ is solved with erroneous coefficients $A - S$. Averaging "over a random population of matrices S " [285, p. 298], Turing found that the "RMS (root mean square) error of coefficients of solution" equals the "RMS error of coefficients of A " times normalizing factors and a value that Turing named the " N -condition number."⁵² Turing's relation is unusual compared to later work because it is an equality and because S is a matrix of random variables. If $Ax = b$ and $(A - S)x' = b$, then the customary inequality that bounds $x' - x$ in terms of A and S was stated later by Wilkinson [329, p. 93, eqn. 12.15],

$$(2.1) \quad \frac{\|x' - x\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|S\|/\|A\|}{1 - \|A^{-1}\| \|S\|}$$

(provided the denominator is positive), where $\|\cdot\|$ is any consistent matrix-vector norm.

For inverting A , Turing [285, pp. 302–305] made an a posteriori forward error analysis (modern terminology) of Gauss–Jordan elimination. That method reduces A to an identity matrix while simultaneously constructing A^{-1} . Because the experience of von Neumann and Goldstine and later researchers was that deriving a subexponential error bound for Gaussian elimination in the general case is intractable, Turing necessarily took his analysis in another direction. Assuming the largest errors encountered in reducing A and building A^{-1} to be bounded by δ and δ' times the largest entries in A and A^{-1} , respectively, he bounded the maximum error in the computed inverse in terms of δ , δ' , and a value he named the " M -condition number." Turing [285, p. 305] invited comparison with the error bound in the inversion paper.

⁵¹The letter [302] is shown in Figure 2.6. Footnote 38 describes von Neumann's 1946–1947 speeches.

⁵²Bargmann, Montgomery, and von Neumann [21, p. 473] had earlier analyzed inverting a matrix M_1 plus a random perturbation matrix H . Statistical analysis was natural for Turing because probability was the subject of his Cambridge dissertation [339], and because he had taken a probabilistic approach to cryptanalysis [141, pp. 196–198, 344]. Probabilistic methods for rounding errors had been discussed in von Neumann's letter [302], in the inversion paper {pp. 1026–1027; p. 1036, par. (e); pp. 1047–1049, sec. 3.5}, and in Rademacher's paper [225] at the Harvard symposium.

Turing coined the name “condition number” [260, p. 144] for measures of sensitivity of problems to error, and the acronym “ LDU ” [140, p. 184] for the general decomposition. Textbooks tend to intimate that Turing introduced modern concepts by introducing the modern nomenclature, but the history is more complex.⁵³ Algorithms had been described with matrix decompositions before Turing’s paper; in particular, the inversion paper had both LDU and LDL^t (written CDB and B^tDB). Measures of sensitivity evolved from as early as Wittmeyer [337] in the 1930s to Demmel [76] in the 1980s. Von Neumann [21, p. 462] has a bound for 2-norms of changes to inverse matrices, which Turing [285, p. 301] restated for M -norms, but neither of them stated the bound in the form (2.1) that Wilkinson [329, p. 29] would later require to justify the name “condition number.”

Turing’s biography suggests an adversarial relationship with von Neumann, “a truly powerful Wizard” [141, p. 124], which is not conveyed in their correspondence. Von Neumann politely corrected a manuscript from Turing in 1935, and they were cordially “Alan” and “John” when Turing submitted a manuscript for publication in 1949 [295, 306]. Turing prominently acknowledged von Neumann in four papers [281, 283, 284, 285]. Wilkinson [332, p. 548] speculated that his contemporaries cited von Neumann “as it were, as a prestige symbol,” but this characterization cannot describe Turing, who disliked pretense [59, p. 104]. For example, Turing ridiculed his technically unqualified supervisor at NPL, and he stored in a toolbox his medallion for Officer of the British Empire [141, pp. 317, 338].

2.1.1. Final Years. The von Neumann constant of the IAS computer became 0 in 1951. In a final illustration of the uncertainty that the joke was meant to capture, the official completion date was set at 1952. Over the next few years, von Neumann and Goldstine made the IAS computer available to researchers around the country who used it to perform dozens of first-kind calculations [10, pp. 121–171, Chaps. 6–7, pp. 156–157, tab. 7.1].

Von Neumann became a public figure after the war. As the member of the scientific elite with the most computing experience, he became the unofficial spokesman for the field [10, p. 236].⁵⁴ Throughout he continued his research (see Figure 2.2) and was president of the AMS in 1951–1953. Von Neumann’s wartime experience demonstrated his talent for seeing through to the heart of problems, which increased the demand for his services. He led the panel that planned the first U.S. strategic missiles (ICBMs) and he served on the AEC, a post requiring Congressional approval.

While in Washington Johnny became ill, and he lingered for a terrible year at Walter Reed Hospital, where he died in February 1957. The obituaries of record are in *The New York Times* [271], which included a statement from President Eisenhower, and in *Life* magazine [31], which led: “John von Neumann, a brilliant, jovial mathematician, was a prodigious servant of science and his country. An almost superhuman mind dwelling in supersecret realms of H-bombs and guided missiles is lost with [his] death.”

Shortly thereafter Goldstine left the IAS to head mathematical research at IBM and eventually became one of the fabled IBM fellows. After retirement he served as executive officer of the American Philosophical Society. In 1983 Goldstine received the National Medal of Science for contributions to the invention of the computer (see Figure 2.9). He died at Bryn Mawr near Philadelphia in June 2004.

⁵³The history of decompositions and condition numbers is discussed in sections 3.5.1 and 3.7.

⁵⁴Enrico Fermi [35, p. 383] also strongly advocated using computers in the sciences.



Fig. 2.9 *Herman Goldstine receiving the National Medal of Science from President Reagan in February 1985 “for fundamental contributions to the development of the digital computer, computer programming and numerical analysis.” The year of record for the award is 1983. Courtesy of the Ronald Reagan Library.*

3. Inversion Paper Highlights. This section describes von Neumann and Goldstine’s contributions in the inversion paper through ten sketches corresponding to the topics mentioned in section 2.9 and highlighted in Figure 2.8. Each is self-contained but with cumulative notation and terminology.

3.1. Introduction to Research Summaries. In the terminology of section 2.1, this section is an internalist history of the contributions made in the inversion paper. Because an intellectual history of modern numerical analysis is not available, it is necessary to rely on the primary sources. Some entirely new material is uncovered in the correspondence between Goldstine and von Neumann. In particular, von Neumann’s letter [302] shows that he personally discovered the matrix condition number (later terminology) in rounding error bounds. Other discoveries are (in modern terminology) that von Neumann and Goldstine proposed stability as a unifying analytical concept for understanding error in computations; applied the Courant–Friedrichs–Lewy paper to stability; devised the exact-round principle of ANSI/IEEE arithmetic; established the deleterious effect of the normal equations; formulated the concept of backward error as a means to measure the errors of calculations; and devised the decomposition paradigm to study the errors of elimination algorithms.

3.2. Mathematical Stability. *Von Neumann and Goldstine used stability as a theme to explain the effects of various types of error in scientific computing. They suggested that discretization error could be understood in this manner by applying the paper of Courant, Friedrichs, and Lewy [61] to illustrate the difference between the stability of a problem and of its numerical approximation.*

Error in Scientific Computing. The inversion paper identified four sources of error in computational science: (\mathcal{A}) physical theory, (\mathcal{B}) data preparation or measurement, (\mathcal{C}) truncation and discretization, and (\mathcal{D}) rounding error {pp. 1024–1025}.⁵⁵ Alston Householder⁵⁶ [145, p. 15] included the basic four errors (\mathcal{A})–(\mathcal{D}) in his early textbook, and they have appeared in some form in all numerical analysis books ever since.

Setting the original presentation apart are the relationships drawn between the errors. Von Neumann and Goldstine dismissed theory (\mathcal{A}) as no concern to mathematicians. Ditto the data errors (\mathcal{B}), except that “their influence on the result is the thing that really matters. In this way, their analysis requires an analysis of this question: What are the limits of the change of the result, caused by changes of the parameters (data) of the problem within given limits? This is the question... somewhat more loosely worded, of the *mathematical stability of the problem*” {p. 1027, original emphasis}. Von Neumann and Goldstine went on to suggest that the errors (\mathcal{C}), (\mathcal{D}), like (\mathcal{B}), should also be interpreted through stability. The inversion paper thus identified stability as a theme with which to interpret all the approximations of disparate origin in scientific computing.

Stable and Stability. Studies of fluid and solid motion in the 19th century transformed the attribute “stable” into the concept of “stability” which was to be investigated in its own right.⁵⁷ For physical systems, stability may result from, but is not the same as, governing principles such as least action. Stability is characterized as exhibiting changes that are uniformly and commensurately small in response to perturbations. The definition of stability by uniformly bounded trajectories for dynamical systems is due to Lyapunov at the end of the 19th century [174]. The introduction of stability concepts to the social sciences is attributed to von Neumann and Morgenstern in the 20th century [179].

Numerical analysis acquired many “stabilities” after World War II, some of which originated with or were inspired by von Neumann. The theory of approximate solutions to ODEs is based on Lyapunov stability [65], but the stability condition for non-stiff equations is based on “von Neumann polynomials” [197], in analogy to PDEs. Von Neumann began developing a stability theory of approximate solutions to PDEs in seminars at Los Alamos in 1947 while the inversion paper was being written [115, p. 288].

Courant Criterium and von Neumann Criterion. Because discretizing differential equations replaces one problem with another, von Neumann and Goldstine realized that the effect of (\mathcal{C}) on stability must be understood. The inversion paper only alludes to the developing theory for PDEs in citing a readily available result: “That the stability of the strict problem need not imply that of an arbitrarily close approximant was made particularly clear by some important results of R. Courant, K. Friedrichs, and H. Lewy” {pp. 1028–1029}.⁵⁸ Von Neumann and Goldstine referred to a paper that considered the existence of solutions to differential equations by examining whether solutions to finite difference equations converge [170, p. 14]. In the same year as the

⁵⁵These errors were discussed previously: Lonseth [177, p. 332] mentioned (\mathcal{B}), (\mathcal{D}), Mauchly [186] discussed (\mathcal{C}), (\mathcal{D}), and Goldstine and von Neumann [119, pp. 15–16] earlier included machine malfunctions and human mistakes.

⁵⁶Householder 1904–1993 was a student of Bliss and a mathematical biologist who became a numerical analyst. He remains the only person to be president of both the ACM and SIAM.

⁵⁷Research for viscous flows focused on the instabilities observed in a few basic cases [68], whereas Routh [240] discussed stability itself in an Adams prize paper selected by no less than J. C. Maxwell and G. G. Stokes.

⁵⁸Biographies of C., F., and L. have been written by Reid [232, 233, 234].

inversion paper, Richard Courant [60] also described his work with Friedrichs and Lewy when he surveyed “theoretical questions, basic for numerical procedure.” He explained the “time-steps must be sufficiently small relative to the space-mesh” for the discrete solutions to converge to the continuum solution. Courant did not relate this result to a concept of stability, which appears to have been entirely von Neumann’s contribution. At Los Alamos von Neumann promoted a “Courant criterium” that became the CFL stability condition [115, p. 288].

Von Neumann also developed his own methods to ensure stable approximate solutions [170, pp. 4–5]. Von Neumann’s method of analyzing stability—Fourier stability analysis of finite difference methods—finally appeared in a paper in 1950:

[The spatial and temporal spacing] Δs and Δt must be small in comparison to the space and time scales of the physically relevant motions. But this does not alone insure accuracy; the small-scale motions for which there is inevitably a large distortion may possibly be amplified in the course of computation to such an extent that they will totally obscure the significant large-scale motions. In the following we shall derive the criteria which insure that no such amplification will take place.

— Charney, Fjörtoft, and von Neumann [53, p. 273]

Von Neumann’s method of achieving stability—artificial viscosity for shock capturing—appeared in a paper written with Richtmyer [316] at the same time. These techniques had great social impact because they enabled early calculations in meteorology and shock physics (see Figure 3.1). Eventually, a theory built around the Lax equivalence theorem [171], [237, p. 25] would show that stability criteria such as von Neumann’s are necessary for convergence of the approximate solutions; Lax acknowledged von Neumann’s contributions to that theory:⁵⁹

... at Los Alamos, [where] we met again, he was always available. When I latched on to some of his ideas like von Neumann criterion, stability, shock capturing, he was interested in my ideas and I profited from his ideas.

— Peter Lax [170, p. 9]

Stability for approximate solutions of PDEs remains an active area of research; see Gustafson, Kreiss, and Olinger [134, Chaps. 5, 12] and Trefethen and Embree [277, Chap. 7].

And Rounding Errors Similarly. Von Neumann and Goldstine expected the computers then being built would perform calculations of unprecedented length [119, p. 29]. Indeed, a few years later the first weather prediction to be made electronically required one million multiplications and divisions, and more were anticipated [53, p. 275]. Since the rounding errors in such calculations “will be individually small [and] appear in large numbers,” von Neumann and Goldstine believed “the decisive factor that controls their effect” should again be a “stability phenomenon” {p. 1030}. They expected that new methods of analysis would be needed to examine this behavior.

Any analysis would be difficult because, in contrast to data errors (\mathcal{B}) and discretization errors (\mathcal{C}), the rounding errors (\mathcal{D}) reoccur differently with each and every arithmetic operation. Moreover, the “Courant criterium” revealed that obscure relationships in the numerical problem could govern the amplification of errors. Von

⁵⁹Lax received the Abel prize in part for this work [208]. In addition to Lax [170], the history of stability criteria is mentioned by Aspray [10, p. 105], Dalmedico [66, pp. 399–400], Goldstine [115, pp. 287–288], Oran and Boris [213, p. 240], Richtmyer [237, pp. v–viii], and Taub [266, Vol. 5, p. 664].



Fig. 3.1 Von Neumann stability analysis was described in the paper [53] where it determined the time step for the first barotropic weather forecast. Marking the accomplishment are visitors (*) from the Weather Bureau [218] and some members of the IAS meteorology project [115, p. 301]. From left to right: H. Weiler*, J. von Neumann, M. H. Frankel*, J. Namias*, J. Freeman, R. Fjørtoft, F. Reichelderfer*, and J. Charney. They are surrounded by ENIAC, which made the forecast at Aberdeen Proving Ground in April 1950. Courtesy of the MIT Museum.

Neumann and Goldstine {pp. 1030–1031} thus proposed to undertake a demonstration analysis. The problem considered should (i) be one “in which the difficulties due to (\mathcal{D}) are a maximum,” (ii) have “a high order iterative character” with potentially “very large numbers of ‘elementary’ operations,” and (iii) have “inherently low, or rather insecure, stability,” whereby “noise introduced in an early stage... should be exposed to a possibility of considerable amplification.” For this example they chose “the well known *elimination method*, or rather, a new variant of it that we shall develop.” After a long demonstration analysis of the errors of elimination they returned to the sources of error at the very end of their paper. Further discussion of this topic is deferred until section 3.8.

3.3. Machine Arithmetic. Von Neumann and Goldstine introduced a formal model of machine arithmetic as a basis for analyzing errors in algorithms. Today this characterization is codified in ANSI/IEEE arithmetic standards. They also suggested double precision operations to reduce arithmetic error, and they analyzed matrix multiplication to demonstrate the advantages of mixing precisions.

Exact-Round Principle (this paper’s terminology). Von Neumann and Goldstine formalized the relationship between real arithmetic and machine arithmetic. They introduced special notation for machine-representable numbers (the overbars which are still used today), and for machine arithmetic operations (called *pseudo* as opposed to

true).⁶⁰ Relations between true and pseudooperations were derived from the assumption that the machine operates exactly and then rounds consistently.

The transition from the true operations to their pseudo-operations is effected by any one of the familiar methods of *round off*.

— inversion paper {p. 1026, original emphasis}

Von Neumann and Goldstine are here proposing that computers should be thought of as performing each arithmetic operation exactly, introducing error only when the resulting real number is rounded to a machine-representable number.

The explicit relationship between true and machine arithmetic illustrates von Neumann's renowned use of the axiomatic method.⁶¹ It allows machine errors to be characterized by one quantity, a "roundoff unit" u , while avoiding details of machine implementation.⁶² This principle provided a basis for the type of analysis of mechanized algorithms that is now called rounding error analysis. Von Neumann and Goldstine used the principle to model fixed point arithmetic {p. 1026},

$$\text{computed}(\bar{x} \bullet \bar{y}) = \bar{x} \bullet \bar{y} + e \quad \begin{cases} |e| \leq u & \text{for } \times \text{ and } \div, \\ e = 0 & \text{for } + \text{ and } -, \end{cases}$$

where u is the roundoff unit. Among prominent early numerical analysts, Householder [145] and Givens [109] adopted von Neumann and Goldstine's fixed point model.⁶³

The extension to floating point arithmetic was circuitous. The earliest computers implemented floating point operations through software that strictly enforced the exact-round principle. Nevertheless, Carr [49] was unsuccessful in modeling floating point arithmetic. Wilkinson [326, p. 321] succeeded, noting "all the operations... give the correctly rounded results" and then introducing without comment the device of interpreting the roundoff error as a relative quantity:

$$\text{computed}(\bar{x} \bullet \bar{y}) = (\bar{x} \bullet \bar{y})(1 + e), \quad |e| \leq u.$$

When floating point hardware finally became available as an expensive "scientific" option, then manufacturers forsook the exact-round principle for economy and speed. Wilkinson [329, pp. 11–13] had to modify his model with different bounds for each operation. Anomalies in the arithmetic supplied by manufacturers eventually became so troublesome that von Neumann and Goldstine's principle was enforced by establishing an arithmetic standard for microprocessors:⁶⁴

every operation... shall be performed as if it first produced an intermediate result correct to infinite precision..., and then rounded that result according to one of the modes in this section.

— ANSI/IEEE Std 754-1985 [153, p. 10]

⁶⁰This terminology was used earlier in an IAS report [42, p. 56].

⁶¹Von Neumann's talent for axiomatization is attested to by Halmos [135, p. 394] and Ulam [286, p. 10]. Von Neumann is said to have met his greatest challenge in axiomatizing quantum mechanics [167].

⁶²The roundoff unit has varying names and notation. The inversion paper and Wilkinson [326, 329] introduced none. For fixed point arithmetic Householder [146, p. 236] had the "unit rounding error" ϵ . For floating point arithmetic Forsythe and Moler [97, p. 91] had the "unit round-off" u . Another name is "machine epsilon."

⁶³James Wallace Givens 1910–1993 was a president of SIAM.

⁶⁴Some of the trouble with mainframe arithmetic is described by Higham [140, pp. 43–44] and Knuth [163, p. 223]. The origin of the microprocessor standard is described by the principal author, William Kahan [248].

Arithmetic Organ. Von Neumann and Goldstine applied their exact-round principle to round-to-nearest rounding of s fractional digits in base β . Their machine-representable numbers were proper fractions whose arithmetic incurred an absolute error of at most $u = \beta^{-s}/2$ whenever a true result exceeded s fractional digits {pp. 1026–1027}. The inversion paper did not choose values for β and s .⁶⁵

Both the inversion paper and the plans for the IAS computer arithmetic organ (now arithmetic unit, or ALU) were written concurrently. One instance can be documented where the inversion paper suggested operations that are still used today. Since the inverse of a representable matrix might not have all its entries in the interval $(-1, +1)$, von Neumann and Goldstine devised a scale-down operation, *iterated halving*, that performed consecutive divisions, each separately rounded {pp. 1039–1041}:

$$\text{computed}(\bar{x}/2^p) = \text{rounded}(\underbrace{\text{rounded}(\dots \text{rounded}(\bar{x}/2) \dots /2)/2}_p).$$

The earliest plan for the IAS computer [42, p. 75] had unary operations for scaling. “The immediate uses are arithmetical but... [there are] other, combinatorial uses” [120, pp. 84–85], so $\times 2$ and $\div 2$ became bit shifts in the IAS computer, which led to the bit shift operations in modern computers [163, p. 226].

Double Precision. Burks, Goldstine, and von Neumann had several precedents to suggest a computer should be capable of the full set of arithmetic operations on numbers of double precision (original terminology) and higher multiples of the native arithmetic format. Hand computers had been advised to retain only the digits they considered significant (see section 2.5) so they varied the number of digits during a calculation. Experience with mechanical calculators suggested using instead high precision to avoid rounding errors. Because desk calculators performed long multiplication, the register in which the product accumulated had as many digits as the multiplier and multiplicand combined. If the accumulator were not cleared after each product, then a sum of products could accumulate to full precision. The ability to sum products without error was considered “theoretically more accurate” because it gave “the approximation resulting from a number of operations rather than the combination of approximations resulting from the operations” [82, p. 112]. Von Neumann personally wanted multiple precision arithmetic for very high precision. He used ENIAC to calculate thousands of digits of e and π to investigate the digits in transcendental numbers [196].⁶⁶

The IAS reports made the intellectual leap from mechanical calculators (see Figure 3.2) to an electronic arithmetic unit capable of multiple precision calculations. Burks, Goldstine, and von Neumann considered this capability so important, and the complexity of providing it was so high, that it was among the reasons they gave for neglecting floating point arithmetic. The IAS reports described the first double precision hardware [163, p. 278], and the inversion paper illustrated its use to reduce arithmetic error.⁶⁷

⁶⁵The inversion paper identified likely values for β and s {p. 1026, fn. 2; p. 1032, fn. 13}. The IAS reports advocated $\beta = 2$ [42, sec. 5.2, pp. 41–43], which was definitive in promoting binary arithmetic [163, p. 202]. The arithmetic of the IAS computer was 40-digit, fixed-point, binary, and 2’s-complement (the last is modern terminology).

⁶⁶A first-person account of von Neumann’s fascination with the calculations is [27, pp. 44–45].

⁶⁷The desirability of multiple precision is explained in the reports [42, pp. 50, 145] and [120, pp. 127–128, 142–151]. Floating point arithmetic is commented on in the inversion paper {p. 1035, fn. 18} and in the same reports [42, pp. 43, 73–74] and [120, p. 113].

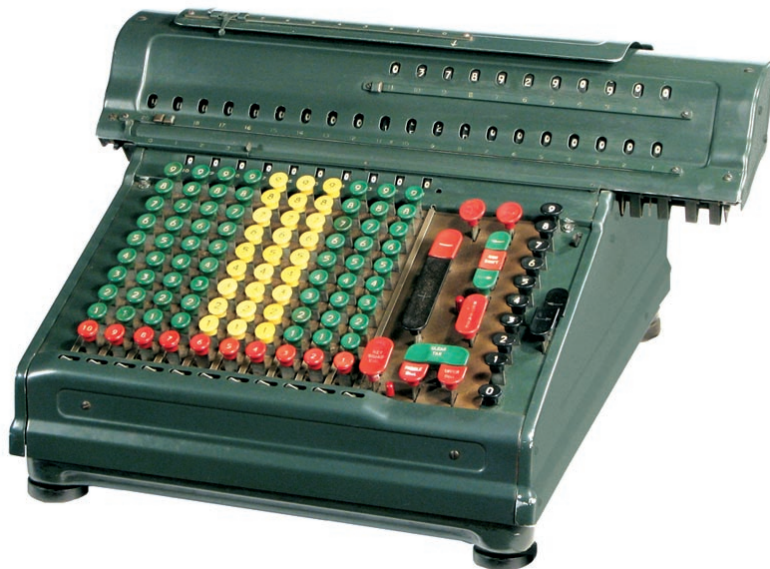


Fig. 3.2 High precision arithmetic was an artifact of long multiplication in machines such as the one pictured here and its manual predecessors. The moving carriage holds an 11-digit multiplier register and a 20-digit accumulator (visible at top); the stationary base holds a 10-digit multiplicand register (barely visible above the uppermost keys). Internal mechanisms varied but the operating principles were the same: the carriage moves to align the multiplicand with each multiplier digit. The multiplicand is entered through the keyboard. The multiplier keyboard is the single column of black keys on the far right. Pressing a multiplier key adds the multiplicand to the accumulator the chosen number of times, while recording the multiplying digit in the multiplier register; the carriage then moves to the next place. This Marchant “Silent Speed” Model 10ACT had over 4,000 parts in a unique continuous-drive mechanism that rendered it unmatched for quiet, reliability, and speed. It was built in Oakland, California, in the 1930s and 1940s. Photo courtesy of the John Wolff Web Museum <http://home.vicnet.net.au/~wolff/>.

The several examples of mixed precision arithmetic in the inversion paper are based on inner products. If \bar{x} and \bar{y} are machine-representable column vectors of length n , then the error in computing their inner product, $\bar{x}^t \bar{y}$, is nominally $n\beta^{-s}/2$ because only the multiplications have fixed point errors. Alternatively, the true products have just $2s$ digits, so they may be summed in double precision before a final rounding to single precision, which is then the only error {p. 1037, eqns. 2.7.b' and 2.7.b''}:

$$|\text{computed}(\bar{x}^t \bar{y}) - \bar{x}^t \bar{y}| \leq \begin{cases} n\beta^{-s}/2, & \text{single precision,} \\ \beta^{-s}/2, & \text{double precision accumulation.} \end{cases}$$

Goldstine [113, p. 4] drafted a complete analysis of Gaussian elimination using “this stunt,” but the published inversion paper only remarked that mixed precision could reduce the error bounds {p. 1038, top}. Much later Wilkinson [329, pp. 6–7, 23–25] advocated mixed precision to reduce error. Double and higher precisions continue to be used and studied, for example by Pope and Stein [221] and more recently by Bailey [13], Demmel and Hida [75], and Lang and Bruguera [169].

3.4. Matrix Norms. *The inversion paper systematically described the properties of norms in matrix algebra and thereby introduced matrix norms to numerical analysis. Von Neumann and Goldstine also described the properties of the matrix lower bound, which is perhaps the only topic in the paper not well known today.*

How von Neumann Thought About Norms. Von Neumann [292, 297] was an expert on norms, having begun the abstract study of Hilbert spaces and their operators and having characterized the unitarily invariant matrix norms. He apparently thought of norms as upper and lower bounds,

$$(3.1) \quad \text{upper bound: } |A| = \max_{|x|=1} |Ax| \quad (= \|A\|_2 \text{ in modern matrix notation}),$$

$$(3.2) \quad \text{lower bound: } |A|_\ell = \min_{|x|=1} |Ax| \quad (= \|A^{-1}\|_2^{-1} \text{ for nonsingular } A).$$

Von Neumann used this terminology and both of these formulas (original notation) in personal correspondence [302], in IAS manuscripts [119], and in the inversion paper {p. 1042}. Indeed, he originally defined (what is now) the matrix condition number as $|A|/|A|_\ell$ (see Figure 2.6).

Numerical analysts appear to have learned about matrix norms from the inversion paper. The properties of norms were known: von Neumann [297] had studied the unitarily invariant matrix norms, Wittmeyer [337] had catalogued some properties of matrix norms, and Rella [235] had stated the axioms sometimes used to characterize multiplicative norms [38]. For these reasons, or perhaps because (3.1) is the definition of an operator norm given by Banach [15], the inversion paper remarked, “the properties of these entities are too well known to require much discussion” {p. 1042}. Nevertheless, Forsythe⁶⁸ [95] cited the inversion paper for properties of norms and Wilkinson [333, p. 192] stated “there had been no previous systematic analysis of . . . the use of norms in matrix algebra.” Norms were not needed to make calculations, but once von Neumann and Goldstine proved theorems that could only be expressed using matrix norms, then norms were clearly needed to study calculations.

No Standard Notation. The influence of the inversion paper can be traced through a curious notation for what today is the Frobenius norm. The notation and terminology for matrices were “far from agreed” at the time of the inversion paper.⁶⁹ Early users of what now are matrix norms included Frobenius [105], [262, p. 71]. Frobenius defined the *Spannung* as $\vartheta(A) = \sum_{i,j} |A_{i,j}|^2$, established a product rule for $\vartheta(A)$, and attributed to Jordan an additive rule for $[\vartheta(A)]^{1/2}$. Wedderburn [319, pp. 125, 171] chose the latter as his *absolute value*. He wrote it with thickened bars, $\mathbf{|A|}$, to avoid notation for determinants, $|\cdot|$, and for matrices, which were then enclosed by (\cdot) , $[\cdot]$, $\{\cdot\}$, or $\|\cdot\|$. A few years later Hotelling [143, p. 11] deliberately called Wedderburn’s quantity *the norm* and wrote $N(A)$.

Of course, von Neumann knew that in functional analysis the notation was $|\cdot|$ and the name was either *Absolutwert* in German [292, p. 54] or *norme* in French [15, p. 53]. There being no agreed notation for matrices, he followed Hotelling’s lead. Thus, the inversion paper had $|A|$ and $N(A)$, but called only the latter a norm {p. 1042}. Statisticians adopted N from Hotelling [85] and, because von Neumann endorsed it, numerical analysts appear to have adopted it from him. For example, Turing [285, p. 297], Householder [145, pp. 39, 40], and Bodewig [34, p. 43] had B , M , N and M ,

⁶⁸George Forsythe 1917–1972 [162] was instrumental in the development of computer science.

⁶⁹The lack of standardization is remarked upon in [104, pp. 1–2] and [180, p. 4].

b , N and $|\cdot|$, m , N , respectively.⁷⁰ There was no common practice except the N , which Todd [272] and Givens [109] also used.⁷¹

Matrix Lower Bound. Gene Golub [125] observed that the lower bound in (3.2) is frequently overlooked in the inversion paper, although it is important for perturbation analysis. Von Neumann and Goldstine exhibited many familiar inequalities for the upper bound paired with a full set of parallel or curiously reversed inequalities for the lower bound {pp. 1042–1044}. For example, they compared the triangle inequality for upper bounds (norms) to the following inequality for lower bounds (their notation):

$$(3.3) \quad |A|_{\ell} - |E| \leq |A + E|_{\ell} \leq |A|_{\ell} + |E|.$$

A lower bound for all matrices that is compatible with operator norms, as well as generalizations for von Neumann and Goldstine's properties, are given in [129] and excerpted in [29].

3.5. Decomposition Paradigm. *Von Neumann and Goldstine helped develop one of the “top ten” algorithmic principles of all time [261], that linear algebra calculations perform matrix decompositions, and they are responsible for the corollary, that the decompositions can be used to study the errors in the algorithms.*

3.5.1. Triangular Factoring. *Von Neumann and Goldstine were the last of several authors to interpret Gaussian elimination as triangular factoring (the decomposition paradigm), and they discovered the LDU form.*

History of Gaussian Elimination. The history of Gaussian elimination in Europe consists of three lines of development.⁷² The oldest development began as a textbook explanation for solving simultaneous equations that by 1800 had evolved into a standard collection of lessons for solving linear equations.

The second development began when Gauss [107] formulated new calculations for the method of least squares. Gauss took advantage of the symmetry in the normal equations to reduce the work and, equally important, he made the work purely arithmetical so that algebraic equations were no longer manipulated. Hand computers “universally adopted” Gauss's approach [54, p. 514]. Thereafter professional methods evolved in response to changing needs and technologies. Methods were invented for use with multiplication tables by Doolittle [80], or for use with mechanical calculators by Cholesky [26], and to efficiently use calculators with other than normal equations by Crout [63].

The third line of development began in the 1930s when several authors offered matrix descriptions of algorithms within a short span of time. The citation analysis in Figure 3.3 identifies six groups or individuals who made unprecedented contributions: Banachiewicz; Frazer, Collar, and Duncan; Jensen; Dwyer; Satterthwaite; and von Neumann and Goldstine.

1. The earliest author to equate calculations with decompositions appears to have been Tadeusz Banachiewicz. His “Cracovian calculus” [164] was a non-Cayleyan matrix algebra of column-by-column products that he used for least squares calculations.

⁷⁰In modern notation the three norms are, respectively, $\|\cdot\|_2$, the maximum magnitude, and $\|\cdot\|_F$.

⁷¹The present notation appears to have arisen as follows. Ostrowski [214] reserved N to represent a general norm in 1955, which Newman and Todd [205, p. 227] adopted in 1962 by moving Hotelling's N to F for *Frobenius*. Mathematics gradually settled on $\|\cdot\|$, which was also adopted by Householder [147] and Wilkinson [329]. The notation $\|\cdot\|_F$ appeared no later than 1973 [259, p. 173].

⁷²See [128] for elaboration and for biographical material about many of the people mentioned here.

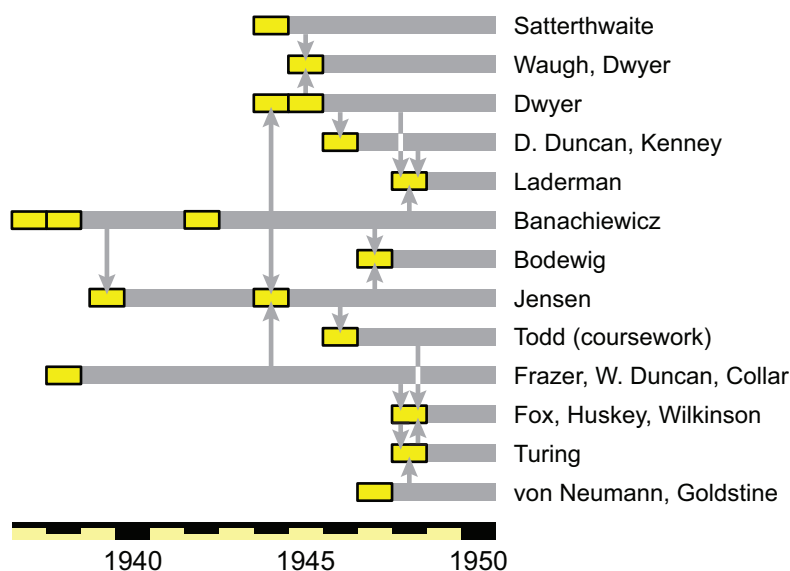


Fig. 3.3 Citation patterns among the earliest authors who mention Gaussian elimination in the form of a matrix decomposition. Arrows show the flow of information to the citing papers at the head from the cited oeuvre at the tail. For example, Dwyer in 1944 cited one or more works of Banachiewicz. Boxes indicate published papers, except in the case of Todd, where a course that he taught in 1946–1947 is indicated. The papers are Banachiewicz [16, 17, 18, 19, 20], Bodewig [32], Duncan and Kenney [81], Dwyer [83, 84], Fox, Huskey, and Wilkinson [103], Jensen [154, 155], Laderman [168], von Neumann and Goldstine [314], Satterthwaite [243], Turing [285], and Waugh and Dwyer [318]. Todd's course is described by Fox, Huskey, and Wilkinson [103, p. 159] and by Taussky and Todd [268].

2. Robert Frazer, Arthur Collar, and William Duncan used matrix notation to formulate structural problems in a precursor to the finite element method at the NPL [93]. They viewed elimination as “building up the reciprocal matrix in stages by elementary operations,” which could produce a triangular matrix “such that its reciprocal can be found easily,” but they did not write out a matrix decomposition [104, pp. 96–99].
3. Inspired by Banachiewicz, Henry Jensen used multiplication by the elementary matrices of Frazer, Collar, and Duncan to show that the methods of Gauss and Cholesky calculated triangular decompositions. Jensen [155, pp. 15, 22] used pictograms to emphasize that the methods amounted to expressing the coefficient matrix of the normal equations as $N = \Delta \nabla$. Taussky and Todd [268] report that Jensen's matrix form of Cholesky's method was taught to Leslie Fox, who communicated it to his colleagues at the NPL, including Alan Turing and James Wilkinson.
4. Paul Dwyer [83, 84] showed that his “abbreviated” or “compact” Doolittle method was an “efficient way of building up” some “so called triangular” matrices from which he developed a “square-root” method (Cholesky's method). Laderman [168] reports that Dwyer's square-root method became popular in the U.S. and was used at the NBS Computation Laboratory, which was the official name for the Mathematical Tables Project shortly before its demise [132, p. 290].

5. An earlier paper of Dwyer inspired Franklin Satterthwaite [243, p. 374] to calculate inverse matrices from the triple “factorization” $A = (R_1 + I) S_1 (I + T_1)$, where S_1 is diagonal and R_1 and T_1 are, respectively, “pre-” and “postdiagonal” (strictly triangular). Satterthwaite appears to have had little influence because his paper received only ten citations, of which four were by Dwyer.

The Combination of Two Tricks. Von Neumann and Goldstine independently discovered triangular decomposition. Whereas many of the earlier sources treated general matrices in a supplementary way because they were primarily concerned with SPD matrices, instead von Neumann and Goldstine began with a nonsymmetric decomposition. They described the three-term LDU form, and were the first to explain how the algorithm taught to school children produces these factors {pp. 1052–1056, secs. 4.3–4.5}.

Specifically, von Neumann and Goldstine described the elimination algorithm as producing a sequence of ever-smaller matrices and vectors from $Ax = y$ (original notation),

$$A = A^{(1)}, A^{(2)}, A^{(3)}, \dots, A^{(n)}, \quad y = y^{(1)}, y^{(2)}, y^{(3)}, \dots, y^{(n)},$$

where the rows and columns of $A^{(k)}$ and $y^{(k)}$ are numbered from k to n {p. 1051, eqns. 4.10–4.11}. For $k = 1, \dots, n-1$, the computation is {p. 1051, eqns. 4.6–4.7},

$$(3.4) \quad A_{i,j}^{(k+1)} = A_{i,j}^{(k)} - A_{i,k}^{(k)} A_{k,j}^{(k)} / A_{k,k}^{(k)} \quad \text{for } i, j > k,$$

$$(3.5) \quad y_i^{(k+1)} = y_i^{(k)} - (A_{i,k}^{(k)} / A_{k,k}^{(k)}) y_k^{(k)} \quad \text{for } i > k.$$

Next, the algorithm solves by backward substitution the equations $B'x = z$, where the entries of B' and z are chosen from the reduced matrices and vectors (the first row of each). For a matrix C of the multipliers $A_{i,k}^{(k)} / A_{k,k}^{(k)}$ with $i \geq k$ (note the unit diagonal), von Neumann and Goldstine summed (3.5) over k to give $Cz = y$. From this equation and $B'x = z$ they concluded $CB' = A$ {p. 1053, eqns. 4.18–4.20}:

We may therefore interpret the elimination method as . . . the combination of two tricks: First, it decomposes A into a product of two semi-diagonal matrices. . . [and second] it forms their inverses by a simple, explicit, inductive process.

— inversion paper {p. 1053}

This passage is the clearest statement of the invention of triangular factoring in any of the primary sources. Von Neumann and Goldstine found a lack of symmetry in the elimination algorithm because the first factor always had 1's on its main diagonal. They divided the second factor by its diagonal to obtain $B' = DB$, hence $A = CDB$, which they called “a new variant” of Gaussian elimination {p. 1031} that is now written LDU .

Lower and Upper. Von Neumann and Goldstine also contributed to the terminology of factoring. Their contemporaries used various adjectives to distinguish the two kinds of triangular matrices: Frazer, Duncan, and Collar [104, pp. 96–99] had the “handedness” of a triangular matrix. A curious variant occurred with Jensen [155, p. 6], who placed the *zeros* “under” and “over” the diagonal. He was imitated by Bodewig [32, Part I, p. 444] who again placed the *zeros* “links” and “rechts.” Instead, von Neumann and Goldstine described the *nonzero* location as “lower” and “upper,” which is still done today. Their name for triangular—“semidiagonal” {p. 1046, bot.}—obviously did not catch on but was noted by Householder [145, p. 65].

Paradigm Shift. “Paradigms” [166] arise when many researchers adopt similar methods. Von Neumann and Goldstine were not the first to describe triangular factoring, but the use they made of it assured they were the last. The factors calculated by elimination were notational conveniences for the earlier discoverers, but von Neumann and Goldstine used them to both characterize the algorithm and analyze its rounding errors. They proved the first theorems about existence and uniqueness {p. 1052, Thm. 4.14; p. 1056, Thm. 4.33}, and they used the decomposition to express aggregate relationships among computed quantities that could be analyzed with higher mathematics. In their correspondence (see section 2.9) they examined different groupings of the triangular and diagonal factors before ultimately proving that the upper and lower bounds of (what now is called) the Cholesky factor appear in bounds for the rounding errors when inverting SPD matrices {p. 1061, eqn. 5.11; p. 1062, eqn. 5.13; p. 1093, eqn. 7.5}. They called the Cholesky factor the “truly fundamental” quantity {p. 1062} because its upper and lower bounds are tied to the matrix by equality relations. Von Neumann and Goldstine thus related numerical errors to properties of the matrix through the decomposition. They share credit with many people for matrix factoring, but the use of the factorization to analyze rounding error was their discovery.

3.5.2. Small A Priori Error Bounds. *Von Neumann and Goldstine found that the factorization of SPD matrices has an error bound that is a polynomial of the matrix order, not an exponential as Hotelling had suggested. The bound does not depend on computed values, so it bounds the errors “a priori” (see section 1.4). SPD matrices remain the most significant class for which triangular factoring has such an error bound. Von Neumann and Goldstine also discovered that inverting SPD matrices has a similar error bound that is linear in a quantity, ℓ (see Figure 2.6), which is now called the matrix condition number.*

Enabling Technology. Applying the decomposition paradigm to derive error bounds for mechanized Gaussian elimination requires numerous properties of SPD matrices.⁷³ The first property is that the singular values (modern terminology) of the particular triangular factor $D^{1/2}U$ (Cholesky factor) are directly related to those of $A = U^t D U$ {p. 1062, eqns. 5.13.a, b},

$$(3.6) \quad \|D^{1/2}U\|_\ell = \|A\|_\ell^{1/2} \quad \text{and} \quad \|D^{1/2}U\|_2 = \|A\|_2^{1/2},$$

where $\|\cdot\|_\ell$ is the lower bound of (3.2). Second {p. 1058, Lem. 5.6},

$$(3.7) \quad \text{if } A^{(k)} \text{ is symmetric and positive definite, then so is } A^{(k+1)}$$

and, moreover, the lower and upper bounds are nested:

$$(3.8) \quad \|A^{(k)}\|_\ell \leq \|A^{(k+1)}\|_\ell \quad \text{and} \quad \|A^{(k+1)}\|_2 \leq \|A^{(k)}\|_2.$$

Third, any SPD matrix satisfies {p. 1058, eqn. 5.4.d}

$$(3.9) \quad \max_i A_{i,i} = \max_{i,j} |A_{i,j}|.$$

Finally {p. 1061, Lem. 5.8},

$$(3.10) \quad \max_{i,j} |A_{i,j}| \geq \max_{i,j,k} |A_{i,j}^{(k)}|.$$

⁷³Deriving these properties was second nature to von Neumann, who had needed similar relationships in his study of unitarily invariant matrix norms [297].

Schwarz [245, Satz 1.7] rediscovered (3.7). Wilkinson [330, p. 629, Thm. 1] rediscovered the right, or upper, inclusion of (3.8). Equation (3.10) was rediscovered by and has been attributed to Wilkinson [242, p. 76], [328, pp. 285–286].

Later researchers evidently found it more convenient to work backward from the success of the inversion paper than to read its proofs. Many facts about SPD matrices pertinent to factoring were rediscovered, so the cumulative impression arose that von Neumann and Goldstine were only the first of many researchers who assembled well-known facts to bound the errors of triangular factorization.⁷⁴ As a result, von Neumann and Goldstine's achievement in discovering that SPD matrices allow error bounds to be proved is not appreciated.

Fixed Point Arithmetic. Von Neumann and Goldstine found that reordering the matrix was needed to obtain an error bound for fixed point arithmetic. Their analysis of matrix decomposition was for the LDL^t factorization (or U^tDU as they preferred to write it) of a machine-representable SPD matrix \bar{A} of order n .

The main step in the algorithm, (3.4), evaluates ab/c , where a and b are entries and c is the pivot (not their terminology) in a reduced matrix. The grouping $(ab)/c$ divides the proper fraction c into the error that fixed point rounding adds to ab . This division would amplify the error by the factor c^{-1} {p. 1063, eqn. 6.3.a}, which is at least 1 and potentially much larger. Thus von Neumann and Goldstine had to use the other grouping, $a(b/c)$. For b/c to be a valid fixed point operation it must be a proper fraction, so “positioning for size” (modern full or complete reordering or pivoting) was needed to make the magnitude of c the largest in the reduced matrix.

Under mild assumptions about \bar{A} and using (3.7), they showed that the computed, reduced matrices are SPD {p. 1066, eqn. 6.16}, so by (3.9) the largest entry occurs on the diagonal, hence the reordering is symmetric. Thus with symmetric reordering the algorithm proceeds to its conclusion and produces factors \bar{D} and \bar{U} for which von Neumann and Goldstine showed {p. 1069, eqn. 6.34}

$$(3.11) \quad \|\bar{A} - \bar{U}^t \bar{D} \bar{U}\|_2 \leq \mathcal{O}(n^2 \beta^{-s}),$$

where n is the matrix order and where $\beta^{-s}/2$ is the largest error of rounding for the fixed point arithmetic described in section 3.3.

The analysis of matrix inversion was performed for the product of the inverses of the factors, $\bar{U}^{-1} \bar{D}^{-1} (\bar{U}^t)^{-1}$. Because a matrix whose entries are in the interval $(-1, +1)$ may not have a representable inverse, von Neumann and Goldstine instead computed a scaled inverse $\bar{W}^{(q_0)}$ that was reduced in magnitude by the factor $1/2^{q_0}$ for an integer $q_0 \geq 0$. Much of the analysis involved the choice of q_0 . Again under mild assumptions on \bar{A} , they produced an approximate inverse with {p. 1093, eqn. 7.5'}

$$(3.12) \quad \|\bar{A} (\bar{W}^{(q_0)} 2^{q_0}) - I\|_2 \leq \underbrace{\mathcal{O}(\|\bar{A}\|_2 \|\bar{A}^{-1}\|_2)}_{\ell} n^2 \beta^{-s}.$$

Equation (3.11) was the first error analysis of Gaussian elimination, and (3.12) was the first error analysis of matrix inversion. Both error bounds are essentially unchanged to this day, even for floating point arithmetic. SPD matrices remain the most significant class for which Gaussian elimination has been proved to have a priori error bounds. Diagonally dominant matrices [328, pp. 288–289] and M -matrices

⁷⁴For example, see [126, p. 145], [140, p. 209], and [330, p. 629].

[36] also have bounded growth of entries in their factors (a prerequisite for deriving error bounds), but their literature lacks complete proofs of error bounds comparable to those for SPD matrices.

Floating Point Arithmetic. A measure of the inversion paper's impact is that many years later Wilkinson [328] reprised inversion, even though, by then, it was clear that inverse matrices have comparatively little use. Wilkinson was only partly successful in his plan to overcome the inversion paper's limitations of SPD matrices, fixed point arithmetic, and reordering.

For SPD matrices and fixed point arithmetic, reordering is needed (von Neumann and Goldstine's case) unless mixed precision arithmetic is available (Wilkinson's case). Wilkinson [328, pp. 295] presumed computers could form products errorlessly and divide single precision numbers into them. Thus $(ab)/c$ could incur only the error of rounding to single precision, so this formula could be used to avoid the grouping $a(b/c)$ that necessitated reordering. He analyzed the inner product form of the Cholesky factorization in this way [328, pp. 305–307].

For SPD matrices and floating point arithmetic, Wilkinson [330] found that neither reordering nor mixed precision are needed. Because floating point rounding errors are multiplicative rather than additive, dividing into a perturbed quantity only slightly increases the relative error. Demmel [78] restated Wilkinson's floating point analysis of the SPD case.

General Matrices. Independent of whether the arithmetic is fixed or floating point, there are two impediments to analyzing triangular factoring of non-SPD matrices: (i) the inability to relate a general matrix and its triangular factors through the magnitudes of their entries or norms (von Neumann and Goldstine explained that this problem stymied them, although they conjectured that such a relation might exist with reordering {p. 1056}); (ii) the computed reduced matrices do not have the backward errors of later steps, so the computed reordering is made with incomplete knowledge and hence might not obey the reordering rule.⁷⁵ Wilkinson found a solution to problem (i) in Hadamard's inequality, but that solution required complete or full reordering or pivoting, so he could not avoid problem (ii). Wilkinson acknowledged the gap in a remark that is easy to overlook in a long paper [328, p. 300, above eqn. 15.1], so his proof failed [140, p. 165, after Thm. 9.5].⁷⁶



3.5.3. A Posteriori Bounds. In later work that is now considered definitive, Wilkinson analyzed triangular factoring of general matrices in the context of solving linear equations. The first part of his analysis much resembles von Neumann and Goldstine's because Wilkinson assumed no particular computer arithmetic [329, pp. 94–96, sec. 3.14]. Both calculations begin with matrices $[\bar{a}_{i,j}^{(1)}] = \bar{A}$ of order n , and they proceed through elimination steps $k = 1, \dots, n - 1$ to produce matrices $[\bar{a}_{i,j}^{(k)}]$. Still, it is necessary to account for different decompositions and notation (see Table 3.1). Since von Neumann and Goldstine had a symmetric matrix, their calculation was {pp. 1064–5, eqns. 6.3, before 6.11}

$$(3.13) \quad \bar{a}_{i,j}^{(k+1)} = \bar{a}_{i,j}^{(k)} - \bar{a}_{k,i}^{(k)} \bar{U}_{k,j} + \eta_{i,j}^{(k)}, \quad k < i \leq j.$$

⁷⁵The final perturbed matrix is always the true product of the computed factors. Because the reordering is chosen before all perturbations are created, the factors might not obey the reordering rule—with respect to the matrix that has all the perturbations (their true product)—so the reordering cannot be invoked to guarantee a relationship between the computed factors and their product.

⁷⁶Many numerical analysts mistakenly believe Wilkinson established an a priori bound; for example, see [101, p. 682, eqn. 12].

Table 3.1 Differences between the error analyses of von Neumann and Goldstine and of Wilkinson for triangular decomposition. Wilkinson did not use overbars, but they are applied here for consistency of notation. The order of the matrix is n .

	1947, von Neumann and Goldstine [314], inversion paper	1963, Wilkinson [329], <i>Rounding Errors in Algebraic Processes</i>
use of the factorization	inverting matrices	solving equations
type of matrix	SPD \bar{A}	general \bar{A}
factorization	$\bar{A} \approx \bar{U}^t \bar{D} \bar{U}$	$\bar{A} \approx \bar{L} \bar{U}$
diagonal treatment	all 1's on diagonal of \bar{U}	all 1's on diagonal of \bar{L}
k th reduced matrix	$[\bar{a}_{i,j}^{(k)}]$ for $1 \leq k \leq i, j \leq n$ where $[\bar{a}_{i,j}^{(1)}] = \bar{A}$	
factors	$\bar{D}_{k,k} = \bar{a}_{k,k}^{(k)}$ $\bar{U}_{k,j}$ = computed $(\bar{a}_{k,j}^{(k)} / \bar{a}_{k,k}^{(k)})$	$\bar{L}_{i,k}$ = computed $(\bar{a}_{i,k}^{(k)} / \bar{a}_{k,k}^{(k)})$ $\bar{U}_{k,j} = \bar{a}_{k,j}^{(k)}$
error at i, j in step k	first called $\eta_{i,j}^{(k)}$, then $\theta_{i,j}^{(k)}$	$\epsilon_{i,j}^{(k+1)}$
total error at i, j	$\zeta_{i,j}$	$e_{i,j}$
final error bound on	$ \zeta_{i,j} $	$ e_{i,j} $
type of error bound	a priori	a posteriori
type of arithmetic	fixed point	at first any, then floating

Von Neumann and Goldstine did not calculate the lower triangle, but in order to treat the lower part symmetrically in the analysis, they replaced (3.13) by

$$(3.14) \quad \text{VNG: } \bar{a}_{i,j}^{(k+1)} = \bar{a}_{i,j}^{(k)} - \bar{U}_{k,i} \bar{D}_{k,k} \bar{U}_{k,j} + \theta_{i,j}^{(k)}, \quad k < i, j,$$

in which $\theta_{i,j}^{(k)}$ includes both the error $\eta_{i,j}^{(k)}$ and the error of replacing $\bar{a}_{k,i}^{(k)}$ by $\bar{U}_{k,i} \bar{D}_{k,k}$. Wilkinson's calculation is [329, p. 95, eqns. 15.1, 15.4–15.7]

$$(3.15) \quad \text{W: } \begin{cases} \bar{a}_{i,j}^{(k+1)} = \bar{a}_{i,j}^{(k)} - \bar{L}_{i,k} \bar{U}_{k,j} + \epsilon_{i,j}^{(k+1)}, & k < i, j, \\ 0 = \bar{a}_{i,k}^{(k)} - \bar{L}_{i,k} \bar{U}_{k,k} + \epsilon_{i,k}^{(k+1)}, & k < i. \end{cases}$$

Note that Wilkinson's second equation defines the error of computing the multiplier $\bar{L}_{i,k}$.⁷⁷

Von Neumann and Goldstine summed the error equations. Their terminating condition was {p. 1068, eqn. 6.26}

$$(3.16) \quad 0 = \bar{a}_{i,j}^{(k)} - \bar{U}_{k,i} \bar{D}_{k,k} \bar{U}_{k,j} + \theta_{i,j}^{(k)}, \quad k = \min\{i, j\},$$

which defined the error of normalizing the entries of \bar{U} . For fixed i, j , summing equations (3.14) and (3.16) for $k = 1, \dots, \min\{i, j\}$ gave {p. 1068, eqn. 6.28}

$$(3.17) \quad \text{VNG: } \bar{A}_{i,j} = \sum_{k=1}^{\min\{i,j\}} \bar{U}_{k,i} \bar{D}_{k,k} \bar{U}_{k,j} + \zeta_{i,j},$$

⁷⁷These equations are in the original papers except (a) for readability, notation for entries of reduced matrices has been replaced by symbols for computed factors when possible; (b) to save space, von Neumann and Goldstine's step index k has been introduced to Wilkinson's analysis (he preferred ellipses); and (c) again for readability, some auxiliary subscripts by which von Neumann and Goldstine accounted for symmetry have been removed.

where {p. 1068, after eqn. 6.28}

$$(3.18) \quad \text{VNG: } \zeta_{i,j} = \sum_{k=1}^{\min\{i,j\}} \theta_{i,j}^{(k)}.$$

Wilkinson [329, p. 95] emphasized the additive nature of the error terms; accordingly, he summed the errors in the same manner as von Neumann and Goldstine. For fixed i, j , summing (3.15) for $k = 1, \dots, \min\{i-1, j\}$ gives [329, pp. 95–96, eqns. 15.2 and 15.8]

$$(3.19) \quad \begin{aligned} \bar{U}_{i,j} &= \bar{A}_{i,j} - \sum_{k=1}^{i-1} \bar{L}_{i,k} \bar{U}_{k,j} + e_{i,j} & i \leq j, \\ 0 &= \bar{A}_{i,j} - \sum_{k=1}^j \bar{L}_{i,k} \bar{U}_{k,j} + e_{i,j} & j < i, \end{aligned}$$

where the total error is [329, pp. 95–96, eqns. 15.3 and 15.9]

$$(3.20) \quad \text{W: } e_{i,j} = \sum_{k=1}^{\min\{i-1,j\}} \epsilon_{i,j}^{(k+1)}.$$

Since $\bar{L}_{i,i} = 1$, Wilkinson's equation (3.19) for the upper triangle is identical to the equation for the lower triangle, so they can be stated as one:

$$(3.21) \quad \text{W: } 0 = \bar{A}_{i,j} - \sum_{k=1}^{\min\{i,j\}} \bar{L}_{i,k} \bar{U}_{k,j} + e_{i,j}.$$

Thus far Wilkinson's final or definitive analysis in (3.21) and (3.20) resembles von Neumann and Goldstine's earlier analysis in (3.18) and (3.17). Both produced approximate decompositions,

$$\begin{aligned} \text{VNG: } \bar{A} &= \bar{U}^t \bar{D} \bar{U} + [\zeta_{i,j}], & \zeta_{i,j} &= \sum_{k=1}^{\min\{i,j\}} \theta_{i,j}^{(k)}, \\ \text{W: } 0 &= \bar{A} - \bar{L} \bar{U} + [e_{i,j}], & e_{i,j} &= \sum_{k=1}^{\min\{i-1,j\}} \epsilon_{i,j}^{(k+1)}, \end{aligned}$$

in which the error terms $\zeta_{i,j}$ and $e_{i,j}$ sum the errors committed at matrix position i, j over all elimination steps. Wilkinson's analysis diverges from von Neumann and Goldstine's in bounding the errors.

Von Neumann and Goldstine proved that all the numbers in their calculation are valid fixed point quantities of magnitude at most 1 provided the matrix is SPD. Since no scale factors are needed to adjust magnitudes, the errors $\theta_{i,j}^{(k)}$ in (3.14) are at worst a small multiple, 2, of the roundoff unit, so the error $\zeta_{i,j}$ in (3.17) grows at worst linearly with the number of elimination steps at position i, j .

Lacking a bound on the entries of the reduced matrices (as explained at the end of section 3.5.2), Wilkinson instead adopted Turing's approach [285, p. 305, top] of retaining computed values in the error bound. (Turing's paper is described in section

2.10.) If the calculation reorders rows so the divisors $\bar{a}_{k,k}^{(k)}$ have largest magnitude in their columns (modern row pivoting or row reordering), then for floating point arithmetic Wilkinson [329, p. 98, eqn. 16.9] showed that the error $\epsilon_{i,j}^{(k+1)}$ in (3.15) is at worst a small multiple, roughly 2, of the roundoff unit times the largest magnitude, g , in any of the reduced matrices. As in von Neumann and Goldstine's case, the bound for $e_{i,j}$ in (3.20) grows at worst linearly with the number of elimination steps at position i, j .⁷⁸

Wilkinson's error bound is additionally scaled by the computed quantity g . The bound is therefore "a posteriori," because it can be evaluated only after the calculation. Wilkinson [329, p. 97, bottom] explained that sufficient computational experience had accumulated by the 1960s to suggest g remains small for row reordering, although there are contrived examples of exponential growth.⁷⁹ Thus, although von Neumann and Goldstine's conjecture is rendered moot by computational experience, it remains open: Can a small a priori error bound be proved for any reordering rule?

3.6. Normal Equations. *Von Neumann and Goldstine discovered that the problem to which Gauss applied Gaussian elimination, the normal equations, is qualitatively different from directly inverting a matrix because ℓ^2 (what is now the condition number squared) appeared in its error bound.*

Trumping Hotelling. Von Neumann and Goldstine's purpose was to disprove the conjecture of Hotelling [143, p. 11] that errors could grow exponentially:

The reasons for taking up this subject are...the rather widespread revival of mathematical interest in numerical methods, and...that considerably diverging opinions are now current as to the extremely high or extremely low precisions which are required when inverting matrices of orders $n \geq 10$.

— inversion paper {p. 1022}

At first, von Neumann and Goldstine had tried to bound the errors of forming the inverse directly from the triangular factors. When they were unable to bound the errors of this algorithm except for SPD matrices, they then used a more complicated algorithm to supply the missing error bound for inverting unsymmetric matrices:

Since we have solved the problem of inverting a matrix in the case when it is definite, we wish to replace the problem of inverting A_I by that of inverting an appropriate definite matrix.

— inversion paper {p. 1081}⁸⁰

This approach was chosen by von Neumann [302], who described it in the letter to Goldstine from his holiday in Bermuda.

Aitken's Formula. The IAS researchers proposed to use $A_I^{-1} = (A_I^t A_I)^{-1} A_I^t$.⁸¹ The paper actually used $A_I^{-1} = A_I^t (A_I A_I^t)^{-1}$, which is stated only late in the proof {pp. 1081, 1085}. Hotelling gave no names to the formulas. Von Neumann used the phrase "a modified method" in his letter [302] and in the IAS manuscript based on his lectures [119, p. 14]. Today the formulas are called normal equations by analogy with least squares problems.

⁷⁸Wilkinson [329, p. 98, eqn. 16.13] uses an elaborate tableau to show the $\min\{i, j-1\}$ dependence of the bound on i, j .

⁷⁹As further evidence that g is likely bounded, with row and column reordering, Wilkinson [329, p. 97, eqn. 16.6] stated his bound from [328, p. 284, eqn. 4.15] which assumes true arithmetic.

⁸⁰The subscript I = indefinite is used in the inversion paper to denote a general matrix.

⁸¹The formula appeared in the earlier IAS report [21, p. 432], in von Neumann's letter from Bermuda, and in the inversion paper {p. 1056}. It was found in Hotelling [143, p. 11], who got it from Dwyer [82, p. 127], both of whom credited it to Aitken [5].

Von Neumann and Goldstine proved that using the modified or normal equations method to invert a general matrix \bar{A}_I of order n , had the following error bound {p. 1093, eqn. 7.5'}:

$$(3.22) \quad \|\bar{A}_I (\bar{W}^{(q_1)2^{q_1}}) - I\|_2 \leq \underbrace{\mathcal{O}(\|\bar{A}_I\|_2^2 \|\bar{A}_I^{-1}\|_2^2)}_{\ell_I^2} n^2 \beta^{-s},$$

where ℓ_I is the figure of merit (modern condition number) for \bar{A}_I . Again it was necessary to compute a scaled inverse $\bar{W}^{(q_1)}$ that was reduced in magnitude by a factor $1/2^{q_1}$ for some integer $q_1 \geq 0$. The significant change was that the figure of merit appears squared in (3.22). Much would be made of this discovery, later, but at the time it hardly seemed to matter because the polynomial bound trumped Hotelling, because any bounds were suspected of being overestimates, and because there were no computers with which to test the bounds.

Good Enough for Gauss. Von Neumann and Goldstine remarked that ways other than using the normal equations might yet be found to disprove Hotelling's claim:

For these reasons, which may not be absolutely and permanently valid ones, we shall restrict the direct application of the elimination method...to definite matrices.

— inversion paper {p. 1056}

Despite von Neumann and Goldstine's hesitation to use the normal equations, Bodewig [33], in reviewing the inversion paper, inserted a more proactive voice: "an arbitrary matrix M must first be transformed into a definite one $M^t M$." In this way the discovery of (3.22) led to the situation that von Neumann and Goldstine are often remembered as having "unfortunately" "recommended" the normal equations.⁸² This interpretation distorts the historical record because von Neumann and Goldstine stated their ambivalence about normal equations, above, and also because they conjectured that an a priori error bound might be established for nonsymmetric elimination with reordering.

Von Neumann's choice of what formula to analyze was auspicious because solving normal equations for least squares problems was the main use for Gaussian elimination during most of its existence [128]. The contrasting results in (3.12) and (3.22) raised the question of whether normal equations have inherent drawbacks. The squared quantity in (3.22) is von Neumann and Goldstine's figure of merit for the matrix $(\bar{A}_I)^t \bar{A}_I$. This led to theoretical investigations by Taussky [267] and as late as 1965 by Olkin [184] into whether, for other definitions, a condition number of $A^t A$ or AA^t necessarily exceeds that of A .

3.7. Condition Numbers. *Von Neumann and Goldstine inspired studies of quantitative measures for their concept of the stability of mathematical problems to perturbations. Eventually their figure of merit, ℓ , was recognized as "the" matrix condition number.*

Sensitivity Analysis of Matrix Calculations. In the 20th century, before the inversion paper, several authors examined how errors in preparing coefficients or performing arithmetic might affect the solving of linear equations.⁸³ These considerations may have been important from the earliest days. Human computers were advised to keep

⁸²For example of this usage, see [89, p. 11] and [217, p. 23].

⁸³Authors include Etherington [91] in 1932, Hotelling [143] in 1943, Lonseth [177] in 1942, Moulton [199] in 1913, Tuckerman [280] in 1941, and Wittmeyer [337] in 1936.

Table 3.2 Measures of sensitivity of the inverse of A , or of solutions of linear equations with coefficient matrix A , to perturbations of the entries or to rounding error during the calculation. (Some values are restated in modern notation, in particular, $\lambda(A)$ and $\sigma(A)$ are the eigenvalues and singular values of A .)

DATE	SOURCE	NAME	VALUE	REFERENCES
1942 1947	Lonseth	measure of sensitivity	$\sum_{i,j} \text{cofactor}_{i,j}(A)/\det(A) $ $= \sum_{i,j} (A^{-1})_{i,j} $	[177, p. 336] [178, p. 198]
1947	von Neumann and Goldstine	figure of merit, ℓ	$\ A\ _2/\ A\ _\ell$ $= \sigma_{\max}(A)/\sigma_{\min}(A)$ $= \ A\ _2 \ A^{-1}\ _2$	see Figure 2.6, [119, p. 14], [302], {p. 1093, eqn. 7.5}
1948	Turing	N -condition number	$\ A\ _F \ A^{-1}\ _F / n$	[285, p. 298]
		M -condition number	$n \max_{i,j} A_{i,j} \max_{i,j} A_{i,j}^{-1} $	[285, p. 298]
1949	Todd	P -condition number	$\max \lambda(A) / \min \lambda(A) $	[272, p. 470]

only the digits judged significant (see section 2.5) lest they get “six-figure results from three-figure data” [244, p. 1].

Various numerical indicators of the difficulty of performing accurate matrix calculations were proposed in the 1940s (see Table 3.2).

1. Lonseth identified a “measure of sensitivity” to perturbations in the coefficients. Originally formulated with cofactors, he later recognized his value was the absolute sum of the entries in the inverse. Other determinant formulas continued to be developed into the 1970s [30].
2. Von Neumann and Goldstine’s a priori error bounds for matrix inversion contained a factor, ℓ , which they called the “figure of merit” for the “amplification” of rounding errors (see sections 2.6, 2.7, and 2.10).
3. Turing gave the name “condition number” to two values of his own (see section 2.10). His “ M -condition number” occurred in an a posteriori bound on the rounding errors of Gauss–Jordan elimination. His “ N -condition number” appeared in an estimate for the standard deviation of the change in the solution of linear equations caused by random fluctuations in the coefficients.

Turing’s name *condition number* came from *ill-conditioned*, which he said was an epithet for difficult matrix calculations [285, p. 297].⁸⁴ Turing’s description of these difficulties,

... small percentage errors in the coefficients may lead to large percentage errors in the solution.
— Alan Turing [285, p. 298]

matched von Neumann and Goldstine’s earlier definition of stability:

... the limits of the change of the result, caused by changes of the parameters (data) of the problem within given limits
— inversion paper {p. 1027}

⁸⁴Evidence for this usage prior to Turing’s paper is scant; see Fox, Huskey, and Wilkinson [103, pp. 154–155] in 1948 and Mallock [182, pp. 461, 474] in 1933. In survey triangulations, the notion of ill-conditioned *triangles* appears to be older; see Ledger in 1882, Rankine in 1872, and Tait in 1875 [172, p. 3], [227, p. 24], [265, p. 21].

Goldstine reports that he and von Neumann “discussed our ideas” [115, p. 291] with Turing, who acknowledged they “communicated them to the writer” [285, p. 288] during his visit to Princeton.

The Matrix Condition Number. Von Neumann and Goldstine’s figure of merit became “the” matrix condition number through the work of John Todd. He gave Turing’s name to von Neumann and Goldstine’s value, calling it the P -condition number.⁸⁵ Todd and many others investigated the properties of Gaussian elimination by evaluating the N - and P -condition numbers of several matrices; see references in [274]. Todd eventually abandoned Turing’s values to use only von Neumann and Goldstine’s quantity:

... because the most detailed error analyses (of which we know) have been carried out in terms of this quantity, and because our experience with both fixed and floating point routines convince us that this condition number is a reasonable one to use.

— Newman and Todd [204, p. 467] in 1958

Even after Turing’s values were neglected, von Neumann and Goldstine’s figure of merit, ℓ , continued to be known as the P -condition number in work of Davis, Haynsworth, and Marcus [72], [183] as late as 1962. Thereafter, von Neumann and Goldstine’s quantity was accepted as simply *the* matrix condition number.

Perturbation Analysis of Matrix Inversion. It was not until some years after the papers of von Neumann and Goldstine and of Turing that matrix condition numbers were exhibited in perturbational inequalities. For example, Bargmann, Montgomery, and von Neumann [21, p. 473] in 1946 derived

$$\|X^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2 \|X - A\|}{1 - \|A^{-1}\| \|X - A\|},$$

in which no condition number is apparent, whereas Bauer [24] in 1959 noted

$$\frac{\|X^{-1} - A^{-1}\|}{\|A^{-1}\|} \leq \|A\| \|A^{-1}\| \frac{\|X - A\|}{\|X\|},$$

in which the condition number is visible. The similar inequality (2.1) for linear equations was stated by Wilkinson [329, p. 29] in 1963. The advantage of having an agreed value to reproduce was invaluable in guiding the development of theories of computational error. The “figure of merit” or matrix condition number continues to inspire research to quantify the difficulty of accurate calculations, such as that by Dedieu [74] and Demmel [76, p. 252], [77].

3.8. Backward Errors. *Von Neumann and Goldstine introduced the concept of (what today is called) backward error, and they inspired Wallace Givens to invent (what today is known as) the method of backward rounding error analysis to establish bounds on these errors. This type of analysis is attributed to Givens [326, p. 319], [332, p. 554], but it has not been realized that his work rested on the prior invention of the backward error by von Neumann and Goldstine.*

⁸⁵Todd chose the P for Princeton [272, p. 470], [274, p. 144]. Von Neumann originally stated the ratio of singular values (see Figure 2.6), but because the matrices of early interest were usually SPD, the P -condition number often was defined as a ratio of eigenvalues. For example, Forsythe used the ratio of eigenvalues [98, p. 10], [99, p. 340].

Invention of Backward Errors (Modern Terminology). This discussion continues from section 3.2. Von Neumann and Goldstine had explained the effects of the measurement error (\mathcal{B}) by introducing the idea of stability as a measure of the sensitivity of the result of a calculation to changes in the initial data {p. 1027}. They alluded to related stability concepts for discretization errors (\mathcal{C}) by discussing the CFL condition {pp. 1028–1029}. Whereas errors (\mathcal{B}), (\mathcal{C}) were committed only once, the rounding errors (\mathcal{D}) were “continuously injected” each time an arithmetic operation was performed. Nevertheless, because the rounding errors “will be individually small [and] appear in large numbers,” von Neumann and Goldstine believed “the decisive factor that controls their effect” should again be a “stability phenomenon” {p. 1031}. This application of stability required a novel interpretation of numerical accuracy:

An approximate inverse of a matrix \bar{A} might be defined as one which lies close to the exact inverse \bar{A}^{-1} . From the standpoint of computational procedure it seems more appropriate, however, to interpret it as the inverse $(A')^{-1}$ of a matrix A' that lies close to \bar{A} , that is, to permit an uncertainty of, say ϵ , in every element of \bar{A} . Thus we mean by an approximate inverse of \bar{A} a matrix $W = (A')^{-1}$, where all elements of $A' - \bar{A}$ have absolute values not greater than ϵ .

—inversion paper {p. 1092}, variable names altered

Von Neumann and Goldstine are here proposing that rounding errors should be understood in terms of compensating, one-time perturbations to the initial data of a calculation (modern backward errors). They were so taken with this idea that they used it as a rule to interpret mathematical properties numerically: a property is true for some numerical data if and only if all nearby data have that property. Von Neumann and Goldstine illustrated this concept in two elaborate examples for nonsingularity and positive definiteness {pp. 1092–1094}.

Invention of Backward Error Analysis. A few years after the inversion paper, Givens referred to the sources of error by the letters (\mathcal{A})–(\mathcal{D}) that von Neumann and Goldstine had used. He restated von Neumann and Goldstine’s view that data preparation errors (\mathcal{B}) should be taken into account when assessing the rounding error (\mathcal{D}). Givens wanted to extend the concept of data perturbations to problems other than matrix inversion:

[A] method of inverting the problem of round-off error is proposed. . . which suggests that it may be unwise to separate the estimation of round-off error from that due to observation and truncation (cf. (\mathcal{B}), (\mathcal{C}), and (\mathcal{D}) of {pp. 1024–1025} for a clear discussion of these terms).

—Wallace Givens [109, pp. 1–3]

He then used von Neumann and Goldstine’s notation (bars for machine-representable data, \bar{d} , and primes for perturbed data, d') to describe his method. The inversion paper inspired him to invent a method of solving the inverse problem for data perturbations, $d' - \bar{d}$. Givens called his method “inverse error analysis” (modern backward error analysis).⁸⁶

Givens [109, pp. 57–58] described the use for his method as follows. A numerical calculation is an approximation, \mathcal{N} , of a theoretical function, \mathcal{T} (original notation and terminology). Rounding errors make \mathcal{N} discontinuous and therefore impossible to study by continuum methods. If a d' can be found for which $\mathcal{T}(d') = \mathcal{N}(\bar{d})$, then

⁸⁶The word “inverse” had previously been used in this way by Scarborough [244, p. 21] for the maximum errors allowable in measurements to achieve a prescribed accuracy in a function of the measurements.

the error of the calculation is

$$\mathcal{N}(\bar{d}) - \mathcal{T}(\bar{d}) = \mathcal{T}(d') - \mathcal{T}(\bar{d}),$$

so the continuity of \mathcal{T} can be used to bound the error, provided $d' - \bar{d}$ is small. Today, $\mathcal{N}(\bar{d}) - \mathcal{T}(\bar{d})$ is called forward error and $d' - \bar{d}$ is called backward error (see section 1.4).

Actual Errors of Matrix Inversion. Von Neumann and Goldstine did not transform (3.12) and (3.22) into bounds on the forward or backward error. They actually commented on this omission, saying: “We leave the working out of the details, which can be prosecuted in several different ways, to the reader” {p. 1093}. Householder [150, p. 55] gave the relevant bounds in his textbook,

$$(3.23) \quad \begin{aligned} \text{forward error:} \quad & \frac{\|X - A^{-1}\|_2}{\|A^{-1}\|_2} \leq \|AX - I\|_2, \\ \text{backward error:} \quad & \frac{\|A' - A\|_2}{\|A\|_2} \leq \frac{\|AX - I\|_2}{1 - \|AX - I\|_2}, \end{aligned}$$

where $X \approx A^{-1}$ and where A' is the matrix whose true inverse is X . With these formulas, von Neumann and Goldstine’s bounds in (3.12) and (3.22) become bounds on the forward and backward errors.

3.9. Accuracy Criterion. *Von Neumann and Goldstine proposed the accuracy criterion that compares backward errors to perturbations (such as measurement errors) created when preparing input data for a calculation. This criterion was later rediscovered by Werner Oettli and William Prager.*

Principia. At the end of his famous book on the laws of motion, Isaac Newton made some predictions about the trajectories of comets. When his predictions did not exactly agree with observations, Newton proposed a criterion for the acceptable level of agreement:

From all this it is plain that these observations agree with the theory, so far as they agree with one another. (Congruunt igitur hae observationes cum theoria quatenus congruunt inter se.)

— Isaac Newton [206, p. 505]

Here Newton refers to the fact that it is pointless to ask for agreement beyond the measurement error in the observations.

Inversion Paper. Von Neumann and Goldstine applied Newton’s criterion—with a simple but unexpected twist—to check the accuracy of calculations based on data from inexact observations. They supposed it possible to perturb the input data of the calculation by backward errors (modern terminology) of size ϵ to account for the rounding errors of the calculation (see section 3.8). Let M be the size of Newton’s observational errors of measurement or other preparation errors in the input data. While M comes with the physical problem and cannot be changed, ϵ is controlled by the choice of algorithm. “It seems reasonable to take this into consideration, when we analyze what concept and what degree of approximation [in the algorithm] are significant” {p. 1092, fn. 33}. Von Neumann and Goldstine are suggesting:

If the backward errors are less than the data errors, $\epsilon \leq M$, then the problem is solved to the extent the problem is known, so the calculation must be accepted.

Oettli and Prager. Oettli and Prager [211] independently discovered the accuracy criterion and extended it with a further twist. They used optimization theory to find the minimal size, μ (in a certain norm), of the backward errors for linear equations. The criterion $\mu \leq M$ not only is sufficient but also is necessary for accuracy in von Neumann and Goldstine's sense.⁸⁷ Wilkinson edited Oettli and Prager's paper and coauthored one of Oettli's two subsequent papers on the subject [210, 212]. Despite Wilkinson's endorsement and the wealth of similar results, von Neumann and Goldstine's and Oettli and Prager's accuracy criterion remains to be fully exploited.

3.10. Stochastic Linear Algebra. *Von Neumann and Goldstine used the probability distribution of matrix condition numbers to assess the suitability of computers for matrix calculations.*

Probability Distributions. The immediate purpose of the inversion paper was to decide whether the first computers would be able to perform useful calculations once they were built. Von Neumann and Goldstine concluded the paper by determining a range of matrix orders for which a random matrix could likely be inverted accurately using the arithmetic precision of these computers. The final sentence of the inversion paper is:

This might provide a basis for estimating what degrees of precision are called for [when solving] this problem by various possible types of procedures and equipment.

— inversion paper {p. 1099}

Mathematical Reviews [33] summed up the inversion paper, saying: "By this result the usual opinions as to the [extreme] precision required when inverting matrices of high order are refuted." Matrices of order 100 were large at that time.

Some estimates for the orders of matrices that could be inverted had appeared in the earlier IAS report about matrix calculations [21, p. 477]. The rigorous error bounds in the inversion paper showed that $\ell = \|A\|_2 \|A^{-1}\|_2$ merited serious study. The theory of distributions derived from random matrices originated in quantum mechanics [191] and multivariate statistics [336]. Von Neumann and Goldstine's application to numerical analysis started a special line of investigation that continues today in stochastic linear algebra in the work of Edelman [88, 89, 90] and Smale [253, pp. 92–94, 118–119]. For a summary of relevant statistical research, see Johnstone [157].

General Matrices Inverted by the Normal Equations. The inversion paper quoted results from Bargmann {p. 1048, fn. 24; p. 1097} that general, invertible matrices A_I of order n have a distribution for $\|A_I\|_2 \|A_I^{-1}\|_2$ that peaks near n for large n . Bargmann's results were never prepared for publication [89, p. 12, fn. 6], but similar results were obtained by Edelman in a dissertation that won the Householder prize. If A_I has entries from a standard normal distribution, then the probability distribution for $\|A_I\|_2 \|A_I^{-1}\|_2 / n$ quickly approaches a limit for large n that peaks near 1.71. From the tail of the distribution it can be shown that [89, p. 23, Tab. 2.1; p. 69, Thm. 7.1]

$$(3.24) \quad \Pr [\|A_I\|_2 \|A_I^{-1}\|_2 / n \leq 10] = 0.80.$$

⁸⁷Although Oettli and Prager were inspired by backward error analysis [211, p. 405], they evidently did not know Givens had invented it, so they were unaware that their work led through Givens to von Neumann and Goldstine.

Table 3.3 Probable value of the error bounds for inverting matrices of order n .

MATRIX TYPE	PROVEN ERROR BOUND	ℓ CONFIDENCE INTERVAL	PROBABLE ERROR BOUND
general, \bar{A}_I	(3.22), $\mathcal{O}(\ell^2 n^2 \beta^{-s})$	(3.24), $\ell \leq \mathcal{O}(n)$	(3.25), $\mathcal{O}(n^4 \beta^{-s})$
SPD, \bar{A}	(3.12), $\mathcal{O}(\ell n^2 \beta^{-s})$	(3.26), $\ell \leq^* \mathcal{O}(n^2)$	

The estimate (*) was assumed by von Neumann and Goldstine on phenomenological grounds.

This identity gives an 80% chance that $\|\bar{A}_I\|_2 \|\bar{A}_I^{-1}\|_2$ in (3.22) can be replaced by $10n$,

$$\|\bar{A}_I (\bar{W}^{(q_1)} 2^{q_1}) - I\|_2 \leq \mathcal{O}(\|\bar{A}_I\|_2^2 \|\bar{A}_I^{-1}\|_2^2 n^2 \beta^{-s}) \leq \mathcal{O}(1) 10^2 n^4 \beta^{-s}.$$

If this bound is less than 1 (so the computed inverse has at least some accuracy),

$$(3.25) \quad \begin{aligned} \mathcal{O}(1) 10^2 n^4 \beta^{-s} &\leq 1 \\ &\Updownarrow \\ n &\leq (10^{-2} \beta^s / \mathcal{O}(1))^{1/4}, \end{aligned}$$

then by von Neumann and Goldstine's reasoning, a range of n has been found for which a randomly chosen matrix has an 80% chance of being inverted with some accuracy. For their $\mathcal{O}(1) = 14.24$, $\beta = 2$, and $s = 40$, Edelman's range $n \leq 131$ nearly matches von Neumann and Goldstine's estimate $n \leq 150$ {p. 1099, eqn. 7.18.c}. Edelman also determined a 98% confidence interval for which the limit drops to $n \leq 43$.

SPD Matrices Inverted Directly. Von Neumann and Goldstine believed that the same estimates might apply for their other error bound. Since an SPD matrix usually originates as $A = A_I^t A_I$,⁸⁸ von Neumann and Goldstine supposed that the relevant distribution for $\|A\|_2 \|A^{-1}\|_2$ resembles that of $(\|A_I\|_2 \|A_I^{-1}\|_2)^2$ {p. 1098}. If so, then

$$(3.26) \quad \begin{aligned} \Pr [\|A\|_2 \|A^{-1}\|_2 / n^2 \leq 100] &\stackrel{?}{\approx} \Pr [(\|A_I\|_2 \|A_I^{-1}\|_2 / n)^2 \leq 100] \\ &= \Pr [\|A_I\|_2 \|A_I^{-1}\|_2 / n \leq 10] \\ &= 0.80. \end{aligned}$$

The hypothesis gave an 80% chance that $\|\bar{A}\|_2 \|\bar{A}^{-1}\|_2$ in (3.12) could be replaced by $100n^2$, so, as before,

$$(3.27) \quad \begin{aligned} \|\bar{A} (\bar{W}^{(q_0)} 2^{q_0}) - I\|_2 &\leq \mathcal{O}(\|\bar{A}\|_2 \|\bar{A}^{-1}\|_2 n^2 \beta^{-s}) \\ &\leq \mathcal{O}(1) 100 n^4 \beta^{-s} \leq 1 \\ &\Updownarrow \\ n &\leq (100^{-1} \beta^s / \mathcal{O}(1))^{1/4}, \end{aligned}$$

which is the same conclusion reached in (3.25).

John Todd's Counterexample. In summary, von Neumann and Goldstine believed that SPD matrices encountered in practice were likely to have a larger figure of merit, ℓ , than randomly chosen general matrices. This counterbalanced the advantage of the smaller error bound for the inversion method applicable to SPD matrices (see Table 3.3). The different powers, n^1 and n^2 , that bounded confidence intervals for ℓ were

⁸⁸When the inversion paper was written, such products occurred when using the normal equations to solve least squares problems. Note that there appears to be no work on the condition of "random" SPD matrices.

matched with different powers, ℓ^2 and ℓ^1 , respectively, in the error bounds. Thus, the likely error bound for matrix inversion was $\mathcal{O}(n^4\beta^{-s})$, whether or not the matrix was SPD.

Todd [272] challenged this conclusion in 1949. There was no experience with condition numbers until he pointed out that a certain class of SPD matrices, of practical interest, had condition numbers $\ell = \mathcal{O}(n)$ far below the hypothesized $\mathcal{O}(n^2)$ [272, p. 470, sec. 3]. The $\mathcal{O}(n^3\beta^{-s})$ error bound for these matrices suggested an advantage over the general case. Todd's example meant it was significant that different powers of condition numbers appeared in the error bounds that von Neumann and Goldstine had derived for different algorithms. This discovery justified subsequent efforts to find algorithms with smaller error bounds for general matrices. It would be 11 years after the inversion paper before Givens [110] and Householder [148] discovered provably superior orthogonal factorizations, 14 years before Wilkinson [328, pp. 310–318] supplied a rounding error analysis of the factorizations, and 18 years before Golub [124] used Householder's method to solve least squares problems.

3.11. Remembering the Inversion Paper Imperfectly. The matrix inversion paper helped establish the legitimacy of computerized elimination. Surveying methods for solving equations in 1950, Householder [144] wrote “one may agree with von Neumann and Goldstine, Bodewig, and others that the Gaussian method of elimination is the best direct method in general.” Before computers existed to make errors, von Neumann and Goldstine had managed to prove the errors would not grow exponentially.

Soon it would be possible to simply observe the errors directly. For example, as part of the acceptance tests for the Harvard Mark II relay computer in 1948, Mitchell [198] inverted the matrix model of the U.S. economy for which Wassily Leontief would win a Nobel prize [67, p. 20]. Mitchell mistook von Neumann and Goldstine's error bounds for estimates, and noted the “predictions” far exceeded his observations. Some misinterpretations of von Neumann and Goldstine's results even group von Neumann and Goldstine with Hotelling in predicting exponentially large errors; for example, see [289, p. 573].

Wilkinson [332, p. 567] lamented that error analysis does overestimate the errors and explained why the bounds are far from sharp. He ascribed the problem to norms and condition numbers. A norm typically sums many separate errors into bounds consisting of the roundoff unit and polynomials of the matrix order. These bounds have low probability of being sharp when the separate errors are independently distributed [329, p. 102]. Moreover, errors of calculations with triangular matrices are bounded by, but can be independent of, condition numbers [329, pp. 103, 108]. When such triangular matrices appear in calculations, their condition numbers may introduce overestimates into the error bounds for the calculations.

4. Synopsis of the Error Analysis. This section outlines von Neumann and Goldstine's error analysis of the inversion of SPD matrices by Gaussian elimination. This material comprises Chapter 6 of the inversion paper.

4.1. Algorithm. Having found that Gaussian elimination produces a triangular factorization, von Neumann and Goldstine invented and analyzed the algorithm that forms the inverse as the product of the inverse factors {p. 1054, eqn. 4.26}:

$$A = LDU \quad \Rightarrow \quad A^{-1} = U^{-1}D^{-1}L^{-1}.$$

In present terminology (see section 1.4), von Neumann and Goldstine derived an a priori bound on the residual of inverting an SPD matrix A in fixed point arithmetic. The proven bound was for $\|AX - I\|_2$, where X is the computed inverse. Bounds on the forward error $\|X - A^{-1}\|_2$ and on the backward error $\|X^{-1} - A\|_2$ follow from (3.23) but were not in the paper. The inversion paper did not consider solving equations.

Von Neumann and Goldstine had hoped to analyze the similar algorithm for an invertible, general matrix A_I . The impediment was an a priori error bound for the triangular decomposition. Von Neumann and Goldstine believed a bound could be obtained by “reordering for size” (in modern terminology, complete pivoting or full reordering). This conjecture, which is of no use today, may be the oldest and most difficult open problem in numerical analysis. Von Neumann and Goldstine instead analyzed $A_I^{-1} = A_I^t(A_I A_I^t)^{-1}$, forming $(A_I A_I^t)^{-1}$ by the previous method. They obtained an a priori bound with the matrix entries contributing $(\|A_I\|_2 \|A_I^{-1}\|_2)^2$. That analysis is not discussed here.

The algorithm to be studied is as follows:

- Step 1. For an SPD matrix A , factor $A = U^t D U$.
- Step 2. Form U^{-1} .
- Step 3. Form D^{-1} .
- Step 4. Form $A^{-1} = U^{-1} D^{-1} (U^{-1})^t$.

This algorithm was to be performed with the computer arithmetic in section 3.3. Fixed point arithmetic complicates the calculation because inverse matrices may need scaling to place all their entries in the interval $(-1, +1)$. Von Neumann and Goldstine therefore reformulated Steps 2–4 with explicit scale factors consisting of nonnegative integral powers of $1/2$.

- Step 2'. Form $Z = U^{-1} \Delta$, where Δ is a diagonal matrix of scale factors.
- Step 3'. Form $\mathcal{D}^{(q)} = D^{-1} \Delta^{-2}/2^q$, where $1/2^q$ is a scale factor that is readjusted in Step 4'.
- Step 4'. Increase q to q_0 as needed when forming the scaled inverse $W^{(q_0)} = Z \mathcal{D}^{(q_0)} Z^t = A^{-1}/2^{q_0}$.

Von Neumann and Goldstine applied these steps with some additional refinements to a machine-representable, SPD matrix \bar{A} of order n resulting in computed quantities \bar{D} , $\bar{\mathcal{D}}^{(\bar{q}_0)}$, \bar{U} , \bar{Z} , \bar{W} and scale factors $\bar{\Delta}$ and $1/2^{\bar{q}_0}$. The scale factors, $1/2^{\bar{q}_0}$ and $\bar{\Delta}$, are given tildes because they depend on computed quantities, but they are not given overbars because they are not evaluated; the integer exponents indicate the number of halvings to be performed when evaluating formulas containing the scale factors.

4.2. Plan of the Error Analysis. The analysis has steps of two kinds. Some steps, that are indicated here by numerals 1–4', bound the rounding errors created by the algorithmic steps of the corresponding numbers. Other steps, that are indicated by letters A–F, bound various combinations of the computed quantities in terms of the initial matrix \bar{A} (see Table 4.1). Today, Steps A–F are called perturbation analyses and 1–4' are rounding error analyses. Only Steps 1–4' and F are outlined here. The most difficult parts of the inversion paper are the perturbation analysis Steps C–E.

4.3. (Step 1) Triangular Factoring. This portion of the analysis is today the most interesting for several reasons. First, only this part of von Neumann and Goldstine's work can be interpreted as a backward rounding error analysis. Second, theirs is the first rounding error analysis of triangular factoring. Third, as explained in section 3.5.2, this analysis treats the most important class of matrices for which Gaussian

Table 4.1 Steps in the error analysis of inverting SPD matrices. Steps 1–4' are rounding error analyses of the corresponding steps of the algorithm. Steps A–F are perturbation analyses.

STEP	RESULT AND ANALYSIS	SECS.	REFERENCE
1	$\ \bar{A} - \bar{U}^t \bar{D} \bar{U}\ _2 \leq \mathcal{O}(n^2 \beta^{-s})$	backward rounding error analysis of triangular factoring	6.1–6.3 {p. 1069, eqns. 6.33–34}
2'	$\ \bar{U} \bar{Z} - \bar{\Delta}\ _2 \leq \mathcal{O}(n^2 \beta^{-s})$	residual rounding error analysis of inverting the triangular factor	6.4–6.5 {pp. 1072–1073, eqns. 6.38, 6.43}
A	$\ \bar{D}^{1/2} \bar{U}\ _2 \leq \mathcal{O}(\ \bar{A}\ _2^{1/2})$	bound on factored matrix	6.6 {p. 1073, eqn. 6.45.a}
B	$\ (\bar{D}^{1/2} \bar{U})^{-1}\ _2 \leq \mathcal{O}(\ \bar{A}^{-1}\ _2^{1/2})$	lower bound on factored matrix	6.6 {p. 1073, eqn. 6.45.b}
C	$\ \bar{Z} \bar{\Delta}^{-2} \bar{D}^{-1} \bar{Z}^t\ _2 \leq \mathcal{O}(\ \bar{A}^{-1}\ _2)$	bound on factored inverse	6.6 {p. 1074, eqn. 6.48'}
D	$\ \bar{\Delta}^{-2} \bar{D}^{-1}\ _2 \leq \mathcal{O}(\ \bar{A}^{-1}\ _2)$	bound on diagonal in factored inverse	6.6 {p. 1075, eqn. 6.49}
3'	$ \bar{D}_{j,j}^{(q)} - \tilde{D}_{j,j}^{(q)} \leq \beta^{-s}$	backward rounding error analysis of inverting the diagonal	6.7 {p. 1077, ln. 1}
4'	$\ \bar{W}^{(q)} - \bar{Z} \tilde{D}^{(q)} \bar{Z}^t\ _2 \leq \mathcal{O}(n^2 \beta^{-s})$	forward rounding error analysis of assembling the inverse	6.7 {p. 1077, eqn. 6.56}
E	$2\tilde{q}_0 \leq \mathcal{O}(\ \bar{A}^{-1}\ _2)$	bound on the final scale factor	6.7 {p. 1078, eqn. 6.60}
F	$A \bar{W}^{(\tilde{q}_0)} 2\tilde{q}_0 - I = \dots$	expansion of the error of inversion	6.8 {pp. 1079–1080}

elimination has been proved to have a small error bound. For these reasons, the analysis is presented in full.

4.3.1. (Step 1) Factoring Algorithm. This remarkable calculation is done in fixed point arithmetic without scaling. The algorithm is determined by a chain of requirements beginning with a consideration of how to evaluate the quotient in (3.4):

$$\bar{A}_{i,j}^{(k+1)} = \text{computed}(\bar{A}_{i,j}^{(k)} - \bar{A}_{i,k}^{(k)} \bar{A}_{k,j}^{(k)} / \bar{A}_{k,k}^{(k)}), \quad 1 \leq k < i, j \leq n.$$

Let $ab/c = \bar{A}_{i,k}^{(k)} \bar{A}_{k,j}^{(k)} / \bar{A}_{k,k}^{(k)}$. As noted in section 3.5.2, the grouping $(ab)/c$ has uncontrolled error because it divides the proper fraction c into the error of evaluating ab {p. 1063, eqns. 6.3.a and 6.3.b}:

$$(4.1) \quad \begin{aligned} |\text{computed}[(ab)/c] - ab/c| &\leq (1 + |c|^{-1})\beta^{-s}/2 \leq ?, \\ |\text{computed}[a(b/c)] - ab/c| &\leq (1 + |a|)\beta^{-s}/2 \leq \beta^{-s}. \end{aligned}$$

Therefore, von Neumann and Goldstine had to use the second grouping, $a(b/c)$. To make the evaluation of this expression a valid machine calculation, they had to further arrange that b/c is a proper fraction. According to (3.9) this precondition can be met by symmetrically reordering $\bar{A}^{(k)}$ so that $\bar{A}_{k,k}^{(k)}$ is largest. Von Neumann and Goldstine found it interesting that only the computed algorithm requires reordering:

A [symmetric, positive] definite A could indeed be put through the algorithm in the rigorous sense [using true arithmetic], without positioning for size. However, if pseudo-operations are used, no satisfactory... [error bounds] seem to be obtainable, unless positioning for size is also effected. — inversion paper {p. 1061}

This analysis has two parts. The first part proves the chosen calculation proceeds to conclusion. The second bounds the rounding error.

4.3.2. (Step I, Part I) Proof Factoring Succeeds. Von Neumann and Goldstine evaluated the reduced matrices (see section 3.5.1)

$$\bar{A} = \bar{A}^{(1)}, \bar{A}^{(2)}, \bar{A}^{(3)}, \dots, \bar{A}^{(n)}.$$

Without loss of generality they assumed that the first matrix has been ordered so $\bar{A}_{k,k}^{(k)}$ is always the largest in $\bar{A}^{(k)}$, and hence the notation need not account for reordering. Taking care to preserve symmetry, (3.4) becomes {p. 1064, eqn. 6.3}

$$(4.2) \quad \bar{A}_{i,j}^{(k+1)} = \text{computed} \begin{cases} \bar{A}_{i,j}^{(k)} - \bar{A}_{k,i}^{(k)} (\bar{A}_{k,j}^{(k)} / \bar{A}_{k,k}^{(k)}), & k < i \leq j \text{ (upper)}, \\ \bar{A}_{j,i}^{(k+1)}, & k < j < i \text{ (lower)}. \end{cases}$$

Note the ordering makes (4.2) valid for fixed point arithmetic, and strict symmetry is preserved by only operating with values on or above the diagonal.

The following properties are assumed true for the first matrix, $k = 1$, and they are proved for succeeding matrices by induction on k .

Property 1_k. $\bar{A}^{(k)}$ is SPD.

Property 2_k. The entries of $\bar{A}^{(k)}$ lie in the interval $(-1, +1)$.

Property 3_k. $\|\bar{A}^{(k)}\|_\ell > \sum_{m=1}^{n-k} m\beta^{-s}$.

The notation $\|\cdot\|_\ell$ is the matrix lower bound of (3.2). Given these properties for $\bar{A}^{(k)}$, the intended next matrix is $\tilde{A}^{(k+1)}$ {p. 1065, eqn. 6.11.a},

$$\tilde{A}_{i,j}^{(k+1)} = \bar{A}_{i,j}^{(k)} - \bar{A}_{i,k}^{(k)} \bar{A}_{k,j}^{(k)} / \bar{A}_{k,k}^{(k)}.$$

(Von Neumann and Goldstine used tildes for quantities derived exactly from machine data.) Property 1_k and (3.7) assure that $\tilde{A}^{(k+1)}$ is SPD. Equation (4.1) assures this intended matrix differs from the actually computed $\bar{A}^{(k+1)}$ by a small perturbation {p. 1066, before eqn. 6.14}:

$$(4.3) \quad |\tilde{A}_{i,j}^{(k+1)} - \bar{A}_{i,j}^{(k+1)}| \leq \beta^{-s} \quad \Rightarrow \quad \|\tilde{A}^{(k+1)} - \bar{A}^{(k+1)}\|_F \leq (n-k)\beta^{-s}.$$

Property 3_{k+1} now follows {p. 1066, eqn. 6.17}:

$$\begin{aligned} \|\bar{A}^{(k+1)}\|_\ell &\geq \|\tilde{A}^{(k+1)}\|_\ell - \|\tilde{A}^{(k+1)} - \bar{A}^{(k+1)}\|_2 && \text{by (3.3),} \\ &\geq \|\bar{A}^{(k)}\|_\ell - \|\tilde{A}^{(k+1)} - \bar{A}^{(k+1)}\|_2 && \text{by (3.8),} \\ &> \sum_{m=1}^{n-k} m\beta^{-s} - (n-k)\beta^{-s} && \text{by Property 3}_k \text{ and (4.3),} \\ &= \sum_{m=1}^{n-(k+1)} m\beta^{-s}, && \text{which is Property 3}_{k+1}. \end{aligned}$$

Since $\tilde{A}^{(k+1)}$ is an SPD matrix, its minimal eigenvalue is $\lambda_{\min} = \|\tilde{A}^{(k+1)}\|_{\ell}$. Therefore, for any vector $u \neq 0$, by reasoning similar to the above {p. 1065–1066, eqns. 6.11–6.16},

$$\begin{aligned} u^t \bar{A}^{(k+1)} u &= u^t \tilde{A}^{(k+1)} u - u^t (\tilde{A}^{(k+1)} - \bar{A}^{(k+1)}) u \\ &\geq \|\tilde{A}^{(k+1)}\|_{\ell} \|u\|_2^2 - u^t (\tilde{A}^{(k+1)} - \bar{A}^{(k+1)}) u \\ &> (\|\bar{A}^{(k)}\|_{\ell} - \|\tilde{A}^{(k+1)} - \bar{A}^{(k+1)}\|_2) \|u\|_2^2 \\ &\geq \left(\sum_{m=1}^{n-(k+1)} m \beta^{-s} \right) \|u\|_2^2. \end{aligned}$$

So $\bar{A}^{(k+1)}$ is positive definite, which is property 1_{k+1} . For such matrices, as noted in (3.9), the largest entries on the diagonal have the largest magnitudes in the matrix. Therefore, the inequalities

$$\bar{A}_{j,j}^{(k+1)} = \text{computed} [\bar{A}_{j,j}^{(k)} - \bar{A}_{k,j}^{(k)} (\bar{A}_{k,j}^{(k)} / \bar{A}_{k,k}^{(k)})] \leq \bar{A}_{j,j}^{(k)}$$

(the subtracted quantity is not negative) show that the largest magnitude in $\bar{A}^{(k+1)}$ is no larger than the largest in $\bar{A}^{(k)}$, and hence is no larger than 1 by Property 2_k , which is Property 2_{k+1} {p. 1066, eqn. 6.19}. This completes the induction.

The fixed point calculation proceeds to its conclusion because all quantities lie in the interval $(-1, +1)$ and no divisor vanishes. Von Neumann and Goldstine summarize this conclusion in {p. 1067, eqns. 6.20–6.23}. The analysis requires the additional hypothesis consisting of Property 3_k for $k = 1$ {p. 1067, eqn. 6.23.c}:

$$(4.4) \quad \boxed{\|\bar{A}\|_{\ell} > \sum_{m=1}^{n-1} m \beta^{-s} = (n^2 - n) \beta^{-s} / 2} \quad (\text{hypothesis}),$$

where $\|\cdot\|_{\ell}$ is the matrix lower bound of (3.2).

4.3.3. (Step 1, Part 2) Error of Factoring. The second part of the factoring analysis assembles and bounds the net perturbation. The computed factorization is $\bar{A} \approx \bar{U}^t \bar{D} \bar{U}$, where {p. 1068, eqns. 6.30–6.31}

$$\bar{D}_{i,k} = \begin{cases} \bar{A}_{k,k}^{(k)}, & k = i, \\ 0, & k \neq i, \end{cases} \quad \bar{U}_{k,i} = \begin{cases} \text{computed} (\bar{A}_{k,i}^{(k)} / \bar{A}_{k,k}^{(k)}), & k < i, \\ 1, & k = i, \\ 0, & k > i. \end{cases}$$

Note that if a, c are machine representable and a/c is a valid operation, then $(\text{computed}(a/c)) \cdot c$ differs from a by at most $\beta^{-s}/2$. With this in mind it is easy to see that the $i-k-j$ intermediate term of the product $\bar{U}^t \bar{D} \bar{U}$,

$$(\bar{U}^t)_{i,k} \bar{D}_{k,k} \bar{U}_{k,j} = \text{computed} (\bar{A}_{i,k}^{(k)} / \bar{A}_{k,k}^{(k)}) \bar{A}_{k,k}^{(k)} \text{computed} (\bar{A}_{k,j}^{(k)} / \bar{A}_{k,k}^{(k)}),$$

differs from the term subtracted at the k th elimination step,

$$\begin{cases} \text{computed} [\bar{A}_{k,i}^{(k)} (\bar{A}_{k,j}^{(k)} / \bar{A}_{k,k}^{(k)})], & k < i \leq j \text{ (upper triangle)}, \\ \text{computed} [\bar{A}_{k,j}^{(k)} (\bar{A}_{k,i}^{(k)} / \bar{A}_{k,k}^{(k)})], & k < j < i \text{ (lower triangle)}, \end{cases}$$

by at most β^{-s} . Since subtraction is errorless, thus {p. 1068, eqns. 6.24–6.25},

$$(4.5) \quad \bar{A}_{i,j}^{(k+1)} = \bar{A}_{i,j}^{(k)} - (\bar{U}^t)_{i,k} \bar{D}_{k,k} \bar{U}_{k,j} + \theta_{i,j}^{(k)} \quad \text{with} \quad |\theta_{i,j}^{(k)}| \leq \beta^{-s}, \quad k < \min\{i, j\}.$$

There are smaller errors for the leading row and column where one division, either $\bar{A}_{k,i}^{(k)}/\bar{A}_{k,k}^{(k)}$ or $\bar{A}_{k,j}^{(k)}/\bar{A}_{k,k}^{(k)}$, is omitted {p. 1068, eqns. 6.26–6.27}:

$$(4.6) \quad 0 = \bar{A}_{i,j}^{(k)} - (\bar{U}^t)_{i,k} \bar{D}_{k,k} \bar{U}_{k,j} + \theta_{i,j}^{(k)} \quad \text{with} \quad |\theta_{i,j}^{(k)}| \leq \beta^{-s}/2, \quad k = \min\{i, j\}.$$

Von Neumann and Goldstine forgot that the diagonal term omits both divisions so $\theta_{k,k}^{(k)} = 0$. Summing (4.5) and (4.6) for $k = 1, \dots, \min\{i, j\}$, canceling identical terms on opposite sides, and remembering that $\bar{A}^{(1)} = \bar{A}$ give the net perturbation {p. 1068, eqn. 6.28}

$$0 = \bar{A}_{i,j} - \sum_{k=1}^{\min\{i,j\}} (\bar{U}^t)_{i,k} \bar{D}_{k,k} \bar{U}_{k,j} + E_{i,j} \quad \text{and} \quad E_{i,j} = \sum_{k=1}^{\min\{i,j\}} \theta_{i,j}^{(k)},$$

with a small bound for $|E_{i,j}|$,

$$|E_{i,j}| \leq \begin{cases} (\min\{i, j\} - 1/2) \beta^{-s}, & i \neq j, \\ (i - 1) \beta^{-s}, & i = j, \end{cases}$$

that becomes a bound for $\|E\|_F$,

$$\|E\|_F \leq \left(\frac{n^4}{6} - \frac{5n^2}{12} + \frac{n}{4} \right)^{1/2} \beta^{-s} \leq \frac{n^2}{\sqrt{6}} \beta^{-s}.$$

THEOREM 4.1. *For fixed point arithmetic with common rounding of s digits in radix β , and for a representable SPD matrix \bar{A} of order n with $\lambda_{\min}(\bar{A}) > (n^2 - n)\beta^{-s}/2$, von Neumann and Goldstine's factorization with symmetric reordering produces diagonal \bar{D} and upper triangular \bar{U} with 1's on the main diagonal, so that*

$$(4.7) \quad \bar{A} + E = \bar{U}^t \bar{D} \bar{U}, \quad \|E\|_F \leq n^2 \beta^{-s} / \sqrt{6}.$$

The perturbed factorization is {p. 1069, eqn. 6.33}. The bound is {p. 1069, eqn. 6.34}. Von Neumann and Goldstine wrote $A = \bar{U}^t \bar{D} \bar{U} + E$, but they confused signs so they actually derived the formula above, with E (their Z) on the other side. Theorem 4.1 will be applied in the following form:

$$\boxed{\|\bar{A} - \bar{U}^t \bar{D} \bar{U}\|_2 \leq \mathcal{O}(n^2 \beta^{-s})} \quad (\text{Step 1}).$$

4.4. Programming with Scale Factors. After factoring, von Neumann and Goldstine's inversion algorithm continues with inverting. In the upcoming steps, the calculation of scale factors involves capabilities beyond the idealized arithmetic described in the inversion paper. Twice von Neumann and Goldstine had to find, for given machine-representable numbers $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_m$, the smallest nonnegative integer \tilde{p} that would make the following sum machine representable:

$$(4.8) \quad \text{computed}(\bar{a}_1/2^{\tilde{p}} + \bar{a}_2/2^{\tilde{p}} + \bar{a}_m/2^{\tilde{p}}).$$

Because the partial sums need not be representable, further knowledge of the computer is required to choose \tilde{p} , and once it has been chosen, to evaluate the sum.

Historical programming techniques suggest how the scale factors could be chosen. As seen in Figure 2.3, the inversion paper was written before any computers were built, so the paper itself does not contain a program. Within a few years, von Neumann and Goldstine's algorithm was programmed "quite precisely as described in their paper" [130] for the IBM 701 and no doubt for other clones of the IAS computer. These programs seem to have been lost, so the programming must be inferred from other sources. Wilkinson later explained that the programmer could shift the radix point of all the summands so "spillover cannot occur" [329, p. 27]. That is, some precision was sacrificed to allow k whole digits, where $m \leq n < 2^k$. By this artifice a program could sum the \tilde{a}_i and choose the optimum $\tilde{p} \leq k$.

A further refinement is possible on the IAS computer. First, the shifted radix point can be used to choose the \tilde{p} that makes the full sum machine representable, then (4.8) can be evaluated to full precision ignoring any possible overflow in the partial sums. The full sum is correct because the signed, fractional addition of the IAS computer was designed to be unsigned, fractional arithmetic modulo 2 [118, p. 4].

4.5. (Step 2') Inverting the Triangular Matrix. This step forms a scaled inverse for \bar{U} by evaluating recurrences for each column, j , beginning with the diagonal entry and working upward {p. 1071, eqn. 6.35}:

$$(4.9) \quad \bar{V}_{i,j} = \begin{cases} 0, & i > j, \\ 1, & i = j, \\ \text{computed } (-\sum_{k=i+1}^j \bar{U}_{i,k} \bar{V}_{k,j} / 2^{\tilde{p}_{i,j} - \tilde{p}_{k,j}}), & i = j-1, j-2, \dots, 1. \end{cases}$$

The exponent $\tilde{p}_{i,j}$ is selected just prior to the evaluation of $\bar{V}_{i,j}$ to be the smallest nonnegative integer, at least $\tilde{p}_{i+1,j}$, that places $\bar{V}_{i,j}$ in $(-1, +1)$:

$$0 = \tilde{p}_{j,j} \leq \dots \leq \tilde{p}_{i+1,j} \leq \tilde{p}_{i,j} \leq \dots \leq \tilde{p}_{1,j}.$$

The denominator in (4.9) increments the scaling of the k th term to that of the i th term, so the cumulative effect is to apply just the last scale factor, $1/2^{\tilde{p}_{1,j}}$, to all entries in the j th column. Repeated for each column of \bar{U}^{-1} , this produces a column-scaled inverse, \bar{Z} , and a diagonal matrix, $\tilde{\Delta}$, of ultimate scale factors:

$$\bar{U}\bar{Z} \approx \tilde{\Delta}, \quad \tilde{\Delta}_{j,j} = 1/2^{\tilde{p}_{1,j}}, \quad \bar{Z}_{i,j} = \begin{cases} 0, & i > j, \\ \text{computed } (\bar{V}_{i,j} / 2^{\tilde{p}_{1,j} - \tilde{p}_{i,j}}), & i \leq j. \end{cases}$$

Von Neumann and Goldstine bounded the residual associated with the scaled inverse {p. 1072, eqn. 6.38; p. 1073, eqn. 6.43}:

$$\boxed{\|\bar{U}\bar{Z} - \tilde{\Delta}\|_2 \leq \mathcal{O}(n^2\beta^{-s})} \quad (\text{Step 2'}).$$

The inversion paper did not explain how this algorithm might be programmed. The simplest approach starts with $\tilde{p}_{i,j}$ equal to the previous value $\tilde{p}_{i+1,j}$, provisionally forms the i th entry in column j as described in section 4.4, and, if the magnitude exceeds 1, then adjusts $\tilde{p}_{i,j}$ upward as needed. The process only appears to be complicated because of the double subscripts in (4.9). A program would use one variable $p(j)$ to hold the latest scale factor for the column. Fixed point calculations frequently

used a common scale factor for many values. Von Neumann and Goldstine's column scaling is what Wilkinson [329, pp. 26–27, 120–121] later called “block-floating” vectors. Wilkinson [327, p. 48], [203, p. 15] adopted what he called this “practical” and “convenient scheme” of scaling to solve triangular systems. He did not cite von Neumann and Goldstine for the method.

4.6. (Step 3') Inverting the Diagonal. A scale factor, $1/2^{\tilde{q}_0}$, must be chosen to make computable the diagonal matrix $\mathcal{D}^{(\tilde{q}_0)}$ in Step 3'. This requirement is more stringent than placing the entries in $(-1, +1)$. For a scale factor $1/2^q$, let $\tilde{\mathcal{D}}^{(q)}$ be the intended value for a computed diagonal $\bar{\mathcal{D}}^{(q)}$,

$$\tilde{\mathcal{D}}^{(q)} = (\tilde{\Delta}^2 \bar{\mathcal{D}})^{-1} / 2^q,$$

where $\tilde{\mathcal{D}}_{j,j}^{(q)} = 2^{2\tilde{p}_{1,j}-q} / \bar{\mathcal{D}}_{j,j}$.

First, \tilde{r}_j is chosen so {p. 1075, eqn. 6.51}

$$1/2 < 2^{\tilde{r}_j} \bar{\mathcal{D}}_{j,j} \leq 1,$$

which makes $2^{\tilde{r}_j} \bar{\mathcal{D}}_{j,j}$ machine representable, hence $(1/2)/(2^{\tilde{r}_j} \bar{\mathcal{D}}_{j,j})$ are valid machine operations. Then q is restricted so that $q \geq \tilde{q}_{\min}$, where {p. 1075, eqn. 6.52}

$$\tilde{q}_{\min} = \max_j 2\tilde{p}_{1,j} + \tilde{r}_j + 1.$$

With this choice of q the scaled diagonal terms can be written {p. 1076, Lem. 6.53}

$$\begin{aligned} \tilde{\mathcal{D}}_{j,j}^{(q)} &= [(1/2) / (2^{\tilde{r}_j} \bar{\mathcal{D}}_{j,j})] / 2^{(q-2\tilde{p}_{1,j}-\tilde{r}_j-1)}, \\ \bar{\mathcal{D}}_{j,j}^{(q)} &= \text{computed value for the expression above,} \end{aligned}$$

which can be evaluated in the grouping shown by fixed point arithmetic and iterated halving with modest error {p. 1077, ln. 1}:

$$\boxed{|\bar{\mathcal{D}}_{j,j}^{(q)} - \tilde{\mathcal{D}}_{j,j}^{(q)}| \leq \beta^{-s}} \quad (\text{Step 3'}).$$

(The original notation for the computed diagonal, \bar{e} , is here replaced by $\bar{\mathcal{D}}$.)

4.7. (Step 4') Forming the Scaled Inverse. The final step forms a scaled inverse $\bar{W}^{(q)}$ by matrix multiplication {p. 1067, eqns. 6.55 onwards},

$$(4.10) \quad \begin{aligned} \bar{W}^{(q)} &= \text{computed } (\bar{Z} \bar{\mathcal{D}}^{(q)} \bar{Z}^t), \\ \text{where } \bar{W}_{i,j}^{(q)} &= \text{computed } \left(\sum_{k=j}^n \bar{Z}_{i,k} \bar{\mathcal{D}}_{k,k}^{(q)} \bar{Z}_{k,j} \right). \end{aligned}$$

The exponent q must satisfy three conditions:

1. $q \geq \tilde{q}_{\min}$ to form $\bar{\mathcal{D}}^{(q)}$;
2. all entries of $\bar{W}^{(q)} \approx \bar{Z} \bar{\mathcal{D}}^{(q)} \bar{Z}^t = \bar{Z} (\tilde{\Delta}^2 \bar{\mathcal{D}})^{-1} \bar{Z}^t / 2^q$ must be in $(-1, +1)$;
3. $1/2^q$ must not be so small that too much significance is lost in $\bar{W}^{(q)}$.

Before a final q can be chosen, it is necessary to bound the rounding error that (4.10) adds to $\bar{W}^{(q)}$. If it is evaluated with s fractional digits and no limit on whole digits, then {p. 1077, eqn. 6.56}

$$\|\bar{W}^{(q)} - \bar{Z}\tilde{D}^{(q)}\bar{Z}^t\|_2 \leq \mathcal{O}(n^2\beta^{-s}) \quad (\text{Step 4'}) .$$

From this, von Neumann and Goldstine were able to prove that (1) there is a $\tilde{q}_{\max} \geq \tilde{q}_{\min}$, so that (2) the entries of $\bar{W}^{(\tilde{q}_{\max})}$ lie in the interval $(-1, +1)$ and {p. 1078, eqn. 6.60}

$$(4.11) \quad 2^{\tilde{q}_{\max}} = \mathcal{O}(\|\bar{A}^{-1}\|_2) \quad (\text{Step E}) .$$

Equation (4.11) does not itself satisfy condition (3), but it does indicate that accuracy can be assured by investigating a property of the matrix (the size of the smallest singular value).

Since \tilde{q}_{\max} depends on $\|\bar{A}^{-1}\|_2$, which is harder to calculate than \bar{A}^{-1} , von Neumann and Goldstine wanted to avoid the impression its value is needed. Instead, they described the calculation in terms of the smallest exponent, $\tilde{q}_0 \geq \tilde{q}_{\min}$, that places all the entries of $\bar{W}^{(\tilde{q}_0)}$ in the interval $(-1, +1)$. This choice depends on “simple, explicit conditions, which permit forming the \tilde{q}_0 in question” {p. 1078, mid. and fn. 32}. Necessarily $\tilde{q}_{\max} \geq \tilde{q}_0$, but the upper bound is not used to choose \tilde{q}_0 . Again, von Neumann and Goldstine do not explain how to program a computer to find \tilde{q}_0 other than, presumably, by trial and error, as in section 4.5. Beginning from $q = \tilde{q}_{\min}$, if a sum in (4.10) has magnitude greater than 1, then increase q , rescale the sums that have been computed, and continue to the next new sum.

4.8. (Step F) Bounding the Residual of the Inverse. The preceding bounds combine with the perturbation analyses to bound the residual for the unscaled inverse, $\bar{W}^{(\tilde{q}_0)}2^{\tilde{q}_0}$, using a four-part splitting {pp. 1079–1080}:

$$\begin{aligned} \bar{A}\bar{W}^{(\tilde{q}_0)}2^{\tilde{q}_0} - I &= \bar{A}(\bar{W}^{(\tilde{q}_0)} - H^t H / 2^{\tilde{q}_0})2^{\tilde{q}_0} \\ &\quad + (\bar{A} - K K^t) H^t H \\ &\quad + K(K^t H^t - I)H \\ &\quad + K(HK - I)K^{-1} . \end{aligned}$$

This splitting is applied with $H = (\bar{D}^{-1/2}\tilde{\Delta}^{-1}\bar{Z}^t)$ and $K = (\bar{U}^t\bar{D}^{1/2})$ to give

$$\begin{aligned} &\bar{A}\bar{W}^{(\tilde{q}_0)}2^{\tilde{q}_0} - I \\ &= \bar{A}(\bar{W}^{(\tilde{q}_0)} - \bar{Z}\tilde{D}^{(\tilde{q}_0)}\bar{Z}^t)2^{\tilde{q}_0} \\ (4.12) \quad &+ (A - \bar{U}^t\bar{D}\bar{U})(\bar{Z}\tilde{\Delta}^{-2}\bar{D}^{-1}\bar{Z}^t) \quad (\text{Step F}) \\ &+ (\bar{D}^{1/2}\bar{U})^t[\bar{D}^{1/2}(\bar{U}\bar{Z} - \tilde{\Delta})(\bar{D}^{1/2}\tilde{\Delta})^{-1}](\bar{D}^{-1/2}\tilde{\Delta}^{-1}\bar{Z}^t) \\ &+ (\bar{D}^{1/2}\bar{U})^t[(\tilde{\Delta}^{-1}\bar{D}^{-1/2})(\bar{U}\bar{Z} - \tilde{\Delta})^t\bar{D}^{1/2}](\bar{D}^{1/2}\bar{U})^{-t} . \end{aligned}$$

From Steps 4', 1, and 2', the red, green, and blue terms are $\mathcal{O}(n^2\beta^{-s})$, so the expression is small provided the remaining, black terms are no larger than $\bar{A}^{\pm 1}$. Perturbation bounds for those terms are given in Table 4.1. To simplify the bounds, von Neumann

and Goldstine removed higher order terms by introducing the assumption {p. 1073, eqn. 6.41; p. 1081, eqn. 6.67}

$$\boxed{\|\bar{A}\|_\ell \geq 10n^2\beta^{-s}} \quad (\text{final hypothesis}),$$

which supersedes (4.4).

On taking norms and replacing the terms by their bounds, (4.12) becomes {p. 1080, eqn. 6.66}

$$\begin{aligned} \|\bar{A}\bar{W}^{(\tilde{q}_0)}2^{\tilde{q}_0} - I\|_2 &\leq \mathcal{O}(\|\bar{A}\|_2 \|\bar{A}^{-1}\|_2 n^2\beta^{-s}) && (\text{using Steps 4', E}) \\ &+ \mathcal{O}(\|\bar{A}^{-1}\|_2 n^2\beta^{-s}) && (\text{Steps 1, C}) \\ &+ \mathcal{O}(\|\bar{A}\|_2^{1/2} \|\bar{A}^{-1}\|_2 n^2\beta^{-s}) && (\text{Steps A, 2', D, C}) \\ &+ \mathcal{O}(\|\bar{A}\|_2^{1/2} \|\bar{A}^{-1}\|_2 n^2\beta^{-s}) && (\text{Steps A, D, 2', B}) \\ (4.13) \quad &= \mathcal{O}[(\|\bar{A}\|_2 + \|\bar{A}\|_2^{1/2} + 1) \|\bar{A}^{-1}\|_2 n^2\beta^{-s}]. \end{aligned}$$

Von Neumann and Goldstine defer to the very end of the paper a final assumption, a scaling up of the original matrix, that makes the bound a multiple of $\|\bar{A}\|_2$ {p. 1091, eqns. 7.3–7.4}:

$$1/2 \leq |\bar{A}_{i,j}| \text{ for some } i \text{ and } j \Rightarrow 1/2 \leq \|\bar{A}\|_2, \text{ hence } 1 \leq (2\|\bar{A}\|_2)^{1/2}.$$

This assumption can be used to place a factor of $\|\bar{A}\|_2$ into every term of (4.13) {p. 1093, eqn. 7.5'}:

$$(4.14) \quad \boxed{\|\bar{A}\bar{W}^{(\tilde{q}_0)}2^{\tilde{q}_0} - I\|_2 \leq \mathcal{O}(\|\bar{A}\|_2 \|\bar{A}^{-1}\|_2 n^2\beta^{-s})}.$$

4.9. Forward, Backward, or Residual Analysis. An error analysis for a computation is characterized by which one of three quantities receives a bound (see section 1.4):

1. errors in the output (forward analysis);
2. compensating perturbations in the input (backward analysis); or
3. the discrepancy of substituting the output into the equations it is supposed to satisfy (residual analysis).

The most appropriate overall classification for the inversion paper is a residual analysis. The final conclusions in (3.12) or (4.14) and (3.22) are both stated in residual form. Von Neumann and Goldstine consistently expressed intermediate results similarly, as bounds on the deviation in the equations that define the quantities to be computed.⁸⁹

5. Conclusion. “Numerical Inverting of Matrices of High Order” was written before modern computers existed during a curious period in the 1940s when everyone knew what the machines looked like but nobody had one (see Figure 2.3). This situation was due in part to von Neumann and Goldstine’s efforts to distribute in-

⁸⁹Backward error analysis appears in only Step 1 of the 10 steps in the analysis for SPD matrices (see Table 4.1), and it occupies just 2 of 26 pages (Chapter 6, section 3, pp. 1067–1069) in the error analysis chapter. Wilkinson [332, p. 554] cited Step 1 as an example of backward analysis, but being the only example, he improperly described “a great part” of the analysis as being of the backward kind.

formation about computers as widely and rapidly as possible, despite military and proprietary interests in the technology and the limited distribution channels of the time. Their IAS reports, which were distributed around the world, the many visitors to their Princeton computer laboratory, and von Neumann's personal appearances in seminars and lectures in the U.S. were overwhelmingly influential at a critical time during the building of the first modern computers.

At the same time there was apprehension over whether and how the computers then under construction would alter the use of mathematics in the sciences. This unease was manifest in exaggerated concerns that rounding errors might ruin the calculations of unprecedented length that the new machines would make possible. Von Neumann and Goldstine put these concerns to rest by examining the inversion of matrices.

Following the pattern established in their reports on computer architecture and programming, von Neumann and Goldstine saw the inversion paper broadly as a means to prepare for the computer era. In this case their goal was to rejuvenate the neglected subject of numerical analysis (later terminology). From the discovery of quantifiable instabilities inherent in some calculations (the Courant criterium, later CFL condition, of PDEs, and the figure of merit, now condition number, of matrix inversion) and from the new economies of calculation (performing arithmetic being less expensive than memory), von Neumann and Goldstine concluded there was a need for new techniques of calculation. They recognized that devising algorithms for the new machines would require new types of mathematical analysis. To that end they developed a conceptual framework of stability to treat errors in calculations and for rounding error specifically, and they demonstrated that calculations could be studied through the use of aggregating abstractions such as matrices and matrix norms.

The inversion paper was the first modern paper in numerical analysis and the most recent by a person of von Neumann's genius. Whereas his collaborators continued to develop his work on numerical solutions of PDEs, ironically, the finished and published analysis of matrix inversion was less accessible. Wilkinson [332, p. 548] asked whether any of the first postwar generation of numerical analysts fully understood von Neumann's formal proof that matrix inversion by Gaussian elimination was accurate for SPD matrices in fixed point arithmetic. The inversion paper can be understood on several levels, and evidently Givens, Householder, Todd, Turing, and Wilkinson understood enough to grasp the important threads and weave them into the fabric of the modern field.

Epilogue. The subject that von Neumann founded with the inversion paper did not alter the prejudices of academic mathematicians [71], so at his death numerical analysis was already becoming part of the emerging field of computer science [162]. Numerical analysis remained a major part, even the dominant part, of computer science until about 1965 [44], after which Traub [276] and Wilkinson [331] note that this prominence began to wane. The composition of computer science at present is described in [250].

In recent years, an authoritative panel discussion on "models of computation" [251] (as the starting point from which to analyze algorithms) quickly turned instead to the question of who would analyze algorithms in the future. The sciences have always undertaken whatever mathematical research they need, and Parlett [251, pp. 10–11] noted the same was happening for computational mathematics—an outcome von Neumann personally hoped to avoid [64, p. 228].⁹⁰ Golub [251, p. 12] thought the

⁹⁰Parlett [216] presciently had made roughly the same observation many years before.

principal concern should be the unity of a broader field of interdisciplinary scientific computing, to which the mathematical contribution so far has been numerical analysis but may be something else in the future.

Notes on Biography.

Secondary Literature. The only comprehensive analysis of von Neumann's work on computing by a professional historian is [10]; see also [11] and [12, pp. xi–xviii]. Specialized analyses are [112, 256]. Retrospectives of all his work are [39, 111, 136, 286]. The main biography of von Neumann in English is [181]; biographies that focus on aspects of his career besides computing are [138, 222]; and biographical sketches are [79, 230]. Obituaries are [31, 271].

Primary Sources. Accounts of von Neumann by his acquaintances are [115, 135, 286, 287, 317]. A video of von Neumann and his colleagues including Goldstine is [209]. Brief meetings with von Neumann that corroborate biographical descriptions of his personal traits are [86, pp. 13–15] and [130]. Von Neumann is sometimes mentioned in oral histories such as [193, 194, 269, 278]. Collections of von Neumann's work are [12, 266]. Major archives are the John von Neumann papers at the Library of Congress, the Herman Goldstine collection at the American Philosophical Society, and the Records of the Electronic Computer Project in the IAS Archives.

Notes on Computers.

Secondary Literature. Computer history up to the period of Figure 2.3 can be found in [334]. Timelines for some machines in the figure and others not “modern” in the tripartite sense (digital, electronic, programmable) are in [335, p. 10, Fig. 10]. The secondary literature mostly relates specific development efforts organized chronologically. Many computers and their developers are discussed in monographs, collections, and the principal journal for computer history, (*IEEE*) *Annals of the History of Computing*.

Comparatively little has been written about the evolution and spread of concepts, for which see [7, 51, 52, 131, 223, 256]. These articles about intellectual history and the books [10, 132, 324, 334] are especially recommended for further reading.

Primary Sources. Von Neumann, alone among early developers, left a contemporary, extensive, and public record of his ideas, for which see [12, 266]. His 1945 *First Draft* report [299] was so influential that other sources, which by themselves might not have led to the same concept, tend to be interpreted after the fact in terms of von Neumann machines. Other primary literature has been reprinted by the Charles Babbage Institute (CBI) at the University of Minnesota or the MIT Press. Oral histories are at the CBI and the National Museum of American History in the Smithsonian Institution.

Exemplars for nearly all extant hardware, except the earliest one-of-a-kind machines, are at the Computer History Museum in Mountain View. Only Johnniac at this museum remains of all the IAS clones in Figure 2.4. The IAS computer itself is at the Smithsonian.

Notes on Historiography. The chief difficulty in preparing historical material for technologists is the expectation that history is a bibliography of current knowledge organized chronologically. Some examples of the more complicated actual evolution are “Gaussian elimination” and “condition numbers” (sections 1.3, 3.5.1 and 2.10, 3.7, respectively). For discussion about writing the history of mathematics, see [70, 127, 187]. Precautions have been taken in this review to avoid common errors of historical commentary.

The *Matthew effect* [192] is the tendency to assign discoveries to the most prolific authors in a field. Von Neumann wrote only the inversion paper on this subject, so his work is easily confused with the many more accessible later papers (see Figure 1.1). There is a curious hostility toward crediting von Neumann for his contributions, perhaps because he was so successful with apparently so little effort, as evidenced by insincere references to “the Wizard” in [141].

May [187, p. 13] describes *anachronism* as using modern terminology to explain earlier work as though our predecessors endorsed the newer concepts. Because affirmation by historical figures is a convenient pedagogic device, Kuhn [166, p. 137] particularly warns against anachronistic attributions. An example regarding the inversion paper is the comment [332, p. 554] that much of it was a backward error analysis, but concealed, thereby borrowing von Neumann’s reputation yet awarding the discovery to others.

The final danger is the repetition and elaboration of *undocumented sources*, which Katz [159] illustrates in computer science history. For the inversion paper this occurs in the unsubstantiated assertion that von Neumann and Goldstine recommended the normal equations.

Here, information has been assembled from reliable first-hand accounts and reputable biographies. For readability these sources are sometimes only cited in general, but specific references have been provided for direct quotations and for critical or unusual items. The conclusions drawn are supported by the chronology of the archival literature and statements made by the principals or their biographers.

Acknowledgments. I thank the reviewers and editors for comments that improved the paper. I am grateful for help with images, interpretations, sources, and text that was generously given by W. Aspray, M. Benzi, D. Bindel, J. Butcher, A. B. Carr, F. J. Conahan, M. Dickey, W. Gander, G. H. Golub, C. B. Greifenstein, D. Knuth, P. D. Lax, J. J. H. Miller, E. Mosner, R. A. Olshen, R. Otnes, P. E. Saylor, R. Shrake, J. Todd, J. Wolff, and E. Zass.

REFERENCES

- [1] J. F. ADAMS, *Maxwell Herman Alexander Newman: 7 February 1897–22 February 1984*, Biogr. Mem. Fell. R. Soc., 31 (1985), pp. 436–452.
- [2] S. N. AFRIAT, *Bounds for the characteristic values of matrix functions*, Q. J. Math., 2 (1951), pp. 81–84.
- [3] F. AFTALION, *A History of the International Chemical Industry*, 2nd ed., Chemical Heritage Foundation, Philadelphia, 2001; translated from the French by O. T. Benfrey.
- [4] H. H. AIKEN, ED., *Proceedings of a Symposium on Large-Scale Digital Calculating Machinery*, Harvard University Press, Cambridge, MA, 1948. Reprinted by MIT Press, Cambridge, MA, 1985.
- [5] A. C. AITKEN, *Studies in practical mathematics I. The evaluation, with application, of a certain triple product matrix*, Proc. Roy. Soc. Edinburgh Sect. A, 57 (1937), pp. 172–181.
- [6] K. J. ARROW AND E. L. LEHMANN, *Harold Hotelling 1895–1973*, Biogr. Mem. Natl. Acad. Sci., 87 (2006), pp. 220–233.
- [7] W. ASPRAY, *International diffusion of computer technology, 1945–1955*, Ann. Hist. Comput., 8 (1986), pp. 351–360.
- [8] W. ASPRAY, *The origins of John von Neumann’s theory of automata*, in *The Legacy of John von Neumann*, J. Glimm, J. Impagliazzo, and I. Singer, eds., AMS, Providence, RI, 1988, pp. 289–310.
- [9] W. ASPRAY, ED., *Computing Before Computers*, Iowa State University Press, Ames, IA, 1990.
- [10] W. ASPRAY, *John von Neumann and the Origins of Modern Computing*, MIT Press, Cambridge, MA, 1990.

- [11] W. ASPRAY, *The Institute for Advanced Study computer: A case study in the application of concepts from the history of technology*, in *The First Computers—History and Architectures*, R. Rojas and U. Hashagen, eds., MIT Press, Cambridge, MA, 2000, pp. 179–194.
- [12] W. ASPRAY AND A. BURKS, EDS., *Papers of John von Neumann on Computing and Computer Theory*, MIT Press, Cambridge, MA, 1987.
- [13] D. H. BAILEY, *High-precision floating-point arithmetic in scientific computation*, *Comput. Sci. Eng.*, 7 (2005), pp. 54–61.
- [14] M. L. BALINSKI AND H. P. YOUNG, *The quota method of apportionment*, *Amer. Math. Monthly*, 82 (1975), pp. 701–730.
- [15] S. BANACH, *Théorie des Opérations linéaires*, Subwencji Funduszu Kultury Narodowej, Warsaw, 1932.
- [16] T. BANACHIEWICZ, *Sur la résolution numérique d'un système d'équations linéaires*, *Bulletin International de l'Académie Polonaise des Sciences et des Lettres, Série A*, (1937), p. 350.
- [17] T. BANACHIEWICZ, *Zur Berechnung der Determinanten wie auch der Inversen, und zu darauf basierten Auflösung linearer Gleichungen*, *Acta Astronomica*, 3 (1937), pp. 41–67.
- [18] T. BANACHIEWICZ, *Méthode de résolution numérique des équations linéaires, du calcul des déterminants et des inverses, et de réduction des formes quadratiques*, *Bulletin International de l'Académie Polonaise des Sciences et des Lettres, Série A*, (1938), pp. 393–404.
- [19] T. BANACHIEWICZ, *Principes d'une nouvelle technique de la méthode des moindres carrés*, *Bulletin International de l'Académie Polonaise des Sciences et des Lettres, Série A*, (1938), pp. 134–135.
- [20] T. BANACHIEWICZ, *An outline of the Cracovian algorithm of the method of least squares*, *Astr. J.*, 50 (1942), pp. 38–41.
- [21] V. BARGMANN, D. MONTGOMERY, AND J. VON NEUMANN, *Solution of linear systems of high order*, in *John von Neumann Collected Works*, Vol. 5, A. H. Taub, ed., Macmillan, New York, 1946, pp. 421–477.
- [22] D. P. BARTLETT, *General Principles of the Method of Least Squares with Applications*, 3rd ed., private printing, Boston, 1915.
- [23] C. J. BASHE, L. R. JOHNSON, J. H. PALMER, AND E. W. PUGH, *IBM's Early Computers*, MIT Press, Cambridge, MA, 1986.
- [24] F. L. BAUER, *On the definition of condition numbers and their relation to closed methods for solving linear systems*, in *Proceedings of the International Conference on Information Processing* (UNESCO, Paris, 1959), Vol. 1, R. Oldenbourg, Munich and Butterworths, London, 1960, pp. 109–110.
- [25] D. BAXANDALL, *Calculating machines*, in *Encyclopedia Britannica*, Vol. 4, 14th ed., 1929, pp. 548–553.
- [26] BENOIT, *Note sur une méthode ... (Procédé du Commandant Cholesky)*, *Bull. Géodésique*, (1924), pp. 67–77. The author is identified only as Commandant Benoit.
- [27] T. J. BERGIN, ED., *Fifty Years of Army Computing: From ENIAC to MSRC*, Technical Report AR-SL-93, U.S. Army Research Laboratory, Aberdeen, MD, 2000.
- [28] B. C. BERNDT, *Hans Rademacher 1892–1969*, in *The Rademacher Legacy to Mathematics*, G. E. Andrews, D. M. Bressoud, and L. A. Parson, eds., AMS, Providence, RI, 1994, pp. xiii–xxxvi.
- [29] D. S. BERNSTEIN, *Matrix Mathematics: Theory, Facts, and Formulas with Application to Linear Systems Theory*, 2nd ed., Princeton University Press, Princeton, NJ, 2009.
- [30] G. BIRKHOFF, *Two Hadamard numbers for matrices*, *Comm. ACM*, 18 (1975), pp. 25–29.
- [31] C. BLAIR, *Passing of a great mind*, *LIFE*, 42 (8) (1957), pp. 89ff.
- [32] E. BODEWIG, *Bericht über die verschiedenen Methoden zur Lösung eines Systems linear Gleichungen mit reellen Koeffizienten. I through V*, *Indag. Math. Part I*, 9 (1947), pp. 441–452. *Part II*, 9 (1947), pp. 518–530. *Part III*, 9 (1947), pp. 611–621. *Part IV*, 10 (1948), pp. 24–35. *Part V*, 10 (1948), pp. 82–90. Also published in *Koninklijke Nederlandsche Akademie van Wetenschappen, Proceedings. Part I*, 50 (1947), pp. 930–941. *Part II*, 50 (1947), pp. 1104–1116. *Part III*, 50 (1947), pp. 1285–1295. *Part IV*, 51 (1948), pp. 53–64. *Part V*, 51 (1948), pp. 211–219.
- [33] E. BODEWIG, *Review of "Numerical inverting of matrices of high order" (J. von Neumann and H. H. Goldstine)*, *Bull. Amer. Math. Soc.*, 53 (1947), pp. 1021–1099, *Math. Rev.*, 9 (1948), MR0024235; see <http://www.ams.org/mathscinet/>.
- [34] E. BODEWIG, *Matrix Calculus*, 2nd ed., North-Holland, Amsterdam, 1959.
- [35] L. BONOLIS, *Enrico Fermi's scientific work*, in *Enrico Fermi: His Work and Legacy*, C. Bernardini and L. Bonolis, eds., Società Italiana di Fisica, Bologna, Springer-Verlag, Berlin, 2004, pp. 314–394.
- [36] C. DE BOOR AND A. PINKUS, *Backward error analysis for totally positive linear systems*, *Numer. Math.*, 27 (1977), pp. 485–490.

- [37] A. BOREL, *Deane Montgomery 1909–1992*, Proc. Amer. Philos. Soc., 137 (1993), pp. 452–456.
- [38] A. H. BOWKER, *On the norm of a matrix*, Ann. Math. Statist., 18 (1947), pp. 285–288.
- [39] J. R. BRINK AND C. R. HADEN, *Prologue: The computer and the brain. An international symposium in commemoration of John von Neumann*, Ann. Hist. Comput., 11 (1989), pp. 161–163.
- [40] J. R. BROWN, *Philosophy of Mathematics: A Contemporary Introduction to the World of Proofs and Pictures*, Routledge, New York, 2nd ed., 2008.
- [41] A. R. BURKS AND A. W. BURKS, *The First Electronic Computer: The Atanasoff Story*, University of Michigan Press, Ann Arbor, 1988.
- [42] A. W. BURKS, H. H. GOLDSTINE, AND J. VON NEUMANN, *Preliminary discussion of the logical design of an electronic computing instrument*, Part I, Volume 1, in John von Neumann Collected Works, A. H. Taub, ed., Vol. 5, Macmillan, New York, 1963, pp. 34–79.
- [43] V. BUSH AND S. H. CALDWELL, *A new type of differential analyzer*, J. Franklin Inst., 240 (1945), pp. 255–326.
- [44] M. CAMPBELL-KELLY, *The History of the History of Computing*; available from various sources on the web, 2002.
- [45] M. CAMPBELL-KELLY, *The ACE and the shaping of British computing*, in Alan Turing's Automatic Computing Engine, B. J. Copeland, ed., Oxford University Press, Oxford, 2005, pp. 149–172.
- [46] M. CAMPBELL-KELLY, M. CROARKEN, R. FLOOD, AND E. ROBSON, EDs., *The History of Mathematical Tables: From Sumer to Spreadsheets*, Oxford University Press, Oxford, 2003.
- [47] M. CAMPBELL-KELLY AND M. R. WILLIAMS, EDs., *The Moore School Lectures*, MIT Press, Cambridge, MA, 1985. Publication of the 1946 lecture notes.
- [48] B. E. CARPENTER AND R. W. DORAN, EDs., *A. M. Turing's ACE Report of 1946 and Other Papers*, MIT Press, Cambridge, MA, 1986.
- [49] J. W. CARR, *Error analysis in floating point arithmetic*, Comm. ACM, 2 (1959), pp. 10–15.
- [50] U. CASSINA, *Sul numero delle operazioni elementari necessarie per la risoluzione dei sistemi di equazioni lineari*, Boll. Un. Mat. Ital. (3), 3 (1948), pp. 142–147.
- [51] P. CERUZZI, *Crossing the divide: Architectural issues and the emergence of the stored program computer, 1935–1955*, IEEE Ann. Hist. Comput., 19 (1997), pp. 5–12.
- [52] P. CERUZZI, *“Nothing New Since von Neumann”: A Historian Looks at Computer Architecture, 1945–1995*, in The First Computers—History and Architectures, R. Rojas and U. Hashagen, eds., MIT Press, Cambridge, MA, 2000, pp. 195–217.
- [53] J. G. CHARNEY, R. FJÖRTOFT, AND J. VON NEUMANN, *Numerical integration of the barotropic vorticity equation*, Tellus, 2 (1950), pp. 237–254.
- [54] W. CHAUVENET, *Treatise on the Method of Least Squares*, J. B. Lippincott & Co., Philadelphia, 1868.
- [55] S. CHIKARA, S. MITSUO, AND J. W. DAUBEN, EDs., *The Intersection of History and Mathematics*, Birkhäuser, Basel, 1994.
- [56] J. A. CLARKSON, *The von Neumann–Jordan constant for the Lebesgue spaces*, Ann. of Math., 38 (1937), pp. 114–115.
- [57] I. B. COHEN, *Howard Aiken: Portrait of a Computer Pioneer*, MIT Press, Cambridge, MA, 2000.
- [58] S. K. COHEN, ED., *Interview with John Todd*, Oral History Project, California Institute of Technology Archives, Pasadena, CA, 1997; available online from <http://oralhistories.library.caltech.edu/132/>.
- [59] B. J. COPELAND, ED., *Alan Turing's Automatic Computing Engine*, Oxford University Press, Oxford, 2005.
- [60] R. COURANT, *Method of finite differences for the solution of partial differential equations*, in Proceedings of a Symposium on Large-Scale Digital Calculating Machinery, H. H. Aiken, ed., Harvard University Press, Cambridge, MA, 1948, pp. 153–156.
- [61] R. COURANT, K. FRIEDRICHS, AND H. LEWY, *Über die partiellen Differenzengleichungen der mathematischen Physik*, Math. Ann., 100 (1928), pp. 32–74. English translation in IBM J. Res. Develop., 11 (1967), pp. 215–234.
- [62] M. CROARKEN AND M. CAMPBELL-KELLY, *Beautiful numbers: The rise and decline of the British Association Mathematical Tables Committee, 1871–1965*, IEEE Ann. Hist. Comput., 22 (2000), pp. 44–61.
- [63] P. D. CROUT, *A short method for evaluating determinants and solving systems of linear equations with real or complex coefficients*, Trans. Amer. Inst. Elect. Engineers, 60 (1941), pp. 1235–1241.

- [64] K. K. CURTIS, N. C. METROPOLIS, W. G. ROSEN, Y. SHIMAMOTO, AND J. N. SNYDER, *John R. Pasta, 1918–1981. An unusual path toward computer science*, Ann. Hist. Comput., 5 (1983), pp. 224–238.
- [65] G. G. DAHLQUIST, *A special stability problem for linear multistep methods*, BIT, 3 (1963), pp. 27–43.
- [66] A. D. DALMEDICO, *History and epistemology of models: Meteorology (1946–1963) as a case study*, Arch. Hist. Exact Sci., 55 (2001), pp. 395–422.
- [67] G. B. DANTZIG, *Linear programming*, in History of Mathematical Programming: A Collection of Personal Reminiscences, J. K. Lenstra, A. H. G. Rinnooy-Kan, and A. Schrijver, eds., North-Holland, Amsterdam, 1991, pp. 19–31.
- [68] O. DARRIGOL, *Stability and instability in nineteenth-century fluid mechanics*, Rev. Histoire-Sci., 8 (2002), pp. 5–65.
- [69] C. G. DARWIN, *Douglas Rayner Hartree 1897–1958*, Biogr. Mem. Fell. R. Soc., 4 (1958), pp. 102–116.
- [70] J. W. DAUBEN, *Mathematics: An historian's perspective*, in The Intersection of History and Mathematics, S. Chikara, S. Mitsuo, and J. W. Dauben, eds., Birkhäuser, Basel, 1994, pp. 1–14.
- [71] C. DAVIS, *Where did twentieth-century mathematics go wrong?*, in The Intersection of History and Mathematics, S. Chikara, S. Mitsuo, and J. W. Dauben, eds., Birkhäuser, Basel, 1994, pp. 129–142.
- [72] P. J. DAVIS, E. V. HAYNSWORTH, AND M. MARCUS, *Bound for the P-condition number of matrices with positive roots*, J. Res. Natl. Bur. Stand., 65B (1961), pp. 13–14.
- [73] L. S. DEDERICK, *Firing tables*, in Proceedings of a Symposium on Large-Scale Digital Calculating Machinery, H. H. Aiken, ed., Harvard University Press, Cambridge, MA, 1948, pp. 194–199.
- [74] J.-P. DEDIEU, *Approximate solutions of numerical problems, condition number analysis and condition number theorem*, in The Mathematics of Numerical Analysis, J. Renegar, M. Shub, and S. Smale, eds., Lectures in Appl. Math. 12, AMS, Providence, RI, 1996, pp. 263–283.
- [75] J. DEMMEL AND Y. HIDA, *Accurate and efficient floating point summation*, SIAM J. Sci. Comput., 25 (2003), pp. 1214–1248.
- [76] J. W. DEMMEL, *The condition number and the distance to the nearest ill-posed problem*, Numer. Math., 51 (1987), pp. 251–289.
- [77] J. W. DEMMEL, *The probability that a numerical analysis problem is difficult*, Math. Comp., 50 (1988), pp. 449–480.
- [78] J. W. DEMMEL, *On Floating Point Errors in Cholesky*, Technical Report CS-89-87, Department of Computer Science, University of Tennessee, Knoxville, 1989. LAPACK Working Note 14.
- [79] J. DIEUDONNÉ, *Von Neumann, Johann (or John)*, in Dictionary of Scientific Biography, Vol. 14, C. C. Gillispie, ed., Charles Scribner's Sons, New York, 1981, pp. 89–92.
- [80] M. H. DOOLITTLE, *Method employed in the solution of normal equations and in the adjustment of a triangularization*, in Report of the Superintendent of the Coast and Geodetic Survey... Ending with June, 1878, Government Printing Office, Washington, D.C., 1881, pp. 115–120.
- [81] D. B. DUNCAN AND J. F. KENNEY, *On the Solution of Normal Equations and Related Topics*, Pamphlet, Edwards Brothers, Ann Arbor, 1946.
- [82] P. S. DWYER, *The solution of simultaneous equations*, Psychometrika, 6 (1941), pp. 101–129.
- [83] P. S. DWYER, *A matrix presentation of least squares and correlation theory with matrix justification of improved methods of solution*, Ann. Math. Statist., 15 (1944), pp. 82–89.
- [84] P. S. DWYER, *The square root method and its use in correlation and regression*, J. Amer. Statist. Assoc., 40 (1945), pp. 493–503.
- [85] P. S. DWYER AND F. V. WAUGH, *On errors in matrix inversion*, J. Amer. Statist. Assoc., 48 (1953), pp. 289–319.
- [86] D. E. ECKDAHL, I. S. REED, AND H. H. SARKISSIAN, *West Coast Contributions to the Development of the General-Purpose Computer: Building Maddida and the Founding of Computer Research Corporation*, IEEE Ann. Hist. Comput., 25 (2003), pp. 4–33.
- [87] W. J. ECKERT, *Punched Card Methods in Scientific Computation*, Thomas J. Watson Astronomical Computing Bureau, New York, 1940.
- [88] A. EDELMAN, *Eigenvalues and condition numbers of random matrices*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 543–560.
- [89] A. EDELMAN, *Eigenvalues and Condition Numbers of Random Matrices*, Ph.D. dissertation, MIT, Cambridge, MA, 1989.

- [90] A. EDELMAN AND N. R. RAO, *Random matrix theory*, Acta Numer., 14 (2005), pp. 233–297.
- [91] I. M. H. ETHERINGTON, *On errors in determinants*, Proc. Edinb. Math. Soc. (2), 3 (1932), pp. 107–117.
- [92] J. EWING, *Paul Halmos: In his own words*, Notices Amer. Math. Soc., 54 (2007), pp. 1136–1144.
- [93] C. A. FELIPPA, *A historical outline of matrix structural analysis: A play in three acts*, Comput. & Structures, 79 (2001), pp. 1313–1324.
- [94] E. FERMI, J. PASTA, S. ULAM, AND M. TSINGOU, *Studies of Nonlinear Problems*, I, Technical Report LA-1940, Los Alamos Scientific Laboratory, 1955.
- [95] G. E. FORSYTHE, *Review of “Bounds for the characteristic values of matrix functions” by S. N. Afriat*, Q. J. Math., 2 (1951), pp. 81–84, Math. Rev., 12 (1951), MR0041094; see <http://www.ams.org/mathscinet/>.
- [96] G. E. FORSYTHE, *Solving linear algebraic equations can be interesting*, Bull. Amer. Math. Soc., 59 (1953), pp. 299–329.
- [97] G. E. FORSYTHE AND C. B. MOLER, *Computer Solution of Linear Algebraic Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1967.
- [98] G. E. FORSYTHE AND T. S. MOTZKIN, *An extension of Gauss’ transformation for improving the condition of systems of linear equations*, Math. Tables Aids Comp., 6 (1952), pp. 9–17.
- [99] G. E. FORSYTHE AND E. G. STRAUS, *On best conditioned matrices*, Proc. Amer. Math. Soc., 6 (1955), pp. 340–345.
- [100] K. A. FOX, *Agricultural economists in the econometric revolution: Institutional background, literature and leading figures*, Oxford Economic Papers (N.S.), 41 (1989), pp. 53–70.
- [101] L. FOX, *James Hardy Wilkinson. 27 September 1919–5 October 1986*, Biogr. Mem. Fell. R. Soc., 33 (1987), pp. 670–708.
- [102] L. FOX, *Early numerical analysis in the United Kingdom*, in A History of Scientific Computing, S. G. Nash, ed., ACM Press, New York, 1990, pp. 280–300.
- [103] L. FOX, H. D. HUSKEY, AND J. H. WILKINSON, *Notes on the solution of algebraic linear systems of equations*, Quart. J. Mech. Appl. Math., 1 (1948), pp. 150–173.
- [104] R. A. FRAZER, W. J. DUNCAN, AND A. R. COLLAR, *Elementary Matrices and Some Applications to Dynamics and Differential Equations*, Cambridge University Press, Cambridge, UK, 1938.
- [105] F. G. FROBENIUS, *Über den von L. Bieberbach gefundenen Beweis eines Satzes von C. Jordan*, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin, (1911), pp. 241–248. Reprinted in Ferdinand Georg Frobenius. Gesammelte Abhandlungen, Vol. 3, J.-P. Serre, ed., Springer-Verlag, Berlin, 1968, pp. 493–500.
- [106] S. GASS, *The first linear programming shoppe*, Oper. Res., 50 (2002), pp. 61–68.
- [107] C. F. GAUSS, *Disquisitio de elementis ellipticis Palladis*, Commentationes Societatis Regiae Scientiarum Gottingensis Recentiores: Commentationes Classis Mathematicae, 1 (1810), pp. 1–26.
- [108] C. F. GAUSS, *Supplementum theoriae combinationis observationum erroribus minimis obnoxiae*, Commentationes Societatis Regiae Scientiarum Gottingensis Recentiores: Commentationes Classis Mathematicae, 6 (1826), pp. 57–98.
- [109] J. W. GIVENS, *Numerical Computation of the Characteristic Values of a Real Symmetric Matrix*, Report ORNL-1574, Oak Ridge National Laboratory, Oak Ridge, TN, 1954.
- [110] J. W. GIVENS, *The Linear Equations Problem*, Technical Report 3, Applied Mathematics and Statistics Laboratory, Stanford University, Stanford, CA, 1959.
- [111] J. GLIMM, J. IMPAGLIAZZO, AND I. SINGER, EDS., *The Legacy of John von Neumann*, AMS, Providence, 1990.
- [112] M. D. GODFREY AND D. F. HENDRY, *The computer as von Neumann planned it*, IEEE Ann. Hist. Comput., 15 (1993), pp. 11–21.
- [113] H. H. GOLDSTINE, *Letter to John von Neumann on July 19, 1947*, in John von Neumann Papers, Box 4, Goldstine folder, Manuscript Division, Library of Congress, Washington, D.C., 1947.
- [114] H. H. GOLDSTINE, *Letter to John von Neumann on July 23, 1947*, in John von Neumann Papers, Box 4, Goldstine folder, Manuscript Division, Library of Congress, Washington, D.C., 1947.
- [115] H. H. GOLDSTINE, *The Computer from Pascal to von Neumann*, Princeton University Press, Princeton, NJ, 1972.
- [116] H. H. GOLDSTINE, *A History of the Calculus of Variations from the 17th through the 19th Century*, Springer-Verlag, New York, 1977.
- [117] H. H. GOLDSTINE, *Remembrance of things past*, in A History of Scientific Computing, S. G. Nash, ed., ACM Press, New York, 1990, pp. 5–16.

- [118] H. H. GOLDSTINE, J. H. POMERENE, AND C. V. L. SMITH, *Final Progress Report on the Physical Realization of an Electronic Computing Instrument*, Technical Report, Institute for Advanced Study Electronic Computer Project, 1954.
- [119] H. H. GOLDSTINE AND J. VON NEUMANN, *On the principles of large scale computing machines*, in John von Neumann Collected Works, A. H. Taub, ed., Vol. 5, Macmillan, New York, 1963, pp. 1–33.
- [120] H. H. GOLDSTINE AND J. VON NEUMANN, *Planning and coding of problems for an electronic computing instrument*, Part II, Volume 1, in John von Neumann Collected Works, A. H. Taub, ed., Vol. 5, Macmillan, New York, 1963, pp. 80–151.
- [121] H. H. GOLDSTINE AND J. VON NEUMANN, *Planning and coding of problems for an electronic computing instrument*, Part II, Volume 2, in John von Neumann Collected Works, A. H. Taub, ed., Vol. 5, Macmillan, New York, 1963, pp. 152–214.
- [122] H. H. GOLDSTINE AND J. VON NEUMANN, *Planning and coding of problems for an electronic computing instrument*, Part II, Volume 3, in John von Neumann Collected Works, A. H. Taub, ed., Vol. 5, Macmillan, New York, 1963, pp. 215–235.
- [123] H. H. GOLDSTINE AND J. VON NEUMANN, *Numerical inverting of matrices of high order II*, Proc. Amer. Math. Soc., 1 (1951), pp. 188–202.
- [124] G. H. GOLUB, *Numerical methods for solving linear least squares problems*, Numer. Math., 7 (1965), pp. 206–216.
- [125] G. H. GOLUB, *Private communication*, 2007.
- [126] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 2nd ed., The Johns Hopkins University Press, Baltimore, MD, 1989.
- [127] I. GRATTAN-GUINNESS, *The mathematics of the past: Distinguishing its history from our heritage*, Historia Math., 31 (2004), pp. 163–185.
- [128] J. F. GRGAR, *How ordinary elimination became Gaussian elimination*, Historia Math., 38 (2011), pp. 163–218. See also J. F. Grcar, *Mathematicians of Gaussian elimination*, Notices Amer. Math. Soc., 58 (2011), pp. 782–792.
- [129] J. F. GRGAR, *A matrix lower bound*, Linear Algebra Appl., 433 (2010), pp. 203–220.
- [130] J. GREENSTADT, *Recollections of the Technical Computing Bureau*, Ann. Hist. Comput., 5 (1983), pp. 149–152.
- [131] D. A. GRIER, *The ENIAC, the verb “to program” and the emergence of digital computers*, IEEE Ann. Hist. Comput., 18 (1996), pp. 51–55.
- [132] D. A. GRIER, *When computers were humans*, Princeton University Press, Princeton, NJ, 2005.
- [133] H. R. J. GROSCH, *COMPUTER: Bit Slices From a Life*, 3rd ed., Columbia University Computing History Project, 2003; web edition available from <http://www.columbia.edu/acis/history/computer.pdf>. Originally published by Third Millenium Books, Novato, 1991.
- [134] B. GUSTAFSON, H. O. KREISS, AND J. OLIGER, *Time Dependent Problems and Difference Methods*, John Wiley & Sons, New York, 1995.
- [135] P. R. HALMOS, *The Legend of John von Neumann*, Amer. Math. Monthly, 80 (1973), pp. 382–394.
- [136] J. C. HARSANYI, *Banquet Speech*, Nobel Banquet, December 10, 1994; available online from http://nobelprize.org/nobel_prizes/economics/laureates/1994/harsanyi-speech.html
- [137] D. R. HARTREE, *Calculating Instruments and Machines*, University of Illinois Press, Urbana, IL, 1949.
- [138] S. J. HEIMS, *John von Neumann and Norbert Wiener: From Mathematics to the Technologies of Life and Death*, MIT Press, Cambridge, MA, 1980.
- [139] M. HESTENES AND J. TODD, *NBS-INA—The Institute for Numerical Analysis—UCLA 1947–1954*, National Institute of Standards and Technology (NIST) Special Publication 730, U. S. Government Printing Office, Washington, D.C., 1991.
- [140] N. J. HIGHAM, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, 2002.
- [141] A. HODGES, *Alan Turing: The Enigma*, Simon and Schuster, New York, 1983.
- [142] H. HOTELLING, *The teaching of statistics*, Ann. Math. Statist., 11 (1940), pp. 457–470.
- [143] H. HOTELLING, *Some new methods in matrix calculation*, Ann. Math. Statist., 14 (1943), pp. 1–34.
- [144] A. S. HOUSEHOLDER, *Some numerical methods for solving systems of linear equations*, Amer. Math. Monthly, 57 (1950), pp. 453–459.
- [145] A. S. HOUSEHOLDER, *Principles of Numerical Analysis*, McGraw-Hill, New York, 1953.
- [146] A. S. HOUSEHOLDER, *Generation of errors in digital computation*, Bull. Amer. Math. Soc., 80 (1954), pp. 234–247.
- [147] A. S. HOUSEHOLDER, *On Norms of Vectors and Matrices*, Report ORNL-1576, Oak Ridge National Laboratory, Oak Ridge, TN, 1954.

- [148] A. S. HOUSEHOLDER, *Unitary triangularization of a nonsymmetric matrix*, J. Assoc. Comput. Mach., 5 (1958), pp. 339–342.
- [149] A. S. HOUSEHOLDER, *Numerical analysis*, in Lectures on Modern Mathematics, Vol. 1, T. L. Saaty, ed., John Wiley & Sons, New York, 1963, pp. 59–97.
- [150] A. S. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964. Reprinted by Dover, New York, 1975.
- [151] H. D. HUSKEY, *The National Bureau of Standards Western Automatic Computer (SWAC)*, Ann. Hist. Comput., 2 (1980), pp. 111–121.
- [152] A. HYMAN, *Charles Babbage: Pioneer of the Computer*, Princeton University Press, Princeton, NJ, 1982.
- [153] INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, *IEEE Standard for Binary Floating-Point Arithmetic*, ANSI/IEEE Std 754-1985, New York, 1985.
- [154] H. JENSEN, *Herleitung einiger Ergebnisse der Ausgleichsrechnung mit Hilfe von Matrizen*, Meddelelse 13, Geodætisk Institut, Copenhagen, 1939.
- [155] H. JENSEN, *An Attempt at a Systematic Classification of Some Methods for the Solution of Normal Equations*, Meddelelse 18, Geodætisk Institut, Copenhagen, 1944.
- [156] D. JERISON AND D. STROOCK, *Norbert Wiener*, Notices Amer. Math. Soc., 42 (1995), pp. 430–438.
- [157] I. M. JOHNSTONE, *On the distribution of the largest eigenvalue in principal component analysis*, Ann. Statist., 29 (2001), pp. 295–327.
- [158] M. KATO AND Y. TAKAHASHI, *On the von Neumann–Jordan constant for Banach spaces*, Proc. Amer. Math. Soc., 125 (1997), pp. 1055–1062.
- [159] K. KATZ, *Historical content in computer science texts: A concern*, IEEE Ann. Hist. Comput., 19 (1997), pp. 16–19.
- [160] T. H. KJELDSEN, *John von Neumann’s conception of the minimax theorem: A journey through different mathematical contexts*, Arch. Hist. Exact Sci., 56 (2001), pp. 39–68.
- [161] J. R. KLAUDER, *Valentine Bargmann 1908–1989*, Biogr. Mem. Natl. Acad. Sci., 76 (1999), pp. 36–49.
- [162] D. E. KNUTH, *George Forsythe and the development of computer science*, Comm. ACM, 15 (1972), pp. 721–726.
- [163] D. E. KNUTH, *The Art of Computer Programming, Volume 2: Seminumerical Algorithms*, 3rd ed., Addison–Wesley, Menlo Park, CA, 1998.
- [164] J. KOCINSKI, *Cracovian Algebra*, Nova Science Publishers, Hauppauge, 2004.
- [165] H. W. KUHN, *Nonlinear programming: A historical note*, in History of Mathematical Programming: A Collection of Personal Reminiscences, J. K. Lenstra, A. H. G. Rinnooy-Kan, and A. Schrijver, eds., North–Holland, Amsterdam, 1991, pp. 82–96.
- [166] T. S. KUHN, *The Structure of Scientific Revolutions*, 2nd ed., University of Chicago Press, Chicago, 1970.
- [167] J. LACKI, *The early axiomatizations of quantum mechanics: Jordan, von Neumann and the continuation of Hilbert’s program*, Arch. Hist. Exact Sci., 54 (2000), pp. 279–318.
- [168] J. LADERMAN, *The square root method for solving simultaneous linear equations*, Math. Tables Aids Comp., 3 (1948), pp. 13–16.
- [169] T. LANG AND J. D. BRUGUERA, *Floating-point multiply-add-fused with reduced latency*, IEEE Trans. Comput., 53 (2004), pp. 988–1003.
- [170] P. D. LAX, *An Interview with Peter D. Lax*, SIAM, Philadelphia, 2006. Conducted by P. Colella in 2003 and 2004. Available online from http://history.siam.org/pdfs2/Lax_final.pdf.
- [171] P. D. LAX AND R. D. RICHTMYER, *Survey of the stability of linear finite difference equations*, Comm. Pure. Appl. Math., 11 (1956), pp. 267–293.
- [172] E. LEDGER, *The Sun: Its Planets and Their Satellites. A Course of Lectures upon the Solar System*, read in Gresham College, London, in the years 1881 and 1882, pursuant to the will of Sir Thomas Gresham, Edward Stanford, London, 1882.
- [173] J. A. N. LEE, *International Biographical Dictionary of Computer Pioneers*, IEEE Computer Society Press, Los Alamitos, 1995.
- [174] R. I. LEINE, *The historical development of classical stability concepts: Lagrange, Poisson and Lyapunov stability*, Nonlinear Dynam., 59 (2009), pp. 173–182.
- [175] J. K. LENSTRA, A. H. G. RINNOOY-KAN, AND A. SCHRIJVER, EDS., *History of Mathematical Programming: A Collection of Personal Reminiscences*, North–Holland, Amsterdam, 1991.
- [176] LIBRARY OF CONGRESS, *John von Neumann Papers*, Manuscript Division, Washington, D.C.
- [177] A. T. LONSETH, *Systems of linear equations with coefficients subject to error*, Ann. Math. Statist., 13 (1942), pp. 332–337.

- [178] A. T. LONSETH, *The propagation of error in linear problems*, Trans. Amer. Math. Soc., 62 (1947), pp. 193–212.
- [179] X. LUO, *On the foundation of stability*, Econom. Theory, 40 (2009), pp. 185–201.
- [180] C. C. MACDUFFEE, *The Theory of Matrices*, Chelsea, New York, 1946.
- [181] N. MACRAE, *John von Neumann*, Pantheon Books, New York, 1992.
- [182] R. R. M. MALLOCK, *An electrical calculating machine*, Proc. Royal Soc. A, 140 (1933), pp. 457–483.
- [183] M. MARCUS, *An inequality connecting the P-condition number and the determinant*, Numer. Math., 4 (1962), pp. 350–353.
- [184] A. W. MARSHALL AND I. OLKIN, *Norms and inequalities for condition numbers*, Pacific J. Math., 15 (1965), pp. 241–247.
- [185] E. MARTIN, *Die Rechenmaschinen und ihre Entwicklung*, Johannes Meyer, Pappenheim, 1925. Translated and edited by P. A. Kidwell and M. R. Williams as *The Calculating Machines: Their History and Development*, MIT Press, Cambridge, MA, 1992.
- [186] J. W. MAUCHLY, *Accumulation of errors in numerical methods*, in *The Moore School Lectures*, M. Campbell-Kelly and M. R. Williams, eds., MIT Press, Cambridge, MA, 1985, p. 563. See the summary by the editors on the page preceding this article.
- [187] K. O. MAY, *Historiography: A perspective for computer scientists*, in *A History of Computing in the Twentieth Century*, N. Metropolis, J. Howlett, and G.-C. Rota, eds., Academic Press, New York, 1980, pp. 11–18.
- [188] S. MCCARTNEY, *ENIAC*, Walker and Company, New York, 1999.
- [189] W. S. MCCULLOCH AND W. PITTS, *A logical calculus for the ideas immanent in nervous activity*, Bull. Math. Biophys., 5 (1943), pp. 115–133.
- [190] E. J. MCSHANE, *Gilbert Ames Bliss 1876–1951*, Biogr. Mem., Natl. Acad. Sci., 31 (1958), pp. 32–53.
- [191] M. L. MEHTA, *Random Matrices*, 3rd ed., Elsevier/Academic Press, 2004.
- [192] R. K. MERTON, *The Matthew effect in science*, Science, 159 (1968), pp. 56–63.
- [193] R. R. MERTZ, *Interview of Julian Bigelow on January 20, 1971*, in *Computer Oral History Collection*, Archives Center, National Museum of American History, Smithsonian Institution, Washington, D.C.; available online from http://invention.smithsonian.org/resources/fa_comporalhist_list_001.aspx.
- [194] U. C. MERZBACH, *Interview of Franz Alt on February 24, 1969*, in *Computer Oral History Collection*, Archives Center, National Museum of American History, Smithsonian Institution, Washington, D.C.; available online from http://invention.smithsonian.org/resources/fa_comporalhist_list_001.aspx.
- [195] N. METROPOLIS, J. HOWLETT, AND G.-C. ROTA, EDS., *A History of Computing in the Twentieth Century*, Academic Press, New York, 1980.
- [196] N. C. METROPOLIS, G. REITWEISNER, AND J. VON NEUMANN, *Statistical treatment of values of first 2000 digits of e and π calculated on the ENIAC*, Math. Tables Aids Comp., 4 (1950), pp. 109–111.
- [197] J. J. H. MILLER, *On the location of zeros of certain classes of polynomials with applications to numerical analysis*, J. Inst. Math. Appl., 8 (1971), pp. 397–406.
- [198] H. F. MITCHELL JR., *Inversion of a matrix of order 38*, Math. Tables Aids Comp., 3 (1948), pp. 161–166.
- [199] F. R. MOULTON, *On the solution of linear equations having small determinants*, Amer. Math. Monthly, 20 (1913), pp. 242–249.
- [200] F. J. MURRAY, *The Theory of Mathematical Machines*, 2nd ed., King's Crown Press, New York, 1948.
- [201] F. J. MURRAY, *Mathematical Machines*, Vols. 1 and 2, Columbia University Press, New York, 1961.
- [202] S. G. NASH, ED., *A History of Scientific Computing*, ACM Press, New York, 1990.
- [203] NATIONAL PHYSICAL LABORATORY, *Modern Computing Methods*, 2nd ed., Notes Appl. Sci. 16, H. M. Stationary Office, London, 1961.
- [204] M. NEWMAN AND J. TODD, *The evaluation of matrix inversion programs*, J. Soc. Indust. Appl. Math., 6 (1958), pp. 466–476.
- [205] M. NEWMAN AND J. TODD, *Automatic computers*, in *Survey of Numerical Analysis*, J. Todd, ed., McGraw-Hill, New York, 1962, pp. 160–207.
- [206] I. NEWTON, *Philosophiae Naturalis Principia Mathematica*, 3rd ed., William and John Innys, London, 1726.
- [207] A. L. NORBERG, *Table making in astronomy*, in *The History of Mathematical Tables: From Sumer to Spreadsheets*, M. Campbell-Kelly, M. Croarken, R. Flood, and E. Robson, eds., Oxford University Press, Oxford, 2003, pp. 176–207.

- [208] NORWEGIAN ACADEMY OF SCIENCE AND LETTERS, *Lax receives Abel Prize*, Notices Amer. Math. Soc., 52 (2005), pp. 632–633.
- [209] A. NOWAK, *John von Neumann: A Documentary*, MAA Video Classics, Mathematical Association of America, Washington, D.C., 1966.
- [210] W. OETTLI, *On the solution set of a linear system with inaccurate coefficients*, SIAM J. Numer. Anal., 2 (1965), pp. 115–118.
- [211] W. OETTLI AND W. PRAGER, *Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides*, Numer. Math., 6 (1964), pp. 405–409.
- [212] W. OETTLI, W. PRAGER, AND J. H. WILKINSON, *Admissible solutions of linear systems with not sharply defined coefficients*, SIAM J. Numer. Anal., 2 (1965), pp. 291–299.
- [213] E. S. ORAN AND J. P. BORIS, *Numerical Simulation of Reactive Flow*, 2nd ed., Cambridge University Press, Cambridge, UK, 2001.
- [214] A. OSTROWSKI, *Über Normen von Matrizen*, Math. Z., 63 (1955), pp. 2–18.
- [215] J. R. PANNABECKER, *For a history of technology education: Contexts, systems, and narratives*, J. Tech. Educ., 7 (1995).
- [216] B. N. PARLETT, *Progress in numerical analysis*, SIAM Rev., 20 (1978), pp. 443–456.
- [217] B. N. PARLETT, *The contribution of J. H. Wilkinson to numerical analysis*, in A History of Scientific Computing, S. G. Nash, ed., ACM Press, New York, 1990, pp. 17–30.
- [218] N. A. PHILLIPS, *Jule Charney's influence on meteorology*, Bull. Amer. Meteorol. Soc., 63 (1982), pp. 492–498.
- [219] N. A. PHILLIPS, *Jule Gregory Charney 1917–1981*, Biogr. Mem. Natl. Acad. Sci., 66 (1995), pp. 80–113.
- [220] H. POLACHEK, *Before the ENIAC*, IEEE Ann. Hist. Comput., 19 (1997), pp. 25–30.
- [221] D. A. POPE AND M. L. STEIN, *Multiple precision arithmetic*, Comm. ACM, 3 (1960), pp. 652–654.
- [222] W. POUNDSTONE, *Prisoner's Dilemma*, Doubleday, New York, 1992.
- [223] E. W. PUGH AND W. ASPRAY, *Creating the computer industry*, IEEE Ann. Hist. Comput., 18 (1996), pp. 7–17.
- [224] H. RADEMACHER, *On the Accumulation of Errors in Numerical Integration on the ENIAC*, in The Moore School Lectures, M. Campbell-Kelly and M. R. Williams, eds., MIT Press, Cambridge, MA, 1985, pp. 223–237.
- [225] H. A. RADEMACHER, *On the accumulation of errors in processes of integration on high-speed calculating machines*, in Proceedings of a Symposium on Large-Scale Digital Calculating Machinery, H. H. Aiken, ed., Harvard University Press, Cambridge, MA, 1948, pp. 176–185.
- [226] B. RANDELL, *Colossus: Godfather of the computer*, New Scientist, 73 (1977), pp. 346–348.
- [227] W. J. M. RANKINE, *A Manual of Civil Engineering*, 8th ed., Charles Griffin and Company, London, 1872.
- [228] M. REDEI, ED., *John von Neumann: Selected Letters*, AMS, Providence, RI, 2005.
- [229] M. REES, *The Computing Program of the Office of Naval Research, 1946–1953*, Ann. Hist. Comput., 2 (1982), pp. 831–848. Reprinted in Comm. ACM, 30 (1987), pp. 832–848.
- [230] E. REGIS, *Good Time Johnny*, in Who Got Einstein's Office?, Addison-Wesley, Reading, MA, 1987, pp. 97–122.
- [231] C. REID, *Hilbert*, Springer-Verlag, Berlin, 1970.
- [232] C. REID, *Courant in Göttingen and New York*, Springer-Verlag, New York, 1976.
- [233] C. REID, *K. O. Friedrichs*, Math. Intelligencer, 5 (1983), pp. 23–30.
- [234] C. REID, *Hans Lewy 1904–1988*, in Miscellanea Mathematica, P. Hilton, F. Hirzebruch, and R. Remmert, eds., Springer-Verlag, Berlin, 1991, pp. 259–267.
- [235] T. RELLA, *Über den absoluten Betrag von Matrizen*, in Comptes Rendus du Congrès International des Mathématiciens, Oslo 1936, Tome II, Conférences de Sections, A. W. Brøggers Boktrykkeri A/S, Oslo, 1937, pp. 29–31.
- [236] J. RENEGAR, M. SHUB, AND S. SMALE, EDS., *The Mathematics of Numerical Analysis*, Lectures in Appl. Math. 12, AMS, Providence, RI, 1996.
- [237] R. D. RICHTMYER, *Difference Methods for Initial Value Problems*, Interscience Publishers, New York, 1957.
- [238] R. ROJAS, *Die Rechenmaschinen von Konrad Zuse*, Springer-Verlag, Berlin, 1998.
- [239] R. ROJAS AND U. HASHAGEN, EDS., *The First Computers—History and Architectures*, MIT Press, Cambridge, MA, 2000.
- [240] E. J. ROUTH, *A Treatise on the Stability of a Given State of Motion, Particularly Steady Motion*, MacMillan and Co., London, 1877.

- [241] D. E. ROWE, *The Philosophical Views of Klein and Hilbert*, in *The Intersection of History and Mathematics*, S. Chikara, S. Mitsuo, and J. W. Dauben, eds., Birkhäuser, Basel, 1994, pp. 187–202.
- [242] H. RUTISHAUSER, *Lectures on Numerical Mathematics*, Birkhäuser, Boston, 1990. Edited by M. Gutknecht, translated by W. Gautschi.
- [243] F. E. SATTERTHWAITE, *Error control in matrix calculation*, *Ann. Math. Statist.*, 15 (1944), pp. 373–387.
- [244] J. B. SCARBOROUGH, *Numerical Mathematical Analysis*, The Johns Hopkins Press, Baltimore, MD, 1930.
- [245] H. R. SCHWARZ, *Numerik symmetrischer Matrizen*, Teubner, Stuttgart, 1968. Translated and reprinted: H. R. Schwarz, H. Rutishauser, and E. Stiefel, *Numerical Analysis of Symmetric Matrices*, Prentice-Hall, Englewood Cliffs, NJ, 1973.
- [246] D. SEGEL, D. MUKAMEL, O. KRICHEVSKY, AND J. STAVANS, *Selection mechanism and area distribution in two-dimensional cellular structures*, *Phys. Rev. E*, 47 (1993), pp. 812–819.
- [247] J.-P. SERRE, ED., *Ferdinand Georg Frobenius. Gesammelte Abhandlungen*, Springer-Verlag, Berlin, 1968.
- [248] C. SEVERANCE, *IEEE 754: An interview with William Kahan*, *Computer*, 31 (1998), pp. 114–115.
- [249] S. SHAPIRO, *Thinking about Mathematics: The Philosophy of Mathematics*, Oxford University Press, New York, 2000.
- [250] M. SHAW ET AL., *Computer Science: Reflections on the Field, Reflections from the Field*, The National Academies Press, Washington, D.C., 2004.
- [251] M. SHUB, *Panel discussion: Does numerical analysis need a model of computation?*, in *The Mathematics of Numerical Analysis*, J. Renegar, M. Shub, and S. Smale, eds., *Lectures in Appl. Math.* 12, AMS, Providence, RI, 1996, pp. 1–18.
- [252] S. SIMONS, *Minimax theorems and their proofs*, in *Minimax and Applications*, D.-Z. Du and P. M. Pardalos, eds., Kluwer Academic Publishers, Dordrecht, 1995, pp. 1–24.
- [253] S. SMALE, *On the efficiency of algorithms of analysis*, *Bull. Amer. Math. Soc.*, 13 (1985), pp. 87–121.
- [254] SMITHSONIAN INSTITUTION, *Computer Oral History Collection*. Archives Center, National Museum of American History, Washington, D.C.; available online from http://invention.smithsonian.org/resources/fa_comporalhist_list_001.aspx.
- [255] J. M. STAUDENMAIER, *Technology's Storytellers: Reweaving the Human Fabric*, MIT Press, Cambridge, MA, 1985.
- [256] N. STERN, *John von Neumann's influence on electronic digital computers, 1944–1946*, *Ann. Hist. Comput.*, 2 (1980), pp. 349–362.
- [257] N. STERN, *From ENIAC to UNIVAC*, Digital Press, Bedford, 1981.
- [258] N. STERN, *Minutes of 1947 Patent Conference, Moore School of Electrical Engineering, University of Pennsylvania*, *Ann. Hist. Comput.*, 7 (1985), pp. 100–116.
- [259] G. W. STEWART, *Introduction to Matrix Computations*, Academic Press, New York, 1973.
- [260] G. W. STEWART, *Matrix Algorithms, Volume I: Basic Decompositions*, SIAM, Philadelphia, 1998.
- [261] G. W. STEWART, *The decompositional approach to matrix computation*, *Comput. Sci. Eng.*, 2 (2000), pp. 50–59.
- [262] G. W. STEWART AND J.-G. SUN, *Matrix Perturbation Theory*, Academic Press, San Diego, CA, 1990.
- [263] G. STIBITZ, *Introduction to the course on electronic digital computers*, in *The Moore School Lectures*, M. Campbell-Kelly and M. R. Williams, eds., MIT Press, Cambridge, MA, 1985, pp. 3–18.
- [264] W. W. STIFLER, ED., *High-Speed Computing Devices*, McGraw-Hill, New York, 1950.
- [265] P. G. TAIT, *Cosmical astronomy*, *Good Words*, (1875), pp. 19–23.
- [266] A. H. TAUB, ED., *John von Neumann Collected Works*, Vols. 1–6, Macmillan, New York, 1963.
- [267] O. TAUSSKY, *Note on the condition of matrices*, *Math. Tables Aids Comp.*, 4 (1950), pp. 111–112.
- [268] O. TAUSSKY AND J. TODD, *Cholesky, Toeplitz and the triangular factorization of symmetric matrices*, *Numer. Algorithms*, 41 (2006), pp. 197–202.
- [269] M. TERRALL, ED., *Olga Taussky-Todd Autobiography*, History Project, California Institute of Technology Archives, Pasadena, CA, 1980; available online from <http://oralhistories.library.caltech.edu/43/>.
- [270] THE INSTITUTE LETTER, *John von Neumann Centennial*, IAS, 2003.
- [271] THE NEW YORK TIMES, *Dr. von Neumann of AEC, 53, Dies*, February 9, 1957, p. 19.

- [272] J. TODD, *The Condition of Certain Matrices*, I, Quart. J. Mech. Appl. Math., 2 (1949), pp. 469–472.
- [273] J. TODD, ED., *Survey of Numerical Analysis*, McGraw–Hill, New York, 1962.
- [274] J. TODD, *On condition numbers*, in Colloques Internationaux 165, Centre National de la Recherche Scientifique, Programmation en Mathématiques Numériques, Éditions Centre Nat. Recherche Sci., Paris, 1966, pp. 141–159.
- [275] J. TODD, *John Hamilton Curtiss, 1909–1977*, Ann. Hist. Comput., 2 (1980), pp. 104–110.
- [276] J. F. TRAUB, *Numerical mathematics and computer science*, Comm. ACM, 15 (1972), pp. 537–541.
- [277] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, NJ, 2005.
- [278] H. TROPP, *Interview of Mina Rees on October 20, 1972*, in Computer Oral History Collection, Archives Center, National Museum of American History, Smithsonian Institution, Washington, D.C.; available online from http://invention.smithsonian.org/resources/fa_comporalhist.list.001.aspx.
- [279] H. TROPP, *Interview of Mrs. Ida Rhodes on March 21, 1973*, in Computer Oral History Collection, Archives Center, National Museum of American History, Smithsonian Institution, Washington, D.C.; available online from http://invention.smithsonian.org/resources/fa_comporalhist.list.001.aspx.
- [280] L. B. TUCKERMAN, *On the mathematically significant figures in the solution of simultaneous linear equations*, Ann. Math. Statist., 12 (1941), pp. 307–316.
- [281] A. M. TURING, *Equivalence of left and right almost periodicity*, J. London Math. Soc., 10 (1935), pp. 284–285.
- [282] A. M. TURING, *On computable numbers, with application to the Entscheidungsproblem*, Proc. London Math. Soc., 42 (1936), pp. 230–265.
- [283] A. M. TURING, *Finite approximations to Lie groups*, Ann. of Math., 39 (1938), pp. 105–111.
- [284] A. M. TURING, *Proposed Electronic Calculator*, Report, National Physical Laboratory, Teddington, 1946. Title varies in revision. Available in collections of Turing’s work such as A. M. Turing’s ACE Report of 1946 and Other Papers, B. E. Carpenter and R. W. Doran, eds., MIT Press, Cambridge, MA, 1986.
- [285] A. M. TURING, *Rounding-off errors in matrix processes*, Quart. J. Mech. Appl. Math., 1 (1948), pp. 287–308.
- [286] S. ULAM, *John von Neumann, 1903–1957*, Bull. Amer. Math. Soc., 64 (1958), pp. 1–49.
- [287] S. M. ULAM, *Adventures of a Mathematician*, Charles Scribner’s Sons, New York, 1976.
- [288] R. S. VARGA, *Matrix Iterative Analysis*, Prentice–Hall, Englewood Cliffs, NJ, 1962.
- [289] D. VISWANATH AND L. N. TREFETHEN, *Condition numbers of random triangular matrices*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 564–581.
- [290] J. VON NEUMANN, *Herstellung von Naphtazarin und seiner Kondensationsprodukte mit aromatischen Aminen*, Diplomarbeit im Instituts für Organische Technologie, Eidgenössische Technische Hochschule (ETH), Zurich, 1926. Prof. Hans-Eduard Fierz, supervisor.
- [291] J. VON NEUMANN, *Zur Theorie der Gesellschaftsspiele*, Math. Ann., 100 (1928), pp. 295–320. Translated to English in Contributions to the Theory of Games IV, A. W. Tucker and R. D. Luce, eds., Ann. of Math. Stud. 40, Princeton University Press, Princeton, NJ, 1959, pp. 13–42. Reprinted in John von Neumann Collected Works, Vol. 6, A. H. Taub, ed., Macmillan, New York, 1963, pp. 1–26.
- [292] J. VON NEUMANN, *Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren*, Math. Ann., 102 (1929), pp. 49–131. Reprinted in John von Neumann Collected Works, Vol. 2, A. H. Taub, ed., Macmillan, New York, 1963, pp. 3–85.
- [293] J. VON NEUMANN, *Almost periodic functions in a group*, Trans. Amer. Math. Soc., 36 (1934), pp. 445–492.
- [294] J. VON NEUMANN, *Zum Haarschen Mass in topologischen Gruppen*, Compos. Math., 1 (1934), pp. 106–114.
- [295] J. VON NEUMANN, *Letter to A. Turing on December 5, 1935*, King’s College Library, Turing Papers, AMT/D/5, 1935. <http://www.turingarchive.org>. Dated to 1935 by A. Hodges in Alan Turing: The Enigma, Simon and Schuster, New York, 1983, p. 545, n. 2.36.
- [296] J. VON NEUMANN, *Über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes*, Ergeb. Math. Kolloq., 8 (1935–1936), pp. 73–83. Translated and reprinted in John von Neumann Collected Works, Vol. 6, A. H. Taub, ed., Macmillan, New York, 1963, pp. 29–37.
- [297] J. VON NEUMANN, *Some matrix-inequalities and metrization of matrix space*, Toms. Univ. Rev., 1 (1937), pp. 286–300. Reprinted in John von Neumann Collected Works, Vol. 4, A. H. Taub, ed., Macmillan, New York, 1963, pp. 205–219.

- [298] J. VON NEUMANN, *Letter to Dr. J. R. Oppenheimer on August 1, 1944 to report on calculating machines*, Los Alamos National Laboratory Archives, 1944.
- [299] J. VON NEUMANN, *First draft of a report on the EDVAC*, 1945. Reprinted in *Papers of John von Neumann on Computing and Computer Theory*, W. Aspray and A. Burks, eds., MIT Press, Cambridge, MA, 1987, pp. 17–82; *From ENIAC to UNIVAC*, N. Stern, Digital Press, Bedford, 1981, pp. 177–246; and *IEEE Ann. Hist. Comput.*, 15 (1993), pp. 27–75.
- [300] J. VON NEUMANN, *Letter to Maxwell H. A. Newman on March 19, 1946*, in *John von Neumann Papers*, Box 5, Newman folder, Manuscript Division, Library of Congress, Washington, D.C., 1946.
- [301] J. VON NEUMANN, *Letter to Norbert Wiener on November 29, 1946*, in *John von Neumann Papers*, Manuscript Division, Library of Congress, Washington, D.C. Reproduced in *John von Neumann: Selected Letters*, M. Redei, ed., AMS, Providence, RI, 2005.
- [302] J. VON NEUMANN, *Letter to H. Goldstine from Bermuda on January 11, 1947*, in *Herman Goldstine papers in the American Philosophical Society manuscripts collection*, 1947.
- [303] J. VON NEUMANN, *The mathematician*, in *The Works of the Mind*, R. B. Heywood, ed., University of Chicago Press, Chicago, 1947. Reprinted in *John von Neumann Collected Works*, Vol. 1, A. H. Taub, ed., Macmillan, New York, 1963, pp. 1–9.
- [304] J. VON NEUMANN, *First report on the numerical calculation of flow problems*, in *John von Neumann Collected Works*, Vol. 5, A. H. Taub, ed., Macmillan, New York, 1963, pp. 664–712.
- [305] J. VON NEUMANN, *Second report on the numerical calculation of flow problems*, in *John von Neumann Collected Works*, Vol. 5, A. H. Taub, ed., Macmillan, New York, 1963, pp. 713–750.
- [306] J. VON NEUMANN, *Letter to A. Turing on September 13, 1949*, King's College Library, Turing Papers, AMT/D/5, 1949, <http://www.turingarchive.org>.
- [307] J. VON NEUMANN, *Letter to R. Plunkett on November 7, 1949*, in *John von Neumann Papers*, Box 6, P miscellaneous folder, Manuscript Division, Library of Congress, Washington, D.C., 1949. Responding to a letter from Plunkett on October 21, 1949 that is kept in the same place. Quoted in *John von Neumann and the Origins of Modern Computing*, W. Aspray, MIT Press, Cambridge, MA, 1990, pp. 102–103.
- [308] J. VON NEUMANN, *Discussion: Shape of metal grains*, in *Metal Interfaces*, C. Herring, ed., Cleveland, 1952, American Society for Metals, pp. 108–110. Remark concerning paper of C. S. Smith, “Grain Shapes and Other Metallurgical Applications of Topology.” Reprinted in *John von Neumann Collected Works*, Vol. 6, A. H. Taub, ed., Macmillan, New York, 1963, pp. 526–527.
- [309] J. VON NEUMANN, *Communication on the Borel notes*, *Econometrica*, 21 (1953), pp. 124–125. Reprinted in *John von Neumann Collected Works*, Vol. 6, A. H. Taub, ed., Macmillan, New York, 1963, pp. 27–28.
- [310] J. VON NEUMANN, *The role of mathematics in the sciences and in society*, address to Princeton alumni, 1954. Reprinted in *John von Neumann Collected Works*, Vol. 6, A. H. Taub, ed., Macmillan, New York, 1963, pp. 477–490.
- [311] J. VON NEUMANN, *On the theory of games of strategy*, in *Contributions to the Theory of Games IV*, A. W. Tucker and R. D. Luce, eds., *Ann. of Math. Stud.* 40, Princeton University Press, Princeton, NJ, 1959, pp. 13–42.
- [312] J. VON NEUMANN, *Theory of self-reproducing automata*, University of Illinois Press, Champaign, IL, 1966.
- [313] J. VON NEUMANN, *The principles of large-scale computing machines*, *Ann. Hist. Comput.*, 10 (1989), pp. 243–256.
- [314] J. VON NEUMANN AND H. H. GOLDSTINE, *Numerical inverting of matrices of high order*, *Bull. Amer. Math. Soc.*, 53 (1947), pp. 1021–1099.
- [315] J. VON NEUMANN AND O. MORGENSTERN, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, 1944.
- [316] J. VON NEUMANN AND R. D. RICHTMYER, *A method for the calculation of hydrodynamic shocks*, *J. Appl. Phys.*, 21 (1950), pp. 232–257.
- [317] N. A. VONNEUMAN, *John von Neumann as Seen by his Brother*, privately published, Meadowbrook, PA, 1987.
- [318] F. V. WAUGH AND P. S. DWYER, *Compact computation of the inverse of a matrix*, *Ann. Math. Statist.*, 16 (1945), pp. 259–271.
- [319] J. H. M. WEDDERBURN, *Lectures on Matrices*, *Colloquium Publ.* 17, AMS, New York, 1934.
- [320] E. WEISS, *Biographies: Eloge: Cuthbert Corwin Hurd (1911–1996)*, *IEEE Ann. Hist. Comput.*, 19 (1997), pp. 65–73.

- [321] N. WIENER, *Letter to John von Neumann on March 24, 1945*, in John von Neumann Papers, Box 7, Wiener folder, Manuscript Division, Library of Congress, Washington, D.C., 1945.
- [322] N. WIENER, *Cybernetics: Or, Control and Communication in the Animal and the Machine*, Technology Press, John Wiley & Sons, New York, 1948.
- [323] M. V. WILKES, *Automatic Digital Computers*, John Wiley & Sons, New York, 1956.
- [324] M. V. WILKES, *Memoirs of a Computer Pioneer*, Series in the History of Computing, MIT Press, Cambridge, MA, 1985.
- [325] M. V. WILKES, D. J. WHEELER, AND S. GILL, *The Preparation of Programs for an Electronic Digital Computer, with special reference to the EDSAC and the use of a library of subroutines*, M. Campbell-Kelly, ed., Charles Babbage Institute reprint series for the History of Computing, Tomash Publishers, Los Angeles, 1982. Originally published by Addison-Wesley, 1951.
- [326] J. H. WILKINSON, *Error analysis of floating-point computation*, Numer. Math., 2 (1960), pp. 319–340.
- [327] J. H. WILKINSON, *Rounding errors in algebraic processes*, in Proceedings of the International Conference on Information Processing (UNESCO, Paris, 1959), Vol. 1, R. Oldenbourg, ed., Munich and Butterworths, London, 1960, pp. 44–53.
- [328] J. H. WILKINSON, *Error analysis of direct methods of matrix inversion*, J. ACM, 8 (1961), pp. 281–330.
- [329] J. H. WILKINSON, *Rounding Errors in Algebraic Processes*, Prentice-Hall, Englewood Cliffs, NJ, 1963.
- [330] J. H. WILKINSON, *A priori analysis of algebraic processes*, in Proceedings of International Congress of Mathematicians (Moscow, 1966), Moscow, 1968, Mir, pp. 629–640.
- [331] J. H. WILKINSON, *Some comments from a numerical analyst*, J. ACM, 18 (1970), pp. 137–147.
- [332] J. H. WILKINSON, *Modern error analysis*, SIAM Rev., 13 (1971), pp. 548–568.
- [333] J. H. WILKINSON, *Error analysis revisited*, IMA Bull., 22 (1986), pp. 192–200.
- [334] M. R. WILLIAMS, *A History of Computing Technology*, 2nd ed., IEEE Computer Society Press, Los Alamitos, 1997.
- [335] M. R. WILLIAMS, *A preview of things to come: Some remarks on the first generation of computers*, in The First Computers—History and Architectures, R. Rojas and U. Hashagen, eds., MIT Press, Cambridge, MA, 2000, pp. 1–13.
- [336] J. WISHART, *The generalized product moment distribution in samples from a normal multivariate population*, Biometrika, 20A (1928), pp. 32–52.
- [337] H. WITTMAYER, *Einfluß der Änderung einer Matrix auf der Lösung des zugehörigen Gleichungssystems, sowie auf die charakteristischen Zahlen und die Eigenvektoren*, ZAMM Z. Angew. Math. Mech., 16 (1936), pp. 287–300.
- [338] T. W. WRIGHT AND J. F. HAYFORD, *The Adjustment of Observations by the Method of Least Squares with Applications to Geodetic Work*, D. Van Nostrand, New York, 1906.
- [339] S. L. ZABELL, *Alan Turing and the central limit theorem*, Amer. Math. Monthly, 102 (1995), pp. 483–494.
- [340] K. ZUSE, *Computerarchitektur aus damaliger und heutiger Sicht*, Abteilung für Informatik, Eidgenössische Technische Hochschule, Zürich, 1992. Parallel German and English texts. Preface by W. Gander.