Determining the Determinant

Daniel E. Otero*

June 12, 2018

Linearity and linear systems of equations are ubiquitous in the mathematical analysis of a wide range of problems, and these methods are central problem solving tools across many areas of pure and applied mathematics. This has given rise, since the late twentieth century, to the custom of including courses in linear algebra and the theory of matrices in the education of students of mathematics and science around the world. A standard topic in such courses is the study of the determinant of a square matrix, a mathematical object whose importance emerged slowly over the course of the nineteenth century (but can be traced back into the eighteenth century, and, even earlier, to China and Japan, although these developments from the Orient were probably entirely independent and unknown to European scholars until very recently). The concept of the determinant coalesced out of a number of mathematical problems. It took shape within the development of linear algebra, which in the 1800s evolved as a branch of a more general algebraic theory of equations and their solutions, one which was already a mainstream program of scholarship with a long history. Within the development of linear algebra, the creation of a theory of matrices came to play a larger and larger role into the 1900s, especially as the advent of electronic computation allowed scientists and engineers to apply methods from these theories to the solution of many important practical problems, especially in analysis.

It is the goal of this project to lead students to an understanding of the concept of the determinant by allowing them to peer over the shoulders of two important contributors to the development of this central idea, to see how they thought and reasoned as they developed its fundamental properties.

1 Cramer's Rule

The modern concept of the determinant emerged from efforts to solve systems of linear algebraic equations. Algebra had become a standard tool for geometry and analysis over the centuries, and the introduction of symbolism in Europe in the 1500s allowed the manipulation of arithmetical operations to be carried out much more simply than had the rhetorical, language-based methods of earlier years. Girolamo Cardano (1501-1576) had worked out a general formula for solving cubic equations; François Viète (1540-1603) championed the use of letters of the alphabet to stand for variables (unknowns) and parameters (known but unspecified quantities) to accompany newly adopted symbols for the arithmetical operations and relations; René Descartes (1596-1650) illustrated the power of uniting algebraic methods with geometrical analysis in his highly influential treatise *La Géométrie* of 1637; and Gottfried Leibniz (1646-1716) introduced the use of indexing symbols with numbers to assist the algebraist in working with numerous equations simultaneously. These represent just a handful of

*Department of Mathematics, Xavier University, Cincinnati, OH, 45207-4441; otero@xavier.edu.

the more important contributors to this field, who laid the groundwork and provided context for the efforts of later mathematicians to tackle the problems that led to the evolution of the determinant.

By the middle of the eighteenth century, a number of mathematicians were writing manuals of algebraic methods. One of these was Gabriel Cramer (1704-1752), who published (in French) his *Introduction a l'Analyse des lignes courbes algébriques* (*Introduction to the Analysis of Algebraic Curved Lines*) [Cramer, 1750] in 1750. Cramer enjoyed a career as a mathematician in his hometown of Geneva, Switzerland, and is best known as a protégé of the eminent Johann Bernoulli (1667-1748), whose completed works he edited. Johann also arranged for Cramer to serve as editor of the completed works of his equally famous mathematician brother, Jakob Bernoulli (1654-1705). Cramer's own career was established with the publication of this *Introduction*, a work that compiled what was known of the algebraic geometry of his day.

In the course of discussing how algebraic curves intersect, Cramer gave a new algebraic proof of a theorem that dates back to the Greek geometer Pappus of Alexandria in the fourth century. It asserts that, given five points in the plane, one can find a unique conic section (a curve described by a polynomial equation in the two variables x and y of degree at most two in each variable) that passes through all five points. To that end, Cramer knew that the general equation of a conic section curve has the form

$$A + By + Cx + Dyy + Exy + xx = 0$$

where the particular parameters A, B, C, D, E identify the specific curve, whether an ellipse (or circle), parabola, or hyperbola. By inserting the values of the coordinates of these five points, $(x_1, y_1), (x_2, y_2), (x_3, y_3),\ (x_4, y_4), (x_5, y_5)$ into the equation of the conic section,[1] he arrived at a system of five linear equations, now in the five unknowns A, B, C, D, E:

$$\begin{aligned}
A + By_1 + Cx_1 + Dy_1^2 + Ex_1y_1 + x_1^2 &= 0\\
A + By_2 + Cx_2 + Dy_2^2 + Ex_2y_2 + x_2^2 &= 0\\
A + By_3 + Cx_3 + Dy_3^2 + Ex_3y_3 + x_3^2 &= 0\\
A + By_4 + Cx_4 + Dy_4^2 + Ex_4y_4 + x_4^2 &= 0\\
A + By_5 + Cx_5 + Dy_5^2 + Ex_5y_5 + x_5^2 &= 0
\end{aligned} \tag{1}$$

Solving for the new unknowns will produce the conic that passes through the five given points – provided there is a unique solution to the system [Cramer, 1750, pp. 57-60]. As he stated it,

> We are able by means of these five equations to find the values of the five coefficients A, B, C, D, E, which determines the eqn: A+By+Cx+Dyy+Exy + xx = 0 of the sought Curve.
>
> The Calculation of it truly will be quite long: One will find it in Appendix N° 1.

In other words, having worked out the details of solving such a system, and noting its complexity, Cramer recognized that he should bury this part of his analysis in Appendices to his *Introduction*

[1] In fact, Cramer used slightly different notation for the coordinates of the five points, but this is of little matter to us here.

so as not to disturb the continuity of the discussion of the problem of the five points. On the other hand, he found that his solution to the linear system is generalizable, allowing him to deal with a different problem altogether, that of obtaining formulas for the solution to any system whatsoever of n linear equations in n unknowns. So he referenced his readers to the Appendices for details, and he simply presented the resulting formulas for the values of A, B, C, D, E in the problem of the five points on a conic in the subsequent text. In the process he found a precise criterion on the coefficients of the system of equations (1) that ensured a *unique* set of values for A, B, C, D, E, that is, a unique conic curve that solved the problem of the five points.

Let us now follow [Cramer, 1750, Appendices I and II, pp. 656-659] to see how Cramer did this. As you read this excerpt, keep in mind that Cramer has relegated to this Appendix his problems with solving the linear system (1); thus, in this discussion, he makes no mention of the problem of the five points which led him here in the first place. As we are far more interested in how he dealt with the system of linear equations than in the problem of the five points, we, like Cramer, will say no more about the geometric problem in what follows.

When a Problem contains many unknowns, of which the relations are so complicated that we find ourselves obligated to form many equations; then, in order to discover the values of these unknowns, we make them all vanish, except one, which combined alone with the known magnitudes gives, if the Problem is determined, a *final Equation*, of which the solution reveals first this unknown, & then by means of it, all the others.

Algebra furnishes for this some Rules, whose success is infallible, provided that we have the patience to follow them. But the Calculation of it becomes extremely long when the number of equations & of unknowns is very great...

Let there be some unknowns† z, y, x, v, &c. & as many equations

$$\begin{array}{rcl} A^1 & = & Z^1z + Y^1y + X^1x + V^1v + \&\text{ c.} \\ A^2 & = & Z^2z + Y^2y + X^2x + V^2v + \&\text{ c.} \\ A^3 & = & Z^3z + Y^3y + X^3x + V^3v + \&\text{ c.} \\ A^4 & = & Z^4z + Y^4y + X^4x + V^4v + \&\text{ c.} \\ & & \&\text{c.} \end{array} \tag{2}$$

where the letters A^1, A^2, A^3, A^4, &c. do not mark, as usual, the powers of A, but rather the first member, assumed to be known, of the first, second, third, fourth &c. equation.‡ Likewise Z^1, Z^2, &c. are the coefficients of z; Y^1, Y^2, &c. those of y; X^1, X^2, &c. those of x; V^1, V^2, &c. those of v; &c. in the first, second, &c. equation.

†Commentator's Note: In case you are wondering why w does not follow x in this list, the letter W did not enter the French alphabet until the middle of the 1800s, some decades after Cramer wrote his book! Even today, it only appears in words borrowed into French from other languages like English and German.

‡Commentator's Note: French writers of Cramer's generation typically referred to the "first member" and "second member" of an equation to mean what we call, respectively, the "left hand" and "right hand" expressions on either side of the equal sign.

This Notation being established, if there is only one equation & only one unknown z, we will have $z = \frac{A^1}{Z^1}$. If there are two equations & two unknowns z & y; we will find $z = \frac{A^1Y^2 - A^2Y^1}{Z^1Y^2 - Z^2Y^1}$, & $y = \frac{Z^1A^2 - Z^2A^1}{Z^1Y^2 - Z^2Y^1}$.

Task 1

(a) Write down what a system of five linear equations in five unknowns would look like in Cramer's notation, as in (2), using letters indexed with superscripts.

(b) Now compare the system in Cramer's notation with that of the system of equations (1) that Cramer would have considered to solve the problem of the five points. Draw up a two-column table that matches the quantities $A^k, Z^k, Y^k, \ldots, z, y, x, \ldots$ with those in system (1) so that variables are associated to variables and parameters to parameters.

(c) You are probably aware that it is standard today to index variables and parameters with subscripts instead of superscripts, but it was not yet so in Cramer's day. Do you find his choice of superscripts to index letters in (2) helpful? Why or why not?

Task 2

Write down what a system of one equation in one unknown would look like in Cramer's notation, as in (2). Solve the equation and compare with the penultimate sentence in the excerpt above. Under what condition would this solution not be possible? Explain.

Task 3

(a) Write down what a system of two linear equations in two unknowns would look like in Cramer's notation. Solve the pair of equations by multiplying each equation by a quantity that allows for the elimination of the variable y when the equations are subtracted one from the other; this reduces the two equations to a single equation in z. Solve for z in terms of the coefficients of the system. Now use a similar procedure (multiplying each of the original equations by a quantity that allows for the elimination of the variable z when the equations are subtracted) to solve for y. Compare your formulas with those of Cramer.

(b) Use Cramer's Rule to solve the system

$$\begin{aligned} 4 &= 7z + 10y \\ 3 &= 5z + 7y \end{aligned}$$

(c) Discuss what happens when you use Cramer's Rule to solve the system

$$\begin{aligned} 10 &= z - 3y \\ -1 &= 4z - 12y \end{aligned}$$

(d) Try to formulate a criterion in terms of the coefficients of a linear system of two equations in two unknowns for when it is possible to solve the system and when it is not possible. Explain what you find in as much detail as you can.

(e) Graph the two equations in (b) in a (z, y)-coordinate plane. Relate the features of your graph to your answer to part (b).

(f) Repeat part (e) using the equations in (c), and note what you observe.

Task 4 (a) Compare the expressions in the numerators and denominators of Cramer's two formulas

$$z = \frac{A^1Y^2 - A^2Y^1}{Z^1Y^2 - Z^2Y^1}, \qquad y = \frac{Z^1A^2 - Z^2A^1}{Z^1Y^2 - Z^2Y^1},$$

that solve a system of two linear equations in two unknowns. What do you notice?

(b) Refer back to the system of two linear equations that gave rise to these formulas. What would happen if we swapped every occurrence of the symbols Z and z in that system with the symbols Y and y, respectively, without affecting the superscripts? Would this alter the system of equations? More to the point, would this change its solution set, i.e., the collection of (z, y) values that satisfy the system?

(c) Now, in Cramer's formula for z swap every occurrence of the symbols Z and z with the symbols Y and y, respectively, without affecting the superscripts. Compare your result with his formula for y. What do you notice? What happens when you perform this same symbol swap on Cramer's formula for y?

(d) Explain how the phenomenon you noticed in (c) above *guarantees* that swapping every occurrence of the symbols Z and z with the symbols Y and y provides a way to derive Cramer's formula for y without having to repeat the algebra steps we used to derive the formula for z as you did in Task 3(a).

In the Appendix of his *Introduction*, Cramer immediately continued with the case of three linear equations in three unknowns.

If there are three equations & three unknowns z, y, & x; we will find

$$z = \frac{A^1Y^2X^3 - A^1Y^3X^2 - A^2Y^1X^3 + A^2Y^3X^1 + A^3Y^1X^2 - A^3Y^2X^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1}$$

$$y = \frac{Z^1A^2X^3 - Z^1A^3X^2 - Z^2A^1X^3 + Z^2A^3X^1 + Z^3A^1X^2 - Z^3A^2X^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1}$$

$$x = \frac{Z^1Y^2A^3 - Z^1Y^3A^2 - Z^2Y^1A^3 + Z^2Y^3A^1 + Z^3Y^1A^2 - Z^3Y^2A^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1}$$

The formulas appear much more daunting here – especially when you see them laid out in the original 1750 typeface! Cramer is not forthcoming with an explanation for *how* he obtained these three formulas (it was not the style – then or now – to be so explicit about such details in formal academic publications), but certainly he realized that this system can be handled in much the same way as the two-unknown system:

Task 5

(a) Write down what a system of three linear equations in three unknowns would look like in Cramer's notation (2).

(b) Use the same algebra 'trick' from Task 3(a) to eliminate the variable x from the *first two* equations of your system in (a); you should then obtain a single equation in z and y having two-term expressions as coefficients. Do the same thing to eliminate the variable x using the *first and third* equations of the original system, and you will get another equation in z and y. You will now have a system of two linear equations in the variables z and y that has the form

$$\begin{aligned} B^1 &= S^1 z + T^1 y \\ B^2 &= S^2 z + T^2 y \end{aligned} \tag{3}$$

where the coefficients $B^1, B^2, S^1, S^2, T^1, T^2$ are each two-term expressions involving the coefficients of the original system A^k, Z^k, Y^k, X^k $(k = 1, 2, 3)$. Note that it has the same form as the two-unknown system you examined in Task 3(a).

(c) Now employ the same method you used in Task 3(a) to eliminate y from the reduced system (3). Before you perform any division to solve the resulting equation for z, substitute for each of $B^1, B^2, S^1, S^2, T^1, T^2$ the two-term expressions in A^k, Z^k, Y^k, X^k $(k = 1, 2, 3)$, which each represents, and expand the multiplications therein. The resulting many-termed equation will allow for some cancellation of terms. You should then be able to divide through the entire equation by X^1. Finally, solving for z should give you Cramer's first formula:

$$z = \frac{A^1Y^2X^3 - A^1Y^3X^2 - A^2Y^1X^3 + A^2Y^3X^1 + A^3Y^1X^2 - A^3Y^2X^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1}. \tag{4}$$

Show all this (admittedly nasty) algebraic work.

Task 6

(a) Recall the result of Task 4(d); we want to import that successful idea into the case of three linear equations in three unknowns (Task 5). What would happen in this three-variable system of equations if we swapped every occurrence of the symbols Z and z with the symbols Y and y, respectively? Does this mean that performing the same swap in the formula (4) will produce a valid formula for y?

(b) Write down the formula for y that comes from the symbol swap described in (a) and compare it with Cramer's formula for y from the Appendix of his *Introduction*.

(c) How might you produce a formula that solves for x in this system? Carry out the corresponding steps, and explicitly describe what you have done. Compare your formula for x with Cramer's.

Rather than take the next obvious step – produce formulas that solve a system of four linear equations in four variables, Cramer leapt directly to the general case:

The examination of these Formulas provides the following general Rule. The number of equations & of unknowns being n, we will find the value of each unknown by forming n fractions of which the common denominator has as many terms as there are diverse arrangements of n different things. Each term is composed of the letters $ZYXV$ &c., always written in the same order, but to which we distribute, as exponents, the first n numbers arranged in all possible ways. Thus, when we have three unknowns, the denominator has $1 \times 2 \times 3 = 6$ terms, composed of the three letters ZYX, which receive successively the exponents 123, 132, 213, 231, 312, 321. We give to these terms $+$ and $-$ signs, according to the following Rule. When an exponent is followed in the same term, mediately or immediately, by an exponent smaller than it, I will call that a *derangement*. Let us count, for each term, the number of derangements: if it is even or zero, the term will have the $+$ sign; if it is odd, the term will have the $-$ sign. For example in the term $Z^1Y^2X^3$ there is no derangement: this term will have therefore the $+$ sign. The term $Z^3Y^1X^2$ has also the $+$ sign, since it has two derangements, 3 before 1 & 3 before 2. But the term $Z^3Y^2X^1$, which has three derangements, 3 before 2, 3 before 1, & 2 before 1, will have the $-$ sign.

The common denominator being thus formed, we will have the value of z by giving to this denominator the numerator which is formed by changing, in all its terms, Z into A. And the value of y is the fraction which has the same denominator & for numerator the quantity which results when we change Y into A, in all the terms of the denominator. And we find in a similar manner the value of the other unknowns.

Cramer's explanation above should allow you, without too much more trouble, to build formulas of your own for z, y, x, v coming from the general system of four linear equations in four unknowns:

$$\begin{aligned} A^1 &= Z^1z + Y^1y + X^1x + V^1v \\ A^2 &= Z^2z + Y^2y + X^2x + V^2v \\ A^3 &= Z^3z + Y^3y + X^3x + V^3v \\ A^4 &= Z^4z + Y^4y + X^4x + V^4v \end{aligned} \tag{5}$$

The following Tasks will organize this work for you.

Task 7 (a) Write down the expression that is "the common denominator" of the formulas that "find the value of each of the unknowns" in the system of three linear equations in three unknowns. It is found in each of the formulas in the source text on p. 5, and is reproduced in formula (4).

(b) Each term in this expression is "composed of the letters ZYX, which receive successive exponents 123, 132, 213, 231, 312, 321." For each of the terms, beginning with $Z^1Y^2X^3$, count the number of derangements. Cramer has already worked out three of these counts for you in the text.

(c) According to Cramer, how does the number of derangements in each term identify the sign ($+$ or $-$) it receives in the "common denominator" expression? Check that this holds for each of the six terms.

Task 8

(a) To handle the system (5) of four equations, Cramer told us that "we will find the value of each unknown by forming n fractions of which the common denominator has as many terms as there are diverse arrangements of n different things." In this case, $n = 4$, so how many terms will the "common denominator" expression in each of the four formulas contain?

(b) "Each term is composed of the letters $ZYXV$," according to Cramer, "to which we distribute, as exponents, the first n numbers arranged in all possible ways." (Again, it is important to note that these exponents are NOT powers, but superscript labels that identify the equation from which the coefficient comes.) Make a list of all these terms (the first of which will be $Z^1Y^2X^3V^4$). Then, by counting derangements in the sequence of exponents that appears in each term, decide which sign ($+$ or $-$) is to be assigned to each. Combine all these terms to produce the "common denominator" expression.

(c) Finally, follow Cramer's instructions in the final paragraph of the above source text to determine the numerators, and thus the fractional formulas for each of z, y, x, and v, thereby solving the general system (5).

2 Cauchy's Determinant: Alternating Symmetric Functions

Augustin-Louis Cauchy was born in Paris on August 21, 1789, just a few weeks after the fall of the Bastille and the turbulent start to the French Revolution. He was named after both the month of his birth and his father, Louis-François Cauchy, an official of the royal government whose collapse was taking place in those very days. The Cauchys were devoted Catholics and a respected bourgeois family, which dictated the circumstances in which the young Augustin-Louis found his place in society. For a time in 1794, during the Reign of Terror, the family escaped central Paris to their country home in Arcueil (today an "inner" Paris suburb), where they struggled to make do: Augustin-Louis contracted smallpox, surviving only after a long convalescence; thereafter, he remained a "timid, frail boy, withdrawn and pensive" [Belhoste, 1991, p. 4].

Later, in 1800, Cauchy's father was made Archivist and Keeper of the Seal of the French Senate, working directly for the Chancellor of the Senate, who was also an accomplished mathematician and member of the Paris Academy of Science, Pierre-Simon de Laplace. Louis-François made provisions for the education of his children (Augustin-Louis and his two younger brothers), and he presented his eldest boy both to Laplace and to Joseph Louis Lagrange, who was also a French Senator and member of the Academy. In the father's office one day, Lagrange was said to have told a gathering of gentlemen: "Now you see that little fellow there, don't you? Well, one day he will replace all of us simple geometers."

After training in some of the best schools in Paris, where he excelled, Augustin-Louis chose to prepare for a career as a civil engineer. He entered the École des Points et Chaussees (School of Bridges and Roads), and completed that course of study in 1810, by which time Napoléon Bonaparte had declared himself Emperor of France. Cauchy was appointed junior engineer and was sent to Cherbourg on the English Channel where he was to assist in the building up of the naval harbor there, which became the largest construction project in the Empire at the time. While there, his interests and talents in mathematics came to dominate his energies; he brought with him to Cherbourg copies of Laplace's *Celestial Mechanics* and Lagrange's *Theory of Analytic Functions*; in 1811 and early 1812 he presented two papers to the Paris Academy of Science on the geometry of polyhedra, which were well received. This success appeared to turn his primary focus to work in mathematics. The next summer, however, he fell ill, supposedly from exhaustion. He returned to Paris, where while recuperating he wrote a third and longer memoir, this time on the topic of symmetric functions. In what follows, we will study selections from this substantial and comprehensive work, a complex and mature piece of mathematical exposition, written by Cauchy... at age 23!

This memoir, with the elaborate title *On functions which take only two values, equal but of opposite sign, by means of transpositions performed among the variables which are contained therein* [Cauchy, 1815], was written in two parts. In the first part, Cauchy investigates a certain kind of multivariable function which he calls an *alternating symmetric function*, and which he defines as one "which may change in sign but not in value, by virtue of ... transpositions"; by a *transposition*, Cauchy understands an operation that swaps all occurrences of any one variable with those of another in a symbolic expression of the function.[2] Thus, as he tells it,

Here I will partition these functions into several orders, according to the number of the quantities that they contain, and I will designate always the n variables which a symmetric function of order n contains by some letters affected with indices, such as

$$a_1,\ a_2,\ \dots,\ a_n.$$

... Thus, ...

$$a_1b_2 + a_2b_3 + a_3b_1 - a_2b_1 - a_1b_3 - a_3b_2,$$

$$\sin(a_1 - a_2)\sin(a_1 - a_3)\sin(a_2 - a_3)$$

will be alternating symmetric functions of the third order.

[2] This is precisely the type of operation we considered in Tasks 4 and 6 above.

Task 9 Let's label Cauchy's two example functions f and g; that is,

$$\begin{aligned} f(a_1,a_2,a_3;b_1,b_2,b_3) &= a_1b_2+a_2b_3+a_3b_1-a_2b_1-a_1b_3-a_3b_2, \\ g(a_1,a_2,a_3) &= \sin(a_1-a_2)\sin(a_1-a_3)\sin(a_2-a_3). \end{aligned}$$

(a) Verify that both functions satisfy the definition he gives for alternating symmetric function. To begin, compute the transposed expression $f(a_2,a_1,a_3;b_2,b_1,b_3)$, in which the quantities with subscripts 1 and 2 are swapped, and relate it to the expression for $f(a_1,a_2,a_3;b_1,b_2,b_3)$. Since there are three different ways to transpose any pair of the quantities a and b, you must repeat this type of computation twice more for f to complete the verification.

(b) Then do likewise for the function g.

Task 10 (a) Give an example of an alternating symmetric function $p(a_1,a_2)$ of the second order, and verify that it has the necessary properties to be so identified.

(b) Give an example of an alternating symmetric function $q(a_1,a_2,a_3)$ of the third order different from the ones given by Cauchy above, and verify that it also has the necessary properties.

The rest of the first part of Cauchy's treatise is devoted to a study of the behavior of symmetric[3] and alternating symmetric functions, especially when those functions are polynomial functions of many variables. While we will need results that Cauchy deduced from this analysis later in this discussion, we will not focus here on the details, since we will be more interested in the material of the second part of the treatise. Indeed, Cauchy ended the first part of his treatise in this way:

I am now going to examine in particular a certain kind of alternating symmetric function which present themselves in a great number of analytic investigations. It is by means of these functions that we express the general values of unknowns that many equations of the first degree contain. ... Mr. Gauss himself is advantageously served by them in his *Recherches analytiques** ..., and he has labeled these same functions with the name *determinants*. I will preserve this terminology, which furnishes an easy method for announcing the results ...

Task 11 What do you think Cauchy meant here by "the general values of unknowns that many equations of the first degree contain"?

[3] Cauchy defines a function as *symmetric* if any transpositions of its variables does not change the value of the function at all; for instance, he notes that the function $F(a_1,a_2)=a_1^2+a_2^2+4a_1a_2$ is a symmetric function.

*Carl Friedrich Gauss, *Disquisitiones Arithmeticae*, Leipzig, 1801.

Cauchy referenced the influence of the work of Carl Friedrich Gauss (1777-1855) on his thinking on the subject of alternating symmetric functions. Gauss, recognized today as one of the greatest mathematicians of all time, published his *Disquisitiones arithmeticae* (whose Latin title Cauchy translated into French as *Recherches analytiques*) just a decade before Cauchy wrote this treatise; Gauss's work contained a wealth of ideas whose elaborations in the theory of numbers led to entirely new fields of mathematics and busied generations of mathematicians for centuries to come. Significantly for us, however, Gauss gave to Cauchy the inspiration to use the term *determinant* for the first time to refer to a very special kind of alternating symmetric function:

[Consider the] alternating symmetric function which ... will be represented by

$$\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}\dots a_{n,n}),$$

the sign $\mathbf{S}$ being relative to the first indices of each letter. Such is the most general form of the functions which I will denote in the following under the name *determinants*. If we suppose successively

$$n = 1, \qquad n = 2, \qquad \dots,$$

we will find

$$\begin{array}{rcl} \mathbf{S}(\pm a_{1,1}a_{2,2}) & = & a_{1,1}a_{2,2} - a_{2,1}a_{1,2}, \\ \mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}) & = & a_{1,1}a_{2,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3} \\ & & -a_{1,1}a_{3,2}a_{2,3} - a_{3,1}a_{2,2}a_{1,3} - a_{2,1}a_{1,2}a_{3,2} \\ \dots\dots\dots & & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \end{array}$$

for the determinants of the second [order], of the third order, etc.

In the excerpt above, Cauchy introduced the symbol $\mathbf{S}(\pm\ \cdot\ )$ as an *antisymmetrizing operator*. That is, the symbol is meant to operate as a function by alternating the signs of expressions obtained from the one within the parentheses, which he called the *indicative term* of the function. This produces a new resulting expression consisting of the sum of all possible expressions obtained from the indicative term by permuting the set of first indices of the variables it contains, with terms appearing with a $+$ sign when that permutation involves an even number of transpositions of indices and with a $-$ sign when it involves an odd number of transpositions.

We see this most easily with the first of the examples he provided: the expression $\mathbf{S}(\pm a_{1,1}a_{2,2})$ is formed by performing every possible permutation to the set of first indices of the variables appearing in the indicative term $a_{1,1}a_{2,2}$. Luckily, there are only two possible permutations that can be performed on the set of indices $\{1, 2\}$, the trivial permutation that does nothing to the indices, and the single transposition that swaps the indices. Consequently, there are just two terms in $\mathbf{S}(\pm a_{1,1}a_{2,2})$, the undisturbed indicative term $a_{1,1}a_{2,2}$ and the term obtained by transposing the pair of first indices: $a_{2,1}a_{1,2}$; the first of these terms appears in $\mathbf{S}(\pm a_{1,1}a_{2,2})$ with a plus sign and the second term appears with a minus sign. Thus,

$$\mathbf{S}(\pm a_{1,1}a_{2,2}) = a_{1,1}a_{2,2} - a_{2,1}a_{1,2}.$$

Task 12 Let's see how Cauchy produced the formula for his second example, $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3})$. Here, every possible permutation of the set $\{1, 2, 3\}$ of first indices must be performed. (You may want to refer back to Cramer's analysis of these permutations on page 7, where he lists them all.) Build a table to organize this information: in the first column, list all the possible permutations; in the second, identify a sequence of transpositions that results in this permutation; in the third, note whether the number of the transpositions is even or odd; and in the fourth, write out the term obtained from the indicative term $a_{1,1}a_{2,2}a_{3,3}$ by performing that permutation on its first indices, supplying a plus sign when the number of the transpositions is even and a minus sign if this number is odd. When you are done, the sum of the terms in the last column of your table should agree with Cauchy's formula for $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3})$. Observe that if we were to do this for Cauchy's first example, the table would have the form:

Permutation	Transpositions	Count/Parity	Term
12	none	0/even	$+a_{1,1}a_{2,2}$
21	$12 \to 21$	1/odd	$-a_{2,1}a_{1,2}$

Naturally, the greater the number of indices involved in these determinant expressions, the more complicated the formula for $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3} \ldots a_{n,n})$ becomes, mostly because the number of permutations on n indices grows extremely quickly with n, but also because one may require a succession of many transpositions in order to build up the permutation associated with a given term of the determinant.[4]

However, before Cauchy took on this issue, he made some important observations about the nature of these determinants.

As the quantities affected with different indices must be generally considered as unequal, we see that the determinant of the second order will contain four different quantities, namely:

$$\begin{array}{ll} a_{1,1}, & a_{1,2}, \\ a_{2,1}, & a_{2,2}, \end{array}$$

that the determinant of the third order contains nine of them, namely:

$$\begin{array}{lll} a_{1,1}, & a_{1,2}, & a_{1,3}, \\ a_{2,1}, & a_{2,2}, & a_{2,3}, \\ a_{3,1}, & a_{3,2}, & a_{3,3}. \end{array}$$

[4]The curious student may wonder if, in fact, this algorithm is well-defined: what's to say that a permutation that is built up from a sequence of, say, 5 transpositions might not also be built up from a sequence of 12 transpositions? But if this were the case, then one could not properly identify the parity of a permutation. We invite the reader to think about why all representations of a permutation as a composition of transpositions must preserve the parity of the number of transpositions employed.

Etc., etc. In general, the determinant of the n^{th} order, or

$$\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n}),$$

will contain a number equal to n^2 different quantities, which will be respectively[†]

$$\left\{\begin{array}{lllll} a_{1,1}, & a_{1,2}, & a_{1,3}, & \ldots, & a_{1,n}, \\ a_{2,1}, & a_{2,2}, & a_{2,3}, & \ldots, & a_{2,n}, \\ a_{3,1}, & a_{3,2}, & a_{3,3}, & \ldots, & a_{3,n}, \\ \ldots, & \ldots, & \ldots, & \ldots, & \ldots, \\ a_{n,1}, & a_{n,2}, & a_{n,3}, & \ldots, & a_{n,n}. \end{array}\right. \tag{1}$$

We suppose these same quantities disposed in a square, as we see here, in a number of horizontal lines equal to n and in as many vertical columns, in a manner such that, of the two indices which affect each quantity, the first varies only in each vertical column and the second varies only in each horizontal line; the set of such quantities will form a system that I will call a *symmetric system* of order n. The quantities $a_{1,1}, a_{1,2}, \ldots, a_{2,2}, \ldots a_{n,n}$ will be the different terms of the system and the letter a, stripped of accents, will be its characteristic. Finally, the quantities contained in one and the same line, either horizontal, or vertical, will be in number equal to n, and will form a sequence that I will call, in the first case, a *horizontal sequence* and, in the second, a *vertical sequence*. The index of each sequence will be the one which remains invariable in all the terms of the sequence. Thus, for example, the indices of the horizontal sequences and those of the vertical sequences of the system (1) are respectively equal to

$$1,\ 2,\ 3,\ \ldots\ ,\ n.$$

I will refer to as *conjugate terms* those which we can produce from each other by a transposition operating between the first and the second index; thus $a_{2,3}$ and $a_{3,2}$ are two conjugate terms. There exist some terms which are their own conjugates. These are the terms in which the two indices are equal to each other, namely:

$$a_{1,1},\ a_{2,2},\ \ldots\ ,\ a_{n,n};$$

I will call them *principal terms*: they are all positioned in system (1) on a diagonal of the square formed by the system.

It should be clear from this text that the objects which Cauchy termed *symmetric systems of order n* are what we today call $n \times n$ *matrices*; that what he called *horizontal* and *vertical sequences,* we today simply refer to as the *rows* and *columns* of the matrix; what he called *conjugate terms*, we typically call *symmetric entries* of the matrix; and what he called *principal terms*, we call *diagonal entries* of the matrix. It is important to understand that at the time of this writing, the theory

[†]Commentator's Note: Because this memoir is being excerpted in this Primary Source Project, the numbering of equations found here is not that in the original text but proceeds consecutively within this PSP, beginning here.

of matrices was decades away from being developed, so the terminology Cauchy used here was his own first attempt to describe the mathematical properties of things that were as yet only vaguely understood. Also, what later became conventional terminology had not yet been fixed, so we must keep this in mind as we read Cauchy. In particular, you may have already noticed that the first indices of the terms in Cauchy's symmetric systems are used to enumerate his vertical sequences while the second indices enumerate the horizontal sequences; in today's matrix notation, we also use double subscripts to identify entries of the matrix, but the modern convention is to use the first index to number the row and the second index to number the column (as well as to omit the comma separating the indices); that is, he had it exactly backwards from the common notational conventions of today!

In modern matrix theory, we typically use capital letters to stand for matrices; we might let A denote the "symmetric system" (1). The determinant of the matrix A we denote as $\det A$, so we can now identify

$$\det A = \mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n}).$$

Each term of the determinant

$$\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n})$$

is the product of n different quantities. The second indices which affect these quantities are respectively equal to the numbers

$$1,\ 2,\ 3,\ \ldots\ ,\ n.$$

which we may consider as being, in all of the terms, disposed according to their natural order. As for the first indices, they are again equal to these same numbers; but the order in which they follow each other vary from one term to the other and present, in the different terms, all the possible permutations of the numbers

$$1,\ 2,\ 3,\ \ldots\ ,\ n.$$

... The products that we will be able to form in this way will be in number equal to that of all possible permutations of the indices $1, 2, 3, \ldots n$, that is, this number will equal the product

$$1.2.3\ldots n.$$

Task 13 In the last sentence above, Cauchy gave $1.2.3\ldots n$ (what we would today abbreviate as $n!$, using *factorial* notation) as a general formula for the number of different permutations of the numbers $1, 2, 3, \ldots, n$. Recall from Task 8(a), that Cramer talked about the expression that appeared as the "common denominator" in his formulas for the n unknowns in systems of n linear equations as having "as many terms as there are diverse arrangements of n different things." Were Cramer and Cauchy computing the same thing? Explain.

Cauchy returned to the problem of identifying the sign of the terms in the determinant expression $\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n})$ for an arbitrarily chosen value of n. Recall from Task 12 that one way to do this is to build up the permutation of the first indices associated with the given term as a succession of transpositions – which in general can be a difficult thing to accomplish! – and then find the parity of the number of such transpositions. Cauchy realized that this has drawbacks, but never fear: he had an easier way to determine the sign!

> Given any symmetric product, in order to discover the sign with which it is affected in the determinant
>
> $$\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n}),$$
>
> it will suffice to apply the rule which serves to determine the sign of a term selected at will in an alternating symmetric function. Let
>
> $$a_{\alpha,1}a_{\beta,2}\ldots a_{\zeta,n}$$
>
> be the corresponding symmetric product, and designate by g the number of circular substitutions equivalent to the substitution
>
> $$\begin{pmatrix} 1.2.3.\ldots n \\ \alpha.\beta.\gamma.\ldots\zeta \end{pmatrix}.$$
>
> This product must be affected with the $+$ sign if $n-g$ is an even number, and with the $-$ sign in the contrary case.

Some comments about what Cauchy meant here: the permutation that carries the initial ordering $123\ldots n$ into the jumbled ordering $\alpha\beta\gamma\ldots\zeta$ of the same set of numbers he denoted as

$$\begin{pmatrix} 1.2.3.\ldots n \\ \alpha.\beta.\gamma.\ldots\zeta \end{pmatrix}.$$

He stated that every such permutation can be decomposed into a number of *circular substitutions*, which in modern mathematics we call *cycles*. A circular substitution/cycle is a sequence of the numbers $1,2,3,\ldots,n$, starting at some initial value which follows the action of the permutation from number to number until one returns to the same initial number at which one started. For instance, the permutation

$$\begin{pmatrix} 1.2.3.4.5.6 \\ 6.5.3.2.4.1 \end{pmatrix},$$

of the numbers $1,2,3,4,5,6$ carries 1 to 6 and 6 back to 1; this is one such cycle associated with the permutation. It also carries 2 to 5, 5 to 4, then 4 back to 2, forming a second cycle. Finally, the number 3 is sent to itself, the third cycle for this permutation. Note that the full permutation can be realized as a combination of these three cycles which operate independently on disjoint subsets of the indices $1, 2, 3, 4, 5, 6$. Cauchy denoted by g the number of circular substitutions/cycles in the decomposition; in our example above, $g = 3$.

At the end of the excerpt, he stated that the term associated with this permutation of the indices appears with a plus or minus sign if the parity of the number $n - g$ is, respectively, even or odd. In our example, $n - g = 6 - 3 = 3$ is odd, so the term $a_{6,1}a_{5,2}a_{3,3}a_{2,4}a_{4,5}a_{1,6}$ appears with a minus sign in the sixth order determinant $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4}a_{5,5}a_{6,6})$. Cauchy provided an example of his own:

It is easy to see that the preceding rule will also hold even if we jumbled the order of the factors

$$a_{\alpha,1}a_{\beta,2}\ldots a_{\zeta,n}$$

or, what is the same, the order of the indices involved in the two terms of the substitution

$$\begin{pmatrix} 1.2.3.\ldots n \\ \alpha.\beta.\gamma.\ldots\zeta \end{pmatrix},$$

provided that in this substitution we always place, one above the other, the two indices which in the given symmetric product affect the same quantity.

Suppose, for example, $n = 7$, and we seek what sign the symmetric product

$$a_{1,3}a_{3,6}a_{6,1}a_{4,5}a_{5,4}a_{2,2}a_{7,7}$$

must have in the determinant

$$\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4}a_{5,5}a_{6,6}a_{7,7}).$$

The substitution that we obtain, by the comparison of the indices which affect in the first and second line each of these factors of the product, is

$$\begin{pmatrix} 1.3.6.4.5.2.7 \\ 3.6.1.5.4.2.7 \end{pmatrix},$$

and this substitution is equivalent to [a combination of] the four circular substitutions

$$\begin{pmatrix} 1.3.6 \\ 3.6.1 \end{pmatrix}, \quad \begin{pmatrix} 4.5 \\ 5.4 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 7 \\ 7 \end{pmatrix};$$

here we have, therefore,

$$g = 4, \quad n = 7$$

and hence,

$$n - g = 3.$$

This last number being odd, the given symmetric product must be affected with the $-$ sign in the determinant of seventh order.

Task 14 At the beginning of this last excerpt, Cauchy claimed that "it is easy to see that the preceding rule will also hold even if we jumbled the order of the factors" in the terms of the determinant.[5] Why does it not matter if the factors of the term are jumbled amongst themselves?

Task 15

(a) Use the method Cauchy gave in the last two excerpts for finding the sign of the terms of a determinant to recalculate the entire expression for $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3})$; compare your answer here to what you obtained in Task 12.

(b) Let's use Cauchy's method to write out a complete formula for the fourth order determinant

$$\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4}).$$

We know that this fourth order determinant involves a large number of terms, each of which corresponds to a permutation of the indices $\{1, 2, 3, 4\}$. (How many terms? See Task 13.) It makes sense to generate these terms systematically in order to ensure that we will identify them all and that we won't inadvertently repeat one or more of them.

To this end, build a table like the one we used in Task 12. In each row of the table, keep track of the following in separate columns: (1) the permutation of the second indices represented by the first indices of that term; (2) the decomposition of this permutation into cycles; (3) the value of the quantity $n - g$ for the permutation; and finally, (4) the "symmetric product" with its proper sign. The determinant formula will then be found by summing all the expressions in the fourth column of the table.

To generate the permutations in the first column systematically, use what is called *backwards reverse lexicographic ordering*: list the permutations first in which the largest index appears last, followed by those in which the next largest index appears last, and so on. (This specific ordering is chosen carefully so as to highlight algebraic patterns that Cauchy will exploit later.) Among those permutations in which the largest index appears last, list those first in which the next-to-last index is largest, followed by those in which the next-to-last index is second largest, and so on. It is as though the permutations were n-letter words in a dictionary in which the words were ordered by considering first the *final* letters of the word, that is, backwards with respect to the standard ordering; and that we use reverse alphabetical (or lexicographical) order to do this. Thus, words in this odd dictionary *ending* in Z would come first, followed by all words ending in Y, then those ending in X, etc.; further, the word *glitz* would come *before* the word *blitz* because G *follows* B in the standard lexicographical order, and both words would precede the words *gauzy* and *latex* since their final letters are in reverse lexicographic order.

Here is the beginning of the table:

[5] Mathematicians are famous for using language like this ("it is easy to see...") in their writing. It is used when the writer believes that the details of the upcoming argument don't have to be elaborated in detail for readers, who can supply the details themselves. Of course, the writer can't always be sure who the readers are, so this is sometimes met by readers who are not adept at this sort of mathematical reasoning with great frustration whenever their notion of "easy" is different from that of the author!

Permutation of 1234	Cycle decomposition	$n-g$	Term
1234	(1)(2)(3)(4)	0	$+a_{1,1}a_{2,2}a_{3,3}a_{4,4}$
2134	(12)(3)(4)	1	$-a_{2,1}a_{1,2}a_{3,3}a_{4,4}$
1324	(1)(23)(4)	1	$-a_{1,1}a_{3,2}a_{2,3}a_{4,4}$
3124	(132)(4)	2	$+a_{3,1}a_{1,2}a_{2,3}a_{4,4}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Again, when you complete the table, the sum of the terms in the rightmost column will give the full expansion of the determinant.

(c) Compare your answer in (b) to the one you gave in Task 8(b) using Cramer's notation, and convince yourself that the two expressions are equivalent.

After completing Task 15(b), you were likely convinced that writing out the full expression of the determinant $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}\dots a_{n,n})$ for all but the very smallest values of n is an extremely tedious exercise indeed!

3 A Modern Approach for 2×2 and 3×3 Matrices

In modern expositions of linear algebra, we view the determinant as a quantity that is attached to a square matrix. For instance, we earlier saw that Cauchy defined the determinants of order 2 and 3 to be given by the algebraic formulas

$$\begin{aligned} \mathbf{S}(\pm a_{1,1}a_{2,2}) &= a_{1,1}a_{2,2} - a_{2,1}a_{1,2}, \\ \mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}) &= a_{1,1}a_{2,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3} \\ &\quad -a_{1,1}a_{3,2}a_{2,3} - a_{3,1}a_{2,2}a_{1,3} - a_{2,1}a_{1,2}a_{3,2} \end{aligned} \tag{2}$$

These formulas he then associated with the respective "symmetric systems"

$$\begin{matrix} a_{1,1}, & a_{1,2}, \\ a_{2,1}, & a_{2,2}, \end{matrix}$$

and

$$\begin{matrix} a_{1,1}, & a_{1,2}, & a_{1,3}, \\ a_{2,1}, & a_{2,2}, & a_{2,3}, \\ a_{3,1}, & a_{3,2}, & a_{3,3}. \end{matrix}$$

Cauchy viewed the determinant primarily as an algebraic expression, and the symmetric system to which he associated the determinant was to him an auxiliary device, a visual aid to help him perceive the underlying algebraic structure of the corresponding expression.

This logic is reversed in the modern approach: today we begin with a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \text{ or } \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \tag{3}$$

then associate with it a quantity $\det A$, which is a function of its entries and which evaluates to a single number when the entries have specific numerical values. This modern approach makes it a bit easier to handle the determinant, especially in low-order cases, since there is already a visual frame given by the array of the matrix which makes it possible to identify the various terms of the determinant and their associated signs.

To see this, we introduce some modern terminology. The *principal diagonal* entries of a matrix are defined in the same way Cauchy defined them for a symmetric system: they are the entries of the form a_{kk} for $k = 1, 2, \ldots, n$. There are other clusters of entries that form parallel *diagonals* of the matrix, and are easier to visualize if the first few columns of the matrix are repeated to the right of the original matrix:

$$\begin{array}{c}
\searrow \qquad \searrow \quad \searrow \qquad \nearrow \quad \nearrow \quad \nearrow \\
\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix} \\
\nearrow \qquad \nearrow \quad \nearrow \qquad \searrow \quad \searrow \quad \searrow
\end{array}$$

For in addition to the principal diagonal a_{11}, a_{22}, a_{33}, which is fully contained within the original matrix, there are two other parallel diagonals, one beginning at a_{12} and including the two other entries found by proceeding diagonally downward across the array and ending at a_{31}; and another diagonal that begins at a_{13} and ends at a_{32}. We also define *skew-diagonals* of the matrix, clusters of entries which begin in the bottom row, proceed diagonally upward and end at the top row: the *principal skew-diagonal* starts at entry a_{31} and ends at a_{13}, and there are two other parallel skew-diagonals beginning to its right.

Task 16

(a) Working with the first formula in (2), show that the two-termed determinant $\det A$ of the 2×2 matrix A in (3) can be found by multiplying the entries down the principal diagonal of A and subtracting the term formed by multiplying the entries along the principal skew-diagonal.

(b) Working with the second formula in (2), show that the six-termed determinant $\det A$ of the 3×3 matrix A in (3) can be found by multiplying the entries down the principal – and other – diagonals of A, and subtracting the terms formed by multiplying entries along the principal – and other – skew-diagonals.

(c) Explain why this method of generating diagonals of entries in a matrix cannot be used to produce a formula for the fourth order determinant! (How many terms must the determinant have? How many diagonals and skew-diagonals are there?) Indeed, why does the method we describe above work *only* for determinants of order 2 and 3?

4 Determinants of Upper Triangular Matrices

There are some other types of matrices whose determinants can be easily calculated using Cauchy's methods; we shall consider some of these now.

We say that an $n \times n$ matrix $A = (a_{ij})$ is *upper triangular* if $a_{ij} = 0$ whenever $i > j$; in other words, all the entries below the principal diagonal are 0. Similarly, we call A *lower triangular* if $a_{ij} = 0$ whenever $i < j$. Further, A is said to be a *diagonal* matrix if it is both upper and lower triangular (that is, if its only nonzero entries are on the principal diagonal). In the three examples below, the first matrix is upper triangular, the middle one is lower triangular, and the third is a diagonal matrix.[6]

$$\begin{pmatrix} -1 & 0 & 2 \\ 0 & 3 & -4 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 \\ -8 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -5 \end{pmatrix}.$$

These types of matrices arise repeatedly in linear algebraic problems, and their determinants are much easier to compute than those of random matrices. For instance, Gaussian elimination, which employs the method of row reduction of matrices, has as its goal to transform the augmented coefficient matrix of a given system of linear equations into row echelon form (or row-reduced echelon form);[7] such matrices are necessarily upper triangular.

Task 17 Cauchy has shown us that the nth order determinant of a "symmetric system," what today we would call the determinant of an $n \times n$ matrix $A = (a_{ij})$, is found by forming the sum of signed products of entries of the matrix, terms of the form

$$\pm a_{\alpha 1} a_{\beta 2} \cdots a_{\zeta n}, \tag{4}$$

where this sum includes a single term for every possible permutation

$$\left(\begin{array}{c} 1.2.3. \ldots n \\ \alpha.\beta.\gamma. \ldots \zeta \end{array} \right)$$

of the indices $\{1, 2, \ldots, n\}$, with signs determined by a well-specified procedure.

(a) Suppose that A is upper triangular. Since many of the entries of such a matrix equal 0, lots of the terms (4) in Cauchy's expansion of the determinant will vanish, leaving only those for which every one of the entries $a_{\alpha 1}, a_{\beta 2}, \ldots, a_{\zeta n}$ have their first index less than or equal to their second. Consider the possible values for the indices α, β, etc., in turn that will ensure that none of these entries appears below the principal diagonal of A; remember that since $\alpha\beta\gamma \ldots \zeta$ is a permutation of $123 \cdots n$, none of the values of $\alpha, \beta, \gamma, \ldots, \zeta$ may repeat. Conclude that Cauchy's formula simplifies to a single term, and write out this simplified formula for $\det A$.

(b) Use your result in part (a) to determine $\det I$, where I is the $n \times n$ identity matrix (with ones along the principal diagonal and all entries equal to 0 off the principal diagonal).

(c) Explain why a (square) matrix with an entire row of zeros must have a zero determinant.

[6] Non-square matrices can also be upper or lower triangular, but as only square matrices have determinants, we shall not consider such objects here.

[7] For definitions of the terms mentioned in this sentence, see Section 11 below, or consult any linear algebra textbook.

(d) Define the $n \times n$ matrix $E[i, j; c]$ to be the matrix that is identical to the identity matrix except that the entry in row i and column j is replaced with the number c. For instance, where $n = 3$, we have

$$E[1, 1; -2] = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E[2, 3; 4] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}.$$

Such matrices are called *elementary* matrices because they can be obtained by performing a single elementary row operation on the identity matrix I. For instance, the matrix $E[1, 1; -2]$ above can be found by multiplying the first row of I by the factor -2, and the matrix $E[2, 3; 4]$ can be obtained from I by replacing its second row by its sum with 4 times its third. Use the result in part (a) above to find $\det E[i, j; c]$.

5 Conjugate Systems and Transposed Matrices

Cauchy turned his attention next to investigating more general properties of determinants relative to the "symmetric systems" (matrices) with which they are associated.

§II. If in each of the terms of system (1) we replace the first index by the second, and reciprocally, we will then have a new system relative to which the first index will remain invariable in each vertical sequence and the second index in each horizontal sequence. The system thus formed will be

$$\left\{ \begin{array}{lllll} a_{1,1}, & a_{2,1}, & a_{3,1}, & \ldots, & a_{n,1}, \\ a_{1,2}, & a_{2,2}, & a_{3,2}, & \ldots, & a_{n,2}, \\ a_{1,3}, & a_{2,3}, & a_{3,3}, & \ldots, & a_{n,3}, \\ \ldots, & \ldots, & \ldots, & \ldots, & \ldots, \\ a_{1,n}, & a_{2,n}, & a_{3,n}, & \ldots, & a_{n,n}. \end{array} \right. \tag{5}$$

I will say that the systems (1) and (5) are respectively *conjugate* to one another. For brevity, I will designate henceforth each of these two systems by the final term of the first horizontal sequence contained between two parentheses; thus, system (1) will be designated by $(a_{1,n})$, and the system (5) by $(a_{n,1})$.

The symmetric products of the systems $(a_{1,n})$ and $(a_{n,1})$ are clearly equal to each other. . . . Consequently, in the expression

$$\mathbf{S}(\pm a_{1,1} a_{2,2} \ldots a_{n,n}),$$

we may assume indifferently either that the sign $\mathbf{S}$ corresponds to the first indices, or that it corresponds to the second; this removes all uncertainty with the value of the expression in question.

Task 18 In the above excerpt, Cauchy stated, "The symmetric products of the systems $(a_{1,n})$ and $(a_{n,1})$ are clearly equal to each other." Can you fill in the details of this argument in your own words to prove that

$$\det A^T = \det A \tag{6}$$

for any square matrix A, where A^T is the *transpose* of the matrix A. Be sure to describe the relationship between what Cauchy called "conjugate systems" and the modern notion of the transpose of a matrix.

6 Laplace Expansion

In the next section of Cauchy's treatise, he dealt with a powerful property of determinants that expresses a determinant of order n in terms of determinants of order $n-1$. This method, known today as *Laplace expansion*, was eventually named for an older contemporary of Cauchy, Pierre-Simon de Laplace (1789-1857), a friend of Cauchy's father and someone who assisted the young Cauchy in his education as a mathematician. This procedure for calculating a determinant turns out to be generally useful only in special cases: when n is small, or when the underlying matrix is sparse – that is, when it has many zero entries. But as small-n situations and sparse matrices are still quite common in linear algebraic settings, it is worthwhile to know how the method works.

§III. We designate by D_n the determinant of order n, of either of the two systems $(a_{1,n})$, $(a_{n,1})$, so that we have

$$D_n = \mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n-1,n-1}a_{n,n}),$$

and we suppose, in order to fix the ideas, that the sign $\mathbf{S}$ is relative to the first indices. The determinant of order $n-1$, or D_{n-1}, will be given by the equation

$$D_{n-1} = \mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n-1,n-1}).$$

If we multiply the latter by $a_{n,n}$, we will have the algebraic sum of the symmetric products in the determinant D_n which have $a_{n,n}$ as factor, these products being taken alternately with the $+$ sign and with the $-$ sign in the sum in question. Moreover, it is easy to see that these same products are affected with the same signs in the determinant D_n as in the determinant D_{n-1} multiplied by $a_{n,n}$. Indeed, the principal product

$$a_{1,1}a_{2,2}a_{3,3}\cdots a_{n,n}$$

is found in both cases affected with the $+$ sign, and the sign of any of the other products is determined in either case by the number of transpositions that we are required to effect on the first indices $1, 2, 3, \ldots, n-1$ in order to derive it from the principal product. Hence, the expression

$$a_{n,n}\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n-1,n-1}),$$

considered as function of the first indices $1, 2, 3, \ldots, n$, will not be, in general, an alternating symmetric function. But if, in this same function, we transpose in all possible ways the indices in question, we will obtain a sequence of values, many of which will be different from one another, and if, from the sum of the different values obtained through an even number of transpositions, we subtract the sum of the different values obtained through an odd number of transpositions, then we will have an alternating symmetric function that we can denote by

$$\mathbf{S}[\pm a_{n,n}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}\ldots a_{n-1,n-1})],$$

the new [outer] sign $\mathbf{S}$ being always relative to the first indices. In order to obtain the various values of the product

$$a_{n,n}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}\ldots a_{n-1,n-1}),$$

it will clearly suffice to change successively the first index n of the factor $a_{n,n}$ into the indices $1, 2, 3, \ldots n-1$ which affect in the first row [of (5)] those quantities contained in the second factor

$$\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}\ldots a_{n-1,n-1}).$$

Having done this, the first factor will become successively equal to each of the quantities

$$a_{n,n}, \quad a_{n-1,n}, \quad a_{n-2,n}, \quad \ldots, \quad a_{2,n}, \quad a_{1,n};$$

and, if we represent respectively by

$$b_{n,n}, \quad -b_{n-1,n}, \quad -b_{n-2,n}, \quad \ldots, \quad -b_{2,n}, \quad -b_{1,n},$$

the corresponding values of the second factor, we will have

$$\begin{aligned}&\mathbf{S}[\pm a_{n,n}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}\ldots a_{n-1,n-1})]\\ &\quad = a_{n,n}b_{n,n} + a_{n-1,n}b_{n-1,n} + \cdots + a_{2,n}b_{2,n} + a_{1,n}b_{1,n}.\end{aligned}$$

By the preceding, we will have

$$\left\{\begin{aligned} b_{n,n} &= \mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n-1,n-1}) = D_{n-1},\\ b_{n-1,n} &= \mathbf{S}(\mp a_{1,1}a_{2,2}\ldots a_{n,n-1}),\\ &\ldots\\ b_{2,n} &= \mathbf{S}(\mp a_{1,1}a_{n,2}\ldots a_{n-1,n-1}),\\ b_{1,n} &= \mathbf{S}(\mp a_{n,1}a_{2,2}\ldots a_{n-1,n-1}),\end{aligned}\right. \tag{7}$$

the sign $\mathbf{S}$ in all these equations being considered as relative either to the first or to the second index.

To help make sense of this last excerpt, let's review Cauchy's statements in the particular case $n = 4$.

Task 19 (a) Return to the work you did in Task 15. There in part (a) you explicitly wrote out the multi-term algebraic expressions for $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3})$ and in part (b) the much longer one for $D_4 = \mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4})$. Write out the full expression for $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4})$ once again (yes, the full expression!). Then, write out the expression for $a_{4,4}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3})$ and <u>underline</u> in the first, longer expression the terms that appear in the second. Are they all "affected with the same signs in the determinant D_n as in the determinant D_{n-1} multiplied by $a_{n,n}$," as Cauchy observed? That is, confirm that $a_{4,4}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3})$ is a *sub-expression* of $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4})$. Note that these common terms are precisely the ones that were listed first in the table you constructed in Task 15(b).

(b) Consider again the expression $a_{4,4}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3})$ which you wrote out in part (a). Now write out the multi-term expression obtained from this one by performing the transposition that swaps any 4 with a 3 (and vice versa) whenever it appears as a first subscript. This should produce the expression $a_{3,4}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{4,3})$. Locate these terms (Cauchy calls them "products") in the expression for $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4})$ and mark them with an $\overline{\text{overscore}}$. How do the signs of these terms in $a_{3,4}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{4,3})$ compare with the signs of the same terms in $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4})$? Note that these terms formed the second group of terms listed in your table from Task 15(b), those which contained the factor $a_{3,4}$.

(c) Once again, consider the expression $a_{4,4}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3})$ from part (a). Now write out the multi-term expression obtained from this one by performing the transposition that swaps any 4 with a 2 (and vice versa) whenever it appears as a first subscript. This will produce the expression $a_{2,4}\mathbf{S}(\pm a_{1,1}a_{4,2}a_{3,3})$. Locate these terms in the expression for $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4})$ and mark them $\boxed{\text{with a box}}$. How do the signs of the terms in $a_{2,4}\mathbf{S}(\pm a_{1,1}a_{4,2}a_{3,3})$ compare with the signs of the same terms in $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4})$? Note that these terms formed the third group of terms listed in your table from Task 15(b), those which contained the factor $a_{2,4}$.

(d) You should now be able to guess what's coming next! Consider $a_{4,4}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3})$ again and write out the multi-term expression obtained from this by performing the transposition that swaps any 4 with a 1 (and vice versa) whenever it appears as a first subscript. This will produce the expression $a_{1,4}\mathbf{S}(\pm a_{4,1}a_{2,2}a_{3,3})$. Locate these terms in the expression for $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4})$ and mark them by circling them. How do the signs of the terms in $a_{1,4}\mathbf{S}(\pm a_{4,1}a_{2,2}a_{3,3})$ compare with the signs of the same terms in $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4})$? Note that these terms form the last group of terms listed in your table from Task 15(b), those containing the factor $a_{1,4}$.

What we have done in this last Task is, as Cauchy relates in the last excerpt, to "change successively the first index n of the factor $a_{n,n}$ into the indices $1, 2, 3, \ldots, n-1$" in the product $a_{n,n}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}\ldots a_{n-1,n-1})$, whence "the first factor will become successively equal to each of the quantities

$$a_{n,n}, \quad a_{n-1,n}, \quad a_{n-2,n}, \quad \ldots, \quad a_{2,n}, \quad a_{1,n},"$$

while simultaneously changing in the same way "those quantities contained in the second factor $\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}\ldots a_{n-1,n-1})$."

Task 20 (a) Let's summarize the work of Task 19. What equation expresses the relationship between the four expressions

$$a_{4,4}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}), \quad a_{3,4}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{4,3}), \quad a_{2,4}\mathbf{S}(\pm a_{1,1}a_{4,2}a_{3,3}), \quad a_{1,4}\mathbf{S}(\pm a_{4,1}a_{2,2}a_{3,3})$$

and the single expression

$$\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4})\ ?$$

Tell how the signs you use for the first four expressions in this equation are related to the ways in which "we transpose in all possible ways the indices in question" so that "from the sum of the different values obtained through an even number of transpositions, we subtract the sum of the different values obtained through an odd number of transpositions."

(b) From the equation you have discovered in part (a), show that

$$\mathbf{S}[\pm a_{4,4}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3})] = \mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}a_{4,4}).$$

(c) Simplify this last equation by substituting

$$\begin{aligned} b_{4,4} &= \mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}), \\ -b_{3,4} &= \mathbf{S}(\pm a_{1,1}a_{2,2}a_{4,3}), \\ -b_{2,4} &= \mathbf{S}(\pm a_{1,1}a_{4,2}a_{3,3}), \\ -b_{1,4} &= \mathbf{S}(\pm a_{4,1}a_{2,2}a_{3,3}), \end{aligned}$$

as Cauchy instructs. Why did Cauchy attach negative signs to $b_{3,4}, b_{2,4}, b_{1,4}$, but not to $b_{4,4}$? (This explains the use of the $\mp$ signs in equations (7).)

From what we have just said, it is easy to conclude that, if we develop the alternating symmetric function

$$\mathbf{S}[\pm a_{n,n}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}\ldots a_{n-1,n-1})],$$

all the terms of the development will be symmetric products of order n, which will have unity as coefficient. These terms, therefore, will be respectively equal to those which we obtain by developing the determinant

$$D_n = \mathbf{S}(\pm a_{1,1}a_{n,2}\ldots a_{n,n}),$$

and as the principal product $a_{1,1}a_{2,2}\ldots a_{n,n}$ is positive on both sides, we will have necessarily

$$\begin{aligned} D_n &= \mathbf{S}[\pm a_{n,n}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}\ldots a_{n-1,n-1})] \\ &= a_{n,n}b_{n,n} + a_{n-1,n}b_{n-1,n} + \cdots + a_{2,n}b_{2,n} + a_{1,n}b_{1,n}. \end{aligned}$$

In modern terms, if

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}, \tag{8}$$

then $\det A$ is an algebraic expression in the sixteen entries of the matrix that evaluates to a number when the entries are numbers. Note that our modern convention of writing a_{ij} to denote the entry of the matrix A in row i and column j differs from Cauchy's method of using $a_{i,j}$ to denote a quantity within a symmetric system (see (5)) in *vertical sequence* i and *horizontal sequence* j. The two notational conventions are transposes of each other, in the matrix sense of that term. However, since we have also seen that $\det A^T = \det A$ (Task 16), the resulting expressions meant the same in the context of evaluating determinants to both Cauchy as it does to us moderns.

In the last excerpt we considered above, Cauchy stated his powerful identity

$$\mathbf{S}(\pm a_{1,1}a_{n,2}\ldots a_{n,n}) = \mathbf{S}[\pm a_{n,n}\mathbf{S}(\pm a_{1,1}a_{2,2}a_{3,3}\ldots a_{n-1,n-1})],$$

which he also wrote in the form

$$D_n = a_{n,n}b_{n,n} + a_{n-1,n}b_{n-1,n} + \cdots + a_{2,n}b_{2,n} + a_{1,n}b_{1,n}. \tag{9}$$

Let us recast these equations, in the special case $n = 4$ on which we have been focusing, but now in modern notation. We would normally write D_4 today as $\det A$ where A is the matrix (8), and the quantities $a_{4,4}, a_{3,4}, a_{2,4}, a_{1,4}$ of Cauchy's symmetric system which appear in (9) would then correspond to the entries $a_{44}, a_{43}, a_{42}, a_{41}$ appearing in the fourth *row* of A. What remains is to interpret the quantities $b_{4,4}, b_{3,4}, b_{2,4}, b_{1,4}$ in (9), quantities which in modern linear algebra are called the *cofactors* of the respective entries $a_{44}, a_{43}, a_{42}, a_{41}$, because of how they appear in equation (9).

Task 21

(a) Consider equations (7); use these to expand formulas for the four quantities $b_{4,4}, b_{3,4}, b_{2,4}, b_{1,4}$ in terms of the elements $a_{i,j}$ in Cauchy's notation.

(b) Write down four copies of the 4×4 matrix A, as in (8). In each successive copy, circle a different one of the entries $a_{44}, a_{43}, a_{42}, a_{41}$, then strike out the entire row and entire column of the entry you circled. We will denote the four respective 3×3 matrices which are left behind by this process with the symbols $A_{44}, A_{43}, A_{42}, A_{41}$; in modern linear algebra, whenever we remove any number of rows, and the same number of columns, from a square matrix A, the resulting smaller square matrix is called a *minor* of A. Compute the four determinants of these 3×3 minors of A (recall Section 3 above).

(c) Compare the quantities $b_{4,4}, b_{3,4}, b_{2,4}, b_{1,4}$ with the quantities $\det A_{44}, \det A_{43}, \det A_{42}, \det A_{41}$, remembering that column indices are displayed first in Cauchy's notation while row indices come first in modern notation! What do you notice?

(d) Rewrite equation (9) for $n = 4$ in modern notation, in terms of the entries $a_{44}, a_{43}, a_{42}, a_{41}$ and the determinants $\det A_{44}, \det A_{43}, \det A_{42}, \det A_{41}$.

You should have observed by now that exactly half of the cofactors of the 4×4 matrix A are precisely equal to the determinants of the minors obtained by striking out the row and column of the entry for which it is a cofactor, while the other half are the negatives of the determinants of their corresponding minors. The next Task will answer the natural question: "What is the pattern of these signs?"

Task 22

(a) Use your work in Task 21(d) to verify that for each of $j = 1, 2, 3, 4$, the cofactor of the entry a_{4j} is identical to the determinant of its corresponding 3×3 minor A_{4j}, but with a factor of $(-1)^{4+j}$ attached.

(b) Use your formula from part (b) to verify that

$$\det \begin{pmatrix} 1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 4 \\ -1 & 2 & 8 & 5 \\ 3 & -1 & -2 & 3 \end{pmatrix} = 3.$$

Cauchy stated next that the expansion of a determinant as the sum of the products of entries of its last row with their corresponding cofactors still works if we replace the entries of the last row with those of *any* row.

In general, if we designate by μ any one of the indices $1, 2, 3, \ldots, n$, we will find in the same manner

$$D_n = \mathbf{S}[\pm a_{\mu,\mu}\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{\mu-1,\mu-1}a_{\mu+1,\mu+1}\ldots a_{n,n})].$$

Let ν be another index different from μ. We represent by $b_{\mu,\mu}$ the coefficient of the factor $a_{\mu,\mu}$ in the indicative term of the alternating symmetric function

$$\mathbf{S}[\pm a_{\mu,\mu}\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{\mu-1,\mu-1}a_{\mu+1,\mu+1}\ldots a_{n,n})]$$

and by $-b_{\nu,\mu}$ that which $b_{\mu.\mu}$ becomes when we replace the first index ν by μ. If we suppose successively

$$\nu = 1, \quad \nu = 2, \quad \ldots, \quad \nu = \mu - 1, \quad \nu = \mu + 1, \quad \ldots, \quad \nu = n,$$

then $b_{\mu,\mu}$ will become

$$-b_{1,\mu}, \quad -b_{2,\mu}, \quad \ldots, \quad -b_{\mu-1,\mu}, \quad -b_{\mu+1,\mu}, \quad \ldots, \quad -b_{n,\mu}$$

and the value of D_n will be given by the equation

$$D_n = a_{1,\mu}b_{1,\mu} + a_{2,\mu}b_{2,\mu} + \cdots + a_{\mu,\mu}b_{\mu,\mu} + \cdots + a_{n,\mu}b_{n,\mu}.$$

If, in this equation, we give successively to μ all the values

$$1, \quad 2, \quad 3, \quad \ldots, \quad n,$$

then we will obtain the following equations:

$$\begin{cases} D_n = & a_{1,1}b_{1,1} + a_{2,1}b_{2,1} + \cdots + a_{n,1}b_{n,1}, \\ D_n = & a_{1,2}b_{1,2} + a_{2,2}b_{2,2} + \cdots + a_{n,2}b_{n,2}, \\ \quad \ldots & \\ D_n = & a_{1,n}b_{1,n} + a_{2,n}b_{2,n} + \cdots + a_{n,n}b_{n,n}, \end{cases} \tag{10}$$

in which we must suppose, in general, that

$$\begin{cases} b_{\mu,\mu} = & \mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{\mu-1,\mu-1}a_{\mu+1,\mu+1}\ldots a_{n,n}), \\ b_{\nu,\mu} = & \mathbf{S}(\mp a_{1,1}a_{2,2}\ldots a_{\mu-1,\mu-1}a_{\mu+1,\mu+1}\ldots \\ & \qquad a_{\nu-1,\nu-1}a_{\mu,\nu}a_{\nu+1,\nu+1}\ldots a_{n,n}). \end{cases} \tag{11}$$

Equations (10) and (11) state that the same patterns we observed in expanding the determinant formula using entries of the bottom row of a matrix (9) hold when using entries of an arbitrary row. In modern notation, this means that expansion of a determinant using entries from row i of a matrix have cofactors which are determinants of the corresponding minors with signs attached such that the cofactor of entry a_{ij} will be $(-1)^{i+j} \det A_{ij}$.

Task 23

(a) Rewrite the second of the equations in (10) for the case $n = 4$ in modern notation, just as you did in Task 21(d). That is, express the determinant of the matrix in (8) in an expansion that uses signed products of entries of its second row with the determinants of their corresponding 3×3 minors.

(b) Use the result of part (a) to repeat the computation of the determinant of the matrix presented in Task 22(c).

We saw in Section 4 above that a square matrix and its transpose have the same determinant. It follows that the row-expansion formulas Cauchy gave in (10) can be extended to provide similar column-expansion formulas.

Task 24

(a) Following the patterns above, express the determinant of the matrix in (8) via an expansion that uses signed products of entries of its third column with the determinants of the corresponding 3×3 minors.

(b) Now check that your column-expansion formula from part (a) can be used to repeat the computations from Task 22(c) and Task 23(b).

Task 25

State a general formula for Laplace expansion of a determinant of an $n \times n$ matrix A along any row or column of A.

7 The Classical Adjoint

The quantities $b_{1,1}$, $b_{1,2}, \ldots$ in the preceding equations [that is, in (11)] are as numerous as the quantities $a_{1,1}, a_{1,2}, \ldots$ against which they multiply, that is, they are equal in number to n^2. They can therefore be disposed in a square in such a way as to form two new systems of order n which are respectively conjugate to one another. One of these systems will be the following:

$$\left\{\begin{array}{llll} b_{1,1}, & b_{1,2}, & \ldots, & b_{1,n}, \\ b_{2,1}, & b_{2,2}, & \ldots, & b_{2,n}, \\ \ldots, & \ldots, & \ldots, & \ldots, \\ b_{n,1}, & b_{n,2}, & \ldots, & b_{n,n}, \end{array}\right. \tag{12}$$

which I will denote by $(b_{1,n})$ according to the established convention: by replacing in the one here the first indices by the second, and reciprocally, we will obtain the other system [in question], which will be represented by $(b_{n,1})$.

...[T]he system $(b_{1,n})$ will be said to be *adjoint to the system* $(a_{1,n})$ and the system $(b_{n,1})$ *adjoint to the system* $(a_{n,1})$.

In modern linear algebra, we still retain this terminology from Cauchy: where $A = (a_{ij})_{i,j=1}^{n}$ is any $n \times n$ matrix, we refer to the transpose of the matrix of cofactors of entries of A as the *classical adjoint*, or *adjugate*, of A, and it is denoted $\operatorname{adj} A$.[8]

Task 26

(a) Return again to the 4×4 matrix example from Task 22(c), which we label A in this Task. For each choice of index pairs $i, j = 1, 2, 3, 4$, let b_{ij} be the cofactor of the corresponding entry a_{ij}. Find the values of all 16 cofactors for the example matrix, then write down $\operatorname{adj} A$. (Don't forget that the classical adjoint is the *transpose* $(b_{ji})_{i,j=1}^{n}$ of the matrix of cofactors!)

(b) Multiply A with $\operatorname{adj} A$ – in both directions; that is, compute $A \cdot (\operatorname{adj} A)$ and $(\operatorname{adj} A) \cdot A$. What do you notice?

It is our next objective to explain why the striking result you found in Task 25(b) is not a coincidence!

[8] As it turns out, the word *adjoint* is also used to define a different concept in linear algebra, so the use of the adjective *classical* helps to clarify the term; indeed, this is also why the word *adjugate* is often now preferred.

8 The Effect of Interchanging Rows/Columns on the Determinant

Task 27 (a) In Task 15(b), you generated a systematic method for writing out the formula for the fourth order determinant, equivalent to finding the determinant of the matrix (8). Consider now the matrix

$$A' = \begin{pmatrix} a'_{11} & a'_{12} & a'_{13} & a'_{14} \\ a'_{21} & a'_{22} & a'_{23} & a'_{24} \\ a'_{31} & a'_{32} & a'_{33} & a'_{34} \\ a'_{41} & a'_{42} & a'_{43} & a'_{44} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}, \tag{13}$$

obtained by interchanging the second and third rows of the original matrix A. One of the many terms of this determinant is $a'_{12}a'_{23}a'_{31}a'_{44}$. Use the method of Task 15(b) to determine the sign of this term in the expansion of $\det A'$. That is, (i) note the specific permutation of the row indices[9] 1234 given by the column indices which uniquely identifies this term in the expansion; (ii) find the cycle decomposition of this permutation; (iii) compute the associated quantity $n - g$, where $n = 4$ and g is the number of cycles in the decomposition; and finally, (iv) note that the sign of the term is $+$ or $-$ according as $n - g$ is even or odd – that is, the sign is given by $(-1)^{n-g}$.

(b) Now use equation (13) to rewrite $a'_{12}a'_{23}a'_{31}a'_{44}$ in terms of the entries of A, i.e., in terms of the un-primed a_{ij}s. Then repeat the four steps listed in part (a) above to determine the sign of this term in the expansion of the determinant of A. You should discover that the sign of the term in $\det A$ is the opposite of the sign of $a'_{21}a'_{32}a'_{13}a'_{44}$ in $\det A'$ which you found in part (a).

(c) Repeat the analysis of parts (a) and (b), but this time begin with the term $a'_{14}a'_{21}a'_{32}a'_{43}$ from $\det A'$: use Cauchy's four-step method to determine its sign in $\det A'$, then, by rewriting it in terms of the entries a_{ij} from the matrix A, find the sign of the corresponding term in the expansion of $\det A$. Again, you should find that the sign of the latter term in $\det A$ is the opposite of the sign of $a'_{14}a'_{21}a'_{32}a'_{43}$ in $\det A'$.

In Task 27 above, notice that the row indices 2 and 3 – the very indices which were interchanged in the matrix A to obtain the matrix A' – appeared *in the same cycle* of the decomposition of the permutations that corresponded to the terms we selected, both in parts (a-b) and in part (c). In each case, the effect of the interchange of the two indices on the underlying permutations was to split the cycle that contained both 2 and 3 into two cycles that each contained one of the two elements. Specifically, in parts (a-b), the permutation 2314 associated with $a'_{12}a'_{23}a'_{31}a'_{44}$ has cycle decomposition (123)(4) while the interchanged permutation 3214 associated with the term $a_{13}a_{22}a_{31}a_{44}$ has cycle decomposition (13)(2)(4). Similarly, in part (c), the permutation 4123 associated with $a'_{14}a'_{21}a'_{32}a'_{43}$ has cycle decomposition (1432) while the interchanged permutation 4132 associated with the term $a_{14}a_{21}a_{33}a_{42}$ has cycle decomposition (142)(3). In both cases, the effect of the interchange of the two elements was to replace the single cycle containing the elements 2 and 3 with a pair of cycles each containing only one of the two elements.

[9] Recall that Cauchy used the "second indices" of his "symmetric products" to identify the rows of his "symmetric system," and the "first indices" to identify the columns in which they were found in the "system."

We now consider the effect of this interchange of rows on terms in which the two elements lie *in distinct cycles* of the decomposition of the permutation.

Task 28 Take the term $a'_{12}a'_{21}a'_{34}a'_{43}$ from $\det A'$, as in Task 27. Use Cauchy's four-step method to find the sign of this term in the expansion of the formula for $\det A'$. Note that the cycle decomposition of the underlying permutation finds the elements 2 and 3 in distinct cycles. As in Task 27, rewrite $a'_{12}a'_{21}a'_{34}a'_{43}$ in terms of the un-primed a_{ij} entries using equation (13), then identify the sign of that term in the expansion of $\det A$. Note that the cycle decomposition of the underlying permutation for this term in $\det A$ finds the elements 2 and 3 now in the same cycle! Consequently, the term $a'_{12}a'_{21}a'_{34}a'_{43}$ from $\det A'$ and its corresponding equivalent term in $\det A$ once again have opposite signs.

The results of Tasks 27 and 28 suggest that something more general can be concluded about the effect on the determinant of an interchange of row entries in a matrix. What we will prove is this:

Theorem. *Suppose that A is an $n\times n$ matrix, and that A' is a matrix obtained from A by interchanging some two rows, say the rows with indices μ and ν. We know that each term in $\det A'$ corresponds to some permutation of $\{1, 2, \ldots, n\}$ that carries the row indices of the factors in that term, factor by factor, to their associated column indices, and that this term is identical to some particular term in $\det A$ corresponding to the same permutation of its indices except for the additional interchange of μ with ν. Then the signs of these terms as they appear in $\det A'$ and $\det A$ are always opposite. Consequently,*

$$\det A' = -\det A.$$

Indeed, by virtue of the fact that $\det A^T = \det A$, we immediately have this result:

Corollary. *Suppose that A is an $n \times n$ matrix, and that A' is a matrix obtained from A by interchanging some pair of columns. Then*

$$\det A' = -\det A.$$

In order to prove the above Theorem, we follow the lead of the examples we considered in the last two Tasks. Suppose that the single term

$$a'_{1\alpha}a'_{2\beta}a'_{3\gamma}\cdots a'_{n\zeta}$$

is taken from $\det A'$; it corresponds to the permutation which in Cauchy's notation would be represent

$$\left(\begin{array}{c} 1.2.3.\ldots n \\ \alpha.\beta.\gamma.\ldots\zeta \end{array}\right).$$

The sign of the term $a'_{1\alpha}a'_{2\beta}a'_{3\gamma}\cdots a'_{n\zeta}$ in $\det A'$ is the sign of the quantity $n-g$ corresponding to the cycle decomposition of the permutation above. This term is also identical to some term in the expansion of $\det A$ whose associated permutation is the same except that elements μ and ν have been interchanged, and the sign of this term in $\det A$ is similarly the sign of the quantity $n - g$ corresponding to the cycle decomposition of this second permutation. But since n is the same for both terms, our Theorem follows if we can show that the number of cycles in the cycle decomposition of the first permutation differs by 1 from the the number of cycles in the second, which was the pattern we noticed in our examples. To accomplish this, we consider two cases.

First, suppose that μ and ν lie in the same cycle of the first permutation (relative to A'). Let's write down the cycle, starting with element μ, in the form $(\mu \cdots i\nu \cdots j)$. The effect of interchanging μ and ν causes the new permutation (relative to A) to send i to μ and j to ν, whence it must contain the two cycles $(\mu \cdots i)$ and $(\nu \cdots j)$. Since the other cycles of the first permutation contain elements which are unaffected by the interchange, the cycles in which they appear relative to the first permutation also appear unchanged in the cycle decomposition of the second permutation. Thus, the number of cycles in the cycle decomposition of the first permutation is exactly 1 more than the the number of cycles in the second.

In the second case, we assume that μ and ν lie in distinct cycles in the decomposition of the first permutation; then we may assume that among the cycles of the decomposition we find $(\mu \cdots i)$ and $(\nu \cdots j)$. The effect of interchanging μ and ν then causes the new permutation to send i to ν and j to μ, which requires replacing this pair of cycles in the decomposition of the first permutation with the single cycle $(\mu \cdots i\nu \cdots j)$ in the decomposition of the second. Again, the other cycles of the first permutation contain elements which are unaffected by the interchange, so the cycles in which they appear also appear unchanged in the cycle decomposition of the second. Thus, the number of cycles in the cycle decomposition of the original permutation is exactly 1 less than the the number of cycles in the new one. Therefore, our proof is complete.

Task 29 Use the Theorem and Corollary above to prove the following

Theorem. *Suppose that A is an $n \times n$ matrix having either two identical rows or two identical columns. Then*

$$\det A = 0.$$

[Hint: What is the determinant of the matrix obtained from A by interchanging these two rows (or columns)?]

9 The Laplace Adjoint Formula and a Derivation of Cramer's Rule

In the next portion of Cauchy's memoir which we will study, he began by summarizing a few results, first, the theorem just proved in Task 29 and second, a restatement of formula (9).

...We have shown ...[that] D_n will be reduced to zero if, in the expression of this determinant, we replace one of the indices which occupy the second place by another index – for example, if we replace the terms

$$a_{1,\mu}, \quad a_{2,\mu}, \quad \ldots, \quad a_{n,\mu}$$

by

$$a_{1,\nu}, \quad a_{2,\nu}, \quad \ldots, \quad a_{n,\nu},$$

ν being different from μ. Besides, the value of D_n, expressed by means of the terms $a_{1,\mu}, a_{2,\mu}, \ldots, a_{n,\mu}$ is, by the preceding,

$$D_n = a_{1,\mu}b_{1,\mu} + a_{2,\mu}b_{2,\mu} + \cdots + a_{n,\mu}b_{n,\mu}; \tag{14}$$

we will have therefore generally

$$0 = a_{1,\nu}b_{1,\mu} + a_{2,\nu}b_{2,\mu} + \cdots + a_{n,\nu}b_{n,\mu}. \tag{15}$$

This last equation will be satisfied whenever ν and μ are two numbers different from one another.

...If we give successively to μ and to ν, in equations (14) and (15), all the integer values, from 1 to n, we will have the following system of equations:

$$\left\{\begin{array}{rl}
D_n = & a_{1,1}b_{1,1} + a_{2,1}b_{2,1} + \cdots + a_{n,1}b_{n,1}, \\
0 = & a_{1,1}b_{1,2} + a_{2,1}b_{2,2} + \cdots + a_{n,1}b_{n,2}, \\
\ldots, & \\
0 = & a_{1,1}b_{1,n} + a_{2,1}b_{2,n} + \cdots + a_{n,1}b_{n,n}; \\
0 = & a_{1,2}b_{1,1} + a_{2,2}b_{2,1} + \cdots + a_{n,2}b_{n,1}, \\
D_n = & a_{1,2}b_{1,2} + a_{2,2}b_{2,2} + \cdots + a_{n,2}b_{n,2}, \\
\ldots & \\
0 = & a_{1,2}b_{1,n} + a_{2,2}b_{2,n} + \cdots + a_{n,2}b_{n,n}; \\
\ldots & \\
0 = & a_{1,n}b_{1,1} + a_{2,n}b_{2,1} + \cdots + a_{n,n}b_{n,1}, \\
0 = & a_{1,n}b_{1,2} + a_{2,n}b_{2,2} + \cdots + a_{n,n}b_{n,2}, \\
\ldots & \\
D_n = & a_{1,n}b_{1,n} + a_{2,n}b_{2,n} + \cdots + a_{n,n}b_{n,n}.
\end{array}\right. \tag{16}$$

This assembles the necessary tools allowing us to prove the formula hinted at in Task 25(b):

Task 30 Use equations (16) to prove

Theorem (Laplace's Adjoint Formula). *If A is any $n \times n$ matrix and I is the $n \times n$ identity matrix, then*

$$A \cdot (\operatorname{adj} A) = (\operatorname{adj} A) \cdot A = (\det A)I.$$

As a first application of this formula, Cauchy was able to give a full derivation of Cramer's Rule!

§IV. The principal theorems contained in the preceding paragraph have been demonstrated for the first time in a general manner by Mr. Laplace... He has shown how we would be able to apply them to discover the values of the unknowns that many equations of the first degree contain. Let

$$x_1, \quad x_2, \quad \dots, \quad x_n$$

be the n unknowns in question. We suppose that we have among them as many equations of the first degree and that the coefficients of the unknowns are represented respectively by the different terms of the system $(a_{1,n})$. The equations in which we are concerned will be of the form

$$\left\{\begin{array}{rl} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = & m_1, \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = & m_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \dots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n = & m_n. \end{array}\right. \tag{17}$$

Thus, the relations established by equations (16) among the terms of the system $(a_{1,n})$ and those of the adjoint system $(b_{1,n})$ furnish an easy means to obtain the value of any unknown, chosen at will. Thus, for example, if we multiply the first of these equations (17) by $b_{1,1}$, the second by $b_{2,1}$, ..., and finally, the last by $b_{n,1}$ and if we add them together, having regard to equations (16), we will have, so as to determine the value of x_1, the following equation:

$$D_n x_1 = m_1 b_{1,1} + m_2 b_{2,1} + \cdots + m_n b_{n,1}.$$

In general, if we multiply the first of equations (17) by $b_{1,\mu}$, the second by $b_{2,\mu}, \dots,$ and finally, the last by $b_{n,\mu}$, and if next we add them together, we will have, by virtue of equations (16),

$$D_n x_\mu = m_1 b_{1,\mu} + m_2 b_{2,\mu} + \cdots + m_n b_{n,\mu}. \tag{18}$$

Hence, the general value of any one of the unknowns will be

$$x_\mu = \frac{m_1 b_{1,\mu} + m_2 b_{2,\mu} + \cdots + m_n b_{n,\mu}}{D_n} = \frac{m_1 b_{1,\mu} + m_2 b_{2,\mu} + \cdots + m_n b_{n,\mu}}{a_{1,\mu} b_{1,\mu} + a_{2,\mu} b_{2,\mu} + \cdots + a_{n,\mu} b_{n,\mu}}.$$

Thus, his value is presented in the form of a fraction which has as denominator the determinant

$$D_n = \mathbf{S}(\pm a_{1,1} a_{2,2} \dots a_{n,n})$$

and as numerator what this determinant becomes when we replace the coefficients of the unknown taken in the equations (17) by the right sides of these same equations.

Task 31 Explain in detail and in your own words how Cauchy combines equations (17) with (16) to obtain the formulas (18), for $\mu = 1, 2, \dots, n$. Then explain why this is equivalent to Cramer's Rule, which solves in general systems of n linear equations in n unknowns.

10 Determinants and Matrix Multiplication

It is generally recognized (see Feldmann [1962]) that the mathematical theory of matrices was inaugurated by British mathematician Arthur Cayley in the 1850s in a series of papers in which he adopted the word *matrix* as coined by his colleague J. J. Sylvester to describe these rectangular arrays of quantities. In his papers, Cayley also explained how the arithmetic of these objects worked, noting especially that multiplication of matrices was *not* commutative. We now invite the reader to witness in our next excerpt from Cauchy's memoir (published some 40 years before Cayley's work!) how the young Frenchman nonetheless prefigured this new system of arithmetic by describing an early, raw form of matrix multiplication:

§V. We now consider a system of equations of the form

$$\left\{\begin{array}{rl}
\alpha_{1,1}a_{1,1}+\alpha_{1,2}a_{1,2}+\cdots+\alpha_{1,n}a_{1,n}= & m_{1,1},\\
\alpha_{2,1}a_{1,1}+\alpha_{2,2}a_{1,2}+\cdots+\alpha_{2,n}a_{1,n}= & m_{1,2},\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, & \\
\alpha_{n,1}a_{1,1}+\alpha_{n,2}a_{1,2}+\cdots+\alpha_{n,n}a_{1,n}= & m_{1,n};\\
 & \\
\alpha_{1,1}a_{2,1}+\alpha_{1,2}a_{2,2}+\cdots+\alpha_{1,n}a_{2,n}= & m_{2,1},\\
\alpha_{2,1}a_{2,1}+\alpha_{2,2}a_{2,2}+\cdots+\alpha_{2,n}a_{2,n}= & m_{2,2},\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, & \\
\alpha_{n,1}a_{2,1}+\alpha_{n,2}a_{2,2}+\cdots+\alpha_{n,n}a_{2,n}= & m_{2,n};\\
 & \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots; & \\
 & \\
\alpha_{1,1}a_{n,1}+\alpha_{1,2}a_{n,2}+\cdots+\alpha_{1,n}a_{n,n}= & m_{n,1},\\
\alpha_{2,1}a_{n,1}+\alpha_{2,2}a_{n,2}+\cdots+\alpha_{2,n}a_{n,n}= & m_{n,2},\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, & \\
\alpha_{n,1}a_{n,1}+\alpha_{n,2}a_{n,2}+\cdots+\alpha_{n,n}a_{n,n}= & m_{n,n}.
\end{array}\right. \tag{19}$$

This system of equations contains three systems of symmetric quantities, namely: $(a_{1,n}), (\alpha_{1,n})$ and $(m_{1,n})$.

Moreover, it is easy to see that, with regard to equations (19), those which are found in the same group form a collection of symmetric equations in which the quantities involved of the form $a_{\mu,1}, a_{\mu,2}, \ldots, a_{\mu,n}$ have as coefficients the terms of the system $(\alpha_{1,n})$.

...I will call a set of equations (19) a *system of symmetric equations*. Each equation of this system can be put into the form

$$\mathbf{S}^n(\alpha_{\nu,1}a_{\mu,1}) = m_{\mu,\nu}, \tag{20}$$

where the symbol $\mathbf{S}$ is relative to the indices 1 that occupy the second position in $\alpha_{\nu,1}$ and $a_{\mu,1}$, and I will denote the entire system of these equations symbolically by

$$\sum \left[\mathbf{S}^n(\alpha_{\nu,1}a_{\mu,1}) = m_{\mu,\nu}\right]. \tag{21}$$

I will call the two systems of quantities $(a_{1,n}), (\alpha_{1,n})$ whose terms are involved in such equations *bound systems*, and the system of quantities $(m_{1,n})$ whose terms are set apart in their second members a *free system*. I will also denote the two bound systems under the name *component systems*, and the free system under the name *resultant system*. Finally, I will mention that in equations (19), represented by the symbol (21), the resultant system $(m_{1,n})$ is symmetrically determined by means of the two component systems $(a_{1,n})$ and $(\alpha_{1,n})$.

Cauchy organized the n^2 equations presented in (19) in a particular way, arranging them into n groups of n equations each (note that in (19), the groups are separated by semicolons) so that "those which are found in the same group form a collection of symmetric equations in which the quantities involved of the form $a_{\mu,1}, a_{\mu,2}, \ldots, a_{\mu,n}$ have as coefficients the terms of the system $(\alpha_{1,n})$." In the language of modern linear algebra, we would rather say that each of these n groups is a system of linear equations in the variables $a_{\mu,1}, a_{\mu,2}, \ldots, a_{\mu,n}$ with matrix of coefficients $U = (u_{ij})$ with $u_{ij} = \alpha_{i,j}$. For instance, the second group of these equations, corresponding to $\mu = 2$, is

$$\begin{aligned}
\alpha_{1,1}a_{2,1} + \alpha_{1,2}a_{2,2} + \cdots + \alpha_{1,n}a_{2,n} &= m_{2,1},\\
\alpha_{2,1}a_{2,1} + \alpha_{2,2}a_{2,2} + \cdots + \alpha_{2,n}a_{2,n} &= m_{2,2},\\
&\vdots\\
\alpha_{n,1}a_{2,1} + \alpha_{n,2}a_{2,2} + \cdots + \alpha_{n,n}a_{2,n} &= m_{2,n}.
\end{aligned} \tag{22}$$

Each one of these equations has the form

$$\alpha_{\nu,1}a_{\mu,1} + \alpha_{\nu,2}a_{\mu,2} + \cdots + \alpha_{\nu,n}a_{\mu,n} = m_{\mu,\nu}$$

for $\mu = 2$ and for a particular value of $\nu = 1, 2, \ldots, n$, which allowed Cauchy to employ his symmetrizing operator (the exponent n on the symbol $\mathbf{S}$ signaling the number of symbols being symmetrically summed) to represent these equations more briefly in the form (20). He collected all these groups of equations, for $\mu = 1, 2, \ldots, n$, into one in (21), using the capital sigma symbol not in the sense of summation (which only became a standard mathematical notation for addition after Cauchy published this memoir) but rather to suggest the idea of a "system."

Because we are familiar with the language of matrices, we can rewrite the system (22) in matrix-vector form as

$$\begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,n} \\ & & \vdots & \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n} \end{pmatrix} \cdot \begin{pmatrix} a_{2,1} \\ a_{2,2} \\ \vdots \\ a_{2,n} \end{pmatrix} = \begin{pmatrix} m_{2,1} \\ m_{2,2} \\ \vdots \\ m_{2,n} \end{pmatrix}.$$

In fact, for any value of $\mu = 1, 2, \ldots, n$, we obtain n matrix-vector equations of the form

$$\begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,n} \\ & & \vdots & \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n} \end{pmatrix} \cdot \begin{pmatrix} a_{\mu,1} \\ a_{\mu,2} \\ \vdots \\ a_{\mu,n} \end{pmatrix} = \begin{pmatrix} m_{\mu,1} \\ m_{\mu,2} \\ \vdots \\ m_{\mu,n} \end{pmatrix},$$

which can be put together into a single matrix equation

$$\begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,n} \\ & & \vdots & \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n} \end{pmatrix} \cdot \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{n,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{n,2} \\ & & \vdots & \\ a_{1,n} & a_{2,n} & \cdots & a_{n,n} \end{pmatrix} = \begin{pmatrix} m_{1,1} & m_{2,1} & \cdots & m_{n,1} \\ m_{1,2} & m_{2,2} & \cdots & m_{n,2} \\ & & \vdots & \\ m_{1,n} & m_{2,n} & \cdots & m_{n,n} \end{pmatrix}. \tag{23}$$

We have already denoted the first matrix on the left in (23) as U, so let's define $V = (v_{ij})$ and $W = (w_{ij})$ by

$$v_{ij} = a_{j,i} \qquad \text{and} \qquad w_{ij} = m_{j,i}.$$

(Notice the orders of the indices – this is quite deliberate!) Now, (23) is greatly simplified by writing

$$UV = W. \tag{24}$$

In this final excerpt we will consider from his memoir, Cauchy carefully argued that if D_n, δ_n, M_n are, respectively, the determinants of the "symmetric systems" $(a_{1,n}), (\alpha_{1,n}), (m_{1,n})$, then the relationship (21) that makes $(m_{1,n})$ the "resultant system" of the two "component systems" $(a_{1,n})$ and $(\alpha_{1,n})$ also guarantees that $D_n \cdot \delta_n = M_n$. In matrix terms, since

$$D_n = \det U, \quad \delta_n = \det V, \quad M_n = \det W,$$

the relation (24) allows us to state this result in the form:

Theorem. *Suppose that U, V, W are $n \times n$ matrices. If $UV = W$, then*

$$\det U \cdot \det V = \det W. \tag{25}$$

In other words, the determinant of the product of two $n \times n$ matrices is the product of their determinants.

We denote respectively by

$$D_n, \quad \delta_n, \quad M_n$$

the determinants of the three systems

$$(a_{1,n}), \quad (\alpha_{1,n}), \quad (m_{1,n});$$

we obtain

$$\begin{cases} D_n &= \mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n}), \\ \delta_n &= \mathbf{S}(\pm \alpha_{1,1}\alpha_{2,2}\ldots \alpha_{n,n}), \\ M_n &= \mathbf{S}(\pm m_{1,1}m_{2,2}\ldots m_{n,n}). \end{cases} \tag{26}$$

In each of the three preceding equations, the symbol $\mathbf{S}$ can be considered as relative to the indices that occupy the first position, or to those that occupy the second. Thus,

$$\mathbf{S}(\pm m_{1,1}m_{2,2}\ldots m_{n,n})$$

is an alternating symmetric function with respect to both of the species of indices which affect the terms of the system $(m_{1,n})$. While equations (21) are all included in the general formula

$$\mathbf{S}^n(\alpha_{\nu,1}a_{\mu,1}) = m_{\mu,\nu},$$

the two indices which in each of the equations affect the letter m are respectively equal to the first indices which in these same equations affect the two letters a and α, that is, the terms of the two systems $(a_{1,n})$ and $(\alpha_{1,n})$. It follows from this remark that if in the right side of the equation

$$M_n = \mathbf{S}(\pm m_{1,1}m_{2,2}\ldots m_{n,n}),$$

we substitute for the terms of the system $(m_{1,n})$ those values in a and α derived from equations (19), we will obtain as a result a function of the systems $(a_{1,n})$ and $(\alpha_{1,n})$, which will be alternating symmetric with respect to the indices that occupy the first position in a and with respect to those that occupy the first position in α. Since each of the quantities m is of the first degree with respect to the quantities a and with respect to the quantities α, each term in the development of

$$\mathbf{S}(\pm m_{1,1}m_{2,2}\ldots m_{n,n})$$

will evidently be of the form

$$\pm\alpha_{1,\mu}\alpha_{2,\nu}\ldots\alpha_{n,\pi}a_{1,\mu}a_{2,\nu}\ldots a_{n,\pi}.$$

As this development must be an alternating symmetric function with respect to the indices that occupy the first position in a and with respect to those that occupy the first position in α, it cannot contain the aforementioned term without also at the same time containing the product

$$\pm\mathbf{S}(\pm\alpha_{1,\mu}\alpha_{2,\nu}\ldots\alpha_{n,\pi})\mathbf{S}(\pm a_{1,\mu}a_{2,\nu}\ldots a_{n,\pi}),$$

so it will therefore be equivalent to one or more products of this type.

If, in the preceding product, we assume that the indices $\mu,\nu,\ldots,\pi$ are all different from each other, we will have

$$\mathbf{S}(\pm\alpha_{1,\mu}\alpha_{2,\nu}\ldots\alpha_{n,\pi}) = \pm\mathbf{S}(\pm\alpha_{1,1}\alpha_{2,2}\ldots\alpha_{n,n}) = \pm\delta_n.$$

Likewise, if we assume that the indices $\mu',\nu',\ldots,\pi'$ are all different from each other, we will have

$$\mathbf{S}(\pm a_{1,\mu'}a_{2,\nu'}\ldots a_{n,\pi'}) = \pm\mathbf{S}(\pm a_{1,1}a_{2,2}\ldots a_{n,n}) = \pm D_n.$$

Moreover, we cannot suppose that two of the indices $\mu,\nu,\ldots,\pi$ are equal to each other without having

$$\mathbf{S}(\pm\alpha_{1,\mu}\alpha_{2,\nu}\ldots\alpha_{n,\pi}) = 0,$$

nor two of the indices $\mu',\nu',\ldots,\pi'$ equal to each other without having

$$\mathbf{S}(\pm a_{1,\mu'}a_{2,\nu'}\ldots a_{n,\pi'}) = 0.$$

Therefore, the development of M_n will be reduced to one or more products of the form

$$\pm D_n \delta_n;$$

so we have

$$M_n = c D_n \delta_n,$$

c being a constant quantity. To determine the constant, it suffices to observe that the preceding equation will have to hold identically if we suppose in general that

$$\alpha_{\mu,\mu} = 1, \qquad a_{\mu,\mu} = 1, \qquad \alpha_{\mu,\nu} = 0, \qquad a_{\mu,\nu} = 0;$$

but then we have in equations (19)

$$m_{\mu,\mu} = 1, \qquad m_{\mu,\nu} = 0,$$

and consequently that

$$M_n = 1, \qquad D_n = 1, \qquad \delta_n = 1;$$

so then we also have that

$$c = 1,$$

and therefore we have in general that

$$M_n = D_n \delta_n. \tag{27}$$

This equation contains a quite remarkable theorem which we express in the following manner:

When a system of quantities is symmetrically determined by means of two other systems, the determinant of the resultant system is always equal to the product of the determinants of the two component systems.

11 Determinants and Row Reduction

We have looked at three procedures for computing values of nth order determinants:

- Cramer's method of summing terms – one for each permutation of indices $\{1, 2, \ldots, n\}$ (which we viewed as superscripts in his notation) attached to products of coefficients of a system of n linear equations in n unknowns, or what in modern parlance are the entries of a square matrix of the system's coefficients – in which the signs of the terms being summed are given by the parity of the number of derangements appearing in each permutation [Section 1 above];

- Cauchy's expansion, the sum of products of n entries chosen from an $n \times n$ matrix so that their index pairs run through all possible permutations of the second set of indices given by the first set, and in which the signs of the terms of the sum are determined by the parity of the quantity $n - g$, where g is the number of cycles that appear in the cycle decomposition of the permutation associated with the given term [Section 2]; and also

- Laplace's expansion, realizing the determinant as the sum of products formed by multiplying the entries along a single row (or column) of the matrix with their cofactors, which are themselves determinants of the minors formed by eliminating the row and column of that entry of the original matrix, with additional signs of these products determined by the parity of the sum of the row and column index; this reduces the computation of the nth order determinant to the computation of (in general) n determinants of order $n-1$ [Section 6].

Now in practice, none of these methods is particularly computationally efficient, especially in the case of matrices that are either large or not very sparse. Instead, with the advent of programmable computing in the twentieth century, mathematicians realized that a much more efficient way to compute a determinant was to exploit the Gaussian elimination method of row reduction to echelon form, which was engineered for the purpose of solving systems of linear equations. We close this project by explaining how the method of row reduction yields an efficient method for computing determinants, by making critical use of the determinant product formula (25).

Recall that, given a system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= c_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= c_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= c_m, \end{aligned} \tag{28}$$

of m linear equations in n unknowns with matrix of coefficients

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

in order to solve for the unknowns

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

we form an *augmented* matrix by attaching the column

$$C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

to the right of A, then subjecting this augmented matrix to a sequence of *elementary row operations*, which are of three simple types:

I. interchange row i with row j;

II. multiply the entries of row i by a constant k;

III. replace row i by its sum with k times the corresponding entries of row j.

The goal of this row reduction process is to bring the augmented matrix to *row echelon form*; a matrix is in this form when

1. all rows in the matrix containing only zeros appear below any rows containing nonzero entries; and

2. the *leading entry* in each row, that is, the first nonzero entry, reading left to right, lies below and to the right of the leading entries of the rows above it.

If in addition, we require that

3. all leading entries equal 1; and

4. all entries above leading entries equal 0,

then we say that the matrix is in *reduced row echelon form.* Because the application of elementary row operations does not change the set of solutions of the original system of linear equations, bringing the augmented matrix to reduced row echelon form is valuable, because it is then possible to read off the solutions to the system immediately from the row reduced augmented matrix. We assume that the reader is familiar with the row reduction process so we shall not review it here. We are interested in row reduction in the computation of determinants.

To this end, take A to be the matrix whose determinant we seek. In particular, we must have $m = n$, since A must be a square matrix. Suppose we apply an elementary row operation to A resulting in the $n \times n$ matrix B; we denote this by writing

$$A \xrightarrow{I} B, \quad A \xrightarrow{II} B \quad \text{or} \quad A \xrightarrow{III} B,$$

depending on whether the row operation performed is of type I, II, or III.

Task 32

(a) Use the Theorem on page 31 to show that if $A \xrightarrow{I} B$, then $\det B = -\det A$.

(b) Suppose that $A \xrightarrow{II} B$. Show that $B = E[i, i; k] \cdot A$ for some choice of index i and constant k. Use the product formula (25) and the result of Task 17(d) to conclude that $\det B = k \cdot \det A$.

(c) Now assume that $A \xrightarrow{III} B$. Show that $B = E[i, j; k] \cdot A$ for some pair of indices i and j and constant k. Conclude as in part (b) that $\det B = \det A$.

The result of Task 32 is significant: if we execute a sequence of elementary row operations upon a given square matrix A to bring it into row echelon form as the matrix A' then

$$\det A' = p \cdot \det A, \tag{29}$$

where the value of p can be determined by keeping track of the various row operations employed along the way: p is the product of all the constant factors k that arose from operations of type II which were performed, modified by the factor $(-1)^q$, where q is the number of row operations of type I performed. Since the matrix A' is in assumed to be in row echelon form, it is also upper triangular, so the result of Task 17(a) can be used to find $\det A'$. Equation (29) thereby gives an effective, and algorithmically efficient, way to calculate $\det A$.

Task 33 Use the row reduction method described above to once more find the determinant of the matrix given in Task 22(c).

References

Bruno Belhoste. *Augustin-Louis Cauchy: a biography.* Springer-Verlag, New York, Berlin, 1991.

Augustin-Louis Cauchy. Functions qui ne peuvent obtenir que deux valeurs égales et de signes contraires par suit des transpositions opérées entre les variables qu'elles renferment. *Journal de l'École Polytechnique*, 1815. Also *Oeuvres de Cauchy*, Series II, Tome 1, 91-169. All English translations in this PSP by Daniel E. Otero, 2017.

Gabriel Cramer. *Introduction à l'Analyse des Lignes Courbes Algébriques.* Cramer & Philibert, Geneve, 1750. English translations from the French by Richard J. Pulskamp, unpub., 2008.

Richard W. Feldmann, Jr. I. Arthur Cayley – founder of matrix theory. *The Mathematics Teacher*, 55(6):482–484, October 1962.

Notes to Instructors

PSP Content: Topics and Goals

The standard first course in linear algebra[10] typically begins with units on solving systems of linear equations via Gaussian elimination and a study of the structure of the solution sets of such systems. This is followed by an introduction to vectors, matrices and their arithmetic, which provide a powerful symbolic language in which to describe these concepts algebraically as well as to discover new ones. By this time, students in a semester-long course are near the midsemester date. At this point most courses will introduce the matrix determinant as a mysterious number associated with a square matrix, together with a handful of intricate formulas for computing determinants, formulas that (strangely) seem to depend on the size of the matrix. The goal of this discussion seems to be, at least at first, the theorem that the determinant of an $n \times n$ matrix is nonzero precisely when a system of equations for which it is the coefficient matrix has a unique solution in $\mathbf{R}^n$. Later, the student may learn (often in a vector calculus course) that the determinant is a measure of the dilation of length, area or volume in the cases $n = 1, 2, 3$, and then by extension to higher order Euclidean spaces, produced by the linear transformation of $\mathbf{R}^n$ given by multiplication of vectors by this matrix.

This PSP is designed to replace the standard textbook chapter on determinants. We begin by presenting Cramer's Rule – in Cramer's own words – to first connect the student's knowledge of solving systems of linear equations to the algebraic expression of the determinant, at least in low-degree cases. Then, using Cauchy's masterful memoir, the general determinant is laid out and its fundamental properties described, the most important of which being the product rule $\det AB = \det A \cdot \det B$. Surprisingly, even though Cauchy's treatment was written decades before the appearance of matrices, he defines equivalent objects called "symmetric systems" which allows him to state (and prove) the product law.

Student Prerequisites

Before tackling this project, students should already be familiar with

- Gaussian elimination procedures, the three row reduction operations used to determine solutions to linear systems of equations;
- vectors and matrices, and their arithmetic; and
- the matrix transpose.

PSP Design, and Task Commentary

The project is divided into 11 sections:

1. **Cramer's Rule**: In an effort to use algebraic methods to solve Pappus' problem of five points in his *Introduction a l'Analyse des lignes courbes algébriques* (1750), Cramer is led to a system of five linear equations in five unknowns, for which he would like formulas to compute their values. This naturally leads him to an analysis of the general problem of solving a linear system of n equations in n unknowns. His choice of notation for describing such a system

[10]There are literally hundreds of textbook presentations of this first course in linear algebra (a course that only became standard in the undergraduate mathematics curriculum in the 1950s!). This project should work in concert with most of them.

alerts him to patterns in the expressions of the formulas he derives, each of which is a quotient of similar polynomial expressions. These expressions will be recognized later as determinants. They are presented here so that students meet determinantal expressions in a familiar context, and so that they gain experience in manipulating the complicated but symmetrical forms that determinants take on. Tasks 3, 4 and 5 are central to this discussion, as they allow students to enter into Cramer's unfamiliar notational universe in the context of what should be the familiar situation of two linear equations in two unknowns. Cramer develops a method for writing down these expressions that introduces the concept of derangements of a permutation. Task 8 should only be undertaken by the hearty, as it works out the 24-term determinant of order 4.

2. **Cauchy's Determinant: Alternating Symmetric Functions**: The rest of this project immerses students in Cauchy's masterful memoir on the determinant, beginning here with his invention of an operator $\mathbf{S}(\pm\,\cdot\,)$ that generates alternating symmetric algebraic expressions from a single "indicative term." The focus is immediately trained on determinantal expressions. Task 12 is crucial for understanding how Cauchy's operator works and to appreciate the interplay between terms of the expression and the permutations of the indices of the variables that appear there. Cauchy's "symmetric systems" are equivalent to our matrices, making it possible for students to read Cauchy's analysis of the determinant without much difficulty. Task 15, the most important of the entire project, allows the student to identify the parity of each term of the order-4 determinant using Cauchy's cycle-decomposition method.

3. **A Modern Approach for 2×2 and 3×3 Matrices**: A pause from Cauchy's memoir is taken so that students can recognize the popular formulas for 2×2 and 3×3 determinants using diagonal and skew diagonal products of entries (and to see why these techniques cannot be extended to higher order determinants).

4. **Determinants of Upper Triangular Matrices**: Students can begin to identify important properties of determinants by dealing with the common classes of matrices used in further study. Task 17 shows students that the determinant of an upper (or lower) triangular matrix is the product of its diagonal entries. It also introduces students to the important class of elementary matrices, which will show up again in the last section of the project.

5. **Conjugate Systems and Transposed Matrices**: Cauchy explains why $\det A^T = \det A$.

6. **Laplace Expansion**: In what is the most difficult section of the project, Cauchy shows that the nth order determinant formula can be realized as a sum of the products of entries of a single row of the matrix with cofactor expressions, each of which is an appropriately signed determinant of degree $n-1$. Tasks 19 and 20 are engineered to help the student deconstruct the Laplace expansion in the order 4 case; Task 21 puts the deconstructed elements back together. There is a lot of detail work here. Expect students to become frustrated with it. In addition to having to worry about the signs attached to certain expressions, Cauchy's notation for entries in his symmetric systems is reversed to modern notation for entries of matrices: the element $a_{i,j}$ in a symmetric system places it in "vertical sequence" i and "horizontal sequence" j ! This adds another layer of difficulty to the reinterpretation of Cauchy's results into modern form. This is why a single example matrix in Task 22(b) is used repeatedly to test the various ways

in which students learn how to compute a determinant. The discussion culminates in Task 25, where students are asked to state the general result.

7. **The Classical Adjoint**: Cauchy defines the classical adjoint, or adjugate, $\operatorname{adj} A$ of a matrix A. Task 26 poses the question "What is the significance of the matrix product $A \cdot (\operatorname{adj} A)$?" This question is properly addressed in section 9.

8. **The Effect of Interchanging Rows/Columns on the Determinant**: Tasks 27 and 28 give the student intuition that helps them to follow the proof of the theorem that interchanging a pair of rows (or columns) of a given matrix changes only the sign of the determinant of that matrix.

9. **The Laplace Adjoint Formula and a Derivation of Cramer's Rule**: Cauchy resolves the question from section 7, and uses the adjoint formula to derive Cramer's Rule in general form.

10. **Determinants and Matrix Multiplication**: We leave Cauchy's memoir with his statement and proof of the determinant product rule. This result will be used to great effect in the next section.

11. **Determinants and Row Reduction**: In Task 32, students use the product formula to ascertain the effect on the determinant of a matrix of performing an elementary row operation on the matrix. This leads to the most effective method for computing determinants, one that uses row reduction to bring the matrix to upper triangular form.

Suggestions for Classroom Implementation

The Sample Implementation Schedule provided below illustrates that the entire PSP should take about three full weeks to work through, not counting advance reading before the first class period or final "mop-up" homework after the last. Instructors are encouraged to modify this schedule to fit their curricular needs by using a portion (or portions) of the PSP. See the section below on Possible Modifications of the PSP for my suggestions.

Students in a first linear algebra course tend to have only minimal experience with proving theorems, so I have worked to scaffold their approach to the small number of theorems that arise in this PSP. I have them work some rather meaty examples (mostly having to do with permutations of the four-element set $\{1, 2, 3, 4\}$) to give them some intuition regarding the kind of order that these theorems bring to the behavior of permutations or of computing determinants. The first theorem of note they encounter is Cramer's Rule, which Cramer merely states for them (but does not prove). I have the students work out the algebraic steps to derive Cramer's Rule for systems of n equations in n unknowns for the cases $n = 1, 2, 3, 4$; the general case will be dealt with near the end of the PSP by Cauchy.

There are tasks in the PSP that are not straightforward exercises

Notice in the Sample Implementation Schedule which Tasks are recommended for use in the classroom; these Tasks are best done in small groups of students, often so that each can oversee the work of the others and minimize the inevitable production of errors that an unmonitored worker tends to make.

Another curious feature of this material is the ubiquity of distinct letters for quantities, and the indices which adorn them – superscripts or subscripts, single or double, separated by commas or not. Your students will likely have never dealt with so many indices! And the level of detail they must handle is considerable in organizing all the permutations of order 4. Caution them that every symbol they encounter matters, especially the tiny ones!

While effort is made to develop multiple methods for deciding whether a permutation is even or odd, the terms "even" and "odd" are not used in the PSP; instead we stick to Cauchy's language of "determining the sign" of the permutation.

Two more specific comments:

- Task 26 require students to compute the adjugate of a 4×4 matrix, and then observe the result of multiplying the original matrix by its adjugate – in both directions. The exercise is instructive, but it will take some time to complete. I suggest assigning it for homework (see the Sample Implementation Schedule below), but in the classroom before they work on this, have them compute the adjugate of a 3×3 matrix M of your choosing, and then also compute $M \cdot (\operatorname{adj} M)$ and $(\operatorname{adj} M) \cdot M$.

- Section 10 contains no student Tasks, but does include Cauchy's argument to prove the determinant product rule. Require students to carefully read the argument before taking time in class to discuss how Cauchy comes to his conclusions (p. 38ff).

LaTeX code of this entire PSP is available from the author by request to facilitate preparation of advanced preparation / reading guides or 'in-class worksheets' based on tasks included in the project. The PSP itself can also be modified by instructors as desired to better suit their goals for the course.

Possible Modifications of the PSP

The simplest way to modify this PSP is to plan to use only a part of it with your students. Here are some suggestions (with estimated timetables):

- Use only Section 1, which addresses Cramer's Rule as a motivation for studying the determinant and as an introduction to the algebraic form of the general determinant formula, but omits the entirely of Cauchy's memoir (two 50-minute classes).

- Omit Sections 1 and 6, following only Cauchy's text, but excluding the long section on Laplace expansions (five 50-minute classes).

- A version that emphasizes the historical approaches taken by Cramer and Cauchy and sets aside my attempts to tie their work into a modern exposition of methods for computing the determinant will jettison Sections 3, 4 and 11 (seven 50-minute classes).

Sample Implementation Schedule (based on a 50 minute class period)

The actual number of class periods spent on each section of this PSP naturally depends on the instructor's goals and on how the PSP is implemented with students. The schedule below assumes that much of the PSP work is completed by students working in small groups during class time.

Day	Preparatory Homework	Classroom Plan
1	Read through §1	Students work through Tasks 1-4 with open discussion as necessary
2	Do Tasks 5-6	Review student questions on Tasks 5-6; students work through Tasks 7-8
3	Complete Tasks 7-8; read §2 and do Tasks 9-10	Discuss Task 10; students work through Tasks 11-12, 15(a)
4	Do Tasks 13, 14, 15(b); read §3, §4, §5	Start Tasks 16, 17, 18
5	Finish Tasks 16, 17, 18, and do Task 19; read pp. 22-25 of §6	Discuss Task 19; students work through Task 20
6	Finish Tasks 19 and 20; read pp. 26-28 of §6	students work through Tasks 21 and 22(a)
7	Do Tasks 22(b) and 23-25; read §7 and §8	Have students compute the adjugate of a 3×3 matrix; students work through Task 27
8	Do Tasks 26, 28, 29; read §9	Discuss Tasks 26 and 29; students work through Tasks 30 and 31, with open discussion as necessary
9	Read §10 (studying Cauchy's argument carefully) and §11	Discuss Cauchy's argument; students work through Task 32
	Final assignment: Do Task 33	

Connections to other Primary Source Projects

Users of this PSP may also be interested in another TRIUMPHS PSP created by Mary Flagg (University of St. Thomas, Houston), "Solving a System of Linear Equations Using Ancient Chinese Methods." In this PSP, students learn Gaussian elimination through the counting board methods of the ancient Chinese.

Acknowledgments

The development of this project has been partially supported by the National Science Foundation's Improving Undergraduate STEM Education Program under Grants No. 1523494, 1523561, 1523747, 1523753, 1523898, 1524065, and 1524098. Any opinions, findings, and conclusions or recommendations expressed in this project are those of the author and do not necessarily reflect the views of the National Science Foundation.

For more information about TRIUMPHS, visit `http://webpages.ursinus.edu/nscoville/TRIUMPHS.html`.