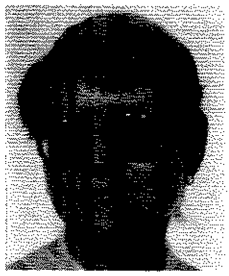
Gauss-Jordan Reduction: A Brief History

STEVEN C. ALTHOEN: I received my B.A. at Kenyon College and my Ph.D. in 1973 at The City University of New York under the direction of Eldon Dyer. I would like to thank Professor Stephen B. Maurer of Swarthmore College for preventing me from crediting the wrong Jordan in my texts by relaying Tucker's observation. My primary research interest is the classification of finite-dimensional real division algebras.



RENATE MCLAUGHLIN: I did my undergraduate work in Germany and received my Ph.D. in complex analysis (1968) at the University of Michigan under the direction of Peter L. Duren. My primary research interest is still complex analysis. But with two children in elementary school, I have also become interested in the mathematics that goes on (or better: does not go on) in elementary school classrooms.

A. W. Tucker [28] has pointed out that it was the geodesist Wilhelm Jordan (1842–1899) and not the mathematician Camille Jordan (1838–1922) who introduced the Gauss-Jordan method of solving systems of linear equations. There is a natural tendency to attribute the method to Camille Jordan, who is justly credited with another linear algebra topic, the Jordan normal form. Recall that in Gaussian Elimination, row operations are used to change the coefficient matrix to an upper triangular matrix. The solution is then found by back substitution, starting from the last equation in the reduced system. In Gauss-Jordan Reduction, row operations are used to diagonalize the coefficient matrix, and the answers are read directly.

This article is about Wilhelm Jordan and the introduction of his method.

1. How do we know it was Wilhelm Jordan? There is little doubt as to the identity of the Jordan referred to in "Gauss-Jordan reduction." It seems that the names were attached to the method by numerical analysts. For example, Householder [13, p. 141] states,

> The Gauss-Jordan method, so-called, seems to have been described first by Clasen [7]. Since it can be regarded as a modification of Gaussian elimination, the name of Gauss is properly applied, but that of Jordan seems to be due to an error, since the method was described only in the third edition of his *Handbuch der Vermessungskunde* prepared after his death.

This reference is clear, but so is the fact that the third edition of Jordan's book was prepared by Jordan himself, well before his death. The foreword to the third edition is dated May 1888 and signed by Jordan. Furthermore, the book review by C. Müller [23] states that Jordan himself prepared the fourth edition, which appeared in 1895. Incidentally, the frontispiece of that particular edition is a picture of Gauss, since he was famous for his geodesy as well as his mathematics and physics. It should also be noted that since Clasen's article [7] appeared in the same year as Jordan's third edition, it seems that their discoveries were independent.

Johnson [15, p. 66] states:

> So far as I know, this method of viewing the reduced normal equations did not appear explicitly in any treatise upon Least Squares prior to the third edition of W. Jordan's *Vermessungskunde*,... .

It should be pointed out that although the first edition contains no hint of Gauss-Jordan reduction, the germ of the idea is already present in the second edition (1877) in the example on pages 34 and 35.

Kunz [18, p. 221] has a section whose title "Gauss-Jordan Method" has the footnote: W. Jordan, "Handbuch der Vermessungskunde." Both Bartlett [5] and Leland [19] mention Jordan's *Handbuch der Vermessungskunde* in their bibliographies. In Gewirtz, Sitomer, and Tucker [10, p. 246] we find:

> Wilhelm Jordan (1819-1904 [sic]) devised the pivot reduction algorithm, known as Gauss-Jordan elimination, for geodetic reasons.

Writers of current linear algebra texts generally apply the name "Gauss-Jordan reduction" without reference to its origins. However, it was through Wilhelm Jordan's *Handbuch der Vermessungskunde* that Gauss-Jordan reduction was introduced to the world. Today Clasen, whom we discuss below, is largely forgotten.

2. Who was W. Jordan? We know about Wilhelm Jordan through several obituaries [26], [16], [24] (see also [11] and [25]). He was born March 1, 1842, in the town of Ellwangen in southern Germany. Following high school, he attended what today would be called an engineering college in Stuttgart. He worked for two years as an engineering assistant on preliminary work for railroad construction and as a "trigonometer" on measuring elevations. He spent two further years as an "Assistant" in geodesy at the college in Stuttgart, and in 1868, when he was only 26 years old, he became full professor of geodesy at the technical college in Karlsruhe.

Jordan was actively involved in surveying several areas of Germany. In 1873 he became editor of the *Zeitschrift für Vermessungswesen* (*Journal for Surveying*) and he remained in this capacity until his death. Jordan was a prolific writer. His major work started in 1873 as *Taschenbuch der Praktischen Geometrie* (*Pocket Book of Practical Geometry*). In later editions this became the *Handbuch der Vermessungskunde* (*Handbook of Geodesy*). By the time of Jordan's death, five editions of this book had appeared, and it had been translated into French, Italian, and Russian.

Jordan's ability to present abstract ideas in lively ways was credited with the wide distribution of this book.

In 1882 Jordan left Karlsruhe and went to the Technical University in Hanover. He continued to be active in surveying and in publishing his works. Apparently he was a first-rate teacher who had a particular talent for bringing out the connections between theory and the real-world problems confronting him. His field trips with students were praised as examples of how one ought to teach.

Although apparently physically fit, Jordan had suffered for years from heart disease and other problems. In 1899, at the age of 57, he died in a state of depression.

3. What was the Problem? Gauss invented the Method of Least Squares to find a best linear function to approximate observed data. The method was naturally attractive to geodesists. Here is a quick description of the mathematics involved, along with an interesting theorem, which was first explicitly stated by Gauss.

Suppose we make m observations, each of which depends on n inputs. Following Jordan, we call the observed values $-l_1, \ldots, -l_n$, but to undertake a modern analysis, it will be useful to label the input vectors

$$(a_{i1}, \ldots, a_{in}), \quad \text{for } i = 1, \ldots, m.$$

Our m observations l_i are a function of the input vectors. It is desirable to have a good linear approximation for that function. Let

$$y = L(y_1, \ldots, y_n) = y_1 x_1 + \cdots + y_n x_n$$

denote a general linear function with coefficients $x_1, \ldots, x_n$. Let v_i denote the error that arises from the linearity assumption:

$$\begin{aligned} v_i &= L(a_{i1}, \ldots, a_{in}) - (-l_i) \\ &= a_{i1}x_1 + \cdots + a_{in}x_n + l_i, \quad \text{for } i = 1, \ldots, m. \end{aligned} \tag{1}$$

We select $x_1, \ldots, x_n$ to minimize the sum of the squared errors

$$E = v_1^2 + \cdots + v_m^2.$$

Thus, the Method of Least Squares reduces to a simple problem of multivariate calculus in which we solve the system of *normal equations*:

$$\partial E/\partial x_j = 0, \qquad \text{for } j = 1, \ldots, n.$$

Let $A = (a_{ij})$, $v = \text{col}(v_1, \ldots, v_m)$, $l = \text{col}(l_1, \ldots, l_m)$, and $x = \text{col}(x_1, \ldots, x_n)$, where col means *column vector*. System 1) can then be written as

$$v = Ax + l.$$

Direct multiplication yields:

$$\begin{aligned} E &= v_1^2 + \cdots + v_m^2 \\ &= \sum a_{i1}a_{i1}x_1^2 + 2\sum a_{i1}a_{i2}x_1x_2 + \cdots + 2\sum a_{i1}a_{in}x_1x_n \\ &\quad + 2\sum a_{i1}l_i x_1 + \cdots, \end{aligned}$$

where we have listed the terms involving x_1 and all the summations run from $i = 1$ to $i = m$. Then

$$\partial E/\partial x_1 = 2\sum a_{i1}a_{i1}x_1 + 2\sum a_{i1}a_{i2}x_2 + \cdots + 2\sum a_{i1}a_{in}x_n + 2\sum a_{i1}l_i = 0,$$

or

$$\sum a_{i1}a_{i1}x_1 + \sum a_{i1}a_{i2}x_2 + \cdots + \sum a_{i1}a_{in}x_n + \sum a_{i1}l_i = 0.$$

Gauss denotes $\Sigma a_{ij}a_{ik}$ by $[a_j a_k]$ and $\Sigma a_{ij}l_i$ by $[a_j l]$. Thus, he would write this equation as

$$[a_1a_1]x_1 + \cdots + [a_1a_n]x_n + [a_1 l] = 0.$$

The remaining $n - 1$ equations are obtained similarly.

It is now easy to verify that the entire system of normal equations can be represented in matrix form as

$$A^tAx = -A^tl$$

(where A^t denotes the transpose of A) with solution

$$x = -(A^tA)^{-1}A^tl.$$

Note that A^tA is a symmetric matrix.

Here is a modern statement and proof of a theorem that occurs in Article 13 of Gauss's *Disquisitio de Elementis ellipticis Palladis* **[8]**.

GAUSS'S THEOREM. *If Gauss-Jordan reduction is used to diagonalize the symmetric matrix*

$$\begin{bmatrix} A^tA & -A^tl \\ (-A^tl)^t & l^tl \end{bmatrix},$$

then the number in the lower right corner is the sum of the squared errors incurred by using the coefficients of the solution vector x as coefficients for the linear approximation and the given data points. That is, after triangularization, the lower right corner will contain

$$E = v_1^2 + \cdots + v_m^2 = v^tv.$$

Proof. View the system

$$\begin{bmatrix} A^tA & -A^tl \\ (-A^tl)^t & 0 \end{bmatrix}$$

as the linear programming problem:

$$\text{Maximize } w = (-A^tl)^tx \tag{2}$$

subject to

$$A^tAx = -A^tl.$$

Then the number in the lower right after triangularization will be the negative of the value of the objective function $(A^tl)^tx$ at the solution to the system. So if we begin with l^tl in that corner, we will obtain

$$l^tl + (A^tl)^t x,$$

where $x = -(A^tA)^{-1}A^tl$.
On the other hand,

$$\begin{aligned} v^tv &= (Ax + l)^t(Ax + l) \\ &= (x^tA^t + l^t)(Ax + l) \\ &= x^tA^tAx + x^tA^tl + l^tAx + l^tl \\ &= x^tA^tA\left[-(A^tA)^{-1}A^tl\right] + x^tA^tl + l^tAx + l^tl \\ &= -x^tA^tl + x^tA^tl + l^tAx + l^tl \\ &= l^tl + (A^tl)^t x. \blacksquare \end{aligned}$$

Thus, if we triangularize the symmetric system

$$\begin{bmatrix} A^tA & -A^tl \\ (-A^tl)^t & l^tl \end{bmatrix},$$

we obtain not only the solution to the normal equations, but also the value of the sum of the squared errors for the particular observations.

4. What was Gauss's Method and Notation? In 1810, in his *Disquisitio de Elementis ellipticis Palladis*, Gauss [8], [9, p. 123] sets out to determine details ("elliptical elements") about the orbit of Pallas, the second-largest asteroid of the solar system. He obtains a system of linear equations in six unknowns, where not all equations can be satisfied simultaneously. Hence he needs to determine values for the unknowns that will minimize the total squared error. Instead of merely solving the problem at hand, Gauss digresses and introduces a method for dealing with such systems of linear equations in general. This is where his characteristic notation appears for the first time.

To save space, we use the symbols of Section 3 above (Gauss lets $p, q, r, \ldots$ denote the variables and writes $\Omega = w^2 + w'^2 + w''^2 + \cdots$ for the sum of the squared errors, etc.). Gauss's starting point is the system

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n + l_1 = v_1, \\ \quad \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n + l_m = v_m. \end{cases}$$

He merely states that "it is easy to see" [*facile quidem perspicitur*] that in order for the total squared error $E = v_1^2 + \cdots + v_m^2$ to be a minimum, the following conditions must be satisfied:

$$\begin{cases} a_{11}v_1 + a_{21}v_2 + \cdots + a_{m1}v_m = 0, \\ \quad \cdots \\ a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{mn}v_m = 0. \end{cases}$$

Gauss does not use the words "method of least squares" in the section where he considers this system of equations, but earlier in the paper [8, p. 16] he indicates that he has long used "certain tricks" [*quaedam artificia practica*] that make the application of the method of least squares more convenient. At any rate, the last system is (except for a common factor of 2) identical to the system of normal equations $\partial E/\partial x_1 = 0, \ldots, \partial E/\partial x_n = 0$. Now Gauss introduces his characteristic notation

$$[a_1 l] = a_{11}l_1 + a_{21}l_2 + \cdots + a_{m1}l_m,$$
$$[a_i a_k] = a_{1i}a_{1k} + a_{2i}a_{2k} + \cdots + a_{mi}a_{mk},$$

and so forth. Hence the unknowns $x_1, \ldots, x_n$ need to be determined from the equations

$$\begin{cases} [a_1a_1]x_1 + \cdots + [a_1a_n]x_n + [a_1l] = 0, \\ \qquad\cdots \\ [a_na_1]x_1 + \cdots + [a_na_n]x_n + [a_nl] = 0. \end{cases} \tag{3}$$

Gauss now uses a procedure (but not matrix notation) that is essentially equivalent to what today is known as Gaussian elimination. He expresses E explicitly in terms of $x_1, \ldots, x_n$ and shows that $E - R_1^2/[a_1a_1]$ is independent of x_1. Here R_1 denotes the left-hand side of the first row of the system above. Next he eliminates x_2 from $E^{(1)} = E - R_1^2/[a_1a_1]$, and so forth. In this way he obtains a representation

$$E = \frac{(A_1(x_1, \ldots, x_n))^2}{a_1} + \frac{(A_2(x_2, \ldots, x_n))^2}{a_2} + \cdots + \frac{(A_n(x_n))^2}{a_n} + A,$$

where $a_1, \ldots, a_n$ are positive numbers. Here A represents the minimum value of E, and the unknowns are determined by back substitution from the equations

$$A_1(x_1, \cdots, x_n) = 0,$$
$$A_2(x_2, \cdots, x_n) = 0,$$
$$\cdots$$
$$A_n(x_n) = 0.$$

This is what we called Gauss's Theorem in Section 3 above.

5. What was Jordan's Method and Notation? On page 83 of the third edition of his *Handbuch der Vermessungskunde* [17], Jordan presents a numerical example, derived from a least squares application in geodesy, to illustrate the method that has come to be known as Gauss-Jordan reduction. The particular system he considers would now be written as

$$\begin{cases} 17.50x - 6.50y - 6.50z = 2.14, \\ -6.50x + 17.50y - 6.50z = 13.96, \\ -6.50x - 6.50y + 20.50z = -5.40, \\ -2.14x - 13.96y + 5.40z = w - 100.34, \end{cases} \tag{4}$$

where the w in the last line is from equation (2). However, since all systems of normal equations are symmetric, Jordan adopts an abbreviated representation:

$$
\begin{array}{rrrrl}
\underline{17.50x} - & 6.50y - & 6.50z - & 2.14 & = 0 \\
+ & \underline{17.50y} - & 6.50z - & 13.96 & = 0 \\
& + & \underline{20.50z} + & 5.40 & = 0 \\
& & + & \underline{100.34}, &
\end{array}
$$

where the number $l'l$ just "floats" at the lower right. The modern method for solving system (4) uses row operations as follows.

$$
\begin{bmatrix}
17.50 & -6.50 & -6.50 & 2.14 \\
-6.50 & 17.50 & -6.50 & 13.96 \\
-6.50 & -6.50 & 20.50 & -5.40 \\
-2.14 & -13.96 & 5.40 & -100.34
\end{bmatrix}
$$

$$
\begin{array}{l}
\\
R_2 + (6.5/17.5)R_1 \\
R_3 + (6.5/17.5)R_1 \\
R_4 + (2.14/17.5)R_1
\end{array}
\begin{bmatrix}
17.50 & -6.50 & -6.50 & 2.14 \\
0 & 15.09 & -8.91 & 14.75 \\
0 & -8.91 & 18.09 & -4.61 \\
0 & -14.75 & 4.61 & -100.08
\end{bmatrix}
$$

$$
\begin{array}{l}
R_1 + (6.5/15.09)R_2 \\
\\
R_3 + (8.91/15.09)R_2 \\
R_4 + (14.75/15.09)R_2
\end{array}
\begin{bmatrix}
17.50 & 0 & -10.34 & 8.49 \\
0 & 15.09 & -8.91 & 14.75 \\
0 & 0 & 12.83 & 4.10 \\
0 & 0 & -4.10 & -85.66
\end{bmatrix}
$$

$$
\begin{array}{l}
R_1 + (10.34/12.83)R_3 \\
R_2 + (8.91/12.83)R_3 \\
\\
R_4 + (4.10/12.83)R_3
\end{array}
\begin{bmatrix}
17.50 & 0 & 0 & 11.79 \\
0 & 15.09 & 0 & 17.60 \\
0 & 0 & 12.83 & 4.10 \\
0 & 0 & 0 & -84.35
\end{bmatrix}
$$

$$
\begin{gathered}
x = 11.79/17.50 = 0.67 \\
y = 17.60/15.09 = 1.17 \\
z = 4.10/12.83 = 0.32 \\
\text{Total squared error} = 84.35
\end{gathered}
$$

Jordan's presentation uses the same arithmetic in a different arrangement. In the first place, he takes advantage of the symmetry that remains as the reduction proceeds. Secondly, he uses different size type so that numbers to be added can be placed conveniently one above the other without confusion. Finally, to achieve this convenience, the entries above the diagonal are placed to the right of the matrix. Here then is Jordan's layout of his method. We have inserted letters next to several of the numbers so that they can be identified in the notes that follow. Of course these letters were not included by Jordan.

a	b	c	l			
$+17.50$	-6.50	-6.50	-2.14			
	$+17.50$	-6.50	-13.96			
	$-2.41a$	$-2.41a$	$-0.79a$			
		$+20.50$	$+5.40$			
		$-2.41a$	$-0.79a$			
			$+100.34$			
			$-0.26b$			
	$+15.09c$	$-8.91c$	$-14.75c$	$-6.50d$		
		$+18.09e$	$+4.61e$	$-+.50f$		
		$-5.26g$	$-8.71g$	$-3.84h$		
			$+100.08i$	$-2.14j$		
			$-14.42k$	$-6.35h$		
		$+12.83l$	$-4.10l$	$-10.34m$	$-8.91n$	
			$+85.66o$	$-8.49m$	$-14.75p$	
			$-1.31q$	$-3.30r$	$-2.85s$	
			$+84.35t$	$-11.79u$	$-17.60v$	$-4.10w$
			$=[vv]$	-17.50	-15.09	-12.83 neg. denominator
				$+0.67$	$+1.17$	$+0.32$
				$=x$	$=y$	$=z$

(a) $\frac{6.5}{17.5}R_1$

(b) $\frac{2.14}{17.5}R_1$

(c) $R_2 + \frac{6.5}{17.5}R_1$

(d) (1, 2)-entry of original system

(e) $R_3 + \frac{6.5}{17.5}R_1$

(f) (1, 3)-entry of original system

(g) $\frac{8.91}{15.09}R_2$

(h) $\frac{6.5}{15.09}R_2$

(i) $R_4 + \frac{2.14}{17.5}R_1$

(j) (1, 4)-entry of original system

(k) $\frac{14.75}{15.09}R_2$

(l) $R_3 + \frac{8.91}{15.09}R_2$

(m) $R_1 + \frac{6.5}{15.09}R_2$

(n) (2, 3)-entry after first reduction

(o) $R_4 + \frac{14.75}{15.09}R_2$

(p) (2, 4)-entry after first reduction

(q) $\frac{4.10}{12.83}R_3$

(r) $\frac{10.34}{12.83}R_3$

(s) $\frac{8.91}{12.83}R_3$

(t) $R_4 + \frac{4.10}{12.83}R_3$

(u) $R_1 + \frac{10.34}{12.83}R_3$

(v) $R_2 + \frac{8.91}{12.83}R_3$

(w) (3, 4)-entry after second reduction

To describe problems in general, Jordan uses and extends the notation of Gauss. He attempts, in the fashion of the times, to give recursive formulae without the help of an index or subscripts. Thus, the input data are labeled a, b, c, etc. and l. In [17, p. 77] he writes the general normal equations as:

$$\begin{aligned}\underline{[aa]x} + [ab]y + [ac]z + [al] &= 0,\\ \underline{[bb]y} + [bc]z + [bl] &= 0,\\ \underline{[cc]z} + [cl] &= 0,\\ \underline{[ll]}.\end{aligned}$$

Then, Gauss-Jordan reduction is given by the tableau [17, p. 82]

$$\begin{array}{ccccc|ccccc}
[aa]_3 & [ab]_0 & [ac]_0 & [ad]_0 & [al]_0 & & & & & \\
 & [bb] & [bc] & [bd] & [bl] & & & & & \\
 & & [cc] & [cd] & [cl] & & & & & \\
 & & & [dd] & [dl] & & & & & \\
 & & & & [ll] & & & & & \\
\hline
 & [bb.1]_3 & [bc.1]_1 & [bd.1]_1 & [bl.1]_1 & [ab]_0 & & & & \\
 & & [cc.1] & [cd.1] & [cl.1] & [ac]_0 & & & & \\
 & & & [dd.1] & [dl.1] & [ad]_0 & & & & \\
 & & & & [ll.1] & [al]_0 & & & & \\
\hline
 & & [cc.2]_3 & [cd.2]_2 & [cl.2]_2 & (ac.1) & [bc.1]_1 & & & \\
 & & & [dd.2] & [dl.2] & (ad.1) & [bd.1]_1 & & & \\
 & & & & [ll.2] & (al.1) & [bl.1]_1 & & & \\
\hline
 & & & [dd.3]_3 & [dl.3]_3 & (ad.2) & (bd.2) & [cd.2]_2 & & \\
 & & & & [ll.3] & (al.2) & (bl.2) & [cl.2]_2 & & \\
\hline
 & & & & [ll.4] & (al.3) & (bl.3) & (cl.3) & [dl.3]_3 & \text{Numer.} \\
 & & & & & [aa]_3 & [bb.1]_3 & [cc.2]_3 & [dd.3]_3 & \text{Denom.} \\
\cline{6-10}
 & & & & & -x & -y & -z & -t & \text{Quotient}
\end{array}$$

The entries in this tableau are determined by formulae that yield the usual Gauss-Jordan reduction. For example,

$$[bb.1] = [bb] - \frac{[ab]}{[aa]}[ab],$$

$$(ac.1) = [ac] - \frac{[bc.1]}{[bb.1]}[ab].$$

6. What about Clasen? In the quotation from Householder given above, we find the claim that the so-called Gauss-Jordan method seems to have been first described by Clasen in [7]. In [22], Thomas Muir also writes about this same article by Clasen:

> In the solution of a set of linear equations with arithmetical coefficients considerable latitude is available in the choice of the series of derived equations which is to end up with the value of the unknowns. The choice made by the writer ... is made with real skill and deserves attention.

Clasen's name appears as "Clasen (abbé B.-I.), curé-doyen d'Echternach (Grand-Duché de Luxembourg)" in the membership lists of The Société Scientifique de Bruxelles from 1887–1901 [2]. His name appears on the list of deceased members in 1902 [3] and it would seem from the historical volume [4, p. 62] that he died that year. He published only one article in the Annales de la Société Scientifique de Bruxelles [4, pp. 121 and 234]. In this section we take a closer look at that article.

First, it is worth noting that Clasen gives no references, and the name of Gauss (or Jordan) is never mentioned. The only other mathematician in the article is Paul Mansion, who apparently served as referee. Clasen thanks him in a brief footnote. In today's notation, Clasen's method applied to the system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

becomes

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ & & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{bmatrix} \to \begin{bmatrix} m & 0 & a_{13}^{(1)} & \cdots & b_1^{(1)} \\ 0 & m & a_{23}^{(1)} & \cdots & b_2^{(1)} \\ a_{31} & a_{32} & a_{33} & \cdots & b_3 \\ & & \cdots & & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & b_n \end{bmatrix} \to$$

$$\begin{bmatrix} m & 0 & a_{13}^{(1)} & a_{14}^{(1)} & \cdots & b_1^{(1)} \\ 0 & m & a_{23}^{(1)} & a_{24}^{(1)} & \cdots & b_2^{(1)} \\ 0 & 0 & R & a_{34}^{(1)} & \cdots & b_3^{(1)} \\ & & \cdots & & & \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} & \cdots & b_n \end{bmatrix} \to \begin{bmatrix} R & 0 & 0 & a_{14}^{(2)} & \cdots & b_1^{(2)} \\ 0 & R & 0 & a_{24}^{(2)} & \cdots & b_2^{(2)} \\ 0 & 0 & R & a_{34}^{(1)} & \cdots & b_3^{(1)} \\ & & & \cdots & & \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} & \cdots & b_n \end{bmatrix}, \text{etc.}$$

Here,

$$m = a_{11}a_{22} - a_{21}a_{12} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

and

$$R = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

The arrows indicate the obvious row operations.

Clasen points out repeatedly that his method does not involve division—a definite sign of precalculator days. He also inserts extra equations as checks.

Clasen himself does not use matrix notation, rather he writes the successively occuring equations in a sequence that seems rather laborious. For example, if he were to solve a 3×3 system, he would proceed as follows.

$$\begin{array}{l} X_1: a_{11}x + a_{12}y + a_{13}z + b_1 = 0 \\ Y_1: a_{21}x + a_{22}y + a_{23}z + b_2 = 0 \\ \underline{Z_1: a_{31}x + a_{32}y + a_{33}z + b_3 = 0} \\ X_1: a_{11}x + a_{12}y \quad + a_{13}z + b_1 = 0 \end{array}$$

$$\begin{array}{rllr} \hline a_{11}Y_1: & a_{11}a_{21}x + a_{11}a_{22}y & +a_{11}a_{23}z & +a_{11}b_2 = 0 \\ -a_{21}X_1: & -a_{21}a_{11}x - a_{21}a_{12}y & -a_{21}a_{13}z & -a_{21}b_1 = 0 \\ Y_2: & \multicolumn{3}{r}{my + (a_{11}a_{23} - a_{21}a_{13})z + a_{11}b_2 - a_{21}b_1 = 0} \\ \hline \end{array}$$

$$\begin{array}{rllr} mX_1: & ma_{11}x + ma_{12}y & +ma_{13}z & +mb_1 = 0 \\ -a_{12}Y_2: & \quad -a_{12}my - a_{12}(a_{11}a_{23} - & a_{21}a_{13})z & -a_{12}(a_{11}b_2 - a_{21}b_1) = 0 \\ & ma_{11}x \quad +[ma_{13} - a_{12}a_{11}a_{23} + a_{12} & a_{21}a_{13}]z & +mb_1 - a_{12}(a_{11}b_2 - a_{21}b_1) = 0 \\ & ma_{11}x & +a_{11}a'_{13}z & +a_{11}b'_1 = 0 \\ X_2: & mx & +a'_{13}z & +b'_1 = 0 \\ \hline \end{array}$$

$$\begin{array}{rllll} mZ_1: & ma_{31}x & +ma_{32}y & +ma_{33}z + & mb_3 = 0 \\ -a_{32}Y_2: & & -a_{32}my & -a_{32}a'_{23}z & -a_{32}b'_2 = 0 \\ -a_{31}X_2: & -a_{31}mx & & -a_{31}a'_{13}z & -a_{31}b'_1 = 0 \\ Z_3: & & & Rz & +b'_3 = 0 \\ \hline \end{array}$$

$$\begin{array}{rlll} RY_2: & Rmy & +Ra'_{23}z & +Rb'_2 = 0 \\ -a'_{23}Z_3: & & -a'_{23}Rz & -a'_{23}b'_3 = 0 \\ Y_3: & Rmy & & +b''_2 = 0 \\ \hline \end{array}$$

$$\begin{array}{rlll} RX_2: & Rmx & +Ra'_{13}z + & Rb'_1 = 0 \\ -a'_{13}Z_3: & & -a'_{13}Rz - & a'_{13}b'_3 = 0 \\ X_3: & Rmx & & +b''_1 = 0 \end{array}$$

At the end of the article, Clasen emphasizes that his method requires fewer calculations than a method using determinants, and he shows how his method can be used to compute determinants.

The same volume of the Annales de la Société Scientifique de Bruxelles that contains Clasen's article also has a review by P. Mansion [20]. Mansion (1844–1919) was a well-known Belgian mathematician and, from 1865, professor at the University of Ghent [1], [27]. Mansion calls Clasen's method "the method of equal coefficients," and he gives a fairly detailed description of Clasen's work. Mansion also gives no reference to Gauss or anyone else.

Later, in 1930, R. Mehmke [21] describes the "traditional and still customary" method of solving systems of linear equations—it is clearly Gaussian elimination, although the name Gauss is never mentioned. Mehmke does observe that back substitution can be troublesome, especially in large systems. He states that Clasen's method is superior to Gaussian elimination, even though the world has completely ignored the former.

The purpose of Mehmke's article is to remind the readers of Clasen's work. But Mehmke introduces a more formalized notation (e.g. arrays of numbers where each variable has a column) that does not appear in Clasen's work. To aid with the computations, Mehmke even uses paper strips! Finally, Mehmke "improves" Clasen's method by adapting it to accelerated elimination, where several variables are eliminated at each step. Mehmke mentions neither Gauss nor Jordan.

7. What about other methods of solving linear systems? In 1947, E. Bodewig [6, p. 931] writes concerning large systems of linear equations that none of the available methods will remove the inherent difficulties and that calculating machines need to be used. He continues, "The newly famous electronic calculating machines solve such systems automatically, so that the method used does not matter. However, the price of such machines is so high that only one specimen of each of the two types exists (or will ever exist). Everyone else in the world must be satisfied with the usual calculating machines for which the method does matter."

The lack of adequate calculators made every little shortcut important. Even though the solution of linear systems was understood in principle, the process was unpleasant. Different methods for dealing with this unpleasantness were popular at various times. At least two survey articles ([6] and [14]) list by name many long forgotten methods of solving linear systems.

REFERENCES

1. L'Académie Royale des Sciences, des Lettres et des Beaux-arts de Belgique, Biographie National Bruylant, Bruxelles, 30 (1959) 540–542.
2. Annales de la Société Scientifique de Bruxelles, 11 (1887) 16, 35; 25 (1901) 19.
3. Annales de la Société Scientifique de Bruxelles, 26 (1902) 46.

4. Annales de la Société Scientifique de Bruxelles, Table Analytique des vingt-cinq premiers volumes 1875–1901, procédée de l'Histoire documentairé de la Société Scientifique et de la Liste générale de ses membres, pp. 62, 121, and 234.
5. D. Bartlett, General Principles of the Method of Least Squares, Boston, 1915.
6. E. Bodewig, Bericht über die verschiedenen Methoden zur Lösung eines Systems linearer Gleichungen mit reellen Koeffizienten, Koninkl. Akad. Wetensch., 50 (1947–48) 930–941, 1104–1116, 1285–1295, and 51 (1948) 53–64, 211–219.
7. B.-I. Clasen, Sur une nouvelle méthode de résolution des équations linéaires et sur l'application de cette méthode au calcul des déterminants, Ann. Soc. Sci. Bruxelles, 12 (1888) 251–281; also Mathesis (Suppl. 2) 9 (1888) 1–31.
8. C. F. Gauss, Werke, Sechster Band, Disquisitio de Elementis ellipticis Palladis ex Oppositionibus Annorum 1803, 1804, 1805, 1807, 1808, 1809, Königliche Gesellschaft der Wissenschaften, Göttingen (1874) 3–24.
9. C. F. Gauss, Abhandlungen zur Methode der Kleinsten Quadrate, Berlin, 1887.
10. A. Gewirtz, H. Sitomer, and A. W. Tucker, Constructive Linear Algebra, Prentice-Hall, Englewood Cliffs, 1974.
11. Harbert, Ruhestätte für Wilhelm Jordan, Zeitschrift für Vermessungswesen, 85 (1960) 138.
12. Helmert, Wilhelm Jordan, Zeitschrift für Vermessungswesen, 28 (1899) 321–328. (The official announcement of Jordan's death (April 17, 1899) is on page 265 of this volume.)
13. A. S. Householder, The Theory of Matrices in Numerical Analysis, Blaisdell, New York, 1953.
14. H. Jensen, An Attempt at a Systematic Classification of the Methods for the Solution of Normal Equations, Geodaetisk Institut, Meddelelse, No. 18 (1944).
15. W. W. Johnson, On Gauss's Method of Substitution, Annals of Mathematics, 7 (1893) 65–69.
16. H. Jordan, Wilhelm Jordan, Lebensbild eines deutschen Geodäten, Zeitschrift für Vermessungswesen, 51 (1922) 577–585.
17. W. Jordan, Handbuch der Vermessungskunde, Erster Band, Dritte verbesserte Auflage, Metzler, Stuttgart, 1888.
18. K. Kunz, Numerical Analysis, McGraw-Hill, New York, 1957.
19. O. M. Leland, Practical Least Squares, McGraw-Hill, New York, 1921.
20. M. P. Mansion, Sur le mémoire de M. l'abbé B.-I. Clasen, Ann. Soc. Sci. Bruxelles, 12 (1888) 50–59.
21. R. Mehmke, Über die zweckmässigste Art, lineare Gleichungen durch Elimination aufzulösen, Z. Angew. Math. Mech., 10 (1930) 508–514.
22. T. Muir, The Theory of Determinants in the Historical Order of Development, Macmillan, London, 1906–1923; Dover reprint, New York, 1960.
23. C. Müller, Bücherschau, Zeitschrift für Vermessungswesen, 51 (1922) 143–144.
24. A. Nüsse, Wilhelm Jordan. Eine Erinnerung, Allgemeine Vermessungsnachrichten, 54 (1942) 66–70.
25. J. C. Poggendorff, Biographisch-Literarisches Handwörterbuch der exakten Naturwissenschaften, Akademie-Verlag, Berlin 7a (1969) 305.
26. C. Runge, Wilhelm Jordan, Zeitschrift für Architektur und Ingenieurwesen 45, (1899) 265–66.
27. E. de Seyn, Dictionnaire biographique des Sciences, des Lettres et des Arts en Belgique, volume 2, L'Avenir, Bruxelles, 1939, p. 713.
28. A. W. Tucker, Private communication.