Abstract awakenings in algebra: Early group theory in the works of Lagrange, Cauchy, and Cayley

Janet Heine Barnett[*]

19 August 2010

Introduction

The problem of solving polynomial equations is nearly as old as mathematics itself. In addition to methods for solving linear equations in ancient India, China, Egypt and Babylonia, solution methods for quadratic equations were known in Babylonia as early as 1700 BCE. Written out entirely in words as a set of directions for calculating a solution from the given numerical coefficients, the Babylonian procedure can easily be translated into the well-known quadratic formula: $x = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$, given that $ax^2+bx+c=0$. But what about a 'cubic formula' for third degree equations, or a 'quartic formula' for equations of degree four? More generally, is it always possible to compute the solutions of polynomial equations of a given degree by means of a formula based on its coefficients, elementary arithmetic operations and the extraction of roots alone?

As basic as this latter question may sound, its answer eluded mathematicians until the early nineteenth century. In connection with its rather surprising resolution (revealed later in this introduction), there emerged a new and abstract algebraic structure known as a *group*. Algebra, understood prior to that time as the study of solution techniques for equations, was forever changed as a result. The goal of this project is to develop an understanding of the basic structure and properties of a group by examining selected works of mathematicians involved in the early phases of its evolution. In the remainder of this introduction, we place these readings in context with a brief historical sketch of efforts to find formulas for higher degree polynomial equations.

In a sense, the search for the solution to higher degree polynomials was also pursued by ancient Babylonian mathematicians, whose repertoire included methods for approximating solutions of certain types of cubic equations. The problem of finding exact (versus approximate) solutions can be said to have originated within the somewhat later geometric tradition of ancient Greece, in which there arose problems such as the 'duplication of a cube' (i.e., constructing a cube with twice the volume of a given cube) which correspond to cubic equations when translated into today's algebraic symbolism (i.e., $x^3 = 2$). In order to construct the line segments which served as geometrical roots of these equations, Greek mathematicians developed various new curves, including the conic sections (i.e., parabolas, hyperbolas and ellipses).[1] The Islamic mathematician and poet Omar Khayyam (1048–1131), who was the first to systematize these geometric procedures, solved several different cubic equation forms by intersecting appropriate conic sections. For example, for a cubic equation of the form $x^3 + d = bx^2$, the points of intersection of the hyperbola $xy = d$ and the parabola $y^2 + dx - db = 0$ correspond to

[*] Department of Mathematics and Physics; Colorado State University-Pueblo; Pueblo, CO 81001-4901; janet.barnett@colostate-pueblo.edu.

[1] Using abstract algebra, it can now be shown that certain of these construction problems, including the duplication of the cube, are impossible using only the Euclidean tools of a collapsible compass and an unmarked straightedge. Likewise, these ideal Euclidean tools are not sufficient to construct a conic section.

the real roots of the polynomial.[2] Because negative numbers were not allowed as coefficients (or roots) of equations at this point, however, the polynomial $x^3 + d = bx^2$ was only of thirteen possible cubic forms, each of which required different conic sections for their solution.

Unlike the Greeks, Khayyam and his fellow medieval Islamic mathematicians hoped to find algebraic algorithms for cubic equations, in addition to geometric constructions based on curves. Although they were unsuccessful in this regard, it was through Islamic texts that algebra became known in Western Europe. As a result, the search for an algebraic method of solution for higher degree equations was next taken up in Renaissance Italy beginning in the 14th century. Some equations studied in this setting were solved through substitutions which reduced the given equation to a quadratic or to a special form like $x^n = a$. Most higher degree equations, however, can not be solved this way. Only in the 16th century, when methods applying to all cubic and quartic polynomials were finally developed, did the Italian search for general solutions achieve some success. The culmination of this search — which we describe in footnote 3 below — is one of the great stories in the history of mathematics.[3]

Although the solution methods discovered in the 16th century are interesting in and of themselves, certain consequences of their discovery were of even greater importance to later developments in mathematics. One such consequence was the discovery of complex numbers, whose eventual acceptance was promoted by the fact that roots of negative numbers often appear during the course of applying the cubic or quartic formulas to specific equations, only to cancel out and leave only real roots in the end.[4] The increased use of symbolism which characterized Western European algebra in the Renaissance period also had important consequences, including the relative ease with which this symbolism allowed theoretical questions to be asked and answered by mathematicians of subsequent generations. Questions about the number and kind of roots that a polynomial equation possesses, for example, led to the Fundamental Theorem of Algebra, the now well-known assertion that an n^{th} degree equation has exactly n roots, counting complex and multiple roots. Algebraic symbolism also allowed questions about the relation of roots to factors to be carefully formulated and pursued, thereby leading to discoveries like the Factor Theorem, which states that r is a root of a polynomial if and only if $(x - r)$ is one of its factors.

Despite this progress in understanding the theory of equations, the problem of finding a solution to equations of degree higher than four resisted solution until the Norwegian Niels Abel (1802–1829) settled it in a somewhat unexpected way. In a celebrated 1824 pamphlet, Abel proved that a 'quintic formula'

[2] To verify this, substitute $y = \frac{d}{x}$ from the equation of the hyperbola into the equation of the parabola, and simplify. Of course, modern symbolism was not available to Khayyam, who instead wrote out his mathematics entirely in words.

[3] Set in the university world of 16th century Italy, where tenure did not exist and faculty appointments were influenced by a scholar's ability to win public challenges, the tale of the discovery of general formulas for cubic and quartic equations begins with Scipione del Ferro (1465–1526), a professor at the University of Bologna. Having discovered a solution method for equations of the form $x^3 + cx = d$, del Ferro guarded his method as a secret until just before his death, when he disclosed it to his student Antonio Maria Fiore (ca. 1506). Although this was only one of thirteen forms which cubic equations could assume, knowing its solution was sufficient to encourage Fiore to challenge Niccolò Tartaglia of Brescia (1499–1557) to a public contest in 1535. Tartaglia, who had been boasting he could solve cubics of the form $x^3 + bx^2 = d$, accepted the challenge and went to work on solving Fiore's form $x^3 + cx = d$. Finding a solution just days before the contest, Tartaglia won the day, but declined the prize of 30 banquets to be prepared by the loser for the winner and his friends. (Poisoning, it seems, was not an unknown occurrence.) Hearing of the victory, the mathematician, physician and gambler Gerolamo Cardano (1501–1576) wrote to Tartaglia seeking permission to publish the method in an arithmetic book. Cardano eventually convinced Tartaglia to share his method, which Tartaglia did in the form of a poem, but only under the condition that Cardano would not publish the result. Although Cardano did not publish Tartaglia's solution in his arithmetic text, his celebrated 1545 algebra text *Ars Magna* included a cubic equation solution method which Cardano claimed to have found in papers of del Ferro, then 20 years dead. Not long after, a furious Tartaglia was defeated in a public contest by one of Cardano's student, Lodovico Ferrari (1522–1565), who had discovered a solution for equations of degree four. The cubic formula discovered by Tartaglia now bears the name 'Cardano's formula.'

[4] The first detailed discussion of complex numbers and rules for operating with them appeared in Rafael Bombelli's *Algebra* of 1560. Their full acceptance by mathematicians did not occur until the nineteenth century, however.

for the general fifth degree polynomial is impossible.[5] The same is true for equations of higher degree, making the long search for algebraic solutions to general polynomial equations perhaps seem fruitless. Then, as often happens in mathematics, Abel's 'negative' result produced fruit. Beginning with the central idea of Abel's proof — the concept of a 'permutation' — the French mathematician Évariste Galois (1811–1832) used a concept which he called a 'group of permutations' as a means to classify those equations which are solvable by radicals. Soon after publication of Galois' work, permutation groups in the sense that we know them today[6] appeared as just one example of a more general group concept in the 1854 paper *On the theory of groups, as depending on the symbolic equation* $\theta^n = 1$, written by British mathematician Arthur Cayley (1821–1895). Although it drew little attention at the time of its publication, Cayley's paper has since been recognized as the inaugural paper in abstract group theory. By the end of the nineteenth century, group theory was playing a central role in a number of mathematical sub-disciplines, as it continues to do even today.

Cayley's ground-breaking paper forms the centerpiece of our study of groups in this project. In keeping with the historical record, and to provide a concrete example on which to base our abstraction, we first study the particular example of permutation groups. We begin in Section 1 with selections from the writings the 18th century French mathematician J. L. Lagrange (1736–1813), the first to suggest a relation between permutations and the solution of equations by radicals. In Section 2, we continue with selections from the writings of another French mathematician, Augustin Cauchy (1789–1857), in which a more general theory of permutations was developed independently of the theory of equations. We then turn to a detailed reading of Cayley's paper in Sections 3 and 4, paying careful attention to the similarities between the theory of permutation groups as it was developed by Cauchy and the modern notion of an abstract group as it was unveiled by Cayley.

1 Roots of unity, permutations and equations: J. L. Lagrange

Born on January 25, 1736 in Turin, Italy to parents of French ancestry, Joseph Louis Lagrange was appointed as a professor of mathematics at the Turin Royal Artillery School by the age of 19. He spent the first eleven years of his professional life teaching there, the next twenty as a mathematical researcher at the Berlin Academy of Sciences, and the final twenty seven as both a teacher and researcher in Paris, where he died on April 10, 1813. Largely self taught, he is remembered today for contributions to every branch of eighteenth century mathematics, and especially for his work on the foundations of calculus. His work in algebra is also recognized as sowing one of the seeds that led to the development of group theory in the nineteenth century.

Lagrange's work on the algebraic solution of polynomial equations was first published as a lengthy article entitled 'Réflexions sur la résolution algébrique des équations' in the *Mémoire de l'Académie de Berlin* of 1770 and 1771. As with all his research, generality was Lagrange's primary goal in his works on equations. In seeking a general method of algebraically solving all polynomial equations, he began by looking for the common features of the known solutions of quadratics, cubics, and quartics. Following a detailed analysis of the known methods of solution, he concluded that one thing they had in common was the existence of an auxiliary (or *resolvent*) equation whose roots (if they could be found) would allow one to easily find the roots of the originally given equation. Furthermore, Lagrange discovered that it was always possible, for any given equation, to find a resolvent equation with roots which are related to the roots of the original equation in a very special way. This special relationship, and a summary of its importance to his work, are nicely described in Lagrange's own words in the

[5] The Italian mathematician Paolo Ruffini published this same result in 1799, but his proposed proof contained a gap. Abel, who was not aware of Ruffini's work until 1826, described it as "so complicated that it is very difficult to decide the correctness of his reasoning" (as quoted in [14, p. 97]).

[6] Galois used the term 'group' in a slightly different way than we do today.

following excerpt from the 1808 edition of his *Traité de la résolution des équations numériques* [11].[7]

∞∞∞∞∞∞∞∞∞∞∞

On the solution of algebraic equations[8]

The solution of second degree equations is found in Diophantus and can also be deduced from several propositions in Euclid's *Data*; but it seems that the first Italian algebraists learned of it from the Arabs. They then solved third and fourth degree equations; but all efforts made since then to push the solution of equations further has accomplished nothing more than finding new methods for the third and the fourth degree, without being able to make a real start on higher degrees, other than for certain particular classes of equations, such as the reciprocal equations, which can always be reduced to a degree less than half [the original degree] ...

In *Mémoire de l'Académie de Berlin* (1770, 1771), I examined and compared the principal methods known for solving algebraic equations, and I found that the methods all reduced, in the final analysis, to the use of a secondary equation called the *resolvent*, for which the roots are of the form

$$x' + \alpha x'' + \alpha^2 x''' + \alpha^3 x^{(iv)} + \ldots$$

where $x', x'', x''', \ldots$ designate the roots of the proposed equation, and α designates one of the roots of unity, of the same degree as that of the equation.

I next started from this general form of roots, and sought *a priori* for the degree of the resolvent equation and the divisors which it could have, and I gave reasons why this [resolvent] equation, which is always of a degree greater than that of the given equation, can be reduced in the case of equations of the third and the fourth degree and thereby can serve to resolve them.

∞∞∞∞∞∞∞∞∞∞∞

Lagrange's emphasis on the *form* of the roots of the resolvent equation is a mark of the more abstract approach he adopted throughout his mathematics, and a critical piece of his analysis in this particular study. Although we will not go through the analysis which led him to this form, we will read several arguments which he based on it in Subsection 1.2 below.[9] Notice for now that the expression for the roots t of the resolvent equation for a polynomial of degree m is actually the sum of only finitely many terms:

$$t = x' + \alpha x'' + \alpha^2 x''' + \alpha^3 x^{(iv)} + \ldots + \alpha^{m-1} x^{(m)}.$$

Lagrange used prime notation $(x', x'', x''', \ldots)$ here instead of subscripts $(x_1, x_2, x_3, \ldots)$ to denote the m roots of the original equation (and not to indicate derivatives). Similarly, $x^{(m)}$ denotes one of these m

[7] As suggested by its title, *Traité de la résolution des équations numériques* primarily addressed the problem of numerically approximating solutions to equations, but it also included a summary of Lagrange's 1770/1771 work on finding exact solutions to polynomial equations as an appendix entitled 'Note XIII: Sur le résolution des équations algébriques' (see [11, pp. 295-327]). The first edition of *Traité de la résolution des équations numériques* appeared in 1798; the revised and extended 1808 edition from which the excerpts in this project were taken is considered the definitive edition.

[8] To set them apart from the project narrative, all original source excerpts are set in sans serif font and bracketed by the following symbol at their beginning and end: ∞∞∞∞∞∞∞∞∞∞∞

[9] As indicated in the excerpt, Lagrange arrived at this expression for the roots of the resolvent equation by examining several specific methods for solving cubics and quartics; that is, his knowledge of this general principle arose from an *a posteriori* analysis of particular instances. An *a priori* analysis, in contrast, is one that starts from a general law and moves to particular instances. The literal meanings of these Latin terms are 'from what comes after' (*a posteriori*) and 'from what comes first' (*a priori*). The arguments we will see below illustrate how Lagrange argued from general algebraic properties in an *a priori* fashion to deduce the degree of the resolvent equation.

roots (and neither a derivative nor a power of x). The symbols $\alpha^2, \alpha^3, \ldots$ here *do* denote powers of α, however, where α itself denotes an m^{th} root of unity: that is, a number for which $\alpha^m = 1$. Of course, $\alpha = 1$ is always a solution of the equation $\alpha^m = 1$. By the Fundamental Theorem of Algebra, however, we know that $\alpha = 1$ is only one of m distinct solutions of the equation $\alpha^m = 1$. Letting $i = \sqrt{-1}$, for example, the four fourth roots of unity are $\{1, i, -1, -i\}$.[10] Although Lagrange specified no restrictions on the value which α may assume in the formula for the resolvent's roots in the preceding excerpt, his 1808 summary later made it clear that for $m > 2$, this formula requires α to be a complex-valued root of unity.

Because complex roots of unity play such a large role in his analysis, Lagrange included a detailed discussion of their properties in his works on the theory of equations. We will consider excerpts of this discussion — especially those parts which relate to group theory concepts — in Subsection 1.1 below. We will also look at methods for computing the values of complex roots of unity in that subsection. In Subsection 1.2, we will then return to the formula Lagrange gave for the roots of the resolvent equation ($t = x' + \alpha x'' + \alpha^2 x''' + \alpha^3 x^{(iv)} + \ldots + \alpha^{m-1} x^{(m)}$), and examine how rearranging (or permutating) the roots $x', x'', x''', \ldots$ within this formula provides us with information about the degree of the resolvent equation.

To set the stage for both these discussions, we first look at a specific example to see how an auxiliary equation, even one of higher degree than the original, can be used to solve a given equation. We note that our primary purpose with this example and its continuation in several later tasks is to provide a context for the group-theoretic connections which emerge from Lagrange's theoretical work on resolvent equations (e.g., roots of unity and permutations), rather than to provide a complete development of that work.

Task 1

This task is based on the analysis in Lagrange's 1770 *Mémoire* [10] of the solution given by Gerolamo Cardano (see footnote 3, page 2) to the cubic equation

$$x^3 + nx + p = 0, \tag{1}$$

where n and p are assumed to be positive real numbers.

In particular, we will see how all three roots of this cubic can be determined from the six roots of the following sixth degree equation:

$$t^6 + 27pt^3 - 27n^3 = 0 \tag{2}$$

Later in this section, after developing some necessary groundwork, we will also see how the roots of these two equations fit into the special form discussed above by Lagrange. Accordingly, we follow Lagrange's terminology throughout this task, and refer to the sixth degree equation (2) as the **resolvent** of the given cubic equation (1).

(a) Verify that the substitution $x = \frac{t}{3} - \frac{n}{t}$, where $t \neq 0$, transforms the cubic equation (1) into the resolvent equation (2).

This means that any solution t of the resolvent which can be expressed in terms of its coefficients, elementary arithmetic operations and the extraction of roots, in turn gives a solution $x = \frac{t}{3} - \frac{n}{t}$ of the original cubic involving only elementary arithmetic operations and the extraction of roots beginning from the coefficients $1, n, p$.

Speculate on how the six roots of the resolvent equation will ultimately result in just three roots for the given cubic, in keeping with the Fundamental Theorem of Algebra.

[10] This can be verified by raising each number to the fourth power, or by solving $(\alpha^2 - 1)(\alpha^2 + 1) = 0$ directly. As a preview of ideas studied later in this project, note that powers of i give all four roots: $i^1 = i$, $i^2 = -1$, $i^3 = -i$, $i^4 = 1$.

Task 1 - continued

(b) Note that the resolvent equation is quadratic in t^3: $\left[t^3\right]^2 + 27p\left[t^3\right] - 27n^3 = 0$.

Let θ_1 and θ_2 denote the roots of $\theta^2 + 27p\theta - 27n^3$.

Use the fact that n, p are positive reals to prove θ_1, θ_2 are distinct real numbers, and explain how it follows from this that $t_1 = \sqrt[3]{\theta_1}$ and $t_2 = \sqrt[3]{\theta_2}$ are distinct real roots of the resolvent which can be expressed in the required form.[11]

How many real roots do you think the cubic (1) will have? Justify your conjecture.

(c) Recall that complex roots of real-valued polynomials come in conjugate pairs.

Thus, either all three roots of the cubic $x^3 + nx + p$ are real (possibly repeated), or this cubic has exactly one real root and two (distinct) complex roots.

In this part of the task, we show that the latter is the case.

Our first step is to express the six roots of the resolvent in a suitable form.

For two of these roots ($t_1 = \sqrt[3]{\theta_1}$ and $t_2 = \sqrt[3]{\theta_2}$), this has already been done.

In his 1770 article, Lagrange expressed the remaining four (complex) roots as products involving t_1, t_2 and the two complex-valued cubic roots of unity.[12]

Do this for one of the four complex roots of the resolvent as follows:

> Let α denote a complex number with $\alpha^3 = 1$, and set $t_3 = \alpha t_1$.
>
> (In Subsection 1.1, we will see how to explicitly write α in the required form.)
>
> Show that t_3 is a complex root of the resolvent equation.[13]
>
> Explain why $x_3 = \frac{t_3}{3} - \frac{n}{t_3}$ must therefore be a complex root of the given cubic.
>
> Finally, explain why the given cubic must have two distinct complex roots and one real root. (You do not need to find the second complex rootj!)

(d) In this part, we resolve an apparent paradox in the preceding results:

> How is it that the resolvent equation has two distinct real roots ($t_1 \neq t_2$), each of which can be used to define a real root for the given cubic via the substitution $x = \frac{t}{3} - \frac{n}{t}$, while the cubic equation itself has only one real root and that single real root itself has a multiplicity of only one?

Begin by setting $x_1 = \frac{t_1}{3} - \frac{n}{t_1}$ and $x_2 = \frac{t_2}{3} - \frac{n}{t_2}$, and recall that $t_1 = \sqrt[3]{\theta_1}$ and $t_2 = \sqrt[3]{\theta_2}$.

Also recall that θ_1 and θ_2 denote the roots of $f(\theta) = \theta^2 + 27p\theta - 27n^3$.

Thus, by the Factor Theorem, $f(\theta) = (\theta - \theta_1)(\theta - \theta_2)$.

Expand this expression for $f(\theta)$ to show that $\theta_1\theta_2 = (-3n)^3$.

Conclude that $t_1 t_2 = -3n$, and use this fact to show that $x_1 = \frac{1}{3}(t_1 + t_2)$.

Proceed similarly to show that $x_2 = \frac{1}{3}(t_1 + t_2)$.

Comment on how this resolves the apparent paradox described above.[14]

[11] You do not need to solve for θ_1, θ_2 to do this or any other part of this task, but you will want to include the fact that the quadratic formula ensures θ_1, θ_2 can be expressed in the required form as part of your explanation.

[12] We could also express the four complex roots of the resolvent directly in terms of i and the roots θ_1, θ_2 of the quadratic $\theta^2 + pt^3 - \frac{n^3}{27}$ by using the substitution $\theta = t^3$ to first rewrite the resolvent as follows, and then applying the quadratic formula: $t^6 + pt^3 - n^3/27 = (t^3 - \theta_1)(t^3 - \theta_2) = (t - \sqrt[3]{\theta_1})(t^2 + \sqrt[3]{\theta_1}t + \sqrt[3]{\theta_1}^2)(t - \sqrt[3]{\theta_2})(t^2 + \sqrt[3]{\theta_2}t + \sqrt[3]{\theta_2}^2)$.

[13] The remaining complex roots of the resolvent equation can be similarly obtained; that is, letting α, β denote the two complex cubic roots of unity, the four complex roots of the resolvent are $t_3 = \alpha t_1$, $t_4 = \beta t_1$ $t_5 = \alpha t_2$ and $t_6 = \beta t_2$.

[14] In Task 10, we will see that the two complex roots of the cubic can also be written as $\frac{1}{3}(t_i + t_j)$ for appropriately chosen values t_i, t_j selected from the four complex roots t_3, t_4, t_5, t_6 of the resolvent equation. We will then use these expressions in Task 14 to establish that Lagrange's form $t = x' + \alpha x'' + \alpha^2 x'''$ holds for this particular example.

1.1 Roots of unity in Lagrange's analysis

As suggested above, roots of unity played an important role in Lagrange's formula for the roots of the resolvent equation, and more generally in the theory of equations. In this subsection, we consider some properties of these special roots as they were described by Lagrange. Our first excerpt on roots of unity also touches on an important point about 'solvability' which we have already raised; namely, the notion of 'solvability' can be defined in a variety of ways. In our context, 'algebraic solvability' specifically requires that the roots of a given equation can be determined from its coefficients by way of a *finite* number of steps involving only elementary arithmetic operations $(+, -, \times, \div)$ and the extraction of roots. Thus, just as Omar Khayyam continued to seek an algebraic algorithm for the roots of cubic equations even after he showed these roots existed by intersecting conic sections in a way that would have satisfied Greek mathematicians, the solution involving trigonometric functions described below by Lagrange would not count as an algebraic solution for today's algebraist (nor would it for Lagrange) *unless* the specific trigonometric values involved could be expressed in the permitted form. We examine this issue further in the following excerpt, taken from Note XIV of Lagrange's *Traité de la résolution des équations numériques* [11].

Although equations with two terms such as

$$x^m - A = 0, \text{or more simply } x^m - 1 = 0$$

(since the former is always reducible to the latter, by putting $x\sqrt[m]{A}$ in for x), are always solvable by trigonometric tables in a manner that allows one to approximate the roots as closely as desired, by employing the known formula[15]

$$x = \cos\frac{k}{m}360^\circ + \sin\frac{k}{m}360^\circ\sqrt{-1}$$

and letting $k = 1, 2, 3, \ldots, m$ successively, their algebraic solution is no less interesting for Analysis, and mathematicians have greatly occupied themselves with it.

Today, this trigonometric formula for the m^{th} roots of unity is written in terms of the radian measure of a circle, 2π, and the now-standard symbol $i = \sqrt{-1}$:

$$x^m - 1 = 0 \quad \Leftrightarrow \quad x = \cos\left(\frac{2\pi k}{m}\right) + i\sin\left(\frac{2\pi k}{m}\right), \text{ where } k = 1, 2, 3, \ldots, m.$$

For example, the solutions of $x^3 - 1 = 0$ are readily obtained from this formula as follows:

$$\begin{array}{lllll}
k=1 \;\Rightarrow\; x = \cos\left(\frac{2\pi\cdot 1}{3}\right) + i\sin\left(\frac{2\pi\cdot 1}{3}\right) & = & \cos\left(\frac{2\pi}{3}\right) + i\sin\left(\frac{2\pi}{3}\right) & = & -\frac{1}{2} + \frac{\sqrt{3}}{2}i \\
k=2 \;\Rightarrow\; x = \cos\left(\frac{2\pi\cdot 2}{3}\right) + i\sin\left(\frac{2\pi\cdot 2}{3}\right) & = & \cos\left(\frac{4\pi}{3}\right) + i\sin\left(\frac{4\pi}{3}\right) & = & -\frac{1}{2} - \frac{\sqrt{3}}{2}i \\
k=3 \;\Rightarrow\; x = \cos\left(\frac{2\pi\cdot 3}{3}\right) + i\sin\left(\frac{2\pi\cdot 3}{3}\right) & = & \cos(2\pi) + i\sin(2\pi) & = & 1
\end{array}$$

Notice that these roots include one real root and two complex roots which are conjugates of each other. Also note that the particular trigonometric values involved here can be expressed in terms of only

[15] A proof of this formula lies outside the scope of this project.

finitely many basic arithmetic operations on rational numbers and their roots (including $i = \sqrt{-1}$).[16] In fact, we could instead solve $x^3 - 1 = 0$ by first factoring $[x^3 - 1 = (x-1)(x^2 + x + 1)]$ and then using the quadratic formula. Thus, the cubic roots of unity are bona fide algebraic solutions of this equation.

In his celebrated doctoral dissertation *Disquisitiones Arithmetica* of 1801, the great German mathematician Frederic Gauss (1777–1855) proved that $x^m - 1 = 0$ can always be solved algebraically. This fact in turn implied Lagrange could legitimately use roots of unity in his formula for the roots of the resolvent and still obtain an algebraic solution. Lagrange commented on this technical point in his 1808 summary, saying of Gauss' work that it was 'as original as it was ingenious' [11, p. 329].

Task 2
Without employing a calculator or computer, use the formula $x = \cos\left(\frac{2\pi k}{m}\right) + i \sin\left(\frac{2\pi k}{m}\right)$ (with $m = 6$) to express all sixth roots of unity in terms of the elementary arithmetic operations on rational numbers and their roots only.

What difficulties do you encounter when you attempt to do this for the fifth roots of unity? How else might you proceed in this case?

Note: Lagrange found the fifth roots of unity to be 1, $\frac{\sqrt{5}-1}{4} \pm \frac{10+2\sqrt{5}}{4} i$, and $-\frac{\sqrt{5}+1}{4} \pm \frac{10-2\sqrt{5}}{4} i$.

As Lagrange commented in the preceding excerpt, this formula for the m^{th} roots of unity in terms of the trigonometric functions was well known by his time. It is straightforward to check its correctness (see Appendix I, Task I.1) using another formula that was then well known:

$$\boxed{\textbf{de Moivre's Formula:} \quad (\cos\theta + i \sin\theta)^n = \cos(n\theta) + i \sin(n\theta)}$$

The French mathematician for which this formula is named, Abraham de Moivre (1667–1754), appears to have had some understanding of it as early as 1707; it seems likely that he discovered it by way of the standard sum identities for trigonometric functions, although he never published a proof. In the first precalculus text ever to be published, *Introductio in analysin infinitorum* (1748), the celebrated mathematician Leonhard Euler (1707–1783) did use trigonometric identities to prove de Moivre's formula in the case where n is a natural number (see Appendix I, Task I.2). As was typical of Euler, he then took matters a step further and used ideas related to power series (see Appendix I, Task I.3) to establish another amazing formula:

$$\boxed{\textbf{Euler's formula:} \quad e^{i\theta} = \cos\theta + i \sin\theta}$$

As a corollary to Euler's formula, de Moivre's formula is easy to prove, although the almost magical ease of this proof makes it seem no less mysterious:[17]

[16] Had the coefficients of the original equation included irrational or imaginary numbers, then any number obtained in a finite number of steps from those coefficients with basic arithmetic operations or extraction of roots would also be allowed. Since the coefficients of $x^3 - 1 = 0$ are integers, these operations allow only rational numbers and their roots.

[17] Another easy but mysterious consequence of Euler's formula, obtained by letting $\theta = \pi$, is **Euler's Identity** relating the five most important mathematical constants in a single equation: $e^{i\pi} + 1 = 0$. After proving Euler's identity to an undergraduate class at Harvard, the American algebraist Benjamin Peirce (1809–1880) is reported in [1] to have said: "Gentlemen, that is surely true, it is absolutely paradoxical; we cannot understand it, and we don't know what it means. But we have proved it, and therefore we know it must be the truth."

$$(\cos\theta + i\sin\theta)^n = (e^{i\theta})^n = e^{i(n\theta)} = \cos(n\theta) + i\sin(n\theta)$$

Additionally, Euler's formula gives us another (more concise) way to represent roots of unity:

$$x^m - 1 = 0 \quad\Leftrightarrow\quad x = \cos\left(\frac{2\pi k}{m}\right) + i\sin\left(\frac{2\pi k}{m}\right) \quad\Leftrightarrow\quad x = e^{\frac{2\pi k}{m}i}, \text{ where } k = 1, 2, 3, \ldots, m.$$

Although Lagrange himself did not make use of this exponential representation in his own writing (despite his familiarity with Euler's identity), we use it in this project as a convenient means to abbreviate notation and simplify computations. For example, the three cubic roots of unity can be written as follows:

$$\begin{aligned} k = 1 \quad&\Rightarrow\quad x = e^{\frac{2\pi\cdot 1}{3}i} = e^{\frac{2\pi i}{3}} \\ k = 2 \quad&\Rightarrow\quad x = e^{\frac{2\pi\cdot 2}{3}i} = e^{\frac{4\pi i}{3}} \\ k = 3 \quad&\Rightarrow\quad x = e^{\frac{2\pi\cdot 3}{3}i} = e^{2\pi i} \end{aligned}$$

We also make use of the geometric representation of roots of unity as points on the unit circle (see Figure 1) which is naturally suggested by the formula $e^{\frac{2\pi k}{m}i} = \cos\left(\frac{2\pi k}{m}\right) + i\sin\left(\frac{2\pi k}{m}\right)$.

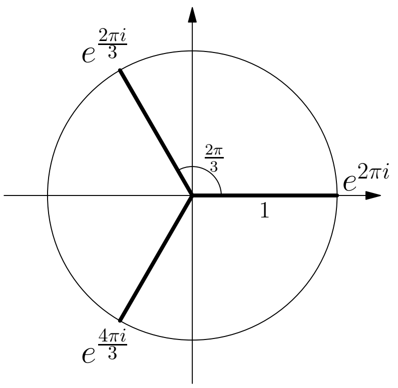

Figure 1: Cubic roots of unity: $e^{\frac{2\pi k}{3}i} = \cos\left(\frac{2\pi k}{3}\right) + i\sin\left(\frac{2\pi k}{3}\right)$, $k = 1, 2, 3$

Task 3

(a) Represent the sixth roots of unity on a unit circle. (See also Task 2.)
Label these in exponential form: $e^{\frac{2\pi k}{6}i} = e^{\frac{\pi k}{3}i}$, where $k = 1, 2, 3, 4, 5, 6$.
Also indicate clearly which of these six roots are real and which are complex.

(b) On the unit circle from part (a), identify which of the sixth roots of unity are also square roots of unity, and which are also cubic roots of unity.

(c) Use a separate unit circle to represent the fifth roots of unity on a unit circle; again label with exponential notation and indicate which are real and which are complex.

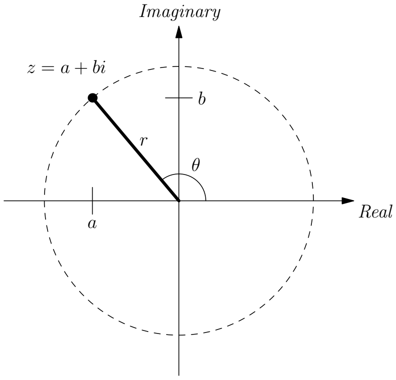

Figure 2: Geometric representation of $z = a + bi = re^{i\theta}$, where $r = \sqrt{a^2 + b^2}$

Although the geometric representation of the set of all complex numbers as points in the two-dimensional plane shown in Figure 2 was first published after Lagrange's work on algebraic solvability,[18] mathematicians of his era were well aware of the connection of roots of unity to the unit circle.[19] In fact, Lagrange explicitly used the geometric representation of roots of unity as points on the unit circle to determine the total number and type (real versus complex) of m^{th} roots of unity, as we read in the following excerpt from his 1770 *Mémoire*.[20] It may be helpful to refer back to the unit circles in Figure 1 and Task 3 (both on page 9) as you read this excerpt.

We first remark concerning this solution that each of the roots of the equation $x^m - 1 = 0$ should be different from each other, since on the circumference there are not two different arcs which simultaneously have the same sine and the same cosine. It is further easy to see that all the roots will be imaginary, with the exception of the last which corresponds to $k = m$ and which will always equal 1, and of that which corresponds to $k = \frac{m}{2}$, when m is even, which will equal -1; since in order for the imaginary part of the expression of x to vanish, it is necessary to have

$$\sin\left(\frac{k}{m}360^\circ\right) = 0,$$

which never occurs unless the arc is equal to 360° or to 180°; in which case we will have either $\frac{k}{m} = 1$ or $= \frac{1}{2}$, and consequently either $k = m$ or $k = \frac{m}{2}$; in the first case, the real part $\cos\left(\frac{k}{m}360^\circ\right)$ will become $\cos 360^\circ = 1$; and in the second it will become $\cos 180^\circ = -1$.

[18] The name of Parisian bookkeeper Jean-Robert Argand (1768–1822) is frequently cited in connection with the development of the complex plane, in recognition of an 1806 pamphlet which he produced on the topic. Although there are indications that Gauss was in possession of the geometric representation of complex numbers as early as 1796, he did not publish on the subject until 1831. Credit for the first publication on the subject instead belongs to the Norwegian Caspar Wessel (1745–1818); unfortunately, Wessel's 1797 paper was written in Danish and went unnoticed until 1897.

[19] It was this relation of the m^{th} roots of unity to points on the unit circle that led to the term 'cyclotomic polynomial' being used for the expression '$x^{m-1} + x^{m-2} + \ldots + 1$' which arises as a factor of $x^m - 1 = (x-1)(x^{m-1} + x^{m-2} + \ldots + 1)$.

[20] Because the following excerpt comes from a different source than that which we have quoted thus far, we have altered the notation in it slightly (replacing, for example, n by m) to be consistent with the notation used in our other excerpts.

Task 4
Compare Lagrange's claim concerning the number and type of m^{th} roots of unity in the preceding excerpt to what you found for $m = 6$ in parts (a) and (b) of Task 3.

How convincing do you find Lagrange's argument in general?

We now read the continuation of Lagrange's comments on roots of unity, in which ideas related to what later came to be known as a 'cyclic group' arise for the first (but not last!) time in this project.

Now, if we let

$$\alpha = \cos\left(\frac{360^\circ}{m}\right) + \sin\left(\frac{360^\circ}{m}\right)\sqrt{-1},$$

we will have ...

$$\alpha^k = \cos\left(\frac{k}{m}360^\circ\right) + \sin\left(\frac{k}{m}360^\circ\right)\sqrt{-1};$$

so that the different roots of $x^m - 1 = 0$ will all be expressed by the powers of this quantity α; and thus these roots will be $\alpha, \alpha^2, \alpha^3, \ldots \alpha^m$, of which the last α^m will always be equal to 1 ....

We pause at this point in our reading of Lagrange to illustrate his central idea for the specific case of $m = 3$. Using exponential notation, we set $\alpha = e^{\frac{2\pi i}{3}}$. Then, as noted by Lagrange, the remaining cubic roots of unity can be obtained simply by taking powers of α, since $\alpha^2 = \left[e^{\frac{2\pi i}{3}}\right]^2 = e^{\frac{4\pi i}{3}}$ and $\alpha^3 = \left[e^{\frac{2\pi i}{3}}\right]^3 = e^{2\pi i} = 1$. Because it is possible to generate all the cubic roots of unity from α in this way, α is called a ***primitive cubic root of unity***.

Task 5
Let $m = 6$ and set $\alpha = \cos\left(\frac{2\pi}{6}\right) + i\sin\left(\frac{2\pi}{6}\right) = e^{\frac{\pi i}{3}}$.
Verify that α is a primitive sixth root of unity by computing α^k for $k = 2, 3, 4, 5, 6$.
Comment on how these results compare to your answer to Task 3(a).

Return now to Lagrange's comments on primitive roots of unity.

It is good to observe here that if m is a prime number, one can always represent all the roots of $x^m - 1 = 0$ by the successive powers of any one of these same roots, excepting only the last; let, for example, $m = 3$, the roots will be $\alpha, \alpha^2, \alpha^3$: if we take the next root α^2 in place of α, we have the three roots $\alpha^2, \alpha^4, \alpha^6$; but, since $\alpha^3 = 1$, it is clear that $\alpha^4 = \alpha$ and that $\alpha^6 = \alpha^3$; so that these roots will be $\alpha^2, \alpha, \alpha^3$, the same as before.

Let us consider what Lagrange is claiming here in more detail. In light of his comments on primitive roots of unity in the first excerpt on page 11, we know the expression 'the last root' is a reference to the real number '1', obtained by taking the 'last' power (α^m) of $\alpha = e^{\frac{2\pi i}{m}}$. We also know from what Lagrange has already said that $\alpha = e^{\frac{2\pi i}{m}}$ is a primitive m^{th} root of unity, since taking powers of α has the effect of cycling through all the m^{th} roots of unity. In the excerpt we have just read, Lagrange has gone beyond this to claim that *every* m^{th} root of unity — other than the last root 1 — behaves in exactly this same way, provided m is prime. In other words:

Theorem
If m is prime and β is a complex m^{th} root of unity, then β is a primitive m^{th} root of unity.

Lagrange's illustration of this theorem for the prime $m = 3$ in the preceding excerpt used the fact that $\alpha = e^{\frac{2\pi i}{3}}$ is already known to be a primitive cubic root of unity; thus, the only other complex root of unity β can be written as a power of α; namely, $\beta = \alpha^2$. Using this notation, we can re-write Lagrange's power computations as follows: $\beta^1 = \alpha^2$, $\beta^2 = \alpha^4 = \alpha^3\alpha = \alpha$ and $\beta^3 = \alpha^6 = [\alpha^3]^2 = 1$. This shows that $\beta = \alpha^2$ is also a primitive root of unity, and the theorem holds in this case.

Task 6
In the continuation of the preceding excerpt, Lagrange next considered the case $m = 5$, where $\alpha = \cos\left(\frac{2\pi}{5}\right) + i\sin\left(\frac{2\pi}{5}\right) = e^{\frac{2\pi i}{5}}$ and the five fifth roots of unity are $\alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^5 = 1$.

Complete the following to prove that α^2, α^3, and α^4 are also primitive fifth root of unity.

(a) Find the first five powers of α^2, and show that these are the same as the five original roots rearranged in the following order: $\alpha^2, \alpha^4, \alpha, \alpha^3, \alpha^5$

(b) Find the first five powers of α^3, and show that these generate the five original roots rearranged in the following order: $\alpha^3, \alpha, \alpha^4, \alpha^2, \alpha^5$.

(c) Determine the order in which the original roots are generated using powers of α^4.

In the next task, the specific case of $m = 6$ is used to show that the restriction to prime numbers in the preceding theorem is necessary; that is, when m is composite, it is no longer the case that every complex root of unity is also a primitive root of unity. A proof of the theorem for the prime case is then outlined in Task 8, followed by further explorations of the composite case in Task 9.

Task 7
In this task, the case $m = 6$ is used to show that the restriction to prime numbers in the preceding theorem is necessary; that is, when m is composite, it is no longer the case that powers of every root of unity can be used to generate all the roots of unity via powers. (You may wish to briefly review Task 2 and parts (a), (b) of Task 3.)
Set $\alpha = \cos\left(\frac{2\pi}{6}\right) + i\sin\left(\frac{2\pi}{6}\right) = e^{\frac{\pi}{3}i}$.

(a) Explain why all the sixth roots of unity are obtained by taking powers of α.

(b) Now consider powers of $\beta = \alpha^2$. Which of the sixth roots of unity do you obtain?

(c) Next consider powers of $\gamma = \alpha^3$. Which of the sixth roots of unity do you obtain?

(d) Which of the powers of α, other than $\alpha^1 = \alpha$, are also primitive roots of unity? That is, which powers of α generate all six of the sixth roots of unity? Justify your response.

Task 8

In this task, we return to the case where m is prime, and sketch a general proof for Lagrange's claim that every complex m^{th} root of unity β except $\beta = 1$ is primitive.

Begin by assuming that m is prime and that $\beta \neq 1$ is a complex m^{th} root of unity.
Also let $\alpha = e^{\frac{2\pi i}{m}}$, and choose $n \in \mathcal{Z}^+$ such that $\beta = \alpha^n$ with $1 \leq n < m$, where $\mathcal{Z}^+$ denotes the set of positive integers. *(How do we know that such a value of n exists?)*

Our goal is to prove that the powers of β generate all possible m^{th} root of unity.
In other words, we wish to show that the list $\beta\,, \beta^2\,, \ldots\,, \beta^m$ consisting of the first m positive integer powers of β corresponds to some arrangement of the list $\alpha\,, \alpha^2\,, \ldots\,, \alpha^m$ of all m^{th} roots of unity.

(a) Begin by explaining why β^s is an m^{th} root of unity for every $s \in \mathcal{Z}^+$ with $1 \leq s \leq m$.

Note: Since different powers of β could produce the same complex number, this proves only that the list $\beta\,, \beta^2\,, \ldots\,, \beta^m$ contains *at most* m distinct m^{th} root of unity.

(b) Use the fact that m is prime to prove the following:
Lemma For all $s \in \mathcal{Z}^+$, $\beta^s = 1$ if and only if m divides s.
Hint? Remember that $\beta = \alpha^n$, where $1 \leq n < m$ and $\alpha = e^{\frac{2\pi i}{m}} = \cos\left(\frac{2\pi}{m}\right) + i \sin\left(\frac{2\pi}{m}\right)$.

Note: This shows that m is the first positive power of β that generates 'the last root' 1; that is, the real root $1 = \beta^m$ appears only once within the list $\beta\,, \beta^2\,, \ldots\,, \beta^m$.
It remains to show that none of the other m^{th} roots of unity are repeated in this list.

(c) Suppose that the list $\beta\,, \beta^2\,, \ldots\,, \beta^m$ contains fewer than m distinct values.
That is, suppose that $\beta^s = \beta^t$ for some integers s, t with $1 \leq s < t \leq m$.
Use the lemma proven in part (b) to derive a contradiction.
Hint? Notice that $1 \leq t - s < m$.

(d) *Optional:* Re-write the proof that there are no repeated elements in the list $\beta\,, \beta^2\,, \ldots\,, \beta^m$ from part (c) without using proof by contradiction.

Task 9

We now return to the case where m is a composite number and consider the number of primitive m^{th} roots of unity in this case.

To this end, let $\alpha = e^{\frac{2\pi}{m} i}$.

(a) Recall that for $m = 4$, the primitive fourth roots of unity are $\alpha = i$ and $\alpha^3 = -i$.
Also review the results which you obtained in Task 7 for the case $m = 6$.

Use this data to develop a general conjecture concerning exactly which powers of α are primitive roots and which are not in the case where m is composite.

(b) Test your conjecture from part (a) in the cases of $m = 8$ and $m = 9$.
Clearly record your evidence that α^s is or is not a primitive root for each value of s.
Refine your conjecture as needed before testing it for the case of $m = 12$.
Continue to refine and re-test it further as needed.
Once you are satisfied with your conjecture, write a general proof for it.
Discuss proof strategies as needed with other students and/or your course instructor.

(c) *Optional:* Modify your conjecture concerning primitive roots of unity in the case where m is composite so that it applies to all values of m, both prime and composite.
Also modify your proof as needed to apply to this more general conjecture.

We close this subsection with a task related to the cubic equation example from Task 1. In Task 14 of the next subsection, we will use the expressions obtained for x, x', x'' below to establish that the special relationship $t = x' + \alpha x'' + \alpha^2 x'''$ which Lagrange claimed will hold between the roots of an equation (x, x', x'') and the roots of its resolvent (t) does indeed hold in this particular example.

Task 10
Recall from Task 1 that the roots of the cubic equation $x^3 + nx + p = 0$, where n, p are positive real numbers, are related to the roots of the sixth degree resolvent equation $t^6 + 27pt^3 - 27n^3 = 0$ by the expression $x = \frac{t}{3} - \frac{n}{t}$.

Also recall that the six (distinct) roots $t_1, t_2, t_3, t_4, t_4, t_6$ of the resolvent can be expressed as products of its two real roots (t_1, t_2) and the three cubic roots of unity. Denoting the cubic roots of unity by $1, \alpha, \alpha^2$, where α is a primitive cubic root of unity, we thus have

$$\begin{array}{lll} t_1 = \sqrt[3]{\theta_1} & t_3 = \alpha t_1 & t_5 = \alpha^2 t_1 \\ t_2 = \sqrt[3]{\theta_2} & t_4 = \alpha t_2 & t_6 = \alpha^2 t_2, \end{array}$$

where θ_1, θ_2 are the two (distinct) real roots of $\theta^2 + 27p\theta - 27n^3$.

In Task 1(d), we used the fact that $t_1 t_2 = -3n$ to show that $\frac{t_1}{3} - \frac{n}{t_1} = \frac{1}{3}(t_1 + t_2) = \frac{t_2}{3} - \frac{n}{t_2}$, concluding that $x' = \frac{1}{3}(t_1 + t_2)$ is the only real root of the given cubic.

Use this same fact $[t_1 t_2 = -3n]$ to show that $\frac{t_4}{3} - \frac{n}{t_4} = \frac{1}{3}(t_4 + t_5) = \frac{t_5}{3} - \frac{n}{t_5}$.
Conclude that $x'' = \frac{1}{3}(t_4 + t_5)$ is one of the two complex roots of the given cubic.

Denoting the second complex root of the given cubic by x''',
write an expression for x''' in terms of t_3, t_6, and verify that your expression is correct.

1.2 Permutations of roots in Lagrange's analysis

We now turn to Lagrange's treatment of general polynomial equations in his 1808 note on this topic. The first excerpt we consider states a relationship between the roots and the coefficients of an equation which was well known to algebraists of his time; following the excerpt, we will see how this relationship leads to the idea of permuting roots.

We represent the proposed equation by the general formula

$$x^m - Ax^{m-1} + Bx^{m-2} - Cx^{m-2} + \ldots = 0,$$

and we designate its m roots by $x', x'', x''', \ldots, x^{(m)}$; we will then have, by the known properties of equations,

$$\begin{array}{l} A = x' + x'' + x''' + \ldots + x^{(m)}, \\ B = x'x'' + x'x''' + \ldots + x''x''' + \ldots, \\ C = x'x''x''' + \ldots \end{array}$$

Task 11

(a) For $m = 2$, note that the general equation becomes $x^2 - Ax + B = 0$, where Lagrange claimed that $A = x' + x''$ and $B = x'x''$. Verify that these formulas for A and B are correct by expanding the factored form of the polynomial: $(x - x')(x - x'')$.

(b) Now write down the formulas for the coefficients A, B, C of the cubic polynomial $x^3 - Ax^2 + Bx - C$ in terms of its roots x', x'', x''', and again verify that these are correct by expanding the factored form of the polynomial.

(c) Use the formulas found in part (b) for the coefficients A, B, C of the cubic polynomial $x^3 - Ax^2 + Bx - C$ to determine the expanded form of the following polynomials *without* multiplying out the given factors.

(i) $(x - 2)(x - 3)(x - 5)$ **(ii)** $(x - 1)(x - (1 + 2i))(x - (1 - 2i))$

In Lagrange's expressions for the coefficients $A, B, C \ldots$, note that the roots $x', x'', x''', \ldots, x^{(m)}$ can be permuted in any way we wish without changing the (formal) value of the expression. For example, if we exchange x' for x'' (and vice-versa) in the case where $m = 2$, we get $A = x'' + x'$ and $B = x''x'$, both clearly equal to the original expressions ($A = x' + x''$ and $B = x'x''$). For $m = 3$ (or higher), more complicated permutations of the roots arise. For example, we could simultaneously replace each occurrence of x' by x''', each occurrence of x'' by x' and each occurrence of x''' by x'' in the original expressions for A, B, C, thereby obtaining the following:

$$\begin{array}{lcl} A = x' + x'' + x''' & \Rightarrow & A = x''' + x' + x'' \\ B = x'x'' + x'x''' + x''x''' & \Rightarrow & B = x'''x' + x'''x'' + x'x'' \\ C = x'x''x''' & \Rightarrow & C = x'''x'x'' \end{array}$$

Again, however, we see that the expressions resulting from this particular permutation of the given roots are formally equivalent to the original expressions. It is similarly straightforward to check that this occurs with every possible permutation of the three roots. (Try it!)

Expressions with the property that every permutation of the variables results in the same formal value are said to be ***symmetric functions***.[21] In contrast, the expression $x_1x_2 + x_3$ is *not* a symmetric function since, for example, exchanging x_1 and x_3 results in a different formal value ($x_3x_2 + x_1$), even though exchanging x_1 and x_2 results in an expression ($x_2x_1 + x_3$) equivalent to the original ($x_1x_2 + x_3$).

Task 12

Determine which of the following are symmetric expressions in x_1, x_2, x_3.

For any which is not, describe a permutation of x_1, x_2, x_3 that changes the formal expression, and (if possible) another permutation which does not change the formal expression.

(a) $(x_1 + x_2 + x_3)^2$ **(b)** $x_1^2 + (x_2 + x_3)^2$ **(c)** $(x_1 + x_2)(x_2 + x_3)$

Returning now to Lagrange, we read two suggestions concerning how one might proceed to find the resolvent equation whose solution would allow us to find an algebraic solution of the original equation.

[21] The symmetric functions given by the coefficients of the polynomial $\prod_{k=1}^{m}(x - x_k)$ are called the ***elementary symmetric polynomials***. An example of a non-elementary symmetric function on three variables is given by $x_1^2 + x_2^2 + x_3^2$.

To obtain the [resolvent] equation . . . , it will be necessary to eliminate the m unknowns x' ,x'' , x''', . . . , $x^{(m)}$ by means of the preceding equations, which are also m in number; but this process requires long calculations, and it will have, moreover, the inconvenience of arriving at a final equation of degree higher than it needs to be.

One can obtain the equation in question directly and in a simpler fashion, by employing a method which we have made frequent use of here, which consists in first finding the form of all the roots of the equation sought, and then composing this equation by means of its roots.

In other words, one can either find the resolvent equation by solving m equations in m unknowns through a series of long calculations involving symmetric functions ...... or one can simplify this process by using the *form* of the roots to obtain the resolvent equation by way of the relation between the roots and the factors given by the Factor Theorem, with each root contributing a factor towards building up the resolvent equation.

In our remaining excerpts from Lagrange's work, we will see how he set out to implement this second plan. The tasks interspersed between these excerpts examine his argument in the specific case $m = 3$. We begin with an excerpt in which Lagrange first reminded his readers about the way in which the roots t of the resolvent appear as a function of the roots $x', x'', \dots x^{(m)}$ of the original equation and the powers of a primitive m^{th} root of unity α. His main goal in this excerpt was to deduce the degree of the resolvent equation, based on the total number of roots which can be formed in this way.

Let t be the unknown of the resolvent equation; in keeping with what was just said, we set

$$t = x' + \alpha x'' + \alpha^2 x''' + \alpha^3 x^{(iv)} + \ldots + \alpha^{m-1} x^{(m)},$$

the quantity α being one of the m^{th} roots of unity, that is to say, one of the roots of the binomial equation $y^m - 1 = 0$. . . . .

It is first of all clear that, in the expression t, one can interchange the roots $x', x'', x''', \dots, x^{(m)}$ at will since there is nothing to distinguish them here from one and another; from this it follows that one obtains all the different values of t by making all possible permutations of the roots $x', x'', x''', \dots, x^{(m)}$ and these values will necessarily be the roots of the resolvent in t which we wish to construct.

Now one knows, by the theory of combinations, that the number of permutations which can be obtained from m things is expressed in general by the product $1.2.3 \ldots m$; but we are going to see that this equation is capable of being reduced by the very form of its roots.

In our next task, we see how permuting the roots of the original equation in the formula for the resolvent's roots produces $m!$ resolvent roots in the specific case of $m = 3$. Of course, since $m! > m$ for $m > 2$, Lagrange's conclusion that an equation of degree m has a resolvent equation of degree $m!$ hardly seems like much progress. In Lagrange's ensuing analysis, however, we will see how 'this [resolvent] equation is capable of being reduced by the very form of its roots' to a lower degree.

Task 13

Consider the case $m = 3$ and let x', x'', x''' denote the three roots of an arbitrary cubic equation and α denote a primitive cubic root of unity. According to Lagrange's analysis in the preceding excerpt, the resolvent for the given cubic will have a total of $3! = 6$ roots, arising from the $3! = 6$ possible permutations of x', x'', x''' in the given formula.

Complete the list of these six roots below.

$$\begin{array}{ccccccc} t & = & x' & + & \alpha x'' & + & \alpha^2 x''' \\ t & = & x' & + & \alpha x''' & + & \alpha^2 x'' \\ t & = & & & & & \\ t & = & & & & & \\ t & = & & & & & \\ t & = & & & & & \end{array}$$

Before returning to Lagrange's analysis, remember what we saw in Task 1: even though the original degree 3 equation in that task had a resolvent equation of degree 6, that resolvent was quadratic in form, allowing us to essentially reduce the degree of the resolvent to 2 by way of a substitution. Lagrange ended the preceding excerpt with the claim that a similar reduction in the degree of the resolvent is always possible, regardless of the specific polynomial given or its degree. Once this reduction is achieved, the next question will be whether the reduced degree is sufficiently small that one could proceed to find an algebraic solution with known methods; if not, then some further reduction in the resolvent's degree would be required to complete the process.

As Lagrange emphasized throughout his work, the key to reducing the degree of the resolvent in the general case will be to consider the *form* of these roots, $t = x' + \alpha x'' + \alpha^2 x''' + \alpha^3 x^{(iv)} + \ldots + \alpha^{m-1} x^{(m)}$, and the effect of permutations on this form. To set the stage for his analysis of the general case, we return to the specific cubic polynomial introduced in Task 1 and complete the proof that the six roots of its resolvent equation assume the required form.

Task 14

In this task, we return to the cubic equation introduced in Task 1, $x^3 + nx + p = 0$, for which n, p are positive reals and the sixth degree resolvent equation is $t^6 + 27pt^3 - 27n^3 = 0$.

Recall from our continuation of that example in Task 10 that the resolvent's roots are:

$$\begin{array}{lll} t_1 & t_3 = \alpha t_1 & t_5 = \alpha^2 t_1 \\ t_2 & t_4 = \alpha t_2 & t_6 = \alpha^2 t_2, \end{array}$$

where t_1, t_2 are the real roots of the resolvent and α is a given primitive cubic root of unity.

Further recall from Task 10 that the three roots of the given cubic can be written as follows:

$$x' = \tfrac{1}{3}(t_1 + t_2) \qquad x'' = \tfrac{1}{3}(t_4 + t_5) \qquad x''' = \tfrac{1}{3}(t_3 + t_6)$$

In this task, we show that the six roots of the resolvent can be obtained via permutations of x', x'', x''' in the expression $t = x' + \alpha x'' + \alpha^2 x'''$.

Task 14 - continued

(a) Begin this task by verifying the following useful fact about sums of powers of primitive cubic roots of unity:

$$1 + \alpha + \alpha^2 = 0$$

- One way to do this is by direct computation, remembering that since α is not specified to be any particular primitive cubic root of unity, there are technically two cases to check: $\alpha = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $\alpha = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$. A geometric diagram may be suggestive of what is happening in this sum, but does not constitute a proof!
- Alternatively, this can be done by first recalling, from Task 11(b), that the coefficients A, B, C of the cubic $x^3 - Ax^2 + Bx - C$ are given by the elementary symmetric functions on the roots x_1, x_2, x_3 of that equation:

$$\begin{array}{rcccccc} A & = & x_1 & + & x_2 & + & x_3 \\ B & = & x_1x_2 & + & x_1x_3 & + & x_2x_3 \\ C & = & & & x_1x_2x_3 & & \end{array}$$

  Then consider the third degree polynomial $x^3 - 1$, for which the coefficient values are $A = 0$, $B = 0$ and $C = 1$ and the roots are the three cubic roots of unity generated by the primitive root $x_1 = \alpha$.

OPTIONAL: State and prove a generalization for the sum of the first m powers of β, where β is a primitive m^{th} root of unity.

(This generalization is not needed in the remainder of this task, but could be used in later tasks. A geometric diagram may again be suggestive.)

(b) Substitute the values for x', x'', x''' and t_3, t_4, t_5, t_6 found in Task 10 (re-stated above) into the expression $x' + \alpha x'' + \alpha^2 x'''$, then simplify using the fact that $1 + \alpha + \alpha^2 = 0$ which was proven in part (a) of this task. Conclude that $x' + \alpha x'' + \alpha^2 x''' = t_1$.

(c) Proceed as in part (b) to show that $x' + \alpha x''' + \alpha^2 x'' = t_2$.

(d) Show that $\alpha(x' + \alpha x'' + \alpha^2 x''') = x''' + \alpha x' + \alpha^2 x''$.
Use this fact along with the result of part (b) to conclude that $x''' + \alpha x' + \alpha^2 x'' = t_3$.

(e) Show that $\alpha^2(x' + \alpha x''' + \alpha^2 x'') = x''' + \alpha x'' + \alpha^2 x'$.
Use this fact along with the result of part (c) to conclude that $x''' + \alpha x'' + \alpha^2 x' = t_6$.

(f) Determine which permutations of the x', x'', x''' in the expression $t = x' + \alpha x'' + \alpha^2 x'''$ give the remaining two resolvent roots, t_4 and t_5.

We now return to Lagrange's argument that is always possible to reduce the degree of the resolvent equation of an m^{th} degree polynomial to a number less than $m!$. We consider the rest of this argument in two separate excerpts. As you read through the first of these two excerpts, remember that he has already established that every permutation of $x', x'', x''', \ldots$ in the expression t will result in a root of the resolvent equation.

One first sees that this expression is an unvariable function of the quantities $\alpha^0 x', \alpha x'', \alpha^2 x''', \ldots$, and also that the result of permuting the roots $x', x'', x''', \ldots$ among themselves will be the same as that of [permuting] the powers of α among themselves.

It follows from this that αt will be the result of the simultaneous permutations of [substituting] x' in for x'', x'' in for x''', $\ldots x^{(m)}$ in for x', since $\alpha^m = 1$. Similarly, $\alpha^2 t$ will be the result of the simultaneous permutations of [substituting] x' in for x''', x'' in for x^{iv}, $\ldots x^{(m-1)}$ in for x' and $x^{(m)}$ in for x'', since $\alpha^m = 1$, $\alpha^{m+1} = \alpha$, and so on.

Task 15

Consider the case $m = 3$, so that $t = x' + \alpha x'' + \alpha^2 x'''$ and $\alpha^3 = 1$.

(a) Write the expression which results from t under the permutation of roots that simultaneously substitutes x' in for x'' , x'' in for x''' and x''' in for x'.

(b) Write the expression which results from t under the permutation of powers of α that simultaneously substitutes α in for α^0 , α^2 in for α , and α^0 in for α^2.

(c) Compare the results of (a) and (b), and comment on how this illustrates that 'the result of permuting the roots $x', x'', x''', \ldots$ among themselves will be the same as that of [permuting] the powers of α among themselves.'

(d) Now determine the product αt and compare it to the results of parts (a) and (b). Explain why this proves that αt is also a root of the resolvent.

(e) Determine the permutation of the powers of α which corresponds to the permutation of roots that simultaneously substitutes x' in for x''', x'' in for x''' and x''' in for x'. How could we obtain this same expression as a product of t by a power of α?

We now continue with the remainder of Lagrange's argument that the degree of the resolvent can always be reduced to something less than m!. Tasks 16 and 17 examine this argument more formally for the specific cases of $m = 3$ and $m = 4$ respectively.

Thus, t being one of the roots of the resolvent equation in t, then αt, $\alpha^2 t$, $\alpha^3 t$, $\ldots \alpha^{m-1} t$ will also be roots of this same equation; consequently, the [resolvent] equation . . . will be such that it does not change when t is replaced there by αt, by $\alpha^2 t$, by $\alpha^3 t$, . . . , by $\alpha^{m-1} t$, from which it is easy to conclude first that this equation can only contain powers of t for which the exponent will be a multiple of m.

If therefore one substitutes $\theta = t^m$, one will have an equation in θ which will be of degree only $1.2.3 \ldots [m-1]$.

Let us pause here to consider exactly what Lagrange has just claimed, and why he believed his claims to be true. Based on Task 15, the first claim in the excerpt should seem quite believable; namely,

> Thus, t being one of the roots of the resolvent equation in t, then αt, $\alpha^2 t$, $\alpha^3 t$, $\ldots \alpha^{m-1} t$ will also be roots of this same equation.

Note that no claim has been made that these m resolvent roots (t , αt , $\alpha^2 t$, $\alpha^3 t, \ldots, \alpha^{m-1} t$) are distinct, but only that this list accounts for m of the $m!$ roots of the resolvent. Lagrange continued by stating, without proof, the following consequence of this fact:

> ...consequently, the [resolvent] equation ...will be such that **it does not change** when t is replaced there by αt, by $\alpha^2 t$, by $\alpha^3 t$, ..., by $\alpha^{m-1} t$;

Lagrange appears to again be considering the *form* of the resolvent equation here. For example, letting $m = 3$ and t be a root of the sixth degree resolvent, we can write the resolvent equation as follows:

$$a_6 t^6 + a_5 t^5 + a_4 t^4 + a_3 t^3 + a_2 t^2 + a_1 t + a_0 = 0.$$

Since αt is also a root of the resolvent for any primitive cubic root α, we can replace t by αt to obtain a second form of this equation:

$$a_6(\alpha t)^6 + a_5(\alpha t)^5 + a_4(\alpha t)^4 + a_3(\alpha t)^3 + a_2(\alpha t)^2 + a_1(\alpha t) + a_0 = 0.$$

But Lagrange's assertion that the resolvent equation 'does not change when t is replaced there by αt' seems to be saying that — despite their apparent initial differences in form — these two equations must ultimately possess an *identical* form. To see how this identical form might be derived, begin with the fact that α is a cubic root of unity (so that $\alpha^3 = 1$) in order to re-write the second equation as follows:

$$a_6 t^6 + a_5 \alpha^2 t^5 + a_4 \alpha t^4 + a_3 t^3 + a_2 \alpha^2 t^2 + a_1 \alpha t + a_0 = 0.$$

Comparing this to the original equation (and remembering that $\alpha \neq 0$), note that setting $a_5 = a_4 = a_2 = a_1 = 0$ does indeed produce an identical form, namely

$$a_6 t^6 + a_3 t^3 + a_0 = 0.$$

Setting $\theta = t^3$, we thus arrive at the following equation of degree $(m-1)! = 2! = 2$ for the resolvent:

$$a_6 \theta^2 + a_3 \theta + a_0 = 0.$$

Looking at the form of the resolvent in this way should make it easier to agree with Lagrange's final conclusions:

> ...it is easy to conclude first that this equation can only contain powers of t for which the exponent will be a multiple of m.
>
> If therefore one substitutes $\theta = t^m$, one will have an equation in θ which will be of degree only $1.2.3 \ldots [m-1]$.

Of course, thinking of the equation $a_6 t^6 + a_5 t^5 + a_4 t^4 + a_3 t^3 + a_2 t^2 + a_1 t^1 + a_0 = 0$ in *numerical* terms, it may not be at all clear that setting the coefficients a_5, a_4, a_2 and a_1 equal to zero is the only way in which to obtain the desired outcome. After all, there are lots of ways in which a sum with non-zero terms can end up being equal to zero, as Lagrange was well aware. Tasks 16 and 17, however, outline how this worry can be laid to rest by employing certain special numerical properties of roots of unity.[22]

[22] The proofs given in these tasks are also designed to foreshadow certain features of permutations which we will study in the next section.

Task 16

This task outlines a rigorous proof of Lagrange's claim that the resolvent 'can only contain powers of t for which the exponent will be a multiple of m' in the case of $m = 3$.

That is, given an arbitrary cubic equation, we use the fact that the roots of its resolvent have the form $x' + \alpha x'' + \alpha^2 x'''$, where x' , x'', and x''' are the roots of the original cubic and α is a primitive cubic root of unity,[23] to show that the resolvent has powers of t^3 only, thereby proving that the resolvent of a cubic equation is necessarily quadratic in form.

We begin by letting t_1, t_2 denote the following two roots of the resolvent equation:[24]

$$\begin{aligned} t_1 &= x' + \alpha x'' + \alpha^2 x''' \\ t_2 &= x' + \alpha x''' + \alpha^2 x'' \end{aligned}$$

(a) Explain why the products αt_1 and $\alpha^2 t_1$ give us two of the four remaining roots of the resolvent equation.[25]

Comment on anything you notice about the formal expressions for t_1, αt_1 and $\alpha^2 t_1$.

(b) Explain why the remaining two roots of the resolvent are αt_2 and $\alpha^2 t_2$.

Comment on anything you notice about the formal expressions for t_2, αt_2 and $\alpha^2 t_2$.

(c) Using the results of parts (a) and (b) in the Factor Theorem, we can write the resolvent equation as follows:

$$(t - t_1)(t - \alpha t_1)(t - \alpha^2 t_1)(t - t_2)(t - \alpha t_2)(t - \alpha^2 t_2) = 0$$

For the purpose of the next part of our argument, group these factors together to get the following cubic functions:

$$g_1(t) = (t - t_1)(t - \alpha t_1)(t - \alpha^2 t_1) \qquad ; \qquad g_2(t) = (t - t_2)(t - \alpha t_2)(t - \alpha^2 t_2)$$

(i) Use the fact that α is a primitive cubic root of unity to show that $g_1(t) = t^3 - t_1^3$.

Hint? Review Task11(c) and also Task 14(a) to see how the elementary symmetric functions can be used to avoid literally multiplying out this expression.

(ii) Proceed as in part (i) to show that $g_2(t) = t^3 - t_2^3$.

(d) Use the results of part (c) to show that the resolvent contains only powers of t for which the exponent is a multiple of 3, and is therefore quadratic in form.

[23] Although there are two primitive cubic roots of unity, $-\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $-\frac{1}{2} - \frac{\sqrt{3}}{2}i$, the argument outlined below does not depend on which of these is used.

[24] Note that we are not assuming that t_1, t_2 are real-valued here! Nor are we assuming that $t_1 \neq t_2$. Rather, we are only assuming that t_1, t_2 are the roots of the sixth degree resolvent equation given by these particular arrangements of x', x'', x''' in the formula for the resolvent's roots. Our choice of which of the six possible arrangements to label as t_1 was completely arbitrary (other than a desire to maintain consistency with the notation used in Tasks 1, 10 and 14 for the specific cubic equation $x^3 + nx + p = 0$). Once t_1 was selected, however, our choice for t_2 had to differ from the arrangements given by t_1, αt_1 and $\alpha^2 t_1$ (for reasons which should become clear later in this task).

[25] In our analysis of the specific cubic equation from Task 1(c), we arrived at these same conclusions by substituting values into the resolvent equation which we already knew to be quadratic in form. We can not do that in this general case, since we are now trying to *prove* the resolvent is quadratic in form.

Task 17
This task outlines a rigorous proof Lagrange's claim that the resolvent 'can only contain powers of t for which the exponent will be a multiple of m' in the case of $m = 4$.

Let x_1, x_2, x_3, x_4 be the four roots of an arbitrary quartic equation.

Let α be a primitive fourth root of unity.[26]

(a) Show that the 4!=24 roots of the resolvent for the given quartic can be partitioned into six disjoint sets of 4 roots each of which has the form $S_i = \{t_i\,,\ \alpha t_i\,,\ \alpha^2 t_i\,,\ \alpha^3 t_i\,\}$, where $i \in \{1, 2, 3, 4, 5, 6\}$ and t_i a particular root of the resolvent.

You might start, for example, by setting

$$t_1 \quad = \quad x' + \alpha x'' + \alpha^2 x''' + \alpha^3 x^{(iv)}\ .$$

Denoting the order in which the four roots appear in this expression by '$1, 2, 3, 4$' in order to abbreviate writing, note that the set S_1 then contains the roots corresponding to the following formal expressions:

$$\begin{array}{rcll} t_1 & = & x' + \alpha x'' + \alpha^2 x''' + \alpha^3 x^{(iv)} & \qquad (1,2,3,4) \\ \alpha t_1 & = & x^{(iv)} + \alpha x' + \alpha^2 x'' + \alpha^3 x''' & \qquad (4,1,2,3) \\ \alpha^2 t_1 & = & x''' + \alpha x^{(iv)} + \alpha^2 x' + \alpha^3 x'' & \qquad (3,4,1,2) \\ \alpha^3 t_1 & = & x'' + \alpha x''' + \alpha^2 x^{(iv)} + \alpha^3 x' & \qquad (2,3,4,1) \end{array}$$

You can then choose t_2 to be any of the remaining 20 expressions obtained by some other permutation of $x', x'', x''', x^{(iv)}$ in the formula for the resolvent roots.
Explain how you can now be sure that the formal expressions for the elements in the set $S_1 = \{t_1\,,\ \alpha t_1\,,\ \alpha^2 t_1\,,\ \alpha^3 t_1\,\}$ are distinct from those in the set $S_2 = \{t_2\,,\ \alpha t_2\,,\ \alpha^2 t_2\,,\ \alpha^3 t_2\,\}$.
Then indicate values that can be used for t_3, t_4, t_5, t_6 (without necessarily writing out all the terms) and explain how you are sure the sets they define are similarly disjoint with respect to formal expressions.

(b) For $i \in \{1, 2, 3, 4, 5, 6\}$, let $g_i(t) = (t - t_i)(t - \alpha t_i)(t - \alpha^2 t_i)(t - \alpha^3 t_i)$.
Show that $g_i(t) = t^4 - t_i^4$.

Hint? To avoid literally multiplying out the four factors in g_i, it will be helpful to write out the elementary symmetric functions which define the coefficients of a quartic $t^4 - At^3 + Bt^2 - Ct + D$ in terms of its roots. (Task 11 may help with this.)
Also remember that α is a primitive fourth root of unity (either i or $-i$), and think about the values of $\alpha^2 + 1$ and $1 + \alpha + \alpha^2 + \alpha^3$ in either case.

(c) Conclude that the resolvent is a function of t^4, and explain why it can therefore be treated as a polynomial of degree 6 only. Why is this not a sufficient reduction to complete the algebraic solution of the original quartic?

[26] Although there are also two primitive fourth roots of unity, i and $-i$, the argument outlined below does not depend on which of these is used. Be careful not to implicitly assume that $\alpha = i$ as you complete it.

Although the fact that the resolvent 'can only contain powers of t for which the exponent will be a multiple of m' reduces the degree of the resolvent from $m!$ to $(m-1)!$, as illustrated in Tasks 16 and 17, some difficulty remains even with relatively small values of m. Granted, a resolvent for a quintic equation undergoes a reduction from degree $5! = 120$ down to degree $4! = 24$ — but 24 is still considerably larger than the original equation's degree of 5. Although a similar problem would seem to arise for quartics, where the initial resolvent degree of $4! = 24$ is reduced only to $3! = 6$, Lagrange used other arguments to show that the resolvent in this case could be further reduced to just a cubic equation.[27] He also explained how this reduction relates to the effect of permuting $x', x'', x''', x^{(iv)}$ in the expression for the resolvent's roots t, again focusing on the *form* of the expression in question. In essence, he showed that the resolvent root t can be written in a sufficiently symmetric way that only 3 'values' arise when $x', x'', x''', x^{(iv)}$ are permuted in all possible ways. More specifically, Lagrange showed that $t = \frac{x'x''+x'''x^{(iv)}}{2}$, and that this (nearly symmetric) expression assumes only three distinct forms when $x', x'', x''', x^{(iv)}$ are permuted in all 24 possible ways — check this if you like!

Despite his success with polynomials of degree 3 and 4, Lagrange suspected (and Abel and Galois later confirmed) that it is not always possible to express the resolvent roots of equations in a form that is sufficiently symmetric to achieve a similar result for polynomials of degree five and higher. Nevertheless, Lagrange's introduction of permutations into the picture was the first significant step forward in the study of algebraic solvability in centuries. It also paved the way for Cauchy's work in the theory of permutations, as we will see in Section 2. Interestingly, Cauchy's further development of the theory of permutations essentially ignored the connection of permutations to the solution of equations, a move towards further abstraction which promoted Cayley's ability to eventually define the notion of a completely abstract group.

2 An independent theory of permutations: A. Cauchy

Augustin Cauchy was born in Paris on August 21, 1789, the year the French Revolution began. His family moved to Arcueil, a town just outside of Paris, to avoid the turmoil of the revolution, and Cauchy spent his earliest days there. He was educated by his father, who counted a number of important scientists and mathematicians, including Lagrange, among his friends. It was Lagrange, in fact, who advised Cauchy's father that his son should obtain a good grounding in languages before starting a serious study of mathematics. Cauchy studied classical languages for two years before being trained as an engineer. He worked as an engineer in Cherbourg, France from 1810–1812, during which time he undertook his first mathematical researches. He then lived and worked as a mathematician in Paris for most of his remaining life, with the exception of eight years (1830–1838) of self-imposed exile from France for political reasons.[28] Even after returning to Paris in 1838, he refused to take an oath of allegiance to the political regime then in power and was unable to regain his various teaching positions. Cauchy's staunch royalism and his equally staunch religious zeal made him contentious, and his relations with other mathematicians and scientists were often strained.[29] Nevertheless, his mathematical contributions were (and still are) widely admired for their depth, their breadth, and their rigor. He is especially remembered for his efforts to reformulate calculus in terms of limits. Cauchy died near Paris in the village of Sceaux on May 23, 1857 after contracting a fever on a trip to the country to help restore his health, which had always been weak.

[27] The strategy of reducing a quartic to a cubic was known to Renaissance algebraists; see footnote 3, page 2.

[28] Interestingly, Cauchy taught for a time in Turin, Italy, where Lagrange had his start, during this period of self-exile.

[29] Cauchy is particularly noted for being unsupportive of young mathematicians. Both Abel and Galois, for example, submitted papers to the French Academy of Sciences which were assigned to Cauchy for review; in each case, Cauchy either failed to return the papers promptly or lost them completely.

Cauchy's research on permutations was completed in two different periods, the first of which occurred around 1812. In that year, he presented a paper entitled *Essai sur les fonctions symétriques* to the French Academy of Sciences, the contents of which were later published in two articles in 1815.[30] He did not publish anything further on permutations until 1844–1846, when his extensive *Mémoire sur les arrangements que l'on peut former avec des lettres données* appeared, in addition to 27 shorter articles. In these later works, Cauchy made no mention of polynomial equations, focusing instead on the systematic development of the algebraic properties of permutations as interesting objects of study in their own right. In doing so, Cauchy established the theory of permutations as an independent branch of mathematics which could, by virtue of its generality, then be applied to a variety of mathematical problems; both Cauchy and Cayley made use of permutations in their work on the theory of determinants, for example.

In the introduction of his earliest manuscript [2], however, Cauchy made it clear that his ideas about permutations were initially stimulated by the specific problem of algebraic solvability, and Lagrange's work in that area in particular. As you read the following excerpt from that introduction, remember that Lagrange had identified the number of distinguishable forms that result from permuting the variables in an expression as a potential tool in studying algebraic solvability.[31] It is this idea that Cauchy extended beyond the realm of formulas for the roots of a resolvent equation, applying it more generally to any function of n variables.

Memoire on the number of forms that a function can assume
by permuting the quantities involved in all possible ways

Messieurs Lagrange and Vandermonde[32] were, I believe, the first to have considered functions of several variables relative to the number of forms they can assume when one substitutes these variables in place of each other. ... Since then, several Italian mathematicians have productively occupied themselves with this matter, and particularly Monsieur Ruffini[33] ... One of the most remarkable consequences of the work of these various mathematicians is that, for a given number of letters, it is not always possible to form a function which has a specified number of forms.

Let us relate the final statement of this excerpt back to our reading of Lagrange. We start with a positive result; namely, that it is possible to find a function $t(x_1, x_2, x_3, x_4)$ of 4 variables which has exactly 3 forms when its variables are permuted in all possible 4!=24 ways. In the context of Lagrange's analysis, the function t gave the roots of a resolvent equation in terms of the roots x_1, x_2, x_3, x_4 of the original quartic polynomial, and the fact that t has just three distinct forms under permutations of these roots meant that the resolvent equation has just three roots.[34] On the other hand, the degree of the resolvent equation for a quintic can not in general be reduced to degree four, due to the fact that it is impossible to form a function of 5 variables $t(x_1, x_2, x_3, x_4, x_5)$ which has exactly 4 forms when the variables are permuted in all the 5!=120 possible ways. As noted earlier, the impossibility of such

[30] See [2, 3] in the bibliography for their titles.

[31] This number can be considered in a sense to be a measure of the amount of symmetry present in the given expression.

[32] Alexandre-Théophile Vandermonde (1735–1796) was a French musician and musical theorist who wrote four mathematical papers between 1771 and 1773. The first of these considered the solvability of algebraic equations, and appeared nearly simultaneously with Lagrange's work on this same problem. Lagrange concluded his own 1808 note on algebraic solvability with a summary of 'the beautiful work done by Vandermonde.'

[33] See footnote 5, page 3, for more information on Ruffini.

[34] See the discussion at the top of page 23 for further detail.

a function was suspected by Lagrange, and proven by his successors. As a natural follow-up to this and similar impossibility statements, Cauchy pursued the following question in his earliest manuscript on permutations: for a given number n of variables, what can be said about the possible number of distinct forms which a function of n variables *can* produce under permutations of those variables?

In developing an answer to this question, Cauchy introduced new notation for permutations which was far superior to that used by Lagrange. (In fact, we have not used Lagrange's notation in this project because it was so unwieldy.) Our next excerpt, taken from Cauchy's 1844 manuscript on the theory of permutations [4], explains this notation.

Memoire on the arrangements that can be formed with given letters,
and on the permutations or substitutions which provide the passage from one arrangement to another

§1^{st} - *General Considerations*

Let $x, y, z, \ldots$ be distinct letters, which we assume to represent independent variables. If we number the places occupied by these variables in a certain function Ω, and then write these variables $x, y, z \ldots$ in the order assigned to the places that they occupy, we obtain a certain *arrangement*

$$xyz \ldots,$$

and when the variables are displaced, this arrangement will be replaced by another, which can be compared to the first by knowing the nature of the displacements.

. . .

We call *permutation* or *substitution* the operation which consists of displacing these variable, by substituting them for each other, in the form given by the function Ω, or in the corresponding arrangement. To denote this permutation, we write the new arrangement that is produced *below* the original, and we close the system of these two arrangements between parentheses. Thus, for example, being given the function

$$\Omega = x + 2y + 3z,$$

where the variables x, y, z occupy respectively the first, the second and the third place, and consequently succeed each other in the order indicated by the arrangement

$$xyz,$$

if we exchange the variables y, z which occupy the two final places, we obtain a new form Ω' from Ω, which will be distinct from the first, and is determined by the formula

$$\Omega' = x + 2z + 3y.$$

Moreover, the new arrangement, corresponding to the new form, will be

$$xzy,$$

and the permutation by which we pass from the first form to the second will be represented by the notation[35]

$$\begin{pmatrix} xyz \\ xzy \end{pmatrix}$$

which indicates sufficiently the manner in which the variables have been displaced.

[35] Cauchy himself wrote the new arrangement *above* the original arrangement; throughout this project, we modify Cauchy's notation slightly so that it will be identical to that used in current texts.

This done, the different forms of a function of n letters correspond evidently to the distinct arrangements that one can form with these n letters. Moreover, the number of these arrangements is, as we know, given by the product

$$1.2.3\ldots n.$$

[W]e . . . put, as an abbreviation,

$$N = 1.2.3\ldots n$$

. . .

Observe that we can, with no inconvenience, erase any letter that appears in the same place in the two terms of a given permutation, thereby indicating that the letter will not be displaced. Thus, in particular, we will have

$$\begin{pmatrix} xyz \\ xzy \end{pmatrix} = \begin{pmatrix} yz \\ zy \end{pmatrix},$$

which sufficiently indicates the manner in which the variables are displaced. The two arrangements xyz, xzy, included in this permutations, form that which we will call its *two terms*, or its *numerator* and its *denominator*.

Task 18
Let $\Omega = x + 2y + 3z$.

(a) Use Cauchy's notation to write the permutation associated with $\Omega'' = y + 2z + 3x$.

(b) Write the form Ω''' obtained from applying the permutation $\begin{pmatrix} xyz \\ zyx \end{pmatrix}$ to Ω .

Note that Cauchy used the term 'arrangement' when referring to an ordered list of variables (e.g., 'xyz', 'xzy'). Today, the term 'permutation' is often used for this purpose; this is especially common usage in probability and combinatorics, where the ordered list 'xzy' is called a permutation of the letters x, y, z. In group theory, the word 'permutation' continues to be used in Cauchy's sense of the word, and refers to the process which changes one arrangement (such as 'xyz') to another arrangement (such as 'xzy'). In other words, a permutation is a *function* which maps one set of objects (in this case, the letters x, y, z) onto the same set in a one-to-one fashion. The notation Cauchy used for permutations is especially well-suited to remind us of this, given its resemblance to a table of function values. Thus, if we let $f = \begin{pmatrix} xyz \\ xzy \end{pmatrix}$, then we would have $f(x) = x$, $f(y) = z$ and $f(z) = y$.

Cauchy himself used the terms 'permutation' and 'substitution' interchangeably in his later work. In the remainder of this project, we will use the term 'permutation' in our translation of his work into English, in keeping with current usage. In Subsection 2.1, we examine excerpts from his 1845 work which explain the notion of permutation multiplication and its basic properties. With these properties in hand, we then turn in Subsection 2.2 to Cauchy's study of the algebraic structure obtained by considering a set of permutations together with the operation of permutation multiplication, a structure known today as a *permutation group*.

2.1 Multiplication of permutations in Cauchy's theory

Although he did not explicitly discuss the fact that permutations are one-to-one functions from a set onto itself,[36] Cauchy was certainly aware of their function nature. This is clear from the following excerpt, in which two types of products involving permutations are discussed. As you read this excerpt, notice that the first type of product — that of an arrangement by a permutation — simply treats the arrangement xyz as an input value for the function f defined by the permutation $\begin{pmatrix} xyz \\ xzy \end{pmatrix}$, with the function values $f(x), f(y), f(z)$ evaluated all at once. The second type of product — that of two permutations — is the operation of *function composition.* It is this second operation — function composition — that mathematicians have in mind when speaking of permutation products today.

∞∞∞∞∞∞∞∞

The *product* of a given arrangement xyz by a permutation $\begin{pmatrix} xyz \\ xzy \end{pmatrix}$ is the new arrangement xzy which is obtained by applying this same substitution to the given arrangement. The *product* of two permutations will be the new permutation that always furnishes, for any arbitrary arrangement, the result obtained by the application of the two [permutations], applied one after the other. The two given permutations are called the two *factors* of the product. The product of an arrangement by a permutation or of a permutation by another [permutation] will be indicated by the [same] notation which serves to indicate the product of two quantities . . . We find, for example,

$$\begin{pmatrix} xyz \\ xzy \end{pmatrix} xyz = xzy$$

and

$$\begin{pmatrix} xyzu \\ yxuz \end{pmatrix} = \begin{pmatrix} xy \\ yx \end{pmatrix}\begin{pmatrix} zu \\ uz \end{pmatrix}$$

There is more; we can, in the second part of the last equation, interchange the two factors with no inconvenience, so that one will have again

$$\begin{pmatrix} xyzu \\ yxuz \end{pmatrix} = \begin{pmatrix} zu \\ uz \end{pmatrix}\begin{pmatrix} xy \\ yx \end{pmatrix}.$$

But this interchange will not always be possible, and the product of two permutations will often vary when the two factors are interchanged. Thus, in particular, we will find

$$\begin{pmatrix} xy \\ yx \end{pmatrix}\begin{pmatrix} yz \\ zy \end{pmatrix} = \begin{pmatrix} xyz \\ yzx \end{pmatrix} \qquad \begin{pmatrix} yz \\ zy \end{pmatrix}\begin{pmatrix} xy \\ yx \end{pmatrix} = \begin{pmatrix} xyz \\ zxy \end{pmatrix}$$

We will say that two permutations *commute*, when their product is independent of the order in which the two factors occur.

∞∞∞∞∞∞∞∞

To verify the products in Cauchy's illustration of the non-commutative nature of permutation multiplication, remember that each 'product' is really a composition function with $(g \circ f)(z) = g(f(z))$. To compute the product of permutations, therefore, we must begin with the *right* most factor (or function), and then move to the left. For example, to determine where z is mapped by the product

[36] A one-to-one onto function is also called a bijection.

$\begin{pmatrix} xy \\ yx \end{pmatrix}\begin{pmatrix} yz \\ zy \end{pmatrix}$, we first look at the effect of the (right-most) permutation $\begin{pmatrix} yz \\ zy \end{pmatrix}$ on the input z, finding that z is mapped to y. We then use y as our input and look at how it is transformed by the (left-most) permutation $\begin{pmatrix} xy \\ yx \end{pmatrix}$, finding that y is mapped to x. We conclude that z is mapped to x by the product permutation, as is indeed the case in Cauchy's example: $\begin{pmatrix} xy \\ yx \end{pmatrix}\begin{pmatrix} yz \\ zy \end{pmatrix} = \begin{pmatrix} xyz \\ yzx \end{pmatrix}$.

Task 19

(a) By tracing the effect which each factor has on individual variables, moving from right to left, explain why x is mapped to y and why y is mapped to z by the product $\begin{pmatrix} xy \\ yx \end{pmatrix}\begin{pmatrix} yz \\ zy \end{pmatrix}$. This completes the verification of the first half of Cauchy's illustration.

(b) Verify the second half of Cauchy's illustration: $\begin{pmatrix} yz \\ zy \end{pmatrix}\begin{pmatrix} xy \\ yx \end{pmatrix} = \begin{pmatrix} xyz \\ zxy \end{pmatrix}$

(c) Compute and compare the two products $\begin{pmatrix} xyu \\ uxy \end{pmatrix}\begin{pmatrix} xzu \\ zux \end{pmatrix}$ and $\begin{pmatrix} xzu \\ zux \end{pmatrix}\begin{pmatrix} xyu \\ uxy \end{pmatrix}$.

Write each in the form $\begin{pmatrix} x\,y\,z\,u \\ *\,*\,*\,* \end{pmatrix}$.

(d) Compute and compare the two products $\begin{pmatrix} xyu \\ uxy \end{pmatrix}\begin{pmatrix} zvw \\ wvz \end{pmatrix}$ and $\begin{pmatrix} zvw \\ wvz \end{pmatrix}\begin{pmatrix} xyu \\ uxy \end{pmatrix}$.

Write each in the form $\begin{pmatrix} x\,y\,z\,u\,v\,w \\ *\,*\,*\,*\,*\,* \end{pmatrix}$.

(e) Based on Cauchy's examples and parts (c) and (d) of this task, what conjectures, if any, do you have about permutations that commute? Explain your reasoning.

Task 20

Recall that function composition is associative, so that $(f \circ g) \circ h = f \circ (g \circ h)$. Accordingly, permutation multiplication is also associative. Verify this in the following particular case by computing the product below in the two ways indicated:

(a) $\left[\begin{pmatrix} xyzu \\ yxzu \end{pmatrix}\begin{pmatrix} xyzu \\ zxuy \end{pmatrix}\right]\begin{pmatrix} xyzu \\ uxyz \end{pmatrix}$ **(b)** $\begin{pmatrix} xyzu \\ yxzu \end{pmatrix}\left[\begin{pmatrix} xyzu \\ zxuy \end{pmatrix}\begin{pmatrix} xyzu \\ uxyz \end{pmatrix}\right]$

Task 21

Cauchy sometimes used indexed letters $x_1, x_2, \ldots$ to denote the objects being permuted. Clearly, this does not affect how permutation products are computed, but only how much writing we do and how easy it is to read the results. To see this, compute the following:

$$\begin{pmatrix} x_1\,x_2\,x_3\,x_4\,x_5 \\ x_3\,x_1\,x_2\,x_5\,x_4 \end{pmatrix}\begin{pmatrix} x_1\,x_2\,x_3\,x_4\,x_5 \\ x_5\,x_4\,x_2\,x_3\,x_1 \end{pmatrix}$$

How might we denote permutations involving indexed letters more concisely?

In our next two excerpts, the results of multiplying a permutation by itself are explored, and an idea related to the algebraic structure of the set of m^{th} roots of unity is introduced. The commentary and project tasks interspersed between and following these two excerpts will elaborate on the details of Cauchy's arguments. Note that Cauchy has used i to denote a natural number, *not* the imaginary square root of -1. Also remember that '1' now represents the multiplicative identity for permutations.

Nothing is lost if we represent the arrangements formed by several variables by simple letters

$$A, B, C, \ldots$$

or by letters affixed with indices

$$A_1, A_2, A_3, \ldots.$$

Then the permutation which has for its terms A and B will be simply represented in the form

$$\begin{pmatrix}A\\B\end{pmatrix}\ldots$$

The total number of permutations relative to a system of n variables . . . will evidently be equal to the number N of arrangements that can be formed with these variables. . . . The permutation for which the numerator and the denominator are the same, can be supposed to reduce to unity, since one can evidently replace it by the factor 1 in products:

$$\begin{pmatrix}A\\A\end{pmatrix}C = C, \qquad \begin{pmatrix}A\\A\end{pmatrix}\begin{pmatrix}C\\D\end{pmatrix} = \begin{pmatrix}C\\D\end{pmatrix}.$$

A permutation $\begin{pmatrix}A\\B\end{pmatrix}$, multiplied by itself several times in a row, gives for successive products its *square*, its *cube*, and generally its different *powers*, which are naturally represented by the notation

$$\begin{pmatrix}A\\B\end{pmatrix}^2, \begin{pmatrix}A\\B\end{pmatrix}^3, \begin{pmatrix}A\\B\end{pmatrix}^4 \ldots,$$

can never include more than N actually distinct permutations. Therefore, in extending this series, we will eventually see the same permutations.

What's more, if we suppose that

$$\begin{pmatrix}A\\B\end{pmatrix}^h = \begin{pmatrix}A\\B\end{pmatrix}^l,$$

h being $< l$, then, in setting, as an abbreviation,

$$l = i + h$$

we will have[37]

$$\begin{pmatrix}A\\B\end{pmatrix}^h = \begin{pmatrix}A\\B\end{pmatrix}^l = \begin{pmatrix}A\\B\end{pmatrix}^{i+h} = \begin{pmatrix}A\\B\end{pmatrix}^i \begin{pmatrix}A\\B\end{pmatrix}^h,$$

consequently

$$\begin{pmatrix}A\\B\end{pmatrix}^i = 1,$$

i being evidently less than l.

[37] In the interest of clarity, a minor modification has been made in the next line of Cauchy's original text.

The central objective of the excerpt we have just read was to prove that for every permutation $\begin{pmatrix} A \\ B \end{pmatrix}$, there exists a natural number $i \leq N$ for which $\begin{pmatrix} A \\ B \end{pmatrix}^i = 1$, where $N = n!$ is the number of all distinct permutations. Task 22 explores the property which defines the number i for a given permutation, while Task 23 further examines Cauchy's proof of the existence of such a number for any permutation.

Task 22
Letting $\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} xyzu \\ uxyz \end{pmatrix}$, compute powers of $\begin{pmatrix} A \\ B \end{pmatrix}$ to determine the *smallest* natural number i for which $\begin{pmatrix} A \\ B \end{pmatrix}^i = 1$. *(Since there are* $4! = 24$ *different permutations of* 4 *variables, we know* $i \leq 24$ *…hopefully, you won't have to go as far as the* 24^{th} *power!)*

Task 23
In the preceding excerpt, Cauchy derived his conclusion that $\begin{pmatrix} A \\ B \end{pmatrix}^i = 1$ directly from the fact[38] that $\begin{pmatrix} A \\ B \end{pmatrix}^h = \begin{pmatrix} A \\ B \end{pmatrix}^i \begin{pmatrix} A \\ B \end{pmatrix}^h$. Setting $X = \begin{pmatrix} A \\ B \end{pmatrix}^h$ and $Y = \begin{pmatrix} A \\ B \end{pmatrix}^i$, we can rewrite this part of his argument as follows:

$$X = XY \;\Rightarrow\; Y = 1, \text{ where 1 represents the identity.}$$

Although this algebraic property always holds in the case where X, Y are non-zero real or complex numbers, it does not hold in all algebraic structures.

(a) Illustrate the failure of this fact in matrix algebra by showing that we have $XY = X$ with $X \neq 0$ and $Y \neq I$ (where I is the 2×2 identity matrix) for the following specific matrices:

$$X = \begin{bmatrix} 0 & 0 \\ 0 & 7 \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}.$$

(b) Now let X, Y be permutations and suppose $X = XY$. According to Cauchy's argument, this is sufficient to conclude, without further ado, that Y must be the identity permutation in this case. Write as convincing an argument as you can to justify this conclusion, based only on what you already know about how permutations behave.

We now look at the continuation of the preceding excerpt, where we will see Cauchy's argument that the powers of $\begin{pmatrix} A \\ B \end{pmatrix}$ behave periodically, in a fashion reminiscent of the roots of unity. Notice that this proof starts from the (already established) fact that there exists a power i with $\begin{pmatrix} A \\ B \end{pmatrix}^i = 1$. Be sure that you can justify each step of the argument that follows from there as you read through it.

[38] Cauchy's argument for this fact is based on the so-called 'Pigeonhole Principle.' With only a finite number ($N = n!$) of different values possible for all the infinitely many powers of $\begin{pmatrix} A \\ B \end{pmatrix}$, some two of these powers (pigeons) must share the same value (hole). That is, there must be some $h \neq l$ such that $\begin{pmatrix} A \\ B \end{pmatrix}^h = \begin{pmatrix} A \\ B \end{pmatrix}^l$.

There is more; if, taking i to be the value determined by the preceding formula, we let l be an arbitrary whole number, k the quotient when l is divided by i, and j the remainder of this division, so that we have

$$l = ki + j,$$

j being less than i, we will find not only that

$$\begin{pmatrix} A \\ B \end{pmatrix}^{ki} = \left[\begin{pmatrix} A \\ B \end{pmatrix}^{i}\right]^{k} = 1^k = 1,$$

but, furthermore, that

$$\begin{pmatrix} A \\ B \end{pmatrix}^{l} = \begin{pmatrix} A \\ B \end{pmatrix}^{ki} \begin{pmatrix} A \\ B \end{pmatrix}^{j} = \begin{pmatrix} A \\ B \end{pmatrix}^{j},$$

and, examining the former formula in the case where the number k is reduced to zero, we will also have

$$\begin{pmatrix} A \\ B \end{pmatrix}^{0} = 1$$

In virtue of these remarks that we have just made, if we prolong indefinitely the series whose terms are

$$\begin{pmatrix} A \\ B \end{pmatrix}^{0} = 1\,, \begin{pmatrix} A \\ B \end{pmatrix}, \begin{pmatrix} A \\ B \end{pmatrix}^{2}, \begin{pmatrix} A \\ B \end{pmatrix}^{3}, \text{ etc}\ldots,$$

the unity [1] will be the first of these terms that will be repeated, and from there the rest of the terms already found will reappear periodically in the same order, so that one will have, for example

$$\begin{array}{ccccccc} 1 & = & \begin{pmatrix} A \\ B \end{pmatrix}^{i} & = & \begin{pmatrix} A \\ B \end{pmatrix}^{2i} & = & \ldots, \\ \begin{pmatrix} A \\ B \end{pmatrix} & = & \begin{pmatrix} A \\ B \end{pmatrix}^{i+1} & = & \begin{pmatrix} A \\ B \end{pmatrix}^{2i+1} & = & \ldots, \\ \begin{pmatrix} A \\ B \end{pmatrix}^{2} & = & \begin{pmatrix} A \\ B \end{pmatrix}^{i+2} & = & \begin{pmatrix} A \\ B \end{pmatrix}^{2i+2} & = & \ldots, \\ \text{etc.} & \ldots & & & & & \end{array}$$

Therefore the number i of distinct terms of the series will always be the smallest natural number value of i which satisfies the formula

$$\begin{pmatrix} A \\ B \end{pmatrix}^{i} = 1$$

The number i thereby determined, or the degree of the smallest power of $\begin{pmatrix} A \\ B \end{pmatrix}$ equivalent to the identity, is what we will call the degree or the *order* of the permutation $\begin{pmatrix} A \\ B \end{pmatrix}$.

Cauchy's definition of the ***order of a permutation*** is still used today. For an example illustrating this concept, look back at Task 22, where you should have found that $i = 4$ is the *smallest* natural number for which $\begin{pmatrix} xyzu \\ uxyz \end{pmatrix}^i = 1$. This means that 4 is the order of the permutation $\begin{pmatrix} xyzu \\ uxyz \end{pmatrix}$.

Task 24
Compute powers of the following permutations to determine the order of each.

(a) $\begin{pmatrix} xyzu \\ zxyu \end{pmatrix}$ **(b)** $\begin{pmatrix} xyzu \\ xzyu \end{pmatrix}$ **(c)** $\begin{pmatrix} xyzuv \\ zxyvu \end{pmatrix}$

As you have been computing products and powers of permutations, you may have noticed that some permutations behave simply by cycling through all the variables in a particular order, as in the case of $\begin{pmatrix} xyzu \\ uxyz \end{pmatrix}$, where we have $x \to u \to z \to y \to x$. Other permutations have more complicated behavior, but still exhibit this sort of cyclic behavior on various subsets of their variables; for example, the permutation $\begin{pmatrix} xyzuv \\ zxyvu \end{pmatrix}$ could be thought of as two separate cycles: $x \to z \to y \to x$ and $u \to v \to u$. In our next excerpt, Cauchy's terminology and notation for permutations with this behavior is introduced.

We now suppose that a permutation reduced to its simplest expression has the form

$$\begin{pmatrix} xy \ldots uvw \\ yz \ldots vwx \end{pmatrix}$$

that is to say that it operates by replacing x by y, then y by z, ..., and so in succession until the final variable w is reached, which is replaced by the variable x with which we began. To carry out this permutation, we can arrange the different variables,

$$x, y, z, \ldots, u, v, w,$$

on the circumference of an *indicator* circle, divided in equal parts, placing the first, the second, the third, ... [variable] on the first, the second, the third, ... point of the division, then to replace every variable with that which comes to take its place when we turn the indicator circle in a certain direction. For this reason we give the permutation in question the name *circular permutation*. We will represent it, as an abbreviation, by the notation

$$(x, y, z, \ldots, u, v, w);$$

and it is clear, in this notation, that any one of the variables

$$x, y, z, \ldots, u, v, w$$

can occupy the first place. Thus, for example, we have the identity

$$(x, y, z) = (y, z, x) = (z, x, y).$$

Today, we use the (shorter) term ***cycle*** for what Cauchy called a *circular permutation.* Although many current texts also omit the commas between elements in a cycle as a notational abbreviation, we follow Cauchy's convention in this regard in the remainder of this project.

Task 25
Which of the following cycles are equal?

$$(x,\, y,\, z,\, u,\, v) \quad ; \quad (y,\, z,\, v,\, x,\, u) \quad ; \quad (z,\, u,\, v,\, x,\, y) \quad ; \quad (u,\, z,\, x,\, y,\, v) \quad ; \quad (v,\, x,\, u,\, y,\, z)$$

Our next excerpt from Cauchy gives a simple rule for finding the order of a cycle. We omit Cauchy's argument for this rule, which involved examining how far the 'indicator circle' must be turned in order to return all the variables to themselves. You should, however, be able to convince yourself of the truth of the rule he gives simply by thinking about the nature of cycles. Cauchy also gave two examples to illustrate the fact that his rule for the order of a cycle gives the same result as would be obtained by taking powers of the cycle until we arrive at the identity; the first of his examples is included in this excerpt and further examined in Task 26; the second of his examples is considered in Task 27.

If we call i the number of variables included in a circular permutation

$$(x, y, z, \ldots, u, v, w),$$

then ...... the order of a circular permutation will be exactly the number i of letters that it contains.

...

If, to fix these ideas, we let $i = 4$, then, in raising the circular permutation

$$(x,\, y,\, z,\, u),$$

to the second and to the third power, we would find

$$(x,\, y,\, z,\, u)^2 = (x,\, z)(y,\, u) \qquad , \qquad (x,\, y,\, z,\, u)^3 = (x,\, u,\, z,\, y)$$

.

Referring to a cycle that involves exactly i letters as an ***i*-*cycle***, Cauchy has just claimed that every i-cycle has order i. In addition to further exploring this claim, the next two tasks direct our attention to similarities in the behavior of an i-cycle and that of a primitive i^{th} root of unity.

Task 26
Consider the 4-cycle $\alpha = (x,\, y,\, z,\, u)$.

(a) Verify Cauchy's claims in the preceding excerpt that $\alpha^2 = (x,\, z)(y,\, u)$ and that $\alpha^3 = (x,\, u,\, z,\, y)$. Then verify that $\alpha^4 = 1$, either by computing $\left[\alpha^2\right]^2$ or $\alpha^3\alpha$.

Explain why this confirms that any arbitrary 4-cycle will have order 4.

(b) Note that the permutation $\alpha^3 = (x,\, u,\, z,\, y)$ is also a 4-cycle and thus has order 4. Determine the order of the permutation $\alpha^2 = (x,\, z)(y,\, u)$, and justify your answer.

(c) Compare the set $\{\alpha, \alpha^2, \alpha^3, 1\}$ with the set of 4^{th} roots of unity $\{i, -1, -i, 1\}$.

Task 27
Consider an arbitrary 6-cycle $\beta = (x,\, y,\, z,\, u,\, v,\, w)$.

(a) Compute the permutations β^k for $k = 2, 3, 4, 5, 6$ to confirm that β has order 6.

(b) Which of the permutations β^k, other than β itself, also have order 6? Justify.

(c) What is the order of the permutations β^k which are not of order 6? Justify.

(d) Compare the set $\{\beta, \beta^2, \beta^3, \beta^4, \beta^5, 1\}$ with the set of 6^{th} roots of unity. (See Task 3.)

From your work in Tasks 26 and 27, it should be clear that the power of a cycle can be a product of two or more cycles, rather than just a single cycle. For instance, for $\alpha = (x,\, y,\, z,\, u)$, we have $\alpha^2 = (x,\, z)(y,\, u)$. Notice also that none of the elements in these factors overlap; in other words, the cycles that appear as factors in this product are ***disjoint***. In the following excerpt, Cauchy explained that every permutation can be written as the product of disjoint cycles in this way. We call this product the ***cycle decomposition*** of the permutation α. As you continue reading, keep the question of whether a cycle decomposition is necessarily unique (up to the order of the factors) in mind.

> Now let A and B be two arbitrary arrangements formed with n variables $x, y, z, \ldots$. To substitute the second arrangement for the first, it will evidently suffice to operate with one or several circular permutations [cycles], that can be readily formed by placing two variables in the order in which the one will be replaced by the other when we pass from the first arrangement to the second. Consequently, the permutation, reduced to its simplest expression, will necessarily be either a circular permutation [cycle], or the product of several circular permutations [cycles]. We find, for example, ...
>
> $$\begin{pmatrix} xyzu \\ uzyx \end{pmatrix} = (x,\, u)(y,\, z) \qquad , \qquad \begin{pmatrix} xyzuv \\ zuvyx \end{pmatrix} = (x,\, z,\, v)(y,\, u)$$
>
> The circular permutations [cycles] of which the arbitrary permutation $\begin{pmatrix} A \\ B \end{pmatrix}$ will be the product are called the *cyclic factors* of $\begin{pmatrix} A \\ B \end{pmatrix}$. Any arbitrary two of these [cyclic factors], being composed of distinct letters, will clearly commute. Thus, all the cyclic factors of a permutation commute with each other, and will represent the permutation in question in any order.

You may already have noticed that disjoint cycles necessarily commute, as noted above by Cauchy. Look back, for example, at your explorations in Task 19, especially part (e). You may also have remarked how useful the commutativity of disjoint cycles can be when computing products or powers. Representing a permutation by its cycle decomposition allows us to take full advantage of this. For example, using commutativity of disjoint cycles, together with associativity of permutation products, we know that

$$[(x,\, u)(y,\, z)]\,[(x,\, u)(y,\, z)] = [(x,\, u)(x,\, u)]\,[(y,\, z)(y,\, z)]\,,$$

so that

$$\begin{pmatrix} xyzu \\ uzyx \end{pmatrix}^2 = [(x,\, u)(y,\, z)]^2 = (x,\, u)^2(y,\, z)^2 = 1$$

Note that this also shows that the permutation $\begin{pmatrix} xyzu \\ uzyx \end{pmatrix}$ has order 2.

Task 28

Use the fact that disjoint cycles commute to simplify computations in this task.

[Remember that multiplication of permutations in general is not commutative!]

(a) Explain why the following permutation has order 2: $\alpha = \begin{pmatrix} xyzu \\ zuxy \end{pmatrix} = (x,\, z)(y,\, u)$

(b) Explain why the following permutation has order 2: $\beta = \begin{pmatrix} xyzuvw \\ zuxyvw \end{pmatrix} = (x,\, z)(y,\, u)(v,\, w)$

(c) What is the order of the following permutation? Justify. $\gamma = \begin{pmatrix} xyzuvw \\ zxyvwu \end{pmatrix} = (x,\, z,\, y)(u,\, v,\, w)$

Task 29

Consider the permutation $\alpha = \begin{pmatrix} xyzuv \\ zuvyx \end{pmatrix} = (x,\, z,\, v)(y,\, u)$

(a) Make a conjecture concerning the order of the permutation α; explain why you think this will be correct based on what we know about the order of cycles.

(b) Find the order of α by computing its powers until the identity 1 is reached.
Use the fact that disjoint cycles commute to simplify the computation.
[Remember that multiplication of permutations in general is not commutative!]

(c) Was your conjecture in part (a) correct? If not, how would you modify it?

Task 30

Products of disjoint cycles are just one representation form for permutations Cauchy studied; this task explores another such product decomposition which remains important today.

We will say that a permutation P is a ***transposition*** if and only if P is a 2-cycle.[39]

Note that a 3-cycle can be written as the product of transpositions in (at least) two ways:

$$(x_1\,,\, x_2\,,\, x_3) = (x_1\,,\, x_2)(x_2\,,\, x_3) \qquad \textit{and} \qquad (x_1\,,\, x_2\,,\, x_3) = (x_1\,,\, x_3)(x_1\,,\, x_2)$$

Also note that both decompositions involve exactly two transpositions.
Although it is not possible to decompose a 3-cycle into fewer than two transpositions, it can be done with more; for example, $(x_1\,,\, x_2\,,\, x_3) = (x_1\,,\, x_2)(x_3\,,\, x_1)(x_1\,,\, x_3)(x_2\,,\, x_3)$.

(a) Using as few transpositions as possible, write the general 4-cycle $(x_1\,,\, x_2\,,\, x_3\,,\, x_4)$ as the product of transpositions in at least two different ways.

(b) Using as few transpositions as possible, write the general n-cycle $(x_1\,,\, x_2\,,\, \ldots\,,\, x_n)$ as the product of transpositions in two different ways.

(c) Explain how to write an arbitrary permutation P as the product of transpositions. Don't forget that the identity 1 is also a permutation!

[39] Cauchy introduced this definition in his earliest works on permutation theory.

Before we turn to Cauchy's writing on systems of permutations in Subsection 2.2, we need one more algebraic idea related to individual permutations, that of an ***inverse permutation***. Cauchy wrote about inverses in several sections of his 1845 manuscript; in the next excerpt, we have selected only those of his comments which are relevant to our immediate purposes of this project.

It is good to observe that if, after substituting for the arrangement A another arrangement B, we wish to return from the arrangement B to the arrangement A, this second operation, the inverse of the first, will be represented not by the notation $\begin{pmatrix} A \\ B \end{pmatrix}$, but by the notation $\begin{pmatrix} B \\ A \end{pmatrix}$. Consequently, it is natural to say that the two permutations

$$\begin{pmatrix} A \\ B \end{pmatrix} , \begin{pmatrix} B \\ A \end{pmatrix}$$

are *inverses* of each other. That said, it is clear that, if the permutation which puts in the place of x another variable y, the inverse permutation will put, oppositely, y in the place of x. ...

. . . .

. . . [I]f we let $P = \begin{pmatrix} A \\ B \end{pmatrix}$, then P^{-1} will be the permutation which, when multiplied by $\begin{pmatrix} A \\ B \end{pmatrix}$, will have as its product 1; that is to say, the permutation $\begin{pmatrix} B \\ A \end{pmatrix}$, inverse of $\begin{pmatrix} A \\ B \end{pmatrix}$. Thus, the notations

$$P, P^{-1}$$

designate in general two permutations which are inverses of each other.

It is perhaps important to remark here that the excerpt you have just read omits the algebraic argument which Cauchy gave concerning why it makes sense to denote the inverse of a permutation P by the (negative) power P^{-1}. You are, of course, already familiar with this notation for inverse functions from calculus.[40] The general algebraic identity $P^{-1}P^{1} = P^{-1+1} = P^{0} = 1$, which is more in keeping with Cauchy's actual argument in favor of this notation, will also be familiar to you.

What may be less familiar is the process of actually finding the inverse of a given permutation. In the first part of this excerpt, Cauchy suggested that one way to find the inverse of the permutation $P = \begin{pmatrix} A \\ B \end{pmatrix}$ is to simply reverse the two rows, to get $P^{-1} = \begin{pmatrix} B \\ A \end{pmatrix}$. For example, if we let

$$P = \begin{pmatrix} xyzu \\ yzux \end{pmatrix}$$

then

$$P^{-1} = \begin{pmatrix} yzux \\ xyzu \end{pmatrix} .$$

If this last permutation looks somehow disarranged to you, it is only because we have been writing the top row as '$xyzu$' in this project; putting the variables of the top row back in the order we have come to expect them and moving the corresponding variables in the bottom row along with them, we get

$$P^{-1} = \begin{pmatrix} y & z & u & x \\ | & | & | & | \\ x & y & z & u \end{pmatrix} = \begin{pmatrix} x & y & z & u \\ | & | & | & | \\ u & x & y & z \end{pmatrix} = \begin{pmatrix} xyzu \\ uxyz \end{pmatrix}$$

[40] Since a permutation is really a one-to-one onto function, it necessarily has an inverse function.

Finding inverses of cycles is particularly easy, since we can simply reverse the order, as shown here:

$$P = \begin{pmatrix} xyzu \\ yzux \end{pmatrix} = (x,\, y,\, z,\, u) \quad \Rightarrow \quad P^{-1} = (x,\, y,\, z,\, u)^{-1} = (u,\, z,\, y,\, x) = (x,\, u,\, z,\, y)$$

There are other procedures for finding the inverse of a given permutation as well. However, the actual procedure for doing this is less important than the general algebraic concept involved; namely, for every permutation P, there is a unique permutation P^{-1} for which $PP^{-1} = P^{-1}P = 1$. Using this concept, for instance, allows us to easily prove that the equation $XY = X$ implies $Y = 1$ in permutation theory by simply multiplying both sides of the equation $XY = X$ on the left by X^{-1}. To more fully appreciate the (algebraic) power offered by the existence of inverses, you may wish to compare this proof to the one which you gave for this same fact in Task 23(b).

Task 31

(a) Find the inverse of the following permutations without resorting to cycle notation.

(i) $R = \begin{pmatrix} xyzuv \\ zxvuy \end{pmatrix}$ **(ii)** $S = \begin{pmatrix} xyzuv \\ zuvyx \end{pmatrix}$

(b) Find the inverse of the following cycles.

(i) $T = (z,\, y,\, x)$ **(ii)** $U = (u,\, v,\, y,\, x,\, z)$

Task 32
In matrix theory, we know that $(MN)^{-1} = N^{-1}M^{-1}$, where M, N are invertible matrices.

(a) Explain why this product inverse rule also holds when M, N are permutations.
Hint? Look at $[MN][N^{-1}M^{-1}]$ and $[N^{-1}M^{-1}][MN]$.

(b) Give an example to show that $(MN)^{-1} = M^{-1}N^{-1}$ may fail to hold for permutations.

(c) Use the rule $(MN)^{-1} = N^{-1}M^{-1}$ to compute the inverses of each of the following. Then write each result as a product of disjoint cycles.

(i) $P = (x,\, u,\, z,\, y)(x,\, s,\, y)$ **(ii)** $Q = (s,\, z,\, y,\, u)(x,\, t,\, u)(s,\, x)$

(d) Compute the inverse of each of the following by first writing the given permutation as the product of disjoint cycles, and then inverting.

(i) $P = (x,\, u,\, z,\, y)(x,\, s,\, y)$ **(ii)** $Q = (s,\, z,\, y,\, u)(x,\, t,\, u)(s,\, x)$

Why is it no longer strictly necessary (but perhaps advisable) to use the permutation product inverse rule $(MN)^{-1} = N^{-1}M^{-1}$ as part of this method?

(e) Which of the inverse methods used in parts (c) and (d) do you prefer, and why?

(f) Find the inverses of each of the following by first writing the given permutations as products of disjoint cycles.

(i) $R = \begin{pmatrix} xyzuv \\ zxvuy \end{pmatrix}$ **(ii)** $S = \begin{pmatrix} xyzuv \\ zuvyx \end{pmatrix}$

Compare your results to those you obtained in Task 31(a).
Comment on which method of finding inverses you prefer, and why.

Task 33
Let P be a permutation of order i, where $i \in \mathcal{Z}^+$. Show that $P^{-1} = P^{i-1}$.

Task 34
This task continues the exploration of transposition decompositions begun in Task 30.

Recall from that task that 'transposition' is another term for a 2-cycle.
Also recall that every permutation can be decomposed into transpositions in multiple ways.

We will say that a permutation is ***even*** (***odd***) if it can be written as the product of an even (odd) number of transpositions.

For example, since every 3-cycle can be written as the product of two permutations (as shown in Task 30), all 3-cycles are even permutations.

(a) Show that the identity permutation 1 is an even permutation.

Note: Although we will not do so here, it is straightforward to prove that the identity permutation can *only* be written as the product of an even number of permutations. (This proof can be found in most undergraduate textbooks on abstract algebra.)

(b) Use your findings from Task 30 to explain why an n-cycle is even when n is odd.

(c) Let P be an arbitrary permutation. Show that P^{-1} is even if and only if P is even.

(d) For the definition of even/odd permutation to be meaningful, we must prove that no permutation is both even and odd.

Use the fact that the identity permutation 1 can only be written as the product of an even number of permutations (see part (a)) to write a careful proof (by contradiction) that no permutation is both even and odd. Begin by assuming P is a permutation which can be written as the product of an even number of permutations and also as the product of an odd number of permutations. Then consider the product PP^{-1}.

2.2 'Systems of conjugate permutations' in Cauchy's theory

Up to this point in his manuscript, Cauchy was considering just one or two permutations at a time. In this subsection, we turn to excerpts in which he considered *sets* of permutations formed in a particular way. The resulting structure, which Cauchy called a 'system of conjugate permutations,' is equivalent to what is today called a *permutation group* or a *group of permutations*.[41] We say more about today's notation for permutation groups at the end of Subsection 2.2. In our excerpts from Cauchy, however, we use his term 'system of conjugate permutations' and his notation in recognition of the fact that the general group concept had not yet emerged at the time he did his work.

Being given one or several permutations which contain n letters $x, y, z, \ldots$ or at least several of these, I will call *derived* permutations all those [permutations] which can be generated from multiplying the one or more of the given permutations, by each other or by themselves, in any arbitrary order; and the given permutations, joined [together] with the derived permutations, will form what I call *a system of conjugate permutations*. The *order* of the system will be the total number of permutations present, including the [identity] permutation which . . . reduces to unity.

When only a single permutation P is given, the derived permutations are the powers of P and form a system of conjugate permutations which has the same order as that of the permutation P.

[41] In group theory today, there is a concept called a 'conjugate' which is related to some of Cauchy's work on permutations, but not directly related to his use of the term here; we will not consider the modern concept in this project.

Let us pause in our reading to consider some examples. Suppose we are given the two permutations

$$P = (x,\, y) \text{ and } Q = (z,\, u).$$

The *derived permutations* here will include all permutations obtained by multiplying P and Q by themselves and with each other, perhaps repeatedly; in general, this will include products like PQ, QP, P^2, Q^2, P^2Q, PQP, QPQ, QP^2QP, P^3, and so on. In this particular example, however, we know that $P^2 = 1$, $Q^2 = 1$ and $PQ = QP$; thus, all these variations reduce to just four *derived permutations*:

$$1\ ,\ P = (x,\, y)\ ,\ Q = (z,\, u) \ \text{ and } \ PQ = (x,\, y)(z,\, u).$$

The *system of conjugate permutations*, which consists of all the given permutations together with all derived permutations, is therefore this same set:[42]

$$\{1\,,\, P\,,\, Q\,,\, PQ\}.$$

Since this set has four elements, we will say (as did Cauchy) that this system has order 4. Notice too that this set is ***closed under products***; that is, the product of any two of the permutations in this set is also in the set. You should convince yourself that this is the case for any system of conjugate permutations (i.e., permutation group), in light of how such a system is formed by collecting together all possible products formed from the given permutations.

Task 35
Suppose we are given the two permutations $P = (x,\, y,\, z)$ and $Q = (u,\, v)$.

(a) Explain why the *derived permutations* have the form P^kQ^m, where $k = 1, 2, 3$ and $m = 1, 2$, so that the *system of conjugate permutations* generated by P, Q is the set

$$\{\, 1\,,\, P\,,\, P^2\,,\, Q\,,\, PQ\,,\, P^2Q\,\}$$

Conclude that the system of conjugate permutations in this example has order 6.

(b) Show that this system is closed under inverses; that is, for every permutation in the system, show that its inverse is also in the system.

(c) Show this system includes two permutations of order 6, two permutations of order 3, and one permutation of order 2. Compare this to the 6^{th} roots of unity. (See Task 3.)

Task 36
Suppose we are given just the one permutation $P = (x,\, y,\, z,\, u,\, v,\, w)$.

(a) Explain why the system of conjugate permutations in this example also has order 6.

(b) Show that this system is closed under inverses, as this is defined in Task 35(b).

(c) Show this system includes two permutations of order 6, two permutations of order 3, and one permutation of order 2. Compare this to the 6^{th} roots of unity. (See Task 3.)

[42] Since permutations on finite sets of variables will always have finite order, the given permutations will also be considered derived permutations. This is not true for permutations on infinite sets, but this case was not considered by Cauchy. Thus, it is unclear why he distinguished between the set of derived permutations and the system of conjugate permutations since these sets are the same in all his examples.

Task 37
Suppose we are given the two permutations $P = (x,\, y)$ and $Q = (x,\, z)$.

Note that P and Q are *not* disjoint cycles, and therefore do not commute! However, since $P^2 = 1$ and $Q^2 = 1$, the *derived permutations* all come from the following forms:[43]

$$P, PQ, PQP, PQPQ, PQPQP, \ldots \qquad \text{and} \qquad Q, QP, QPQ, QPQP, QPQPQ, \ldots$$

(a) Show that $PQP = QPQ$ by computing these two permutations.
This means that the list of forms we need to compute is reduced to the following:

$$P, PQ, PQP, PQPQ, PQPQP, \ldots \qquad \text{and} \qquad Q, QP$$

(b) Show that $PQPQ = QP$ by computing these two permutations.
Thus, the list of forms we need to compute is further reduced to the following:

$$P, PQ, PQP \qquad \text{and} \qquad Q, QP$$

(c) Conclude that the system of conjugate permutations in this example has order 6.
List each of the six permutations as a product of disjoint cycles.
Also show that this system is closed under inverses, as this is defined in Task 35(b).

(d) Show that — unlike the systems in Tasks 35 and 36 — none of the individual permutations in this system of conjugate permutations is of order 6, even though the system itself has order 6.

Note: This shows that, although individual permutations behave in a fashion analogous to a root of unity of the same order as that permutation, a *system* of conjugate permutations may *not* have the same algebraic structure as the set of roots of unity of the same order as the system of conjugate permutations.

We return now to Cauchy's comments on systems of conjugate permutations.

The system of all permutations that one can form with n letters $x, y, z \ldots$ is evidently a system of conjugate permutations. If one names the different arrangements that can be formed with n variables $x, y, z, \ldots$ by

$$A,\ B,\ C, \ldots$$

then the system in question will be

$$(1) \qquad \binom{A}{A},\ \binom{A}{B},\ \binom{A}{C}, \ldots$$

and the number N of these permutations, or the order of the system, will be determined by the formula

$$N = 1.2.3 \ldots n.$$

[43] If P instead had order 3, while Q had order 2, we would also need to consider permutations of the forms P^2, P^2Q; P^2QP, P^2QP^2; P^2QPQ, P^2QP^2Q; P^2QPQP, P^2QPQP^2, P^2QP^2QP, $P^2QP^2QP^2$; $\ldots$ and of the forms Q, QP^2; QP^2Q; QP^2QP, QP^2QP^2; QP^2QPQ, QP^2QP^2Q ....

Now let

$$(2) \qquad 1, P, Q, R \ldots$$

be an arbitrary system of conjugate permutations. According to the very definition of such a system, we will always reproduce the same permutations, only arranged in a different manner, if each of them is multiplied separately by some particular one of them, or if some one of them is multiplied by itself and by the others. Thus, if we let S be some arbitrary one of the permutations in (2), then the distinct terms of the series

$$(3) \qquad S,\ SP,\ SQ,\ SR, \ldots,$$

as well as [the terms] of the series

$$(4) \qquad S,\ PS,\ QS,\ RS, \ldots,$$

are the same as the terms of the series (2) arranged in a new order.

Again we pause in our reading for an example. Consider the system of conjugate permutations of order four consisting of the following permutations:[44]

$$1 \ , \ P = (x,\, y) \ , \ Q = (z,\, u) \ , \ R = (x,\, y)(z,\, u).$$

We wish to (left) multiply each term of this system by some one particular permutation (which Cauchy labeled S) in the system. For the sake of specificity, let us take $S = Q = (z,\, u)$. In this case, the resulting series is given by:

$$\underbrace{(z,\, u)}_{S} \ , \ \underbrace{(z,\, u)(x,\, y)}_{SP} \ , \ \underbrace{(z,\, u)(z,\, u)}_{SQ} \ , \ \underbrace{(z,\, u)(x,\, y)(z,\, u)}_{SR}$$

But notice that we also have the following:

$$\begin{array}{ll} S = (z,\, u) = Q & SP = (z,\, u)(x,\, y) = (x,\, y)(z,\, u) = R \\ SQ = (z,\, u)(z,\, u) = 1 & SR = (z,\, u)(x,\, y)(z,\, u) = (x,\, y)(z,\, u)^2 = (x,\, y) = P \end{array}$$

Thus, the series S, SP, SQ, SR is simply the original system arranged in a new order: Q, R, 1, P.

Tasks 38 and 39 provide some additional concrete examples to solidify this idea. But, as Cauchy remarked, it is expected from the very definition of a system of conjugate permutations that we will get the entire original series back. Since a system of conjugate permutations is closed under products, and S is itself a permutation in the system, multiplying any element of the system by S must give us an element of the system. You may be wondering though how we know these products are necessarily distinct — could we somehow end up with $SP = SQ$ for some $P \neq Q$? If so, then only some (not all) of the permutations in the system would be listed in the series S, PS, QS, $RS, \ldots$, and Cauchy's claim would be false. Using what you know about inverses, however, you should be able to convince yourself that $SP = SQ$ can only occur in a system of conjugate permutations when $P = Q$.

[44] We looked at this example in the final paragraph on page 39; here, we have set $R = PQ = QP$.

Task 38
Consider the example discussed above, on page 41:

$$1 \ , \quad P = (x,\, y) \quad , \quad Q = (z,\, u) \quad , \quad R = (x,\, y)(z,\, u).$$

(a) Again letting $S = Q = (z,\, u)$, determine the series S, PS, QS, RS.

Recall from the discussion on page 41 that multiplication of the elements $1, P, Q, R$ on the left by S re-ordered these elements in the following way: $Q, R, 1, P$.

Compare this to the order in which the elements $1, P, Q, R$ occur in the series obtained from right multiplication by S.

(b) In general, do you think the original permutations will appear in the same order when left-multiplication is used (e.g., $S, SP, SQ, SR, ST, \ldots$) as when right-multiplication is used (e.g., $S, PS, QS, RS, TS, \ldots$)? Why or why not?

Task 39
Consider the system of conjugate permutations of order 6 from Task 37:

$$1 \ , \ P = (x,\, y) \ , \ Q = (x,\, z) \ , \ R = (x,\, z,\, y) \ , \ S = (y,\, z) \ , \ T = (x,\, y,\, z)$$

(a) Verify that the series S, SP, SQ, SR, S^2, ST gives us all six of the original permutations arranged in a new order.

(b) Verify that the series S, PS, QS, RS, S^2, TS gives us all six of the original permutations arranged in a new order.

(c) Compare the order in which the original permutations appear in the two series found in parts (a) and (b). Would you now change your answer to part (b) of Task 38? Why or why not?

Our final excerpt from Cauchy's comments on systems of conjugate permutations (i.e., permutation groups) gives the proof of an extremely important theorem in group theory. In fact, we will see Cayley use it very early in his paper on abstract groups in the next section. Cauchy, of course, was thinking only of permutation groups in his proof. However, his proof strategy works perfectly well for abstract groups, even though it differs somewhat from the proof strategy found in most of today's textbooks. You will encounter the theorem (stated for groups in general) in any current algebra textbook — look for it under the name 'Lagrange's Theorem' — along with a proof reminiscent of Cauchy's approach (but phrased in terms of 'cosets,' a concept we consider in passing later in this project). Although Lagrange did state a result related to the theorem we are about to see, his work was always done in the (more limited) context of the number of forms resulting from permutations of variables of a function. See [12] for more on the history of this theorem, its proof and how it came to have its current name. We say more about the modern statement of the theorem and its proof below.

You should read (and re-read!) Cauchy's proof of Lagrange's Theorem (for permutation groups) in the following excerpt until you feel that you understand its details. We provide some comments on his proof strategy following the excerpt, and introduce the current terminology which corresponds to the ideas he used in it. In Task 41, you will be asked to re-write this proof using this terminology.

Before reading Cauchy's proof for the first time, recall that he has already introduced the following conventions in the preceding excerpt:

- Series (1) designates the system of *all* possible $N = n!$ permutations on n letters;
- Series (2) designates an arbitrary system of conjugate permutations on n variables, with the individual permutations in the series denoted as $1, P, Q, R \ldots$
- The ***order*** of a system of conjugate permutations is its cardinality as a set.[45]

Theorem 1 The order of a system of conjugate permutations in n variables will always be a divisor of the number of arrangements N that one can form with these variables.

Proof We suppose that the given system is given by the series (2), and we let M be the order of this system. If the series (2) is the same as the series (1), then we have exactly $M = N$; otherwise, we designate by $U, V, W, \ldots$ those permutations which are part of the series (1) but do not appear in the series (2). If we call m the number of terms of the series

$$(5) \qquad\qquad 1, U, V, W, \ldots\ldots$$

then the table

$$(6) \qquad \begin{cases} 1, & P, & Q, & R, & \ldots \\ U, & UP, & UQ, & UR, & \ldots \\ V, & VP, & VQ, & VR, & \ldots \\ W, & WP, & WQ, & WR, & \ldots \\ \text{etc,} \end{cases}$$

will give us m horizontal rows each composed of M terms, with all the terms of each row distinct from each other. If, moreover, two different horizontal rows, for example the second and the third, include equal terms, in which case we would have

$$VQ = UP,$$

we would conclude from this that

$$V = UPQ^{-1}$$

or simply

$$V = US,$$

$S = PQ^{-1}$ being one of the terms of series (2). In this case, the first term V of the third horizontal row in the table (6) would be one of the terms of the second [horizontal rows]. Thus, if the first term of each horizontal row is taken from outside [of all of] the preceding series, all the terms of table (6) will be distinct from each other. Granting that this condition is fulfilled, we continuously add new series to table (6), thereby increasing the number m [of rows]. This operation will stop only when the table (6) includes all N terms contained in the series (1); but then we will evidently have

$$N = mM.$$

Thus, M will be a divisor of N.

[45] The cardinality of a set is simply the number of elements that it contains.

Cauchy's basic strategy in his proof of this theorem was to form an m by M array in which every one of the possible $N = n!$ permutations on n letters appears exactly once, and M is the order of the given system of conjugate permutations. The key to doing this is to make sure that:

(i) no row contains repeated elements ; and

(ii) no element of a given row appears in some earlier row.

Cauchy really did not say how he knew condition (i) held for all rows in the array, although (i) clearly holds for the first row in which each of the M distinct permutations of the given system are listed exactly once. To convince yourself that condition (i) holds for the other rows, consider what would happen, for example, if $UP = UT$, remembering that $P \neq T$. Concerning condition (ii), Cauchy gave considerably more detail, arguing essentially that careful selection of the first element of each row permits us to successively add new rows containing exactly M elements, all of which are distinct from the elements in every preceding row, until all $N = n!$ possible permutations on n variables are exhausted. In Task 40 below, you will complete the construction of a table satisfying both conditions in a specific case.

Again, you should reread this part of Cauchy's proof in detail and discuss it with others as needed until you are ready to complete Task 41 in which you will re-write Cauchy's proof using the current terminology of permutation groups. The following list explains this terminology and comments on how it is related to that of Cauchy.

- The ***symmetric group*** S_n is the set of all permutations on n objects.[46]
  - This is what Cauchy called (somewhat long-windedly):
    'the system of all permutations that one can form with n letters $x, y, z \ldots$'.
  - In Cauchy's proof of Theorem 1, series (1) lists the elements of S_n.
  - The order of S_n is $n!$.
    Textbooks use different notation to denote the order of a group (or subgroup).
    A common way to do this would be to write: $|S_n| = n!$.

- A ***subgroup of*** S_n is a subset $H \subseteq S_n$ which consists of a collection of given permutations together with all permutations derived from that collection. In other words, a subgroup of S_n is a non-empty subset $H \subseteq S_n$ which is closed under products.
  - This is what Cauchy called a 'system of conjugate permutations.'
  - In Cauchy's proof of Theorem 1, series (2) lists the elements of a subgroup H.
  - To denote that H is a subgroup of S_n, we write $H \leq S_n$.
  - S_n is always considered a subgroup of itself, since $S_n \subseteq S_n$.

Using this terminology, we now re-state Cauchy's Theorem 1 as follows:

Lagrange's Theorem for the Symmetric Group S_n
If H is a subgroup of S_n, then the order of H divides the order of S_n.

[46] The term 'symmetric' in this definition appears to relate back to Lagrange's original analysis of algebraic solvability, where permutations helped to measure the degree to which a given function (for the resolvent's roots) was symmetric.

Task 40
This task examines Cauchy's proof of Lagrange's Theorem for S_n in a specific example.

Consider S_4 to be the set of all permutations on the four letters x, y, z, u.

Let $H = \{h_1\,,\, h_2\,,\, h_3\,,\, h_4\}$ where the elements h_1, h_2, h_3, h_4 are as follows:

$$h_1 = 1 \quad ; \quad h_2 = (x,\, y,\, z,\, u) \quad ; \quad h_3 = h_2^2 = (x,\, z)(y,\, u) \quad ; \quad h_4 = h_2^3 = (x,\, u,\, z,\, y)$$

(a) Explain how we know that H is a subgroup of S_4.

(b) What are the values of N and M in Cauchy's proof for this specific example?
Use these values to explain why the completed table should have six rows.

(c) In the partially completed table below, a first element (denoted r_2, r_3, r_4 respectively) has been selected for rows 2 - 4.

(i) Assume for now that the selections made for r_2, r_3, r_4 are valid.
WITHOUT COMPUTING ANY ADDITIONAL PRODUCTS, explain why:

(α) $r_4h_3 \neq r_4h_4$ $\qquad$ (β) $r_3h_3 \neq r_4h_3$ $\qquad$ (γ) $r_4h_3 \neq r_2h_4$

Note: Part (γ) corresponds to the section of Cauchy's proof (beginning with the assumption $VQ = UP$) which shows different rows do not include equal terms.

(ii) Explain why the particular values chosen for r_2 and r_3 are valid choices.

(iii) Complete row 3, and explain why the particular value chosen for r_4 is valid.

(iv) Complete row 4, and explain why there are now eight possible choices for the first entry (r_5) of row 5. Select one of these and explain why your choice is valid.
How many possible choices remain for the first entry (r_6) of row 6?
(You do not need to complete these last two rows, buy may do so if you wish.)

$h_1 = 1$	$h_2 = (x,\, y,\, z,\, u)$	$h_3 = (x,\, z)(y,\, u)$	$h_4 = (x,\, u,\, z,\, y)$
$r_2 = (x,\, y)$	$r_2h_2 = (x,\, y)(x,\, y,\, z,\, u)$ $= (y,\, z,\, u)$	$r_2h_3 = (x,\, y)(x,\, z)(y,\, u)$ $= (x,\, z,\, y,\, u)$	$r_2h_4 = (x,\, y)(x,\, u,\, z,\, y)$ $= (x,\, u,\, z)$
$r_3 = (x,\, z)$	$r_3h_2 = (x,\, z)(x,\, y,\, z,\, u)$ $=$	$r_3h_3 = (x,\, z)(x,\, z)(y,\, u)$ $=$	$r_3h_4 = (x,\, z)(x,\, u,\, z,\, y)$ $=$
$r_4 = (x,\, u)$			

(d) Suppose the first element of row 2 in the above table were chosen to be $r_2' = (y,\, z,\, u)$, instead of $r_2 = (x,\, y)$. How would this change the resulting table. Would we have been able to use the same values of r_3, r_4, r_5, r_6 in this case? Why or why not?

Task 41
Write a fully general rigorous proof of 'Lagrange's Theorem for S_n' using Cauchy's strategy of building an $m \times M$ array in which all $N = n!$ elements of S_n appear exactly once and $M = |H|$ for a given subgroup H. Use the current terminology introduced on page 44. In order to do this in full generality, you should introduce indexed variables to denote the elements of H as well as the first element of each row of the array. (See Task 40.) Then formally (and carefully) define the array recursively and explicitly prove that the completed array satisfies both conditions (i) and (ii) discussed on page 44. Also add detail and/or rephrase Cauchy's reasoning where you feel this is needed and/or helpful.

Task 42
As a corollary to his Theorem 1 (aka, Lagrange's Theorem for the Symmetric Group S_n), Cauchy gave the following result, stated here in the terminology currently in use:

> **Corollary I**
> If H is a subgroup of S_n and P is an arbitrary permutation in H,
> then the order of P divides the order of H.

Cauchy proved Corollary I using an array strategy similar to the one he used to prove Lagrange's Theorem for S_n. Rather than go through the details of this strategy again, this task examines the following somewhat less general corollary:

> **Corollary II**
> For every permutation $P \in S_n$, the order of P divides the order of S_n.

To prove Corollary II directly from Lagrange's Theorem, we only need to find a subgroup H which we know has the same order as the permutation P. The natural candidate for this subgroup H is the set which contains all powers of P; that is, let

$$H = \{ P^k \,|\, k \in \mathcal{Z}^+ \}.$$

Today, this set H is called the ***cyclic subgroup generated by*** P, denoted $H = \langle P \rangle$.

(a) Explain why H is a subgroup of S_n.

(b) Explain why the order of subgroup H is equal to the order of the permutation P.

(c) Explain why Corollary II now follows from Lagrange's Theorem.

Cauchy proved much (much!) more about the theory of permutations in his manuscripts of 1844–1845, which regrettably lies beyond the scope of this project. Before we turn to the next part of this project — Cayley's paper on abstract groups — we mention one final thing about how the symmetric group of permutations S_n is typically thought of today. Namely, rather than use a set of n letters as the underlying objects being permuted, it is standard to use permutations of the first n natural numbers: $1, 2, 3, \ldots, n$. This idea was actually first used by Cauchy himself, who employed it in his earliest (1815) paper [2] by first introducing indexed letters (e.g., $x_1, x_2, \ldots, x_n$) to represent n variables, and then representing these variables by their subscripts alone.

Look back, for example, at the product in Task 21:

$$\begin{pmatrix} x_1\,x_2\,x_3\,x_4\,x_5 \\ x_3\,x_1\,x_2\,x_5\,x_4 \end{pmatrix}\begin{pmatrix} x_1\,x_2\,x_3\,x_4\,x_5 \\ x_5\,x_4\,x_2\,x_3\,x_1 \end{pmatrix}.$$

Using subscripts to represent each variable, Cauchy represented this product more simply as

$$\begin{pmatrix} 1\,2\,3\,4\,5 \\ 3\,1\,2\,5\,4 \end{pmatrix}\begin{pmatrix} 1\,2\,3\,4\,5 \\ 5\,4\,2\,3\,1 \end{pmatrix},$$

which can be even more simply written using cycle decomposition:

$$[(1,\,3,\,2)(4,\,5)]\,[(1,\,5)(2,\,4,\,3)]\,.$$

You can check that the final product in this example is the 5-cycle (1, 4, 2, 5, 3). The main point we wish to make here, however, is that mathematicians today think of S_n as a set of permutations on numbers (e.g., $1, 2, 3, 4, 5$), rather then a more cumbersome version of these same permutations involving letters, either indexed or not. In fact, it is something of a mystery why Cauchy did not do this in his later works on permutation theory. This may have been because he was no longer thinking of the letters as variables in functions, but rather treating permutations as interesting mathematical objects in their own right (which they are!).

As illustrations of the notation currently in use, we list the elements of the symmetric groups S_2 and S_3 in this form:

$$S_2 = \{\,(1)(2)\,,\,(1,\,2)\,\}$$

$$S_3 = \{\,(1)(2)(3)\,,\,(1,\,2,\,3)\,,\,(1,\,3,\,2)\,,\,(1,\,2)\,,\,(1,\,3)\,,\,(2,\,3)\,\}\,.$$

Note that the identity permutation is written here as the product of 1-cycles, in order to avoid any possible confusion about whether '1' is the identity permutation or one of the numbers being permuted. In part because of this potential confusion, it has also become standard practice to label the identity permutation as 'ε' (rather than '1'); thus, we might write '$(1,\,3,\,2)^3 = \varepsilon$', or '$XY = X \Rightarrow Y = \varepsilon$'. As you gain practice with permutation computations using the currently standard notation in Task 43 below, watch for ways in which the identity permutation ε can be used to simplify computations — but also be careful to avoid the trap of using commutativity where it doesn't apply!

Task 43

(a) Find the cycle decomposition of the following.

(i) $\begin{pmatrix} 1\,2\,3\,4\,5\,6 \\ 1\,3\,2\,5\,6\,4 \end{pmatrix}$ **(ii)** $\begin{pmatrix} 1\,2\,3\,4\,5\,6\,7 \\ 7\,3\,5\,2\,4\,6\,1 \end{pmatrix}$ **(iii)** $\begin{pmatrix} 1\,2\,3\,4\,5\,6 \\ 6\,2\,5\,1\,4\,3 \end{pmatrix}^{-1}$

(b) Find each of the following products; write each as the product of disjoint cycles.

(i) $(1,\,3,\,2,\,5)(1,\,5,\,4)(2,\,3,\,6,\,4)$ **(iii)** $[(1,\,2,\,5,\,4)(3,\,1,\,4)]^3$

(ii) $(1,\,5,\,4,\,2)^{-1}(4,\,2,\,3)(1,\,5,\,4,\,2)$ **(iv)** $(1,\,2,\,5,\,4)^3(3,\,1,\,4)^3$

(c) Determine the order of each of the following permutations. Justify your answer.

(i) $(7,\,1,\,4,\,5,\,3)$ **(iii)** $(7,\,1,\,4)(5,\,3)$ **(v)** $(3,\,1,\,4)(1,\,5,\,6)$

(ii) $(7,\,1,\,4,\,5,\,3)^{-1}$ **(iv)** $(7,\,1,\,4)(7,\,3)$ **(vi)** $(3,\,1,\,4)(7,\,5,\,6)$

3 The first paper on abstract group theory: A. Cayley

In Section 1, we saw how the concept of a permutation first arose in connection with the algebraic solution of polynomial equations in Lagrange's work. Within that context, we also examined certain algebraic properties of roots of unity. We then saw, in Section 2, how Cauchy's work on permutations moved in a new direction by treating permutations as algebraic objects in their own right. One of the interesting properties that surfaced there is the fact that every individual permutation α of order n behaves in a fashion analogous to a primitive n^{th} roots of unity, even though a system of permutations (or permutation group) might have an algebraic structure quite different from that of the roots of unity of the same order as that system.[47] Having carefully studied Cauchy's work, the British mathematician Arthur Cayley was familiar with these ideas.[48] Cayley was also well versed in the symbolic algebra tradition which dominated British mathematics at the time. By drawing on his understanding of both areas, Cayley became the first to formulate a description of the algebraic structure now known as an abstract group. In this section, we study the celebrated 1854 paper in which he announced this concept: *On the theory of groups, as depending on the symbolic equation* $\theta^n = 1$.

Born in Surrey County, England, on August 16, 1821, Arthur Cayley and his parents lived in St. Petersburg, Russia during the first eight years of his childhood before returning to England to live near London. Cayley began publishing research papers in mathematics while still an undergraduate at Trinity College, Cambridge. Following his graduation in 1842, he taught as a Fellow at Cambridge for four years before training as a lawyer to secure a means of support. Admitted to the bar in 1849, Cayley worked as a lawyer for 14 years while continuing his mathematical research. The approximately 250 mathematical papers he published during this time include both the first paper ever written on matrix theory and the first paper ever written on group theory. Despite a significant decrease in income, Cayley left the legal profession in 1863 to accept an appointment as Sadleirian professor of Pure Mathematics at Cambridge, thereafter devoting his professional life to mathematical research until his death on January 26, 1895.

Cayley's mathematical interests were strongly influenced by the general state of British mathematics at the time. Over the course of the eighteenth century, British mathematics had become cut off from that of the European continent, in part due to a priority controversy concerning the invention of calculus by Gottfried Leibniz (1646–1716) and Isaac Newton (1643–1727). On the continent, the differential techniques and notation of Leibniz allowed mathematicians to make significant advances in the development of calculus. In comparison, British mathematics appeared almost stagnant. By the early nineteenth century, however, changes began to emerge in English society as a result of the Industrial Revolution, and the role of universities in promoting research and science became a subject of concern. One outcome of these changes for mathematics was the formation of 'The Analytic Society' by a group of mathematicians at Cambridge in 1812. As Society members began to import continental methods of symbolic manipulation into the university curriculum, concerns about the foundations of algebra arose, including questions about the meaning of negative and imaginary numbers. The concept

[47] For example, multiplication in a permutation group need not be commutative, whereas roots of unity are complex numbers and, therefore, do have commutative products. Another difference between these two structures is the fact that the set of n^{th} roots of unity always includes one or more primitive roots of order n (as we saw in Tasks 6–9), while a permutation group of order n might not contain any permutations of order n (as we saw in Task 37).

[48] Cayley was familiar with the work of continental mathematicians in general, including Galois' articles on algebraic solvability which were fully published in 1846 (thirteen years after Galois' own death). Cayley's familiarity with the study of algebraic solvability probably dates back to his undergraduate days. In the year he graduated from Cambridge (1842), the Tripos Exam (a notoriously challenging written examination completed by all undergraduate students at Cambridge) included the following question: "All equations whose coefficients are rational are resolvable into simple or quadratic factors. In what sense is it said that no equation above the fourth degree admits of resolution?" Although Galois' work had not yet been published in 1842, the work of Abel (and Ruffini) was known by that time.

of a 'symbolical algebra' was developed in response to these questions.

In contrast to 'arithmetical algebra,' which derives its laws from the actual meaning of operations on numbers, a 'symbolical algebra' begins with formal laws for a given set of symbols and operations on those symbols, and only later interprets these as having particular meaning. A good illustration of symbolic algebra is the work done by George Boole (1815–1864) on the logic of classes (or sets) which eventually developed into the algebraic structure known as a *boolean algebra*. Among the properties of symbolic logic first stated by Boole, one finds such non-standard 'algebra' laws as 'idempotency' [$x^2 = x$] and 'distributivity of addition over multiplication' [$x + yz = (x + y)(x + z)$]. Cayley was certainly familiar with Boole's work on the algebra of logic; in fact, the two were regular correspondents and shared a number of mathematical interests.

Another symbolic algebra system with which Cayley was familiar was that of the *quaternions* — a set of 'numbers' of the form $a + bi + cj + dk$ subject to certain special 'arithmetic' rules discovered in 1843 by William Rowan Hamilton (1805–1865). Hamilton was led to his discovery while seeking an algebraic system that would faithfully represent the three-dimensional geometrical space of physics, in a manner analogous to the interpretation of the algebra of complex numbers $a+bi$ as geometrical objects in a two-dimensional plane.[49] After years of unsuccessful attempts to define appropriate operations on triplets consisting of one real and two imaginary components ($a + bi + cj$), Hamilton came to the following revelation (as quoted in [13, p. 230]):

> And here there dawned on me the notion that we must admit, in some sense, a fourth dimension of space for the purpose of calculating with triplets; ... or transferring the paradox to algebra, [we] must admit a third distinct imaginary symbol k, not to be confounded with either i or j, but equal to the product of the first as multiplier, and the second as multiplicand; and therefore [I] was led to introduce quaternions such as $a + ib + jc + kd$, or (a, b, c, d).

Even after adjoining these three purely imaginary components to one real component, however, Hamilton was only able to faithfully represent the mechanics of three-dimensional physics by defining multiplication on the resulting quaternions in a way that abandoned the property of commutativity.[50]

It was within this setting of new and strange algebraic structures that Cayley's inaugural paper on abstract groups appeared.

3.1 Cayley's Definition of an Abstract Group

We begin our study of Cayley's work by quoting the opening paragraph of his ground-breaking paper [5] in its entirety. In this paragraph, Cayley discussed some basic ideas about operations and their algebraic properties. Many of these properties will sound familiar from Cauchy's theory of permutations, although we will see that Cayley is thinking of them far more generally. As you read through the excerpt, keep an eye out for ideas which we have already encountered in this project, as well as those which seem to be completely news. Following this rather lengthy excerpt (which you should read at least once in its entirety), we will examine specific portions of it to see how Cayley's more general notion of group fits with the theory of permutations.

[49] Although the geometrical interpretation of complex numbers as points (or vectors) in a two-dimensional plane is now considered standard, remember that it was developed only in the early 1800's. Thus, it was a relatively new idea when Hamilton began his search for an algebraic representation of the three-dimensional space of physics.

[50] Soon after Hamilton's discovery, physicists realized that only the 'vector part' $bi + cj + dk$ of a quaternion was needed to represent three-dimensional space, so that the real component a could be dropped from consideration. Although vectors replaced quaternions in physics by the end of the nineteenth century, the vector operation of cross product retains the non-commutativity of quaternions. Compare, for example, the quaternion multiplication rules $ij = k$ and $ji = -k$, with the vector cross products $\vec{i} \times \vec{j} = \vec{k}$ and $\vec{j} \times \vec{i} = -\vec{k}$ which hold in a right-handed coordinate system.

On the theory of groups, as depending on the symbolic equation $\theta^n = 1$

Let θ be a symbol of operation, which may, if we please, have for its operand, not a single quantity x, but a system $(x, y, \dots)$, so that

$$\theta(x, y, \dots) = (x', y', \dots),$$

where x', $y', \dots$ are any functions whatever of $x, y, \dots$, it is not even necessary that x', $y', \dots$ should be the same in number with $x, y, \dots$. In particular, x', y', &c. may represent a permutation of x, y, &c., θ is in this case what is termed a substitution; and if, instead of a set $x, y, \dots$, the operand is a single quantity x, so that $\theta x = x' = fx$, θ is an ordinary functional symbol. It is not necessary (even if this could be done) to attach any meaning to a symbol such as $\theta \pm \phi$, or to the symbol 0, nor consequently to an equation such as $\theta = 0$ or $\theta \pm \phi = 0$; but the symbol 1 will naturally denote an operation which ... leaves the operand unaltered, and the equation $\theta = \phi$ will denote that the operation θ is ... equivalent to ϕ, and of course $\theta = 1$ will in like manner denote the equivalence of the operation θ to the operation 1. A symbol $\theta\phi$ denotes the compound operation, the performance of which is equivalent to the performance, first of the operation ϕ, and then of the operation θ; $\theta\phi$ is of course in general different from $\phi\theta$. But the symbols θ, ϕ, ... are in general such that $\theta.\phi\chi = \theta\phi.\chi$, &c., so that $\theta\phi\chi$, $\theta\phi\chi\omega$, &c. have a definite signification independent of the particular mode of compounding the symbols; this will be the case even if the functional operations involved in the symbols θ, ϕ, &c. contain parameters such as the quaternion imaginaries i, j, k; but not if these functional operations contain parameters such as the imaginaries which enter into the theory of octaves, and for which, e.g. $\alpha.\beta\gamma$ is something different from $\alpha\beta.\gamma$, a supposition which is altogether excluded from the present paper. The order of the factors of a product $\theta\phi\chi\dots$ must of course be attended to, since even in the case of a product of two factors the order is material; it is very convenient to speak of the symbols θ, $\phi\dots$ as the first or furthest, second, third, &c., and last or nearest factor. What precedes may be almost entirely summed up in the remark, that the distributive law has no application to the symbols $\theta\phi\dots$; and that these symbols are not in general convertible, but are associative. It is easy to see that $\theta^0 = 1$, and that the index law $\theta^m.\theta^n = \theta^{m+n}$, holds for all positive or negative integer values, not excluding zero. It should be noticed also that, if $\theta = \phi$, then, whatever the symbols α, β may be, $\alpha\theta\beta = \alpha\phi\beta$, and conversely.

Let us begin again at the beginning, and look back at what Cayley had to say about the behavior of the sort of operations to be considered in his paper. Although it may not be completely clear what he meant by the term 'operation,' it is clear that permutations fall under this category. Cayley will later make it very clear that permutations are not the only operations he had in mind — in fact, he mentioned that 'quaternion imaginaries' can be involved midway through his opening paragraph. Generally speaking, however, this paragraph tells us that the operations of interest with regard to groups will always have certain characteristics that are exemplified by permutations. For example, in thinking of a permutation as the process of substituting one arrangement of objects for another, the idea of adding or subtracting two permutations really does not arise, nor does the concept of an additive identity ('0'). On the other hand, the idea of an identity operation ('1') which leaves every arrangement unaltered arises quite naturally in permutation theory, as does the notion of following one substitution ϕ with another θ to obtain the composition function (or 'compound operation') denoted by the symbol $\theta\phi$. Again quite naturally, commutativity fails to hold when multiplication is interpreted as function composition in this way.

In the sentence following Cayley's declaration that '$\theta\phi$ is of course in general different from $\phi\theta$,' we find his discussion of another algebraic property: associativity. Look back at what Cayley had to say in this regard, and especially how the 'quaternion imaginaries' differ from the 'imaginaries which enter into the theory of octaves' with regard to this property. By stipulating that the former (quaternion imaginaries) were allowed while the latter (octave imaginaries) were 'altogether excluded from the present paper,' Cayley explicitly excluded any non-associative system of operations for which '$\alpha.\beta\gamma$ is something different from $\alpha\beta.\gamma$.' Cauchy, on the other hand, did not even mention associativity in his work on permutations, perhaps because it had not yet occurred to anyone that multiplication could fail to be associative. Stop now to convince yourself that $(f \circ g) \circ h = f \circ (g \circ h)$ for any functions f, g, h. In Task 44, we consider the system of quaternion imaginaries, both to provide a concrete example of a group which does not involve permutations and to further examine the value of associativity. Although octave imaginaries will not form a group due to their non-associative nature, we also consider them briefly in Task 45 as another illustration of why associativity is important.

Task 44

Recall from the introduction to Section 3 (page 49) that a quaternion is a linear expression of the form $a + bi + cj + dk$, where i, j, k are considered to be three (independent) imaginary units and a, b, c, d are real numbers.

By defining addition and scalar multiplication componentwise, the set of quaternions can be viewed as a four-dimensional vector space over the real numbers with basis $\{1, i, j, k\}$.

Although the formal concepts of 'vector space' and 'basis' were defined only much later, Hamilton clearly recognized the essential role played in his system by these four basis elements. In particular, he noted that multiplication of quaternions required only nine basic products (i^2 , ij , ik , ji , j^2 , jk , ki , kj , k^2) to be determined.[51]

This task examines properties of the group formed by considering these four basis elements together with their opposites under the operation of multiplication defined by Hamilton.

To this end, we define the *quaternion group* to be the set

$$G = \{\, 1 \,,\, -1 \,,\, i \,,\, -i \,,\, j \,,\, -j \,,\, k \,,\, -k \,\}$$

under a product with multiplicative identity 1 for which the following rules are given:

$$\begin{array}{ll} i^2 = j^2 = k^2 = -1 & \qquad (-1)i = i(-1) = -i \\ (-1)^2 = 1^2 = 1 & \qquad (-1)j = j(-1) = -j \\ i(jk) = (ij)k = -1 & \qquad (-1)k = k(-1) = -k \end{array}$$

We also assume for the purpose of this task that multiplication is associative.[52] In the following exercises, we use these facts to determine the values of the products ij, ji, ik, ki, jk and kj, and to further explore the properties of associativity and commutativity.

[51] To see why this is true, assume that multiplication is distributive over addition and write out the terms of the product $(a + bi + cj + dk)(a' + b'i + c'j + d'k.)$

[52] For Hamilton, the associativity of quaternion multiplication followed from the geometry of 3-space which he was attempting to model.

Task 44 - continued

(a) By multiplying both sides of the equation '$ijk = -1$' on the right by $-k$, and then simplifying according to the given rules, we get the following:

$$\begin{aligned}(ijk)(-k) = (-1)(-k) \quad &\Rightarrow \quad (ij)k[(k)(-1)] = (-1)[(-1)k] \\ &\Rightarrow \quad (ij)k^2(-1) = (-1)^2k \\ &\Rightarrow \quad (ij)[(-1)(-1)] = (1)k \\ &\Rightarrow \quad ij(1) = k \\ &\Rightarrow \quad ij = k\end{aligned}$$

(i) Notice that by writing 'ijk' without parentheses, we have already made use of associativity in writing the initial equation $ijk = -1$.
Identify all other instances where associativity was used in the derivation above.

(ii) Use a similar strategy to show that $jk = i$ and $ki = j$.
Be careful not to commute elements unless the given rules explicitly allow this!
(For example, $(-1)k = k(-1) = -k$ allows us to commute k and -1.)
Hint? For $ki = j$, multiply '$ijk = -1$' by i on the right and the left of both sides.

(b) From part (a), we know that $j = ki$, $i = jk$ and $k = ij$.
Use these facts to show that $ji = -k$, $kj = -i$ and $ki = -j$.
Conclude that quaternion multiplication is not commutative.

(c) What are the values of $(-i)(-j)$ and $(-i)j$? Use associativity to justify your answers.

(d) Construct an 8 by 8 multiplication table for G using the multiplication facts given or derived thus far, along with any additional derivations needed to fill in all 64 products. Follow the convention that the row denotes the element on the left in a product and the column denotes the element on the right in a product. Comment on patterns.

Task 45

The discovery of an algebraic system involving 8 elements $(1, i, j, k, p, q, r, s)$ and their opposites was discovered independently by two men soon after Hamilton's announcement of the quaternions. One of these men was Cayley; the other was the Irish jurist and mathematician John T. Graves (1806–1870), a close friend of Hamilton.[53] Concerning the properties of octaves, Hamilton declared (as quoted by Crilly, [9, p. 102]):

> To this associative principle, or property of multiplication, I attach great importance. . . . . . The *absence* of associativity appears to me to be a great inconvenience in the octaves or octonomials of Messrs. J. T. Graves and Arthur Cayley.

To see how this inconvenience arises, consider the (partial) multiplication table[54] for the system of octaves given below.

[53] Graves appears to have had precedence in terms of discovery, but Cayley was the first to publish. For this reason, the elements of this system of octaves are sometimes referred to as 'Cayley numbers.'

[54] The complete 16×16 table can be obtained by assuming the system of octaves satisfies two familiar rules of opposites: $(-x)y = x(-y) = -(xy)$ and $(-x)(-y) = xy$ for all x, y.

Task 45 - continued

	1	i	j	k	p	q	r	s
1	1	i	j	k	p	q	r	s
i	i	-1	p	s	$-j$	r	$-q$	$-k$
j	j	$-p$	-1	q	i	$-k$	s	$-r$
k	k	$-s$	$-q$	-1	r	j	$-p$	i
p	p	j	$-i$	$-r$	-1	s	k	$-q$
q	q	$-r$	k	$-j$	$-s$	-1	i	p
r	r	q	$-s$	p	$-k$	$-i$	-1	j
s	s	k	r	$-i$	q	$-p$	$-j$	-1

(a) Following the convention that the row denotes the element on the left in a product and the column denotes the element on the right in a product, the multiplication table for octaves given above indicates that $jk = q$. Use this table to explain why it is impossible in this system to assign a value to the product jkp. Indicate how this relates to the lack of associativity for octaves, and comment on why this is inconvenient.

(b) Now find a specific example to show that octave multiplication is also non-commutative.

We return now to Cayley's opening paragraph, beginning with:

> What precedes may be almost entirely summed up in the remark, that the distributive law has no application to the symbols $\theta\phi\ldots$; and that these symbols are not in general convertible, but are associative.

Since distributivity involves two methods of combination (e.g., addition and multiplication), Cayley's dismissal of the distributivity law $[\theta(\phi + \omega) = \theta\phi + \theta\omega]$ here is in keeping with his interest in studying an algebraic structure with just a single way to combine elements (e.g., multiplication). As you no doubt surmised from the context, 'convertible' was Cayley's term for 'commutative', while 'associative' had the same sense then as now. The properties of exponents ($\theta^0 = 1$ and $\theta^m\theta^n = \theta^{m+n}$) which Cayley next asserted to be 'easy to prove' are also already familiar to you, both from what you knew about real numbers before this project, and from you have learned about permutations in this project. Since Cayley was considering operations in general, and not just permutations, today's mathematician may wonder what sort of proof he had in mind with respect to these properties. Within the context of British symbolic algebra, however, properties such as these had been extensively studied as part of the 'calculus of operations' prior to Cayley's work on groups. For this and other reasons, we will not consider their proofs of in this project.

There is, however, one important consequence of these exponential properties which must be noted. Namely, by declaring that the index law ($\theta^m\theta^n = \theta^{m+n}$) holds 'for all positive or *negative* integer values, *not excluding zero*,' Cayley essentially decreed his interest in only those systems with the following property, referred to today as the ***Inverse Property***:

$$\boxed{\text{For every operation } \theta\text{, there exists an operation } \theta^{-1} \text{ such that } \theta\theta^{-1} = \theta^{-1}\theta = 1.}$$

Again, it is clear from our reading of Cauchy that all permutations necessarily have inverses (as do all non-zero real numbers). For quaternions, it is also easy to check that $-i$ is the inverse of i, and similarly for $-j$, j and $-k$, k. (Do this!) The next task explores the general role played by the Inverse Property in solving group equations.

Task 46
Suppose we are given an arbitrary set of operation symbols and an associative product on this set which possesses an identity element 1 and satisfies the Inverse Property.

Assume $\alpha, \beta, \gamma, \theta$ are operation symbols in this set.

Solve the equations and systems of equations given in parts (a) - (d) below for θ.
Be careful not to assume commutativity, and pay attention to where associativity is used.

For example, we can solve the system $\left\{ \begin{array}{c} \alpha\theta^2 = \beta \\ \theta^3 = 1 \end{array} \right.$ for θ by first right-multiplying both sides of the top equation by θ, and then simplifying as follows:

$$\alpha\theta^2\theta = \beta\theta \;\Rightarrow\; \alpha\theta^3 = \beta\theta \;\Rightarrow\; \alpha{\cdot}1 = \beta\theta \;\Rightarrow\; \alpha = \beta\theta \;\Rightarrow\; \beta^{-1}\alpha = \underbrace{\beta^{-1}\beta}_{\beta^{-1}\beta\ =1}\ \theta \;\Rightarrow\; \boxed{\beta^{-1}\alpha = \theta}$$

(a) $\alpha\theta\beta = 1$ **(b)** $\alpha\theta\beta = \gamma$ **(c)** $\left\{ \begin{array}{c} \theta^2 = \beta^2 \\ \theta^5 = 1 \end{array} \right.$ **(d)** $\left\{ \begin{array}{c} \theta^2\alpha = \beta^{-1}\theta\gamma^{-1} \\ \alpha\gamma\theta = \theta\alpha\gamma \end{array} \right.$

We come now to the last sentence of Cayley's opening paragraph, in which he asserted that a certain statement and its converse were both true:

> It should be noticed also that, if $\theta = \phi$, then, whatever the symbols α, β may be, $\alpha\theta\beta = \alpha\phi\beta$, and conversely.

Beginning with the converse, we obtain a property known today as the ***Cancellation Property***:[55]

$$\boxed{\text{For all operations } \alpha, \beta, \theta \text{ and } \phi, \text{ if } \alpha\theta\beta = \alpha\phi\beta, \text{ then } \theta = \phi.}$$

Do you see how easily we can prove that the Cancellation Property holds in any system which satisfies the Inverse Property? This would have been equally clear to Cayley and his colleagues, given their experience with operations satisfying the Inverse Property. As for how Cayley and his contemporaries would have thought of the original statement — that $\theta = \phi$ implies $\alpha\theta\beta = \alpha\phi\beta$, regardless of whether $\alpha, \beta, \theta, \phi$ represent permutations, or quaternion imaginaries, or some other type of operation altogether — it is again likely that the conception of 'operation' which was standard at the time also made this a straightforward claim to accept. Indeed, you should have little trouble convincing yourself of its truth in the case where $\alpha, \beta, \theta, \phi$ are permutations. In the case where these symbols represent quaternion imaginaries, you probably also have no difficulty believing that $\theta = \phi$ implies $\alpha\theta\beta = \alpha\phi\beta$. In fact, you may be wondering how a product could fail to behave in this way; we illustrate this possibility in Task 47. Today, we say that multiplication which behaves in this way ($\theta = \phi \;\Rightarrow\; \alpha\theta\beta = \alpha\phi\beta$ for all $\alpha, \beta, \theta, \phi$) is **well-defined**. Like Cayley, we also work only with systems in which multiplication is well-defined whenever we are talking of groups.

[55] In fact, it has become standard to distinguish between left cancellation ($\alpha\theta = \alpha\phi \Rightarrow \theta = \phi$) and right cancellation ($\theta\beta = \phi\beta \Rightarrow \theta = \phi$) when stating this property, especially in the context of non-commutative products.

Task 47

In this task, we examine how a 'product' can fail to be well-defined.

(a) We first define a product $*$ on the rational numbers in such a way that it will be possible to have $b = c$ with $a * b \neq a * c$.

Assume a, b are rational numbers and define the product $a * b$ as follows:

$$a * b = \frac{m_a}{n_b}, \text{ where } a = \tfrac{m_a}{n_a}\ , b = \tfrac{m_b}{n_b}, \text{ and } m_a, n_a, m_b, n_b \text{ are integers.}$$

For example, if $a = \frac{3}{5}$ and $b = \frac{9}{10}$, then $a * b = \frac{3}{10}$.

Let $a = \frac{2}{3}$, $b = \frac{3}{4}$, and $c = \frac{6}{8}$. Note that $b = c$.
Show that $*$ is not well-defined by computing $a * b$ and $a * c$.

(b) We now modify the definition of $*$ from part (a) to obtain a well-defined operation. Again assume a, b are rational numbers and define the product $a * b$ as follows:

$$a * b = \tfrac{m_a}{n_b}, \text{ where } a = \tfrac{m_a}{n_a}\ , b = \tfrac{m_b}{n_b}, \text{ and } m_a, n_a, m_b, n_b \text{ are } \textbf{positive} \text{ integers}$$
$$\textbf{with } gcd(m_a, n_a) = gcd(m_b, n_b) = 1.$$

Explain how modifying the definition in this way implies that $*$ is well-defined.
That is, why does $b = c$ now imply that $a * b = a * c$ for any rational numbers a, b, c?

This (finally!) completes our commentary on Cayley's first paragraph, and we return to our reading of his paper. Up to now, we have not heard his full definition of a group. But having set the stage in his first paragraph with a description of the general properties of operation symbols, Cayley was ready for his lead player to make an entrance. As you read his definition of a group below, keep in mind that Cayley has already told us (without explicitly saying so) that multiplication in a group will be well-defined, associative, possess an identity element 1, and satisfy the Inverse Property. This is very nearly the modern definition of a group. The only property missing from this list is one that we have already encountered in our reading of Cauchy: ***closure under products***. After adding closure to the properties already described, Cayley discussed one of its consequences by introducing, for the first time in mathematics, a table similar to the one used in Task 45 to define octave multiplication. We will explore these tables, known today as ***Cayley tables*** or ***group tables***, further in the next subsection. In the rest of this subsection, we focus on the definition of a group as Cayley presented it in this paper.

A set of symbols,

$$1,\ \alpha,\ \beta, \ldots$$

all of them different, and such that the product of any two of them (no matter what order), or the product of any one of them into itself, belongs to the set, is said to be a *group*.[56] It follows that if the entire group is multiplied by any one of the symbols, either as further or nearer factor, the effect is simply to reproduce the group; or what is the same thing, that if the symbols of the group are multiplied together so as to form a table[57] thus:

[56] [Cayley's footnote]: The idea of a group as applied to permutations or substitutions is due to Galois, and the introduction of it may be considered as marking an epoch in the progress of the theory of algebraical equations.

[57] Cayley himself placed the further (right) factors along the top of his table and the nearer (left) factors along its side. We have transposed rows and columns throughout his paper so that tables in this project conform with the now standard practice of placing the product of the i^{th} row factor by the j^{th} column factor in the $(i, j)^{th}$ table position.

	1	α	β	..
1	1	α	β	..
α	α	α^2	$\alpha\beta$	..
β	β	$\beta\alpha$	β^2	..
	:			

that as well each line as each column of the square will contain all the symbols 1, α, $\beta, \ldots$. It also follows that the product of any number of the symbols, with or without repetition, and in any order whatever, is a symbol of the group.

By fitting together the various pieces laid out by Cayley up to this point, we can now fully state the properties which define a group. First and foremost, note that Cayley has moved to treating $1, \alpha, \beta, \ldots$ as completely abstract symbols — not permutations, nor quaternion imaginaries, nor any other concrete instance of an operation. (We use Latin letters a, b, c for this purpose in our own definition.) Consequently, any theorem that can be deduced from the properties stated in this definition will necessarily apply to every particular system which satisfies those properties.

Definition A ***group*** is a set G together with a well-defined binary operation $*$ for which the following properties are satisfied:

Closure under Products: For all $a, b \in G$, $a * b \in G$.
Associativity: For all $a, b, c \in G$, $(a * b) * c = a * (b * c)$.
Identity: There exists $1 \in G$ such that for all $a \in G$, $a * 1 = 1 * a = a$.
Inverses: For all $a \in G$, there exists $a^{-1} \in G$ such that $a * a^{-1} = a^{-1} * a = 1$.

In stating this definition, we have used the term 'binary operation' in almost its modern sense to refer to a method of combining two elements of G (e.g., multiplication) to produce a single (not necessarily new) object — *almost* because the full modern sense requires a binary operation on a set G to produce an object which is also an element of G. In other words, a ***binary operation on*** G is a function $* : G^2 \to G$. This means that closure is really part of what it now means to say $*$ is a (binary) operation on G. Because of this, most textbooks list only our last three properties in their definition of a group. We have listed closure as a separate property partly to conform with Cayley, but also because this is a key property in the study of subgroups (considered further below).

We have also conformed with Cayley's use of '1' to denote the identity, whereas most textbooks today denote it by 'e' or by 'ϵ' as a reminder that $*$ need not actually be multiplication of numbers (even though it is called multiplication) and that the identity need not actually be the number one. By now, we are used to this idea since our main example has been that of permutations, where 'multiplication' is really function composition and '1' is really the identity function which maps each element to itself. Based on our reading of Cauchy's work, it should be clear that the symmetric group S_n on n letters satisfies all four axioms of Cayley's definition of an abstract group. Task 49 below gives other examples of groups, including several for which the group operation is not actually multiplication.

Notice that by referring to 1 as *the*identity, we are implicitly assuming that there is only one element in a group G which 'acts' like an identity. This is in fact true, although Cayley seems to have simply taken it for granted. To prove the uniqueness of the identity, suppose 1_1 and 1_2 both act like group identities, so that for any $x \in G$, we have $1_1 * x = x * 1_1 = x$ and $1_2 * x = x * 1_2 = x$. Then $1_1 = 1_1 * 1_2$ (since 1_2 is an identity) and $1_1 * 1_2 = 1_2$ (since 1_1 is an identity), which implies $1_1 = 1_2$. Inverses are also unique, since given $x \in G$ and elements $y, z \in G$ with $x * y = y * x = 1$ and $x * z = z * x = 1$, associativity allows us to conclude that $y = y * 1 = y * (x * z) = (y * x) * z = 1 * z = z$. In other words, any element that acts like an inverse of x really is *the* inverse of x.

Task 48
In this task, we use the uniqueness of inverses to prove three elementary results about group inverses. Assume G is a group with identity element 1 and let $x, y \in G$.

(a) Use uniqueness of inverses to prove that $(x^{-1})^{-1} = x$.
That is, show that x acts like the inverse of x^{-1}.

(b) Use uniqueness of inverses to prove that $(xy)^{-1} = y^{-1}x^{-1}$.

(c) Prove that $xy = 1 \;\Rightarrow y = x^{-1}$.
This means that we can establish y is the inverse of x simply by checking that $xy = 1$, even in the case where G is non-abelian.

Task 49
In this task, we examine the defining properties of a group by analyzing some additional examples and non-examples. Note that several of the groups in this task are infinite.

(a) Consider the set of non-zero rational numbers $\mathcal{Q}^* = \left\{\frac{m}{n} \,|\, m, n \in \mathcal{Z}^*\right\}$ under multiplication $\times$, where $\mathcal{Z}^*$ denotes the set of non-zero integers. Given that $\times$ is associative, show $(\mathcal{Q}^*, \times)$ is a group by checking that it satisfies the other three group properties. For an arbitrary $q = \frac{m}{n} \in \mathcal{Q}^*$, identify q^{-1} clearly and verify that $q^{-1} \in \mathcal{Q}^*$.

Note: Since rational multiplication is commutative, $(\mathcal{Q}^*, \times)$ is our first example of a ***commutative group***, also referred to as an ***abelian group***.

(b) Explain why the set of non-zero integers $\mathcal{Z}^*$ is not a group under multiplication.
Identify all properties of the group definition that fail.

(c) Now consider $\mathcal{Z}$ under addition. Given that integer addition is closed and associative, explain why $(\mathcal{Z}, +)$ is a group with identity element 0 and $n^{-1} = -n$ for each $n \in \mathcal{Z}$.

Note: $(\mathcal{Z}, +)$ is an example of an ***additive group*** for which it is standard to use the symbols '+,' '0,' and '$-n$' to denote the group operation, identity and inverse of n respectively. $(\mathcal{Z}, +)$ is also an abelian group.

(d) Let $\mathcal{Z}_6 = \{0\,,\,1\,,\,2\,,\,3\,,\,4\,,\,5\}$ and let + denote addition modulo 6.
That is, to find $i + j$, we take the remainder of the sum when divided by 6.
(Or think of arithmetic on a clock with $0, 1, 2, 3, 4, 5$ equally spaced on its face.)[58]
For example: $3 + 4 = 1$; $5 + 5 = 4$; $1 + 2 = 3$; $1 + 5 = 0$; $4 + 0 = 4$

Given that + is associative with identity 0, show that $(\mathcal{Z}_6, +)$ is a group by explaining why $\mathcal{Z}_6$ is closed under +, and then finding the inverse element $-n$ for each $n \in \mathcal{Z}_6$. Is $(\mathcal{Z}_6, +)$ abelian? Why or why not?

(e) Consider the set $M = \{X \,|\, X \text{ is a } 2 \times 2 \text{ matrix}\}$ under matrix multiplication $\times$.
Given that $\times$ is associative, show that there is an identity element I, but that $(M, \times)$ is not a group. How could we modify the definition of M to obtain a group under $\times$?

[58] Compare this to your diagram of the 6^{th} roots of unity in Task 3.

Task 50
This task explores the definition and properties of integer exponents for group elements.

Let $(G, *)$ be a group. For $x \in G$ and $n \in \mathcal{Z}^+$, define $x^n = \underbrace{x * x * \ldots * x}_{n \text{ factors}}$.

- **(a)** Given $x, y \in G$ and $n \in \mathcal{Z}^+$, explain why it is not generally the case that $(xy)^n = x^n y^n$. Is it generally the case that $(xy)^n = y^n x^n$? Why or why not?
- **(b)** Assume G is an abelian group. Prove by induction that $(xy)^n = x^n y^n$ for all $x, y \in \mathcal{Z}^+$.
- **(c)** In this part, assume only that G is a group (not necessarily abelian), but let $x, y \in G$ be such that $xy = yx$; that is, assume that the specific group elements x, y commute with each other, but do not assume that all group elements commute. Indicate how to modify your proof in part (b) to show that $(xy)^n = x^n y^n$ for all $n \in \mathcal{Z}^+$.
- **(d)** Use part (c) and uniqueness of inverses to prove that $(x^n)^{-1} = (x^{-1})^n$.
  That is, show that $(x^{-1})^n$ acts like the inverse of x^n.
- **(e)** Using part (d), we extend the definition of integer powers for a group element $x \in G$ to include negative integer powers by setting $x^{-n} = (x^n)^{-1}$ for all $n \in \mathcal{Z}^+$.
  To complete the definition, we also let $x^0 = 1$ for all $x \in G$.
  Assume that G is a group, $x, y \in G$ and $m, k \in \mathcal{Z}$.
  Use cases (based on the signs of k, m) and the (extended) definition of integer powers to prove the following exponential properties hold in a group:
  - **(i)** $(x^k)^m = x^{km}$
  - **(ii)** $x^k x^m = x^{k+m}$
  - **(iii)** If $xy = yx$, then $(xy)^m = x^m y^m$. (The case $m > 0$ was done in part (c).)

While we are looking at examples of groups, we should also look at some subgroup examples. Cayley himself did not use this term, but his emphasis on the closure property and other aspects of his work tell us that he was aware of the subgroup concept. He was also familiar with Cauchy's work, where the idea of a subgroup arises in Cauchy's notion of a 'system of conjugate permutations.' Recall from our discussion of Cauchy's work that we defined a *subgroup of* S_n to be any subset $H \subseteq S_n$ which is closed under products. Of course, S_n is a very special example of a group. In particular, all of its elements (i.e., permutations on n objects) are of finite order; under this condition, closure under products is sufficient to obtain a subgroup, as we will prove in Task 54.

In general, however, closure of products is not enough to capture Cauchy's full notion of a 'system of conjugate permutations.' Taking our cue from what Cauchy said about these systems, the essential idea will be to ensure that H is a subset of G which is itself a group in its own right — a group living inside another group, so to speak. To ensure this, it may seem that we will need to check that an alleged subgroup H has all four properties from the definition of group. However, since G is already known to be a group and $H \subseteq G$, then we already know that associativity applies to all of G's elements, including those belonging to H. Also, if H is closed under both products and inverses, then given $h \in H$, we'll have $h^{-1} \in H$ (by closure under inverses), so that $1 = hh^{-1} \in H$ (by closure under products). In short, to define the general notion of a subgroup, only two properties from the definition of a group are needed:

Definition Let G be a group and H a non-empty subset of G. Then H is a ***subgroup*** of G iff the following properties are satisfied:

Closure under Products:	For all $a, b \in H$, $a * b \in H$.
Closure under Inverses:	For all $a \in H$, $a^{-1} \in H$.

You may feel we stretched the truth a bit in saying that there are only two properties to check; in fact, H must satisfy four properties in all: (1) H must be a subset of G; (2) H must be non-empty; (3) H must be closed under products; and (4) H must be closed under inverses. It is usually easy to check the first of these (that $H \subseteq G$), while checking the second (that $H \neq \emptyset$) is typically done by showing $1 \in H$. (In fact, $H = \{1\}$ is always a subgroup of a group G, known as the ***trivial subgroup***, so that 1 may be the only element in a subgroup H.)

As for the two closure properties required for a non-empty subset to be a subgroup, note that these are not a matter of checking that multiplication is well-defined, or that inverses exist for all elements. We already know both these things to be true, since G is given as a group. Rather, we must show that *products and inverses formed from elements of H live in the right place* — namely, in H. For example, even though every non-zero integer n has a multiplicative inverse $\frac{1}{n}$, it is not generally the case that $\frac{1}{n}$ is itself an integer; thus, $\mathcal{Z}^*$ is not closed under multiplicative inverses,[59] which in turn implies that $(\mathcal{Z}^*, \times)$ is NOT a subgroup of the group $(\mathcal{Q}^*, \times)$. On the other hand, $(\mathcal{Q}^*, \times)$ IS a subgroup of the group $(\mathcal{R}^*, \times)$ since both the products and the inverses of non-zero rational numbers are necessarily non-zero rational numbers. The remaining five tasks in this subsection will give you some practice with this important idea.

Task 51
In this task, we examine examples and non-examples of subgroups of specific groups.

(a) Consider the group $(\mathcal{Q}^*, \times)$ and let $H = \{\frac{1}{2^m} \,|\, m \in \mathcal{Z}^+\}$, where $\mathcal{Z}^+$ denotes the set of natural numbers. Note that H is a non-empty subset of $\mathcal{Q}^*$. Show that H is closed under products, but is not a subgroup of $\mathcal{Q}^*$.

(b) Again consider the group $(\mathcal{Q}^*, \times)$, but now let $K = \{\frac{1}{2^m} \,|\, m \in \mathcal{Z}\}$.
Show that K is a subgroup of $\mathcal{Q}^*$.

(c) Note that $(\mathcal{R}, +)$ is a group with identity 0, where $\mathcal{R}$ denotes the set of real numbers. Let $J_1 = \{\log(a) \,|\, a \in \mathcal{Q}^+\}$. Determine whether J_1 is a subgroup of $\mathcal{R}$ under $+$, and justify your reply. What about $J_2 = \{\log(a) \,|\, a \in \mathcal{Z}^+\}$?

Task 52
In this task, we examine results related to the special case of a subgroup which is *cyclic.*

(a) Assume that G is a group with identity 1, where G is not necessarily abelian.
Let $\alpha \in G$ and $H = \langle\alpha\rangle = \{\alpha^k \,|\, k \in \mathcal{Z}\}$.
Prove that H is an abelian subgroup of G.
Note: H is called the ***cyclic subgroup generated by*** α.

(b) Identify all cyclic subgroups of $(\mathcal{Z}_6, +)$, where $+$ denotes addition modulo 6.
Note: In this case, we are looking at *additive* subgroups: $\langle\alpha\rangle = \{k\alpha \,|k \in \mathcal{Z}\}$.
(Why we really only need to consider the values of integer k for which $1 \leq k \leq 6$ here?)

(c) If there is an element of G which generates the entire group G, we say that G is cyclic. Prove $(\mathcal{Z}, +)$ is cyclic and find all its generators. Is $(\mathcal{Q}, +)$ cyclic? Explain.

[59] Note that $\mathcal{Z}^*, \times)$ is closed under products.

Task 53
In this task, we examine some more general ways to obtain subgroups of a given group G.

(a) Assume G is an abelian group with identity 1. Let $H = \{\, x \in G \,|\, x = x^{-1} \,\}$.
Prove H is a subgroup of G, indicating clearly where the abelian assumption is used.

(b) Assume G is an abelian group with identity 1. Let $n \in \mathcal{Z}$ and $K = \{\, x \in G \,|\, x^n = 1 \,\}$.
Prove K is a subgroup of G, indicating clearly where the abelian assumption is used.

(c) Assume G is a group with identity 1, but not necessarily abelian.
Define the ***center of G*** to be the set $C = \{\, x \in G \,|\, xa = ax \text{ for all } a \in G \,\}$.
(Thus, C consists of all the elements of G that commute with every element of G.)
Prove that C is a subgroup of G

Task 54
In this task, we prove that closure under products is a sufficient condition for H to be a subgroup of G in the case where G is a group in which every element is of finite order, where the definition of order for permutations generalizes to abstract groups as follows:

> **Definition** Let G be a group with identity element 1 and let $\alpha \in G$. Then the ***order of*** α ***in*** G is the least positive integer r for which $\alpha^r = 1$, provided such a number exists. If such a number exists, we say α has ***finite order***.

Assume G is a group with identity element 1 and that every element of G has finite order.

(a) Let $\alpha \in G$, $\alpha \neq 1$, and suppose α has order $r \in \mathcal{Z}^+$. Explain why $\alpha^{-1} = \alpha^{r-1}$.

(b) Now assume that H is a non-empty subset of G which is closed under products.
Prove that H is a subgroup of G.

Task 55
In this task, we examine a special subgroup of S_n, where $n \in \mathcal{Z}^+$ and $n \geq 2$.

Recall from Tasks 30 and 34 that every permutation can be written as the product of transpositions, where a transposition is a 2-cycle.

Also recall that, although this decomposition is not unique, the parity (even versus odd) of the number of transpositions used is unique, since a permutation can be written either as the product of an even number of transpositions, or as the product of an odd number of transpositions, but never both. The parity (even or odd) of a permutation is defined to be the parity (even or odd) of this number.

(a) Show that the product of two permutations is even if and only if both permutations have the same parity (i.e., both are even or both are odd).

Note: Your proof must consider all three parity cases (even/even, odd/odd, even/odd). Do you see why?

(b) Let A_n be the subset of S_n consisting of all even permutations.
Prove that A_n is a subgroup of S_n. **Note:** A_n is called the *alternating subgroup.*

(c) Explain why the set $S_n - A_n$ of all odd permutations is never a subgroup of S_n.
Give at least two reasons.

(d) Show that S_n contains an equal number of even and odd permutations by letting $A_n = \{\alpha_1\,, \alpha_2\,, \ldots\,, \alpha_k\}$ and $S_n - A_n = \{\beta_1\,, \beta_2\,, \ldots\,, \beta_m\}$, where $k, m \in \mathcal{Z}^+$, and then proving $k = m$. Explain why we can then conclude that $|A_n| = \frac{n!}{2}$.
Hint? Consider the sets $\{\beta_1\alpha_1\,, \beta_1\alpha_2\,, \ldots\,, \beta_1\alpha_k\}$ and $\{\beta_1\beta_1\,, \beta_1\beta_2\,, \ldots\,, \beta_1\beta_m\}$.

Task 56
In this task, we connect the notion of a subgroup to that of a Cayley table.

Consider the set $M_{2,2} = \{X \mid X \text{ is a } 2 \times 2 \text{ matrix, } \det(X) \neq 0\}$. Since matrices with non-zero determinants are invertible, $M_{2,2}$ is a (non-abelian) group under matrix multiplication.

Let $G = \{I, A, B, C, D, E\}$, with I the identity matrix and A, B, C, D, E as follows:

$$A = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \quad C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix} \quad E = \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}$$

(a) Complete the following Cayley table for G, and explain how it shows that G is a subgroup of $M_{2,2}$. In particular, indicate how the table shows that G satisfies the Inverse Property.

	I	A	B	C	D	E
I	I	A	B	C	D	E
A	A		I			
B	B	I				
C	C			I		
D	D				I	
E	E					I

(b) Determine the order of each element of G. (See Task 54 for definition of 'order'.) Which familiar group of order 6 does G remind you of, and why?

3.2 Some Theorems on Groups

Look again at the Cayley table you completed in Task 56, and notice how each element of G appears exactly once in every row and every column. This is exactly the theorem stated by Cayley (without proof) immediately following his definition of group. We re-state this theorem here, somewhat more formally, and label it as 'Cayley's Theorem 1' for future reference.

Cayley's Theorem 1
If $G = \{1\,,\,\alpha\,,\,\beta\,,\ldots\}$ is a group with identity element 1, then the Cayley table for G is such that every row (e.g., $\{\alpha\,,\,\alpha^2\,,\,\alpha\beta\;\,,\ldots\}$) and every column (e.g., $\{\alpha\,,\,\alpha^2\,,\,\beta\alpha\,,\ldots\}$) contains every element of G exactly once.

Here, Cayley's rows and columns correspond to the two series of permutations S, SP, SQ, $SR, \ldots$ and S, PS, QS, $RS, \ldots$ which Cauchy discussed in the excerpt on page 41 of this project. Cauchy also noted in that excerpt (again without proof) that the terms of these series are the same as those of the original given system of permutations (1, P, Q, $R, \ldots$) only 'arranged in a new order'.[60] The major difference between the assertions of the two mathematicians is that Cauchy was speaking only about permutations, while Cayley's claim applied to *any* set of operation symbols that satisfied his definition of a group. As a result, Cayley's result is far more general than that of Cauchy. In Task 57, we outline a proof of Cayley's Theorem 1 using only the properties appearing in Cayley's definition of group. Task 58 then looks at a more formal version of this same theorem, and introduces some current notation and terminology related to it.

Task 57
Let $G = \{1\,,\, \alpha\,,\, \beta\,, \ldots\}$ be a group with identity element 1. Let $\phi \in G$ be arbitrary.

(a) Using only the properties of a group given in the definition of a group on page 56, explain why the row corresponding to ϕ in the Cayley table for G (see below) includes every element of G exactly once. Write your explanation as a formal proof, and clearly indicate which properties of the definition you are using and where they are needed. **Note that there are two things to show here:**

(i) Every element of G appears somewhere in the row corresponding to ϕ.
That is, given any $\theta \in G$, there is some $\gamma \in G$ for which $\theta = \phi\gamma$.
(After introducing θ as an arbitrary element of G, explicitly define γ, indicate how you know γ is an element of G, and verify that $\theta = \phi\gamma$ for this choice of γ.)
Note: Since ϕ is also arbitrary, this corresponds to the formal statement:

$$(\forall \phi \in G)\,(\forall \theta \in G)\,(\exists \gamma \in G)\,(\theta = \phi\gamma)$$

(ii) No element of G appears more than once in the row corresponding to ϕ.
That is, given $\gamma_1, \gamma_2 \in G$ with $\phi\gamma_1 = \phi\gamma_2$, it follows that $\gamma_1 = \gamma_2$.
(Start by introducing γ_1, γ_2 as arbitrary elements of G and assuming $\phi\gamma_1 = \phi\gamma_2$.)
Note: Since ϕ is also arbitrary, this corresponds to the formal statement:

$$(\forall \phi \in G)\,(\forall \gamma_1, \gamma_2 \in G)\,(\phi\gamma_1 = \phi\gamma_2 \Rightarrow \gamma_1 = \gamma_2)$$

	1	α	β	$\ldots$	ϕ	$\ldots$	γ	$\ldots$
1								
α								
β								
$\vdots$								
ϕ	ϕ	$\phi\alpha$	$\phi\beta$	$\ldots$	ϕ^2	$\ldots$	$\phi\gamma$	$\ldots$
$\vdots$								

(b) How would you modify this explanation to apply to the columns of the Cayley table?

[60] You may wish to review Tasks 38 and 39 concerning this part of Cauchy's paper.

Task 58
In this task, we explore a theorem related to Cayley's Theorem 1. We first state this theorem in the more formalized terms that one finds in current textbooks and comment on their meaning and importance. We then outline its proof for you to complete.

> **Cayley's Theorem 1 - A Formalized Variation**
> If $G = \{1\,,\,\alpha\,,\,\beta\,,\ldots\}$ is a group with identity element 1 and $\phi \in G$ is arbitrary, then $G\phi = G = \phi G$, where $G\phi = \{\phi\,,\,\alpha\phi\,,\,\beta\phi\,,\ldots\}$ and $\phi G = \{\phi\,,\,\phi\alpha\,,\,\phi\beta\,,\ldots\}$.

Comments

- The equation $\phi G = G\phi$ does *not* imply that the elements of G commute, but only that these sets contain the same elements, possibly arranged in a different order.[61]
- By using set notation in the conclusion $\phi G = G$ of this theorem, we are no longer explicitly stating that every element of G appears in the form $\phi\gamma$ for only one element $\gamma \in G$. This fact still holds, however, as proved Task 57(a-ii).
- It is important that ϕ is assumed to be an element of the group G in this theorem!

  To see this, remember that Cauchy's proof of Lagrange's Theorem for S_n (discussed in Subsection 2.2) relied on sets of the form $\{U\,,\,UP\,,\,UQ\,,\ldots\}$, which were obtained by (left) multiplying the elements of the set $\{1\,,\,P\,,\,Q\,,\ldots\}$ by the permutation U. In that proof, the set $\{1\,,\,P\,,\,Q\,,\ldots\}$ was specified to be a subgroup of S_n and the permutation U came from *outside* of that subgroup (but inside S_n). Consequently, as Cauchy argued, the two sets $\{U\,,\,UP\,,\,UQ\,,\ldots\}$ and $\{1\,,\,P\,,\,Q\,,\ldots\}$ are disjoint.

  In current adaptations of Cauchy's proof of Lagrange's Theorem for S_n to the case of a general group G, it is standard to refer to the type of set employed in the proof as a (left) coset, denoted and defined by $aH = \{ah | h \in H\}$, where H is a subgroup of G and $a \in G$. In this context, it may or may not be the case that $a \in H$ and the formal variation of Cayley's Theorem 1 that we are looking at in this task may not apply. In the event that $a \in H$, then this theorem does apply, and we can conclude that $aH = H$. In the contrary case, where $a \notin H$, we would be able to show $Ha \cap H = \emptyset$ in exactly the same way that Cauchy established this result for permutation groups.

Proof Sketch: Formalized Variation of Cayley's Theorem 1

Let $G = \{1\,,\,\alpha\,,\,\beta\,,\ldots\}$ be a group with identity element 1. Let $\phi \in G$ be arbitrary.

Define the sets $G\phi = \{\phi\,,\,\alpha\phi\,,\,\beta\phi\,,\ldots\}$ and $\phi G = \{\phi\,,\,\phi\alpha\,,\,\phi\beta\,,\ldots\}$.

(a) Complete the following details to show that $G\phi = G$.

  (i) Explain why $G\phi \subseteq G$.
  That is, given $x \in G\phi$, what property of groups allows us to conclude that $x \in G$?

  (ii) To show that $G \subseteq G\phi$, we let $y \in G$ and show that $y \in G\phi$.
  Explain why this requires us to find some $z \in G$ for which $y = z\phi$.
  Then use group properties to determine z. How do we know that $z \in G$?

(b) Now prove that $\phi G = G$.

[61] It may be helpful to review Tasks 38 and 39 in this regard.

We now return to our reading of Cayley's second paragraph, in which he makes a connection between finite groups of order n and the set of n^{th} roots of unity. This excerpt is packed with theorems! As you read through it, try to identify as many of these as possible. Also notice that Cayley is considering only finite groups in what follows.

Suppose that the group

$$1,\ \alpha,\ \beta, \ldots$$

contains n symbols, it may be shown that each of these symbols satisfies the equation

$$\theta^n = 1;$$

so that a group may be considered as representing a system of roots of this symbolic equation. It is, moreover, easy to show that if any symbol α of the group satisfies the equation $\theta^r = 1$, where r is less than n, then that r must be a submultiple of n; it follows that when n is a prime number, the group is of necessity of the form

$$1,\ \alpha,\ \alpha^2,\ \ldots \alpha^{n-1},\ (\alpha^n = 1);$$

and the same may be (but is not necessarily) the case, when n is a composite number. But whether n be prime or composite, the group, assumed to be of the form in question, is in every respect analogous to the system of the roots of the ordinary binomial equation $x^n - 1 = 0$; thus, when n is prime, all the roots (except the root 1) are prime roots; but when n is composite, there are only as many prime roots as there are numbers less than n and prime to it, &c.

In the remainder of this subsection, we separate the various theorems including in this excerpt, stating each in current terminology and numbering them in the order in which Cayley cited them. Since Cayley himself did not provide proofs, we will also examine why he felt they were so straightforward in light of what we know about permutation groups, and write our own proofs for them using current terminology and notation. We begin with three theorems drawn from the first two sentences of the excerpt.

Cayley's Theorem 2
Let G be a finite group of order n with identity element 1 and let $\alpha \in G$ be arbitrary. Then $\alpha^n = 1$.

Cayley's Theorem 3[62]
Let G be a finite group of order n with identity element 1 and suppose $\alpha \in G$ has order r. Then r is a divisor (submultiple) of n.

Cayley's Theorem 4
Every group G of prime order is cyclic.
That is, if n is prime and $|G| = n$, then $G = \langle\alpha\rangle = \{\, 1\,,\ \alpha\,,\ \alpha^2\,,\ \ldots, \alpha^{n-1}\,\}$ for some $\alpha \in G$.

[62] Notice that we have interpreted Cayley slightly here, since he did not explicitly state that r had to be the *least* number for which $\alpha^r = 1$, but only required that r be less than n. However, $r < n$ is not enough for Cayley's conclusion to hold. For example, if G is a group of order 6 and $\alpha \in G$ has order 2, then we will have $\alpha^4 = (\alpha^2)^2 = 1^2 = 1$ with $4 < 6$, even though 4 is not a divisor of 6. A specific example of this occurs when $G = S_3$ and $\alpha = (1,\, 2)$.

All three of these theorems bear a resemblance to facts we already know about permutation groups. In Task 42, for example, we saw how the following version of Cayley's Theorem 3 follows from Lagrange's Theorem for S_n:

> **Cayley's Theorem 3 - Restricted to S_n**
> For every permutation $P \in S_n$, the order of P divides the order of S_n.

The standard proof given today for the fully general version of Cayley's Theorem 3 (i.e., without the restriction to permutation groups) uses the fully generalized version of Lagrange's Theorem, in a fashion similar to that you employed in Task 42. We will say more about this below (and remind you of Lagrange's Theorem along the way). But first, it will be helpful to look at some facts related to the order of elements in an abstract group, drawing on what we already know about the order of permutations. We begin by restating the definition of order for group elements, and extending it just a bit to include elements which are not of finite order:

> **Definition** Let G be a group with identity element 1 and let $\alpha \in G$. Then the **order of α in** G is the smallest natural number r for which $\alpha^r = 1$, provided such a number exists. If such a natural number exists, α has ***finite order***; otherwise, α has ***order infinity***.

Some groups (e.g., $(\mathcal{Z}, +)$) contain only one element of finite order (i.e., the identity). In contrast, *every* element of the permutation group S_n has finite order. It turns out — as Cayley's Theorem 2 implies — that every group of finite order contains only elements of finite order. This claim is straightforward to prove directly from the Pigeonhole Principle, just as Cauchy did in his proof that every permutation has finite order. We start our proofs of Cayley's theorems by doing this and examining some other facts related to the order of group elements in Tasks 60 − 62. We then use these facts to prove a theorem about cyclic groups in Task 63. Task 59 first provides examples of order computations for group elements in non-permutation groups to provide some concrete context for these various proofs.

Task 59
In this task, you will compute the order of specific elements in non-permutation groups.

(a) Consider $\mathcal{Z}_8 = \{\,0\,,\,1\,,\,2\,,\,3\,,\,4\,,\,5\,,\,6\,,\,7\,\}$ under addition modulo 8.
Determine the order of each element; indicate which of these generate the entire group.
Remember that 0 is now the identity in this additive group!

(c) Find all elements of finite order in the group of non-zero rational numbers $(\mathcal{Q}^*, \times)$.

Task 60
In this task, you will formally prove the following:

> **Order Fact 1**
> If G is a finite group and $\alpha \in G$, then α has finite order.

Begin by assuming G is a finite group of order n with identity element 1 and $\alpha \in G$.

(a) Apply the Pigeonhole Principle (see footnote 33, page 30) to the set
$H = \{\alpha^m \,|\, m \in \mathcal{Z}^+\}$ in order to obtain $h, l \in \mathcal{Z}^+$ such that $\alpha^h = \alpha^l$ and $h < l$.

(b) Use group properties to show that $\alpha^{l-h} = 1$.
Explain how this shows that the set $K = \{k \in \mathcal{Z}^+ \,|\, \alpha^k = 1\}$ is non-empty.

(c) Complete the proof by applying the Well Ordering Principle[63] to the set K.

[63] The Well Ordering Principle states that every non-empty subset of $\mathcal{Z}^+$ has a least element.

Task 61
In this task, you will formally prove the following:

> **Order Fact 2**
> If $\alpha \in G$ has finite order r and $m \in \mathcal{Z}$ with $\alpha^m = 1$, then r is a divisor of m.

Begin by first stating all necessary assumptions, and using the division algorithm to obtain $k, j \in \mathcal{Z}$ such that $m = kr + j$ and $0 \leq j < r$. Then employ the assumption that α has order r to show that $j = 0$.

Note: This proof does not require the group G itself to have finite order.

Task 62
In this task, you will formally prove the following:

> **Order Fact 3**[64]
> The order of an element $\alpha \in G$ is equal to the order of the cyclic subgroup $\langle \alpha \rangle = \{ \alpha^m \,|\, m \in \mathcal{Z} \}$ generated by α.

Begin by assuming G is a group and noting that there are two cases to consider:
(i) $\alpha \in G$ has finite order r, $r \in \mathcal{Z}^+$; and (ii) $\alpha \in G$ has in finite order.

Beginning with the finite case, assume $\alpha \in G$ has finite order r, where $r \in \mathcal{Z}^+$.

(a) Let $m \in \mathcal{Z}$. Use the division algorithm to obtain $k, j \in \mathcal{Z}$ such that $m = kr + j$ and $0 \leq j < r$. Show that $\alpha^m \in \{ 1 \,,\, \alpha \,,\, \alpha^2 \,,\, \ldots \,,\, \alpha^{r-1} \}$. Note that this means that $\langle \alpha \rangle$ contains at most r elements.

(b) Let $h, l \in \{ 0 \,,\, 1 \,,\, \ldots , r-1 \}$ and suppose that $\alpha^h = \alpha^l$. Use the fact that r is the *least* natural number for which $\alpha^r = 1$ to show that $h = l$.

(c) Explain how (a) and (b) allow us to conclude that $\langle \alpha \rangle = \{ 1 \,,\, \alpha \,,\, \alpha^2 \,,\, \ldots , \,,\, \alpha^{r-1} \}$ contains exactly r elements.

(d) At this point in the proof, the following implication has been established:

$$\text{If } \mathrm{ord}(\alpha) = r, \text{ where } r \in \mathcal{Z}^+, \text{ then } |\langle \alpha \rangle| = r.$$

Now prove its converse:

$$\text{If } \langle \alpha \rangle = \{ \alpha^k \,|k \in \mathcal{Z} \} \text{ has finite order } r, \text{ then } \alpha \text{ has order } r.$$

Note: You need to show that (i) $\alpha^r = 1$ and (ii) $\alpha^i \neq 1$ for $i \in \mathcal{Z}^+$ with $i < r$.

(e) Part (d) completes the proof of the finite case. Explain why the case in which α has infinite order now follows directly from the contrapositive of the following statement:

$$\langle \alpha \rangle = \{ \alpha^m \,|\, m \in \mathcal{Z} \} \text{ can be infinite iff there is no } r \in \mathcal{Z}^+ \text{ for which } \alpha^r = 1.$$

[64] Both Cauchy and Cayley hinted at this relationship in their work; in fact, Cauchy explicitly proved it for permutation groups, while Cayley took it for granted as another easy-to-establish result.

Task 63
This task uses ideas similar to those used in the proofs of Order Facts 1 – 3 in order to prove the following theorem about subgroups of cyclic groups:[65]

Every subgroup of a cyclic group is cyclic.

Begin by assuming $G = \langle\alpha\rangle$ is a cyclic group with generator α and H is a subgroup of G. (We do not need to assume that either G or H is finite.)

To show that H is cyclic, we must find a generator. Note that if $H = \{1\}$, then we are done. (*Why?*) Thus, we can assume that H is a non-trivial subgroup.

(a) Explain why the set $K = \{m \in \mathcal{Z}^+ \mid \alpha^m \in H\}$ is non-empty.
Then apply the well-ordering principle to K to obtain the least natural number m_0 such that $\alpha^{m_0} \in H$. Our plan is to show that H is generated by α^{m_0}, or $H = \langle\alpha^{m_0}\rangle$.

(b) Begin with the easy direction, and explain why $\langle\alpha^{m_0}\rangle \subseteq H$.

(c) For the other direction, let $\beta \in H$ and show $\beta \in \langle\alpha^{m_0}\rangle$ as follows:

(i) Explain how we know that there exists some $m \in \mathcal{Z}$ such that $\beta = \alpha^m$.

(ii) Use the division algorithm to obtain $k, j \in \mathcal{Z}^+$ with $m = km_0 + j$ and $0 \leq j < m_0$.

(iii) Use part (ii) and the definition of subgroup to explain why the element α^j must be in H. Then conclude from the definition of m_0 that $j = 0$.

(iv) Complete the proof by explaining why $j = 0$ implies that $\beta \in \langle\alpha_0^m\rangle$.

We are now ready to consider the proof of Cayley's Theorem 3, which states that the order of any element divides the order of the entire group. First, some notation and reminders. Assume G is a group of order n and $\alpha \in G$. By Order Fact 1, we know that α has some finite order r. To distinguish the notion of *group order* (cardinality) from that of *element order*, we will write $|G| = n$ and $\text{ord}(\alpha) = r$. If we knew that $\alpha^n = 1$ (Cayley's Theorem 2), then we could conclude that r divides n by Order Fact 2. But we have not yet established Cayley's Theorem 2 in this project; in fact, we plan to prove Cayley's Theorem 2 as a corollary to Cayley's Theorem 3. Thus, to avoid circular reasoning, we will need to prove that r divides n in some other way. We do this by bringing in the cyclic subgroup $H = \langle\alpha\rangle$ generated by α. As we know from our proof of Order Fact 3, the definition of $\text{ord}(\alpha)$ as the *least* natural number for which $\alpha^r = 1$ implies that $|H| = r$. The only remaining piece needed to prove that $r = \text{ord}(\alpha) = |H|$ divides $n = |G|$ is to recall what Lagrange's Theorem says about the orders of a group and its subgroups:

Lagrange's Theorem for Finite Groups
If H is a subgroup of the finite group G, then the order of H divides the order of G.

Remember that our first statement of Lagrange's Theorem in Subsection 2.2 applied only to the particular group S_n of permutations on n elements. But as we commented at that time, Cauchy's proof strategy for permutation groups is readily adapted to the general case of any finite group, either directly or by re-casting it in the language of cosets. You can look back at the proof you wrote in Task 41 to see how easy this is to do directly from Cauchy's proof; in Appendix II, we also outline the now-standard coset proof of the general theorem. For the rest of this project, however, we simply assume (as Cayley seemed to have done) that Lagrange's Theorem holds for groups in general.

[65] Although this theorem does not involve facts about the order of elements, the similarity of its proof to those of Order Facts 1, 2 and 3 makes it a good exercise with these proof techniques.

Task 64

Use Lagrange's Theorem for Finite Groups and the ideas we sketched out just above its statement to write a complete formal proof of Cayley's Theorem 3:

$$\text{If } G \text{ is a finite group and } \alpha \in G, \text{ then } \operatorname{ord}(\alpha) \,\big|\, |G|.$$

Task 65

Assume G is a group of order n with identity element 1 and let $\alpha \in G$ be arbitrary.

Now that we know (by Cayley's Theorem 3) that $\operatorname{ord}(\alpha)$ divides $|G|$,
it is easy to establish Cayley's Theorem 2 by proving that $\alpha^{|G|} = 1$.

Do this by letting $n = kr$, where $|G| = n$, $\operatorname{ord}(\alpha) = r$ and $k \in \mathcal{Z}^+$.
Write your proof out completely and carefully, explaining clearly how k is obtained.

Of course, it may be that Cayley had some proof of his Theorem 2 in mind which somehow did not require first proving his Theorem 3. Given Cauchy's work on permutations, however, the proof you just completed in Task 65 is a good possibility for how Cayley himself arrived at his Theorem 2; it is also how Cayley's Theorem 2 is typically proven today. Cayley's Theorem 3 also makes it easy to prove Cayley's Theorem 4 — that every group of prime order is cyclic — although this too can be proven directly from Lagrange's Theorem.

Task 66

Assume n is prime and G is a group of order n with identity element 1.

Let $\alpha \in G$ be such that $\alpha \neq 1$. *(How do we know such an element exists?)*
Use the fact that $\operatorname{ord}(\alpha)$ divides $|G|$ (from Cayley's Theorem 3) to show that $G = \langle \alpha \rangle$.

Note: If you prefer, you can instead show that G is cyclic by using Lagrange's Theorem.

Before we return to the remaining theorems stated by Cayley in the previous excerpt from his paper, we comment on one more aspect of Lagrange's Theorem. Clearly, this theorem tells us that the order of a subgroup can assume only certain very specific values; namely, only those which are the divisors of the order of a group. Thus, for instance, a group of order 6 can *not* have a subgroup of order 4 or 5. However, Lagrange's Theorem does *not* say that a group of order 6 *must* have a subgroup of order 1, 2, 3 or 6, even though these are all divisors of its order. As it turns out, every group of order 6 does have subgroups of every possible order, as do all cyclic groups.[66] The Alternating Group A_4 consisting of all even permutations in S_4, on the other hand, is a group of order 12 which contains no subgroup of order 6, even though 6 is a divisor of the order of the group.[67]

[66] For groups of order 6, this follows from a proof we will see in Subsection 4.2 in which Cayley showed that all groups of order 6 necessarily have elements of order 2 and 3, which in turn generate subgroups of order 2 and 3. Although Cayley employed Lagrange's Theorem as a tool in this proof, the existence of these subgroups does not follow directly from Lagrange's Theorem and can be proven without making use of it.

The proof that a cyclic group of order n has a (cyclic) subgroup of order d for every divisor d of n likewise does not require Lagrange's Theorem, but relies instead on basic facts about the order of elements established in Tasks 60–63. (Writing out this proof would serve as a good review and application of those facts!)

[67] A_n was defined in Task 55 and shown there to have order $\frac{n!}{2}$. The claim that A_4 has no subgroup of order 6 can be verified by writing down the 12 even permutations in S_4, and using the closure properties to eliminate the possibility of building a subgroup containing exactly 6 elements from that list. (It may be easier to first write down the 12 odd permutations in S_4 since this can be done using what we know about the parity of n-cycles.) In Task 75, we will prove that A_4 does contain subgroups of order 2, 3 and 4.

Similarly, although Cayley's Theorem 3 tells us that the order of any element in a finite group must divide the order of the group, that theorem does not imply the existence of elements of any particular order. In fact, the only way a group of order n could have an element of order n would be if the group were cyclic. But as Cayley remarked in the preceding excerpt, for a composite number n (such as $n = 6$):

> " it may be (but is not necessarily) the case that the group is of the form
>
> $$1,\ \alpha,\ \alpha^2,\ \ldots\alpha^{n-1},\ (\alpha^n = 1)."$$

Cayley again gave no proof for this, but we have already seen several examples that illustrate its truth. The following task reminds us of these examples for the case $n = 6$.

Task 67

(a) Show that $(\mathcal{Z}_6, +)$ is cyclic and find all its generators. (See Task 52(b).)

(b) Let $U_6 = \left\{e^{\frac{2\pi k}{6}i} \,|\, k = 1, 2, 3, 4, 5, 6\right\}$ denote the set of the 6^{th} roots of unity.
Recall from Task 5 that $\alpha = e^{\frac{2\pi}{6}i} = e^{\frac{\pi}{3}i}$ is a primitive root of unity, meaning that $U_6 = \{\,1\,,\,\alpha\,,\,\alpha^2\,,\,\alpha^3\,,\,\alpha^4\,,\,\alpha^5\,\}$. Thus, $(U_6, \times)$ is a cyclic group generated by α.
Determine all other generators of U_6. (Task 3 may also be helpful.)
What similarities/differences do you notice between the groups U_6 and $\mathcal{Z}_6$?

(c) Consider the group of matrices $H = \{I, A, B, C, D, E\}$ defined in Task 56.
Explain how we know that H is non-cyclic. (Give at least two pieces of evidence.)
What similarities/differences do you notice between H and the groups U_6 and $\mathcal{Z}_6$?

(d) Explain how we know that the symmetric group S_3 is non-cyclic, where

$$S_3 = \{\,(1)(2)(3)\,,\,(1,\,2,\,3)\,,\,(1,\,3,\,2)\,,\,(1,\,2)\,,\,(1,\,3)\,,\,(2,\,3)\,\}\,.$$

What similarities/differences do you notice between S_3 and the groups in (a) – (c)?

Although a group of composite order may or may not be cyclic, Cayley concluded the excerpt from which we are extracting our list of his theorems with the following observations about cyclic groups in general:

> But whether n be prime or composite, the group, assumed to be of the form in question *[that is,* $1,\ \alpha,\ \alpha^2,\ \ldots\alpha^{n-1},\ (\alpha^n = 1)$*]* is in every respect analogous to the system of the roots of the ordinary binomial equation $x^n - 1 = 0$; thus, when n is prime, all the roots (except the root 1) are prime roots; but when n is composite, there are only as many prime roots as there are numbers less than n and prime to it.

Cayley's 'system of the roots of the ordinary binomial equation $x^n - 1 = 0$' is exactly the system of n^{th} roots of unity that we first encountered in our reading of Lagrange. In fact, Lagrange himself explained that this system forms a cyclic group of order n, although not in this exact language. Denoting this set by $U_n = \{e^{\frac{2\pi k}{n}i} \,|\, k = 1, 2, 3, \ldots, n\}$, Lagrange also discussed its generators (or primitive roots) and came to the same conclusions stated by Cayley above. (Cayley referred to the generators as 'prime roots'.) What Cayley has added to Lagrange's insights is the idea that *any* cyclic group of order n will behave in exactly the same manner as U_n. For example, as we have suggested at various points in this project, the additive group Z_6 is essentially the same of the multiplicative group U_6, although with different names and interpretations attached to its elements and operation.

The technical term used today to express the idea that two groups are 'in every respect analogous' is to say that these groups are *isomorphic*. This term comes from the Greek *iso*, meaning 'same,' and *morph*, meaning 'form.' Notice how the following formal definition uses a special sort of bijective function to convey the idea that the underlying group structure (or form) is being maintained, with only the names of the elements changed by the function (e.g., from 'a' to '$f(a)$').

> **Definition** Let $(G, *)$ and $(G', \star)$ be two groups. Then G is a ***isomorphic*** to G' iff there exists a one-to-one onto function $f : G \to G'$ with the following property:
>
> $$\text{For all } a, b \in G,\ f(a * b) = f(a) \star f(b).$$
>
> The function f is called a ***(group) isomorphism***.

The two symbols $*$ and $\star$ are used here to emphasize that the group operation in the domain G need not be the same as the group operation in the range G'. For example, to show that the additive group Z_n is isomorphic to the multiplicative group U_n, we could define $f : Z_n \to U_n$ by $f(k) = e^{\frac{2\pi k}{n} i}$, and then check that f is one-to-one, onto and satisfies $f(k + m) = f(k)f(m)$ for all $k, m \in Z_n$.

Task 68

(a) Use the function f defined above to formally prove $(Z_n, +)$ is isomorphic to $(U_n, \times)$.

(b) Define a function $g : H \to S_3$ which could be used to show that the group of matrices $H = \{I, A, B, C, D, E\}$ defined in Task 56 is isomorphic to the symmetric group S_3. Explain carefully how you decided what each particular function value should be and why these choices will preserve the algebraic structure of H. Is your function the only possible isomorphism between these two groups? Why or why not?

Note: It may be useful to review parts (c) and (d) of Task 67.

To further illustrate the idea that isomorphic groups are really the same groups, but with the names changed to partially disguise this fact, the next task examines some group properties that are preserved by isomorphisms.

Task 69

In this task, we look at some group properties preserved by isomorphisms.

Assume that $(G, *)$ and $(G', \star)$ are isomorphic groups with identities $1, 1'$ respectively.
Let $f : G \to G'$ be a group isomorphism.

(a) Prove that $f(1) = 1'$ by showing that $f(1) \star y = y \star f(1) = y$ for all $y \in G'$.
This shows that identities are preserved by group isomorphisms.

Note: Starting with $y \in G'$ arbitrary, you will need to find some $x \in G$ with $f(x) = y$. Explain carefully how we know that such an element x exists.

(b) Given $x \in G$, prove that $[f(x)]^{-1} = f(x^{-1})$ by showing that $f(x^{-1}) \star f(x) = 1'$.
This shows that inverses are preserved by group isomorphisms.

(c) Given $x \in G$ with finite order r in G, prove that $\text{ord}[f(x)] = r$ in G'
by showing that r is the *least* natural number with $[f(x)]^r = 1'$.
This shows that order is preserved by group isomorphisms.
Hint? Start by using induction to prove that $f(x^k) = [f(x)]^k$ for all $k \in \mathcal{Z}^+$.

(d) Prove that G is abelian if and only if G' is abelian.
This shows that the property of being abelian is preserved by group isomorphisms.

Using the language of isomorphisms, we can restate Cayley's assertion that cyclic groups are 'in every respect analogous to the system of the roots of the ordinary binomial equation $x^n - 1 = 0$' as follows; we also give the modified version of this same which is more likely to be found in current textbooks. The proof of either version proceeds by defining $f(\alpha^k) = \beta^k$ for $0 \leq k \leq n-1$, where α, β are generators of the two groups in question — be sure that you can verify these details if asked!

Cayley's Theorem 5
Let G be a cyclic group of order n, with $G = \langle\alpha\rangle = \{1\,,\,\alpha\,,\,\alpha^1\,,\,\ldots,\alpha^{n-1}\,\}$ for some $\alpha \in G$.
Then G is isomorphic to the group $U_n = \{e^{\frac{2\pi k}{n}i}\,|\,k = 1, 2, 3, \ldots, n\}$ under multiplication.

Cayley's Theorem 5 - A Modern Variation
Every cyclic group of order n is isomorphic to the group Z_n under addition modulo n.

This leaves us with just two more theorems from Cayley's second paragraph, both of which were also discussed by Lagrange for the specific case of the n^{th} roots of unity.

Cayley's Theorem 6
If G is a cyclic group of prime order, then every element of G other than the identity element is a generator.

Cayley's Theorem 7
Let $\phi(n)$ denote the number of natural numbers $k < n$ for which $gcd(k, n) = 1$.
If G is a cyclic group of order n, then G has $\phi(n)$ generators.

The function ϕ in Cayley's Theorem 7 is known as the ***Euler phi function***, and plays an important role in number theory, as well as in algebra. The proof of Cayley's Theorem 7 lies beyond our immediate purposes, but appears in many current algebra textbooks. To illustrate it in the specific case of $n = 6$, first that there are only two natural numbers less than 6 which are relatively prime to 6 (namely, 1 and 5), so that $\phi(6) = 2$. Thus, Cayley's Theorem 7 predicts that $\mathcal{Z}_6$ will have only two generators, a fact that you proved in Task 67(a). Notice also that if n is prime, then every natural number less than n is relatively prime to n, so that $\phi(n) = n - 1$. Thus, one way to prove Cayley's Theorem 6 would be to apply Cayley's Theorem 7 to the particular case where n is prime. Since you have already proved (in Task 66) that $G = \langle\alpha\rangle$ for any *arbitrary* non-identity element $\alpha \in G$ in a group of prime order, however, no further proof of Cayley's Theorem 6 is needed!

Task 70
In this task, we look at Cayley's Theorem 7 in some additional particular cases.

(a) Explain why $\phi(8) = 4$ and identify the four generators of $\mathcal{Z}_8$.

(b) Explain why $\phi(9) = 6$ and identify the five generators of $\mathcal{Z}_9$.

(c) Make a conjecture concerning what the generators will be for $\mathcal{Z}_n$. Test your conjecture for $n = 10$ and $n = 18$ and refine it as needed before writing a formal proof for it.

4 Cayley's Classification of Groups of Small Order

We have now arrived at the heart of Cayley's paper: the classification of groups of finite order according to their form. In fact, by asserting that all groups of prime order have the same form as the cyclic group U_p, Cayley had already accomplished this classification for groups of prime order p. Although we are more likely today to consider the integers mod p under addition (rather than the p^{th} roots of unity under multiplication) as the prototype for cyclic groups of order p, the key idea here is the same:

for any given prime p, there is essentially only one group of order p. As Cayley well knew, however, a group G of order n might not be isomorphic to Z_n when n is composite; in fact, a group of composite order need not even be abelian.

The problem which Cayley therefore began to study at this point of his paper was that of identifying all possible non-isomorphic groups of order n for various composite values of n. If these distinct non-isomorphic group structures could be fully identified, they would in turn serve as prototypes for all groups of that same order, much as $\mathcal{Z}_p$ serves as the prototypical group of order p when p is prime. That is, any group of order n would necessarily be isomorphic to one of the distinct prototypical structures available for that order.

We will see Cayley's classification of all groups of order $n = 4$ and $n = 6$ below; in a later paper, published in 1859, he also gave the classification of groups of order $n = 8$ (see [7]). In time, however, it became clear that the problem of classifying *all* groups of finite order is a very difficult one — so difficult that mathematicians eventually turned to the (presumably more straightforward) task of classifying certain types of finite groups. One particular type of group was considered especially important in this regard, due to the special role it plays within group theory. Namely, much as prime numbers serve as the basic building blocks of all natural numbers, *simple* groups serve as the basic building blocks of all groups. The definition of a 'simple group' is given in Appendix II of this project; its basic concept can be traced back to the work of Galois on algebraic solvability.

The first major results related to the classification of finite simple groups were published in 1870 by the French mathematician Camille Jordan (1838–1922). Yet it was not until 1981, after a coordinated effort by over 100 different mathematicians beginning in 1965, that the search for all finite simple groups was declared complete. Even then, gaps initially remained in the proof — as might well be expected with a proof that stretched out over approximately 500 journal articles totaling 10,000–15,000 pages. Known as the 'Enormous Theorem' in recognition of the complexity of its proof, the Classification of Finite Simple Groups Theorem is now considered by experts in the field to be fully established. The publication of a revised 'second generation proof' expected to fill approximately twelve volumes and 3000–4000 pages is also currently under way by the American Mathematical Society.

4.1 Groups of Order 4

Cayley started his classification efforts, naturally enough, with the first composite number, $n = 4$.

The distinction between the theory of the symbolic equation $\theta^n = 1$, and that of the ordinary equation $x^n - 1 = 0$, presents itself in the very simplest case, $n = 4$. For consider the group

$$1,\ \alpha,\ \beta,\ \gamma,$$

which are a system of roots of the symbolic equation

$$\theta^4 = 1.$$

There is, it is clear, at least one root β, such that $\beta^2 = 1$; we may therefore represent the group thus,

$$1,\ \alpha,\ \beta,\ \alpha\beta,\ (\beta^2 = 1);$$

then multiplying each term by α as the further factor, we have for the group α, α^2, $\alpha\beta$, $\alpha^2\beta$, so that α^2 must be equal either to β or else to 1. In the former case the group is

$$1,\ \alpha,\ \alpha^2,\ \alpha^3\ (\alpha^4 = 1),$$

which is analogous to the system of roots of the ordinary equation $x^4 - 1 = 0$.

For the sake of comparison with what follows, I remark, that, representing the last mentioned group by

$$1,\, \alpha,\, \beta,\, \gamma,$$

we have the table

	1	α	β	γ
1	1	α	β	γ
α	α	β	γ	1
β	β	γ	1	α
γ	γ	1	α	β

If, on the other hand, $\alpha^2 = 1$, then it is easy by similar reasoning to show that we must have $\alpha\beta = \beta\alpha$, so that the group in this case is

$$1,\, \alpha,\, \beta,\, \alpha\beta,\; (\alpha^2 = 1,\, \beta^2 = 1,\, \alpha\beta = \beta\alpha);$$

or if we represent the group by

$$1,\, \alpha,\, \beta,\, \gamma,$$

we have the table[68]

	1	α	β	γ
1	1	α	β	γ
α	α	1	γ	β
β	β	γ	1	α
γ	γ	β	α	1

or, if we please, the symbols are such that

$$\begin{array}{ccccccc} \alpha^2 & = & \beta^2 & = & \gamma^2 & = & 1, \\ \alpha & = & \beta\gamma & = & \gamma\beta, & & \\ \beta & = & \gamma\alpha & = & \alpha\gamma, & & \\ \gamma & = & \alpha\beta & = & \beta\alpha & & \end{array}$$

[and we have thus a group essentially distinct from that of the system of roots of the ordinary equation $x^4 - 1 = 0$].

[68] Notice the double lines that Cayley used to break this table into four smaller squares, with the smaller squares forming two disjoint pairs. One of these pairs corresponds to the subgroup $H = \langle\alpha\rangle = \{1, \alpha\}$. The other corresponds to the set obtained by multiplying the elements of H by β: $\beta H = \{\beta, \beta\alpha\} = \{\beta, \gamma\}$. The set βH is called the left coset of H corresponding to β. See also the comments about cosets in Task 58.

It is easy to see just by looking at their Cayley tables that these two groups are non-isomorphic. Among their distinguishing features, the first is cyclic and has only one element of order 2, while the second is non-cyclic and has three elements of order 2. Each of these characteristics would have been preserved by an isomorphism, if one existed. Of course, Cayley's claim is not just that these are distinct groups of order 4, but that these are the *only* two groups of order 4 possible (up to isomorphism). The analysis given by Cayley in support of this claim began with the assertion "there is, it is clear, at least one root β, such that $\beta^2 = 1$." Given the essential role this fact played in his analysis, let us take some time to carefully prove it, and also to fill in the details of how it implies there are only two distinct groups of order 4. Interestingly, since both these groups are abelian, we can then conclude that every group of order 4 is abelian.

Task 71
This task fills in the details of Cayley's proof that there are exactly two groups of order 4.

Let $G = \{1, \alpha, \beta, \gamma\}$ be a group of order 4.
Thus, the four symbols represent four distinct elements, so that $\alpha \neq 1$, $\beta \neq \gamma$, etc.

(a) Explain why every element of G other than the identity must have order 2 or 4.
If $\text{ord}(\alpha) = 2$, note that Cayley's claim that G contains at least one element of order 2 is clearly true. Suppose therefore that $\text{ord}(\alpha) = 4$, and explain how we know that G must also contain at least one element of order 2 in this case.

In the remainder of this task, we will let β denote an element of order 2 (as Cayley did), since this involves no loss of generality.

(b) Explain why γ must be equal to $\alpha\beta$.
That is, explain how we know that $\alpha\beta \neq 1$, $\alpha\beta \neq \alpha$ and $\alpha\beta \neq \beta$.
Hint? It may be helpful to remember that $\beta^2 = 1$, per part (a).

(c) Since multiplying each element by α will reproduce all the elements of the group simply re-arranged in some different order, we now know that $\{1, \alpha, \beta, \alpha\beta\} = \{\alpha, \alpha^2, \alpha\beta, \alpha^2\beta\}$.

> **Note:** This is simply Cayley's Theorem 1, since the ordered list $(\alpha, \alpha^2, \alpha\beta, \alpha^2\beta)$ is just the row of the Cayley table corresponding to (left) multiplication by α.

In particular, α^2 must be equal to one of the original four elements of the group.
Explain how we know that $\alpha^2 \neq \alpha$ and $\alpha^2 \neq \alpha\beta$.

This leaves only two remaining cases to consider: $\alpha^2 = 1$ and $\alpha^2 = \beta$.

(d) Explain why the assumption $\alpha^2 = \beta$ implies that α has order 4.
This case corresponds to the cyclic group of order 4 in Cayley's first table.

(e) Finally, suppose $\alpha^2 = 1$ and employ 'similar reasoning' to obtain the second group.
That is, multiply through (on the left) by β to reproduce G (Cayley's Theorem 1!), so that $\{1, \alpha, \beta, \alpha\beta\} = \{\beta, \beta\alpha, \beta^2, \beta\alpha\beta\}$.

- Explain why $\beta\alpha$ can not equal 1, α or β, and conclude that $\beta\alpha = \alpha\beta$.
- Use the defining equations obtained for this case ($\beta^2 = 1$, $\alpha^2 = 1$, $\alpha\beta = \beta\alpha$) to explain why $\gamma = \alpha\beta$ also has order 2.

This case corresponds to the non-cyclic group of order 4 in Cayley's second table.

We should remark here again that, although every group of order 4 must have an element of order 2, as you proved in part (a) of the previous task, it is not generally true that a group G will have an element of order d for every divisor d of $|G|$. In fact, Cayley's example of a non-cyclic group of order

4, shows it is possible for a group of order 4 to contain no elements of order 4, despite the fact that 4 is a divisor of the group's order. Similarly, even though 6 is a divisor of $|S_3| = 6$, we know S_3 has no element of order 6. Later in this section, we will also see a group of order 8 which has no elements of order 4 or of order 8. As another counterexample involving a proper divisor of the group's order, recall that the 12-element group A_4 has no subgroup (and hence no element) of order 6, as was noted in an earlier section.[69] Thus, although the order of every element in a finite group must divide the order of the group (Cayley's Theorem 3), the reverse does not generally hold.

There is, however, a partial converse to Cayley's Theorem 3, which we state here without proof:

Cauchy's Theorem
A finite group G contains an element of order p for every prime divisor p of $|G|$.

Despite its name, Cauchy's Theorem was not known in its current form to either Cauchy or Cayley[70] — much as Lagrange's Theorem was not known in its current form to Lagrange — and Cayley could not simply invoke it to conclude that a group of order 4 must have an element of order 2. Rather, Cayley deduced this fact directly from what he knew about the *structure* of 4 element groups. There is, however, an even more direct way to prove that there is only one non-cyclic group of order 4, without using the fact that all groups of order 4 have an element of 2. The following task outlines this proof.

Task 72
In this task, you will give another proof that there is only one non-cyclic group of order 4.

Begin by assuming that $G = \{1, \alpha, \beta, \gamma\}$ is non-cyclic.

First explain why this means that every non-identity element must be of order 2.
This fact gives us the partially completed Cayley table for G shown below.

Apply Cayley's Theorem 1 (i.e., that each row and column contains each group element once and only once) directly to this Cayley table in order to fill in the remaining entries. Keep track of the order in which you fill in each entry, and explain how you know what element of the group belongs there.

Finally, explain how this proves that there is only one non-cyclic group of order 4.

	1	α	β	γ
1	1	α	β	γ
α	α	1		
β	β		1	
γ	γ			1

The non-cyclic group of order 4 which Cayley identified as the second of only two possible 4-element groups later became known as the ***Klein four-group***, after the German mathematician Felix Klein

[69] See footnote 65, page 68.

[70] Cauchy published a proof for permutation groups in 1844 about which Cayley later commented "the whole course of the investigation is peculiar to [permutations], and I see no way of extending it to [groups] in general — otherwise it would be the very theorem which I mentioned . . . I was in want of . . ." (as quoted in [9, p. 248].)

(1849–1925), or as the ***Viergruppe***, where 'Vier' is German for 'four'. This group can be visualized as the symmetry group of a rectangle, where the four elements represent the identity, the rotation through 180°, the vertical reflection, and the horizontal reflection.[71] Cayley's own insights into its structure are especially impressive in light of the fact that he arrived at it by purely formal means at a time when permutations were the only objects that had been explicitly studied as groups by anyone. Not only did Cayley completely describe the algebraic structure of the Viergruppe for the first time ever, he also identified several specific mathematical phenomena other than permutations which exhibit this structure. We examine two of his examples in the following excerpt.

Systems of this form are of frequent occurrence in analysis, and it is only on account of their extreme simplicity that they have not been expressly remarked. For instance, in the theory of elliptic functions, if n be the parameter, and

$$\alpha(n) = \frac{c^2}{n} \quad ; \quad \beta(n) = -\frac{c^2+n}{1+n} \quad ; \quad \gamma(n) = -\frac{c^2(1+n)}{c^2+n},$$

then α, β, γ form a group of the species in question. ...

Again, in the theory of matrices, if I denote the operation of inversion, and tr that of transposition, ... we may write

$$\alpha = I,\ \beta = \text{tr},\ \gamma = I.\text{tr} = \text{tr}.I.$$

Task 73
Let $G = \{\,1,\ \alpha,\ \beta,\ \gamma\,\}$, where 1 is the identity function defined on $\mathcal{R}$ by $1(n) = n$, and α, β, γ are the real-valued functions defined for $n \in \mathcal{R}$ as follows:

$$\alpha(n) = \frac{c^2}{n} \quad ; \quad \beta(n) = -\frac{c^2+n}{1+n} \quad ; \quad \gamma(n) = -\frac{c^2(1+n)}{c^2+n}$$

Using multiplicative notation to denote function composition, verify that G is the Viergruppe by showing that $\alpha^2 = \beta^2 = \gamma^2 = 1$ and $\alpha\beta = \beta\alpha$.

Task 74
Let $M_{n,n}$ denote the set of all $n \times n$ invertible matrices, where $n \in \mathcal{Z}^+$.
Given $X \in M_{n,n}$, let X^T denote the transpose of X.
Let $G = \{\,1,\ \alpha,\ \beta,\ \gamma\,\}$, where 1 is the identity function defined on $M_{n,n}$ by $1(X) = X$ and the functions α, β, γ are defined for $X \in M_{n,n}$ as follows:

$$\alpha(X) = X^{-1} \quad ; \quad \beta(X) = X^T \quad ; \quad \gamma(X) = (\alpha\beta)(X) = \alpha(\beta(X)) = [X^T]^{-1}$$

Again use multiplicative notation to denote function composition, and let $X \in M_{n,n}$.
Then the properties of matrix algebra imply that:

$$(\beta\alpha)(X) = \beta(\alpha(X)) = \beta(X^{-1}) = [X^{-1}]^T = [X^T]^{-1} = \gamma(X)$$

Complete the verification that G is the Viergruppe by showing that $\alpha^2 = \beta^2 = \gamma^2 = 1$.

[71] You may find it interesting to explore what this means further, and to consider groups defined by the symmetries of other geometrical figures, such as a square or an equilateral triangle. (We will see the former of these in Task 83, and have already encountered the latter in a different guise — can you identify it?)

Task 75
This task examines the subgroup structure of the alternating group A_n for $n \geq 4$, thereby providing another specific instance of the Viergruppe.

Recall from Task 55 that A_n is the subgroup of S_n consisting of all $\frac{n!}{2}$ even permutations. Also review the statement of Cauchy's Theorem on page 75.

(a) Explain how Cauchy's Theorem implies A_n contains a cyclic subgroup of order 2. Then explain why the generator of this subgroup can not be a 2-cycle. Finally, give a specific example of such a subgroup in A_4 by exhibiting its generator.

(b) Explain how Cauchy's Theorem implies A_n contains a cyclic subgroup of order 3. Then explain why every 3-cycle in S_n generates such a subgroup. *(Begin by showing that every* 3*-cycle is an element of* A_n*.)* Then give an example of a cyclic subgroup of order 3 in A_6 for which the generator is *not* a 3-cycle.

(c) As noted previously, A_4 contains a subgroup of order 4, but no element of order 4.

(i) Provide convincing evidence that A_4 does not contain an element of order 4. **Note:** One way to do this is to find all elements of A_4 and their orders; another is to find all elements of order 4 in S_4 and show they are odd. You may find some other approach that you prefer to these two.

(ii) Assuming that a 4-element subgroup of A_4 exists, explain why it must be isomorphic to the Viergruppe. Then find a specific subgroup $H = \{1\,,\, \alpha\,,\ \ \beta\,,\, \gamma\}$ of A_4 and verify that it is the Viergruppe by showing that $\alpha^2 = \beta^2 = \gamma^2 = 1$.

(d) Now assume $n > 4$. Show that A_n contains a subgroup of order 4 which is isomorphic to the Viergruppe by explicitly finding such a subgroup. For which values of n will A_n also contain a cyclic subgroup of order 4? Justify.

4.2 Groups of Order 6

We now look at Cayley's analysis of groups of order 6, the next greatest composite number. We consider this analysis in three separate excerpts in order to assimilate it more completely. The first of these gives Cayley's argument that every group of order 6 must have an element of order 3, or as Cayley phrased it, an element of index 3. Cayley again based this argument directly on the structure of the group, examining how elements of a group of order 6 can (and can not) interact with each other. Although he could not simply invoke Cauchy's Theorem, he did know (by Cayley's Theorem 3) that the only possible orders for elements in a group of order 6 are $1, 2, 3$ and 6. As we will see, Lagrange's Theorem concerning the order of subgroups also played a central role in his argument. Although Cayley did not know this theorem by that name — nor did he have a separate term for the concept of a subgroup — it was necessary for him to somehow refer to subgroups in order to incorporate this theorem into his argument. Notice how the language he used to do this reinforces the idea that a subgroup is a group which happens to be part of another group.

I proceed to the case of a group of six symbols,

$$1,\ \alpha,\ \beta,\ \gamma,\ \delta,\ \epsilon,$$

which may be considered as representing a system of roots of the symbolic equation

$$\theta^6 = 1.$$

It is in the first place to be shown that there is at least one root which is a prime root of $\theta^3 = 1$, or (to use a simpler expression) a root having the index 3. It is clear that if there were a prime root, or root having the index 6, the square of this root would have the index 3. It is therefore only necessary to show that it is impossible that *all* roots should have index 2. This may be done by means of a theorem which I shall for the present assume, viz. that if among the roots of the symbolic equation $\theta^n = 1$, there are contained a system of roots of the symbolic equation (or, in other words, if among the symbols forming a group of the order there are contained symbols forming a group of the order p), the p is a submultiple of n. In the particular case in question, a group of the order 4 cannot form part of the group of the order 6. Suppose, then, that γ, δ are two roots of $\theta^6 = 1$, having each of them the index 2; then if $\gamma\delta$ had also the index 2, we should have $\gamma\delta = \delta\gamma$; and $1, \gamma, \delta, \delta\gamma$, which is part of the group of the order 6, would be a group of the order 4. It is easy to see that $\gamma\delta$ must have the index 3, and that the group is, in fact, $1, \gamma\delta, \delta\gamma, \gamma, \delta, \gamma\delta\gamma$, which is, in fact, one of the groups to be presently obtained;

Let us stop here to fill in the details of Cayley's argument (by contradiction) that a group G of order 6 must contain an element of order 3. Note that he first took care of the case where G is cyclic by noting that $\text{ord}(\alpha^2) = 3$ for any generator α of G. Thus, he could assume that G is not cyclic, so that all elements are of order 1, 2, or 3 by virtue of Cayley's Theorem 3. To set up the contradiction, Cayley assumed that every element (other than the identity) has order 2. The crux of his argument was then to show that this leads to a subgroup of order 4, which is impossible by Lagrange's Theorem.

Task 76
This task examines the details of Cayley's argument that a group order 6 contains an element of order 3.

(a) Assume in this part that G is a group of order 6 and that $\gamma, \delta \in G$ are such that the three elements γ, δ and $\gamma\delta$ are of order 2.

- **(i)** Explain why this means that $\gamma^{-1} = \gamma$, that $\delta^{-1} = \delta$ and that $(\gamma\delta)^{-1} = \gamma\delta$.
- **(ii)** Use the facts established in part (a) to prove Cayley's claim that $\gamma\delta = \delta\gamma$
- **(ii)** Explain why $H = \{1\,,\ \gamma\,,\ \delta\,,\ \gamma\delta\}$ must therefore be a subgroup of G.
  That is, how do we know that H is closed under products and inverses?

(b) Use the result of part (a) to write a proof (by contradiction) that it is impossible for every non-identity element to have order 2 in a group G of order 6.

After proving that it is not possible for the three elements γ, δ and $\gamma\delta$ to all have order 2, Cayley remarked in the preceding excerpt that it is 'easy to see' that the assumption that $\text{ord}(\gamma) = \text{ord}(\delta) = 2$ leads to two conclusions. The first conclusion, that $\text{ord}(\gamma\delta) = 3$, is indeed easy to see; the second, that $G = \{1, \gamma\delta, \delta\gamma, \gamma, \delta, \gamma\delta\gamma\}$, is somewhat less obvious. We will not look at the details of how to prove this claim, although you should feel free to do so on your own! Instead, we continue to follow Cayley's analysis of groups of order 6 starting from the now-established fact that all such groups have an element of order 3. Our next excerpt begins with a statement of Cayley's own preference for continuing the analysis in this way. He then invoked the now familiar strategy of multiplying the group by a particular element in order to reproduce the entire group.

I prefer commencing with the assumption of a root having the index 3. Suppose that α is such a root, the group must clearly be of the form

$$1,\, \alpha,\, \alpha^2,\, \gamma,\, \alpha\gamma,\, \alpha^2\gamma,\ (\alpha^3 = 1);$$

and multiplying the entire group by γ as nearer factor, it becomes γ, $\alpha\gamma$, $\alpha^2\gamma$, γ^2, $\alpha\gamma^2$, $\alpha^2\gamma^2$; we must therefore have $\gamma^2 = 1$, α or α^2. But the supposition $\gamma^2 = \alpha^2$ gives $\gamma^4 = \alpha^4 = \alpha$, and the group is in this case $1, \gamma, \gamma^2, \gamma^3, \gamma^4, \gamma^5, (\gamma^6 = 1)$; and the supposition $\gamma^2 = \alpha$ gives also the same group.

Cayley's argument in this excerpt should seem reasonably straightforward. Starting from the previously justified assumption $\text{ord}(\alpha) = 3$, the three elements in the subgroup $H = \langle\alpha\rangle = \{\,1,\, \alpha,\, \alpha^2\,\}$ are certainly distinct. Taking γ to be some element of G which in not in H, you should also able to explain why the elements in the set $H\gamma = \{\gamma\,,\, \alpha\gamma\,,\, \alpha^2\gamma\}$ are both mutually distinct, and distinct from the elements in H.[72] (Pause for a moment to be sure you can do this!) Then since $H \cup H\gamma$ gives us six distinct elements — so that $G = \{1\,,\, \alpha\,,\, \alpha^2\,,\, \gamma\,,\, \alpha\gamma\,,\, \alpha^2\gamma\,\}$ — it is not hard to say why γ^2 must be either 1, α or α^2, or equivalently, why $\gamma^2 \neq \gamma$, $\gamma^2 \neq \gamma\alpha$, and $\gamma^2 \neq \gamma^2\alpha^2$. This leaves us with only Cayley's final claim in the excerpt to explain: why do two of these cases ($\gamma^2 = \alpha^2$ and $\gamma^2 = \alpha$) imply that G is the cyclic group generated by γ? In both these cases, however, simply taking powers of γ readily shows that all six elements of G belong to $\langle\gamma\rangle$. In the case of $\gamma^2 = \alpha^2$, for example, we get:

$$\gamma^1 = \gamma\ \ ;\ \ \gamma^2 = \alpha^2\ \ ;\ \ \gamma^3 = \gamma^2\gamma = \alpha^2\gamma\ \ ;\ \ \gamma^4 = (\gamma^2)^2 = \alpha^4 = \alpha\ \ ;\ \ \gamma^5 = \gamma^4\gamma = \alpha\gamma\ \ ;\ \ \gamma^6 = (\gamma^2)^3 = \alpha^6 = 1$$

Task 77

This task completes the details above, by showing that G is cyclic when $\gamma^2 = \alpha$.

Assume $G = \{1\,,\, \alpha\,,\, \alpha^2\,,\, \gamma\,,\, \alpha\gamma\,,\, \alpha^2\gamma\,\}$ is a group of order 6 with $\alpha^3 = 1$ and $\gamma^2 = \alpha$.
By taking powers of γ, show that $G = \langle\gamma\rangle$.

Cayley now had only one case to consider: $\gamma^2 = 1$. We see in the next excerpt how this case leads to two subcases. Remember that to arrive at this point in his analysis, Cayley used the following fact:

$$G = \{1\,,\, \alpha\,,\, \alpha^2\,,\, \gamma\,,\, \alpha\gamma\,,\, \alpha^2\gamma\,\} = \{\,\gamma\,,\, \alpha\gamma\,,\, \alpha^2\gamma\,,\, \gamma^2\,,\, \alpha\gamma^2\,,\, \alpha^2\gamma^2\}$$

By multiplying the entire group G by γ as a further factors (i.e., on the left), Cayley also had the following fact available to him:

$$G = \{\,\gamma\,,\, \gamma\alpha\,,\, \gamma\alpha^2\,,\, \gamma^2\,,\, \gamma^2\alpha\,,\, \gamma^2\alpha^2\}$$

Again, these six elements are just another arrangement of G's elements; as such, they are mutually distinct. Thus, in assuming that $\gamma^2 = 1$ (the final remaining case), Cayley also assumed that no other element in the list ($\gamma\,,\, \gamma\alpha\,,\, \gamma\alpha^2\,,\, \gamma^2\,,\, \gamma^2\alpha\,,\, \gamma^2\alpha^2$) is equal to 1. In particular, it follows that $\gamma\alpha \neq 1$.[73]

Concerning the possible values for $\gamma\alpha$ among G's five non-identity elements ($\alpha\,,\, \alpha^2\,,\, \gamma\,,\, \alpha\gamma\,,\, \alpha^2\gamma$), our next excerpt begins with Cayley's assertion that there are actually only two possibilities. As was

[72] Recall from Task 58 that sets of the form $H\gamma = \{x\gamma\,|\, x \in H\}$ are called the (right) cosets of H. Note that in this case, the coset $H\gamma = \{\gamma\,,\, \alpha\gamma\,,\, \alpha^2\gamma\}$ is not a subgroup of G — can you see why? We will say more about this and cosets more generally later in this section.

[73] Another way to show that $\gamma\alpha \neq 1$ would be to note that $\gamma^2 = 1$ implies $\gamma^{-1} = \gamma$. Thus, $\gamma\alpha = 1$ would imply $\alpha = \gamma^{-1} = \gamma$, contrary to the fact that α, γ are distinct elements.

often the case, he did so without providing explicit reasons for ruling out the other three possibilities. However, you will have little difficulty by now convincing yourself (and others) that $\gamma\alpha \neq \alpha$, $\gamma\alpha \neq \alpha^2$, and $\gamma\alpha \neq \gamma$, given that α and γ are distinct non-identity elements of G. Following the excerpt, we will look at the details concerning why a cyclic group results from the first subcase discussed, and use these to introduce some ideas related to groups in general.

It only remains, therefore, to assume $\gamma^2 = 1$; then we must have either $\gamma\alpha = \alpha\gamma$ or else $\gamma\alpha = \alpha^2\gamma$. The former assumption leads to the group

$$1,\, \alpha,\, \alpha^2,\, \gamma,\, \alpha\gamma,\, \alpha^2\gamma,\ (\alpha^3 = 1, \gamma^2 = 1, \gamma\alpha = \alpha\gamma),$$

which is, in fact, analogous to the system of roots of the ordinary equation $x^6 - 1 = 0$; and by putting $\alpha\gamma = \lambda$, might be exhibited in the form $1, \lambda, \lambda^2, \lambda^3, \lambda^4, \lambda^5, (\lambda^6 = 1)$, under which this system has previously been considered. The latter assumption leads to the group

$$1,\, \alpha,\, \alpha^2,\, \gamma,\, \alpha\gamma,\, \alpha^2\gamma,\ (\alpha^3 = 1, \gamma^2 = 1, \gamma\alpha = \alpha^2\gamma),$$

and we have thus two, and only two essentially distinct forms of a group of six.

The next task fills in some proof details for Cayley's claims in the preceding excerpt. Tasks 79 and 80 then extend some of the ideas raised in this proof to more general group settings.

Task 78

This task considers Cayley's claim that G is cyclic in the case where $\gamma\alpha = \alpha\gamma$.

Assume (with Cayley) that $G = \{1\,,\,\alpha\,,\,\alpha^2\,,\,\gamma\,,\,\alpha\gamma\,,\,\alpha^2\gamma\}$ is a group of order 6 with $\text{ord}(\alpha) = 3$ and $\text{ord}(\gamma) = 2$. Further assume that $\gamma\alpha = \alpha\gamma$ and let $\lambda = \alpha\gamma$.

Recall from Task 50(e) that, since $\gamma\alpha = \alpha\gamma$, we know $(\alpha\gamma)^k = \alpha^k\gamma^k$ for all $k \in \mathcal{Z}$.

Note: Although we are only assuming that γ and α commute here, you should be able to convince yourself from the way in which G is defined in terms of these two generators that G is, in fact, abelian. This fact will also immediately follow once we prove G is cyclic.

(a) Calculate powers of λ to show that $\text{ord}(\lambda) = 6$.
For example: $\lambda^2 = (\alpha\gamma)^2 = \alpha^2\gamma^2 = (\alpha^2)(1) = \alpha^2$.

(b) Explain why the results of part (a) imply that G is cyclic with generator λ.

Task 79

This task generalizes ideas from the example in Task 78 to obtain general results about the order of group elements.

(a) Give an example of a finite non-abelian group G containing elements γ and α which commute and are of order 2 and 3 respectively. (*Hint?* Think permutations!)

Verify that $\text{ord}(\alpha\gamma) = 6$ in this case.

Note that G is not cyclic, since it is non-abelian.

Why does this not contradict the result of Task 78?

In the remainder of this task, let G be an arbitrary group with identity 1. Also let $\alpha, \gamma \in G$ and assume that α, γ commute.

Note: The group G itself may or may not be abelian, as shown in part (a). We also can *not* assume that α, γ generate G, as was the case in Task 78. By Task 50(e), however, it is still the case that $(\alpha\gamma)^k = \alpha^k\gamma^k$ for all $k \in \mathcal{Z}$.

(b) Suppose that $\text{ord}(\alpha) = 2$ and $\text{ord}(\gamma) = 3$.
Show that $\alpha\gamma \neq 1$ by deriving a contradiction from the assumption $\alpha\gamma = 1$.
Then explain carefully how we know that $(\alpha\gamma)^k \neq 1$ for $k = 2, 3, 4, 5$.
Complete the verification that $\text{ord}(\alpha\gamma) = 6$ by showing that $(\alpha\gamma)^6 = 1$.

(c) Suppose that $\text{ord}(\alpha) = 3$ and $\text{ord}(\gamma) = 4$. Show carefully that $\text{ord}(\alpha\gamma) = 12$.
In particular, indicate clearly how we know that $(\alpha\gamma)^k \neq 1$ for $1 \leq k \leq 11$.

(d) Suppose that $\text{ord}(\alpha) = m$ and $\text{ord}(\gamma) = n$, where $m, n > 1$.
State a conjecture concerning $\text{ord}(\alpha\gamma)$.
Test your conjecture in the case where $m = 2$ and $n = 6$, and modify it as needed.

(e) Consider the elements $\alpha = (2,\ 3,\ 4)$ and $\gamma = (2,\ 4,\ 3)$ from the symmetric group S_4.
Show α, γ commute (even though S_4 is non-abelian) and determine $\text{ord}(\alpha\gamma)$.
Use the result to modify your conjecture from part (d) as needed.
State your final conjecture carefully, and provide a complete proof.

Task 80

This task introduces a general group concept suggested by the preceding task.

Let G, H be groups with identities $1_G, 1_H$ respectively.

Consider the set $G \times H = \{(\alpha, \gamma) \,|\, \alpha \in G, \gamma \in H\}$
under componentwise multiplication, so that $(\alpha, \gamma)(\beta, \delta) = (\alpha\beta, \gamma\delta)$.

Then $G \times H$ is a group known as the **direct product** of G, H.

(a) Verify that the identity of $G \times H$ is the ordered pair $(1_G\ ,\ 1_H)$.

(b) Given $(\alpha, \gamma) \in G \times H$, what is $(\alpha, \gamma)^{-1}$? Explain.

(c) Prove that the direct product of two abelian groups is necessarily abelian.

(d) Write out all six elements of $\mathcal{Z}_2 \times \mathcal{Z}_3$ and show $\mathcal{Z}_2 \times \mathcal{Z}_3$ is cyclic by finding a generator.

(e) For arbitrary groups G, H and $(\alpha\ ,\ \gamma) \in G \times H$, prove that $\text{ord}(\,(\alpha\ ,\ \gamma)\,) = \text{lcm}(\,\text{ord}(\alpha), \text{ord}(\gamma)\,)$.
Compare this to your final conjecture in Task 79(d), and comment on why the relationship between $\text{ord}(\,(\alpha\ ,\ \gamma)\,)$ and $\text{lcm}(\,\text{ord}(\alpha), \text{ord}(\gamma)\,)$ is different in these two cases.

(f) Give an example to show the direct product of two cyclic groups need not be cyclic.
Hint? Look at examples of the form $\mathcal{Z}_m \times \mathcal{Z}_n$, with part (e) of this task in mind.

Although Cayley's analysis of groups of order 6 was complete once he concluded that there are only two distinct groups of order 6, he went on in his paper to display the Cayley tables for these two groups. The first you will recognize as $\mathcal{Z}_6$ — in disguise, of course. Or, if you prefer, this group could be described as $\mathcal{Z}_2 \times \mathcal{Z}_3$ (again in disguise), in keeping with Cayley's emphasis on the ***generating elements*** α, γ, of order 2 and 3 respectively. The equations which Cayley used to describe this group ($\alpha^3 = 1, \gamma^2 = 1, \gamma\alpha = \alpha\gamma$) are referred to as the ***generating equations*** of the group. The generating equations for the second possible group of order 6 ($\alpha^3 = 1, \gamma^2 = 1, \gamma\alpha = \alpha^2\gamma$) lead to the second table in the next excerpt.[74] In the closing task of this section, we provide some practice in working with groups defined by generating equations.

If we represent the first of these two forms, viz. the group

$$1,\, \alpha,\, \alpha^2,\, \gamma,\, \alpha\gamma,\, \alpha^2\gamma,\ (\alpha^3 = 1, \gamma^2 = 1, \gamma\alpha = \alpha\gamma),$$

by the general symbols

$$1,\, \alpha,\, \beta,\, \gamma,\, \delta,\, \epsilon,$$

we have the table

	1	α	β	γ	δ	ϵ
1	1	α	β	γ	δ	ϵ
α	α	β	γ	δ	ϵ	1
β	β	γ	δ	ϵ	1	α
γ	γ	δ	ϵ	1	α	β
δ	δ	ϵ	1	α	β	γ
ϵ	ϵ	1	α	β	γ	δ

while if we represent the second of these two forms, viz. the group

$$1,\, \alpha,\, \alpha^2,\, \gamma,\, \alpha\gamma,\, \alpha^2\gamma,\ (\alpha^3 = 1, \gamma^2 = 1, \gamma\alpha = \alpha^2\gamma),$$

by the general symbols

$$1,\, \alpha,\, \beta,\, \gamma,\, \delta,\, \epsilon,$$

we have the table

[74] As you examine its table, see if you can recognize this second group as isomorphic to S_3, a fact made explicit by Cayley towards the end of the excerpt.

	1	α	β	γ	δ	ϵ
1	1	α	β	γ	δ	ϵ
α	α	β	1	δ	ϵ	γ
β	β	1	α	ϵ	γ	δ
γ	γ	ϵ	δ	1	β	α
δ	δ	γ	ϵ	α	1	β
ϵ	ϵ	δ	γ	β	α	1

An instance of a group of this kind is given by the permutation of three letters; the group

$$1,\ \alpha,\ \beta,\ \gamma,\ \delta,\ \epsilon,$$

may represent a group of substitutions[75] as follows:

$$\begin{array}{cccccc} abc, & abc, & abc, & abc, & abc, & abc, \\ abc & cab & bca & acb & cba & bac \end{array}$$

Another singular instance is given by the optical theorem proved in my paper "On a property of the Caustic[76] by refraction of a Circle, ..."

Task 81

In this task, we translate the permutations of three letters defined by Cayley in the last excerpt into the notation of S_3.

For example, letting $a = 1$, $b = 2$ and $c = 3$, the permutation α which Cayley denoted simply by the two rows $\begin{array}{c} abc \\ cab \end{array}$ corresponds to the cycle $\alpha = \begin{pmatrix} 123 \\ 312 \end{pmatrix} = (1, 3, 2)$.

(a) Write the permutations denoted by Cayley as β, γ, δ and ϵ as cycles in S_3.

(b) Use cycle multiplication to verify the row corresponding to α in the Cayley table. That is, verify that $\alpha^2 = \beta$, $\alpha\beta = 1$, etc. using the cycles of S_3 denoted by 1, α, β, γ, δ, ϵ.

[75] Cayley's actual notation for substitutions (i.e., permutations) followed that of Cauchy (but without the parentheses), and used the bottom row for the original arrangement in the bottom row, and the top row for the arrangement being substituted in its place; we have again reversed this to be consistent with current notation.

[76] A caustic is a curve related to the reflection (or refraction) of light off a surface in the study of optics. Cayley did not say more about this example in his 1854 paper on group theory paper, but did describe it in considerable detail in his paper "On a property of the Caustic by refraction of a Circle." Interestingly, these two papers were submitted for publication in the journal *Philosophical Magazine* on the same day, and some scholars [8] contend that it was Cayley's discovery of this second concrete example of a non-abelian group of order 6 which inspired him to generalize the abstract group concept from that of a permutation group.

Task 82
This task explores some properties of cosets within the specific example of the group S_3.

We assume the elements of S_3 are denoted as in Task 81; for example, $\alpha = (1,\,3,\,2)$. Thus, the Cayley table given in the last excerpt can be used to determine all necessary products.

Recall from Task 58 that the left coset of H corresponding to γ is the set defined by $\gamma H = \{\,\gamma x \,|\, x \in H\}$, where H is a subgroup of G and γ an element of G.

(a) Let $H = \langle \alpha \rangle$.

(i) Use the Cayley table for S_3 to verify that $H = \{1\,,\,\alpha\,,\,\beta\,\}$ and $\gamma H = \{\gamma\,,\,\epsilon\,,\,\delta\,\}$.
Describe where H, γH show up in the Cayley table of S_3.
Also explain why γH is not a subgroup of G.

(ii) Find and compare the left cosets $1H = H$, αH, and βH.
Also find and compare the left cosets γH, δH, and ϵH.
Comment on what you observe.

(iii) Now consider the right coset $H\gamma = \{\,x\gamma \,|\, x \in H\}$.
Show that $\gamma H = H\gamma$ in this particular example.
Do you think it will always be the case that $\gamma H = H\gamma$. Why or why not?

(b) Now consider the subgroup $K = \langle \gamma \rangle = \{\,1\,,\,\gamma\,\}$ in S_3.
Use the Cayley table of S_3 to find all left and right cosets of K and compare these.
Would you now change your answer to the question of whether the left and right cosets corresponding to an element are always equal? Why or why not?

Task 83
In this task, we look at some groups of order 8, starting from their generating equations.

(a) Write the Cayley table for the group $G = \{\,1\,,\,a\,,\,b\,,\,b^2\,,\,b^3\,,\,ab\,,\,ab^2\,,\,ab^3\,\}$
with generating equations $a^2 = 1\;,\; b^4 = 1\;,\; ba = ab^3$.
This group is known as the ***dihedral group*** D_4.[77]
In general, the dihedral group D_n for $n \in \mathcal{Z}^+$ is a group of order $2n$
defined by two generators a, b satisfying $a^n = 1\;,\; b^2 = 1$ and $ba = ab^{n-1}$.

(b) Write the Cayley table for the group $G = \{\,1\,,\,a\,,\,b\,,\,b^2\,,\,b^3\,,\,ab\,,\,ab^2\,,\,ab^3\,\}$
with generating equations $a^2 = b^2\;,\; b^4 = 1\;,\; ba = ab^3$.
Explain why this group is isomorphic to the quaternion group (see Task 44).

(c) Write the Cayley table for the group $G = \{\,1\,,\,a\,,\,b\,,\,c\,,\,ab\,,\,ac\,,\,bc\,,\,abc\,\}$
with generating equations $a^2 = b^2 = c^2 = 1$, $ba = ab\;,\; ca = ac$.
Explain why this group[78] is isomorphic to $\mathcal{Z}_2 \times \mathcal{Z}_2 \times \mathcal{Z}_2$.

(d) Find two other groups of order 8, and explain how we know these groups are distinct from those in parts (a), (b) and (c). *Hint?* Try using direct products.

EXTRA CREDIT:
Are these five groups the only groups of order 8 possible, up to isomorphism?
If so, prove this. If not, find all possible groups of order 8, up to isomorphism.

[77] D_4 can be interpreted as the group of symmetries of a square, in a fashion similar to that in which the Viergruppe represents the group of symmetries of a rectangle, by letting a represent a reflection across one of the two diagonals and b represent a rotation through 90°. Similarly, D_{2n} can be interpreted as the symmetries of a regular polygon with $2n$ sides.

[78] Note that this group contains no elements of order 4 or 8, even though both 4 and 8 divide the order of the group.

5 Concluding Remarks, and Cayley's Theorem Today

With the classification of all groups of order 6, our study of Cayley's ground-breaking paper on groups comes to an end. Cayley himself went on in this paper to describe two non-cyclic groups of higher order — one of order 18 and the other of order 27. Although he apparently never fully classified groups of these orders, his complete classification of all groups of order 8 was published five years later, in 1859.[79] Cayley did not return to the study of groups until 1878, when he published another four papers in group theory. In one of these, he explicitly stated the one theorem named after him in group theory today:

> **Cayley's Theorem:** Every group is isomorphic to a group of permutations.

This theorem tells us that, in a sense, everything there is to know about groups in general is already true of permutation groups in particular. Although it may not surprise you to learn that Cayley stated this theorem without proof, you probably also won't find it very difficult to believe the theorem is true. In essence, each row (or column) of a Cayley table can be used as a (somewhat long) label for the element it represents; that is, the Cayley table gives us a 1-to-1 correspondence between the elements α of a group and the ordered list obtained by multiplying each element of the group by α. But this list of elements (i.e., the row) is really nothing more than a permutation of the elements of the group! Our concluding task outlines a formal proof of this theorem, as it would appear in a current textbook.

Task 84
In this task, we outline a formal proof of Cayley's Theorem.

Let G be a group. We must find a permutation group with the same group structure as G.

To do this, define a function $\pi_\alpha : G \to G$ for each $\alpha \in G$ by setting $\pi_\alpha(x) = \alpha x$.
(This should remind you of how we get rows in the Cayley table!)

Let $G^* = \{\, \pi_\alpha \,|\, \alpha \in G \,\}$. Our plan is to show that G^* is the desired permutation group.

(a) We first must show that the elements of G^* are permutations.
To do this, let $\alpha \in G$ and prove that π_α is one-to-one and onto.

(b) We next prove that G^* is a group under composition, by completing the following:

- **(i)** For arbitrary $\alpha, \beta \in G$, show that $\pi_\alpha \pi_\beta = \pi_{\alpha\beta}$.
(Do this by showing $\pi_\alpha \circ \pi_\beta(x) = \pi_{\alpha\beta}(x)$ for an arbitrary $x \in G$.)
Explain why this proves that G^* is closed under products.
- **(ii)** For arbitrary $\alpha \in G$, show that $\pi_\alpha \pi_1 = \pi_1 \pi_\alpha = \pi_\alpha$.
Explain why this proves that π_1 is an identity for G^*.
- **(iii)** For arbitrary $\alpha \in G$, show that $\pi_\alpha \pi_{\alpha^{-1}} = \pi_1$.
Explain why this proves that $\pi_{\alpha^{-1}}$ is the (group) inverse of π_α.

(c) Finally, prove that G is isomorphic to G^* by showing that the function $\Phi : G \to G^*$ defined by $\Phi(\alpha) = \pi_\alpha$ is an isomorphism. That is, prove that Φ is one-to-one, onto and satisfies $\Phi(\alpha)\Phi(\beta) = \Phi(\alpha\beta)$ for all $\alpha, \beta \in G$.

By the late nineteenth century, when the paper in which Cayley stated the theorem which now bears his name appeared in print, the time was mathematically ripe for mathematicians in general to fully recognize how the concept of group served as a powerful tool in geometry, number theory and the theory of equations. The rest, one might say, is history!

[79] His classification appears in [7], if you'd like to check your answer to the extra credit question in Task 83!

APPENDIX I
Optional exercises on de Moivre's and Euler's Formulas

The following tasks explore proofs of the following famous formulas, discussed in Subsection 1.1:

de Moivre's Formula: $(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)$

Euler's formula: $e^{i\theta} = \cos\theta + i\sin\theta$

Task I.1

Let $x = \cos\left(\frac{2\pi k}{m}\right) + i\sin\left(\frac{2\pi k}{m}\right)$ where $k \in \{1, 2, 3, \ldots, m\}$.
Use de Moivre's formula to show that $x^m = 1$.

Task I.2

This exercise outlines a proof of de Moivre's formula for natural numbers n, based only on trigonometric sum identities and induction.

RECALL: $\cos(x + y) = \cos x\cos y - \sin x\sin y \qquad \sin(x + y) = \cos x\sin y + \cos y\sin x$

(a) Prove de Moivre's formula holds for $n = 2$ by expanding $(\cos\theta + i\sin\theta)^2$ and applying the sum identities with $x = y = \theta$.

(b) Use the result of part (a) to expand $(\cos\theta + i\sin\theta)^3$; then use the sum identities with $x = \theta$ and $y = 2\theta$ to prove de Moivre's formula holds for $n = 3$.

(c) Use the sum identities and induction on n to prove de Moivre's formula in the case where n is a natural number.

Task I.3

Use the power series for e^x, $\sin x$ and $\cos x$ given below to prove Euler's formula.[80]
Recall that $i^2 = -1$, $i^3 = -i$ and $i^4 = 1$.

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!}x^k \qquad \sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}x^{2k+1} \qquad \cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}x^{2k}$$

Note: Recall that all three series are absolutely convergent on $\mathcal{R}$; thus, we can legitimately rearrange their terms, and do not change the values to which they converge by doing so.

[80] Euler's derivation of this formula used a somewhat different approach in which infinitesimals appeared.

APPENDIX II
Cosets, Lagrange's Theorem and Factor Groups

In the latter part of this project, the concept of a *coset* came up in passing at several points. In this appendix, we outline a proof of Lagrange's Theorem for Finite Groups cast in the language of cosets, and introduce the notion of a *factor group* (or *quotient group*) which arises by imposing a binary operation on the collection of all cosets. We begin by formally stating the definition of a coset.

Definition
Let G be a group, H a subgroup of G, and $a \in G$.
The ***left coset of*** H corresponding to a is the set $aH = \{\, ah \,|\, h \in H \,\}$.
The ***right coset of*** H corresponding to a is the set $Ha = \{\, ha \,|\, h \in H \,\}$.

In other words, (left) cosets are formed by collecting together all the products ah, where h ranges over the entire subgroup H and a is fixed. If you think back to the table that Cauchy used to prove Lagrange's Theorem for S_n, you will see that the first row of that table was the given subgroup H (itself the coset corresponding to $a = 1$), and that the rows below it were cosets of H that corresponded to permutations selected from outside of H. In his proof of Lagrange's theorem, Cauchy described and made use of several properties concerning the way in which these cosets (i.e., rows of his array) were interrelated. We state these various properties in the language of cosets here, along with some other properties of cosets which were suggested by our reading of Cayley's paper. As you read through them, consider which of these properties you already believe based on your previous work in this project. The subsequent task outlines a proof for each; the properties are arranged in an order which will make several of these proofs appears as corollaries of earlier ones.

Coset Properties
Let G be a group and H a subgroup of G.

1. For all $a \in G$, $a \in aH$ and $a \in Ha$.
2. $G = \bigcup_{a \in G} aH = \bigcup_{a \in G} Ha$.

   Note: Property 2 formalizes Cauchy's somewhat vague procedure of adding rows to the table used in his proof of Lagrange's Theorem for S_n until all the elements of the group appear somewhere in that table.

3. For all $a \in G$, if $a \notin H$, then $H \cap Ha = \emptyset = H \cap aH$.
4. For all $a, b \in G$, if $aH \cap bH \neq \emptyset$, then $aH = bH$.
   Similarly, if $Ha \cap Hb \neq \emptyset$, then $Ha = Hb$.
5. For all $a, b \in G$, $a \in bH$ if and only if $aH = bH$.
   Similarly, $a \in Hb$ if and only if $Ha = Hb$.
6. For all $a \in G$, $a \in H$ if and only if $Ha = H = aH$.
7. For all $a \in G$, $|aH| = |Ha| = |H|$.
8. If G is abelian, then $Ha = aH$ for all $a \in G$.

Task II.1

In this task, you will prove the eight elementary properties given above for left cosets.[81]

Begin by assuming G is a group, H a subgroup of G, and $a, b \in G$.

(a) Use the definition of (left) coset to prove Coset Property 1.
That is, let $a \in G$ and explain how we know that $a = ah$ for some $h \in H$.

(b) Use Coset Property 1 to prove Coset Property 2.
That is, show that $x \in G$ if and only if $x \in \bigcup\limits_{a \in G} aH$.

(c) Use contradiction to prove Coset Property 3.
That is, suppose $a \notin H$ and $H \cap aH \neq \emptyset$, so that $h_1 = ah_2$ for some $h_1, h_2 \in H$.
Then show that this contradicts the assumption that $a \notin H$.

(d) Use properties of subgroups to prove Coset Property 4.
Begin by assuming $a, b \in G$ with $aH \cap bH \neq \emptyset$, and let $x \in aH \cap bH$.
Taking $h_1, h_2 \in H$ with $x = ah_1$ and $x = bh_2$, explain why $ah_1 = bh_2$ implies $a \in bH$.
Then use the fact that $a = bh$ for some $h \in H$ to show that $y \in aH$ implies $y \in bH$.
This tells us that $aH \subseteq bH$.

Follow a similar plan to show that $bH \subseteq aH$, and conclude that $aH = bH$.

(e) Use Coset Properties 1 and 4 to prove the 'only if' direction of Coset Property 5.
That is, given $a, b \in G$ with $a \in bH$, explain why $aH = bH$.

To prove the 'if' direction, assume $aH = bH$ and explain how $a \in bH$ follows.

(f) Use Coset Property 5 with $b = 1$ to prove Coset Property 6.

(g) Use the function $f : H \to aH$ defined by $f(h) = ah$ to prove Coset Property 7.
That is, show that f is one-to-one and onto.

(h) Use the definition of an abelian group to prove Coset Property 8.
That is, show that $x \in Ha$ if and only if $x \in aH$.

Coset Properties 2 and 4 correspond to two central ideas in Cauchy's proof of Lagrange's Theorem for S_n.[82] Today, we would say that these two properties imply that the collection of all left (or right) cosets of a group forms a *partition* of G. The concept of a partition does not apply only to groups, however, as indicated by the following formal definition of a partition for any general set S.

[81] The left coset proofs that you will write in this task and elsewhere in this appendix are easily adapted to right cosets.

[82] Cauchy himself used Coset Property 3 to explain why the top row of his table did not overlap with any of its other rows; today we would simply consider the top row to be another coset ($H = 1H$), so that only Coset Property 2 is needed to know that there is no overlap between any of the rows.

Definition[83]

Let S be a set and P_α a non-empty subset of S for each $\alpha \in I$, where I is an indexing set. The collection $\mathcal{P} = \{ P_\alpha \}_{\alpha \in I}$ is a ***partition*** of S if and only if every element of S appears in one and only one $P_\alpha \in \mathcal{P}$. That is, if and only if

1. The subsets P_α are mutually disjoint: given $\alpha, \beta \in I$ with $P_\alpha \neq P_\beta$, then $P_\alpha \cap P_\beta = \emptyset$.
2. $S = \displaystyle\bigcup_{\alpha \in I} P_\alpha$.

Applying this definition to the collection of all (left) cosets of a group, you should be able to convince yourself (and others!) that Coset Properties 2 and 4 do indeed prove that $\mathcal{P} = \{ aH \}_{a \in G}$ is a partition of G. Once you're convinced, use this fact to see how easily we can now prove Lagrange's Theorem in the general case where G is any finite group.

Task II.2
Let G be a finite group and H a subgroup of G.

Set $r = |H|$, $n = |G|$ and let k be the number of distinct left cosets of H.
(The number k is called the ***index*** of H in G, denoted $(G : H)$.)

Use the fact that $\mathcal{P} = \{ aH \}_{a \in G}$ is a partition of G along with Coset Property 7 to explain why $n = kr$. Conclude that $|H|$ divides $|G|$, and that $(G : H) = \frac{|G|}{|H|}$.

Although Task II.2 concludes our discussion of how to prove Lagrange's Theorem using the language of cosets, there are some other issues related to cosets which will be useful in your study of more advanced group theory. We comment on these in the remainder of this appendix.

One thing that may have puzzled you about the requirement that the aH be mutually disjoint is the fact that it is often the case that $aH = bH$ for some $a \neq b$. For example, the group table given by Cayley for S_3 makes it clear that $1H = \alpha H = \beta H$ and $\gamma H = \delta H = \epsilon H$. (See Task 82(a-ii)). In fact, Coset Property 6 implies that $Ha = H = Hb$ for any $a, b \in H$. But this doesn't mean that the Ha and Hb are two distinct non-disjoint cosets; rather, it means that 'Ha' and 'Hb' are two different *names* for the same coset. This is exactly what happens with rational numbers where, for instance, $\frac{3}{4}$ and $\frac{6}{8}$ are two different names for just one rational number. Similarly, each individual coset aH is an object in its own right, but with several different representations.[84]

Another question you may have had about cosets is how the left and right cosets corresponding to a particular element are related. This is another of the ideas that you explored in Task 82, which we remind you of here while adding some insights into this example which come from Lagrange's Theorem. Recall first that, in that task, we had $G = S_3 = \{1\,, \alpha\,, \beta\,, \gamma\,, \delta\,, \epsilon\}$, with

$$\alpha = (1,\, 3,\, 2)\ ,\ \beta = (1,\, 2,\, 3)\ ,\ \gamma = (2,\, 3)\ ,\ \delta = (1,\, 3)\ ,\ \epsilon = (1,\, 2).$$

For the first subgroup considered in the task, $H = \langle \alpha \rangle = \{1\,, \alpha\,, \beta\}$, it turned out to be the case that $\gamma H = \delta H = \epsilon H$ and that $\gamma H = H\gamma$. In fact, since $|G| = 6$, $|H| = 3$ and $(G : H) = \frac{6}{3} = 2$, we would

[83] You may wish to read more about partitions in an algebra or other mathematics textbook, especially to see how they are related to *equivalence relations*, another important tool in much of mathematics (including group theory) today.

[84] This idea was also implicit in Task 40(d) where Cauchy's proof of Lagrange's Theorem for S_n was explored in a specific case.

now be able to predict in advance that H has exactly two distinct left cosets. By Coset Property 6, the first of these left cosets will be $H = \alpha H = \beta H$. Since all cosets are the same size (by Coset Property 7), we could then conclude that the second left coset consists of the three remaining elements of G, or $\gamma H = \delta H = \epsilon H = \{\gamma\,,\,\delta\,,\,\epsilon\}$. Similarly, since there are only two distinct *right* cosets, one of which is $H = H\alpha = H\beta$, it must be the case that $H\gamma = H\delta = H\epsilon$. This means that not only does $\gamma H = H\gamma$ hold in this example, as you proved in Task 82, but also that $\delta H = H\delta$ and $\epsilon H = H\epsilon$. Combining this with Coset Property 6, it follows that $\theta H = H\theta$ for *every* $\theta \in G$, again in this particular case.[85]

Again in Task 82, however, we saw how another choice of subgroup led to quite different results. Namely, letting $K = \langle\gamma\rangle = \{1\,,\,\gamma\,\}$, the left and right cosets of an element were generally different. For instance, $\alpha K = \{\alpha\,,\,\delta\,\}$, while $K\alpha = \{\alpha\,,\,\epsilon\,\}$. Or, taking another example for this same subgroup K, $\delta K = \{\delta\,,\,\alpha\,\}$, while $K\delta = \{\delta\,,\,\gamma\,\}$. Thus, we see that, in general, the left and right cosets corresponding to a particular element may not equal and, furthermore, that this depends on the particular *subgroup* involved, more so than the particular group.

These comments lead us to the following definition of a special type of subgroup, which we feel compelled to state for both its beauty and its importance.

Definition
Let G be a group and H a subgroup of G.
Then H is a ***normal subgroup*** if and only if $aH = Ha$ for every $a \in G$.

The notion of a normal subgroup is precisely what is needed to define the concept of a **simple group** which was discussed at the beginning of Section 4; namely, a a nontrivial group is simple if and only if its only normal subgroups are the trivial group and the group itself. Another reason why normal subgroups are important is that they allow us to impose group structure on the collection of all left (or right) cosets. This is done by first defining coset multiplication in the following natural way:

Definition
Given a group G with $a, b \in G$ and H a subgroup of G, the ***product of the left cosets*** aH and bH is defined as

$$(aH)(bH) = (ab)H.$$

That is, the product of the left cosets corresponding to $a, b \in G$ is the left coset corresponding to the product ab.

Naturally, the first thing to worry about here is whether coset multiplication is *well-defined.* After all, each individual left coset, just like any given rational numbers, can be represented by several different names. Thus, we would want to be sure that given $aH = cH$ and $bH = dH$, we will also have $(aH)(bH) = (cH)(dH)$. It turns out that this is the case, provided the subgroup H is normal. We omit the proof of this, as well as the proof of the theorem in the following statement, both of which you can find it in any current algebra textbook.

Definition and Theorem
Let G be a group, H a normal subgroup of G, and $G/H = \{\,aH \,|\, a \in G\}$.
Then G/H is a group under coset multiplication, called the ***factor group*** of G by H.

G/H is also called a ***quotient group***, and turns out (not surprisingly) to have an algebraic structure intimately related to that of G. We encourage you to study this relationship further — look for it under the name ***Fundamental Homomorphism Theorem*** in any group theory textbook, and enjoy!

[85] Remember that $\gamma H = H\gamma$ does not imply that the element γ commutes with the elements of H. This is clearly not true in this example, since $\gamma\alpha = \epsilon$, while $\alpha\gamma = \delta$. Rather, $\gamma H = H\gamma$ means that the list of elements in γH is the same as the list of elements in $H\gamma$, only arranged in a different order.

References

[1] Archibald, R.C., *benjamin Peirce, 1809–1880: Biographical Sketch and Bibliography*, 1925.

[2] Cauchy, Augustin, Mémoire sur le nombre des valeurs qu'une fonction peut acquérir, lorsqu'on y permute de toutes les manières possibles les quantités qu'elle renferme, *Journal de l'École Polytechnique*, Cahier XVII, Tome X (1815), 1–28, and in *Œuvres complètes de Augustin Cauchy*, Série 2, Tome 1 (1905), 64–90.

[3] Cauchy, Augustin, Mémoire sur les fonctions que ne peuvent obtenir que deux valeuers égales et de signes contraires par suite des transpositiosn opérées entre les variable qu'elles renferment, *Journal de l'École Polytechnique*, Cahier XVII, Tome X (1815), 29–117, and in *Œuvres complètes de Augustin Cauchy*, Série 2, Tome 1 (1905), 91–169.

[4] Cauchy, Augustin, Mémoire sur les arrangements que l'on peut former avec des lettres données, et sur les permutations ou substitutions à l'aide desquelles on passe d'un arrangement à un autre, in *Exercises d'Analyse et de Mathématiques Physiques*, Tome III (1844), 151–242, and in *Œuvres complètes de Augustin Cauchy*, Série 2, Tome 13 (1933), 171–282.

[5] Cayley, Arthur, On the theory of groups, as depending on the symbolic equation $\theta^n = 1$ - Part I, *Philosophical Magazine,* 7 (1854), 40–47, and in *The Collected Mathematical Papers of Arthur Cayley*, Cambridge: Cambridge University Press, vol. 2 (1889), 123–130.

[6] Cayley, Arthur, On the theory of groups, as depending on the symbolic equation $\theta^n = 1$ - Part II, *Philosophical Magazine,* 7 (1854), 408–409, and in *The Collected Mathematical Papers of Arthur Cayley*, Cambridge: Cambridge University Press, vol. 2 (1889), 131–132.

[7] Cayley, Arthur, On the theory of groups, as depending on the symbolic equation $\theta^n = 1$ - Part III *Philosophical Magazine,* 18 (1859), 34–37, and in *The Collected Mathematical Papers of Arthur Cayley*, Cambridge: Cambridge University Press, vol. 3 (1889), 594–602.

[8] Chakraborty, Sujoy and Chowdhury, Munibur Rahman, Arthur Cayley and the Abstract Group Concept, *Mathematics Magazine*, vol. 78, vol. 4 (Oct. 2005), 269–282.

[9] Crilly, Tony, *Arthur Cayley: Mathematician Laureate of the Victorian Age*, Baltimore: Johns Hopkins University Press, 2006.

[10] Lagrange, J. L., Réflexions sur la résolution algébrique des équations, *Mémoire de l'Académie de Berlin*, 1770-1771, pp..

[11] Lagrange, J. L., *Traité de la résolution des èquations numériques de tous les degrés, avec des notes sur plusieurs points de la théorie des équations algébriques*, Paris: Courcier, 1808.

[12] Roth, Richard L., A History of Lagrange's Theorem on Groups, *Mathematics Magazine*, vol. 74, no. 2 (April 2001), 99–108.

[13] van der Waerden, B. L., Hamilton's Discovery of Quaternions, *Mathematics Magazine*, Vol. 49, No. 5 (Nov., 1976), 227–234.

[14] Wussing, Hans, *The Genesis of the Abstract Group Concept: A Contribution to the History of the Origins of Abstract Group Theory*, Cambridge and London: MIT Press, 1984.