Winter 2022

Gauss and the First “Rigorous” Proof of the Fundamental Theorem of Algebra

Sarah Hagen

Alan Kappler

Gauss and the First "Rigorous" Proof of the Fundamental Theorem of Algebra

Sarah Hagen* and Alan Kappler

February 10, 2023

The German mathematician Carl Friedrich Gauss (1777–1855) had perhaps the longest and most illustrious career of any historical mathematician.[1] Among his many groundbreaking accomplishments, he is generally given credit for the first rigorous proof of the Fundamental Theorem of Algebra. This theorem, which states that any nonconstant polynomial of degree m has m complex roots counted with multiplicity, had been known in some form since at least the early 15^{th} century. However, despite its importance in algebra and number theory, a fully rigorous proof eluded mathematicians for centuries. The first published attempt at a proof was given by Jean-Baptiste le Rond d'Alembert (1717–1783) in 1746. Further attempts were made that century by Leonhard Euler (1707–1783), François Daviet de Foncenex (1734–1799), and Joseph-Louis Lagrange (1736–1813), among others. In the introduction to Gauss's 1799 doctoral dissertation, he surveyed the attempts by each of these mathematicians in turn, and proceeded to make an account of the serious flaws or gaps in each. He then turned to giving the reader what he promised in the title of his dissertation: *New Proof of the Theorem That Every Algebraic Rational Integral Function in One Variable Can Be Resolved Into Real Factors of the First or the Second Degree.*[2]

It is common in modern textbooks to treat the Fundamental Theorem of Algebra as a consequence of Liouville's Theorem in complex analysis (named after Joseph Liouville (1809–1892)). While this high-powered theorem certainly does the job, Liouville's Theorem itself wasn't first proved until nearly 50 years after Gauss's dissertation (and was named after a man who was not even born when Gauss's dissertation was published). So not only does the standard presentation of the Fundamental Theorem of Algebra misrepresent the historical development of the theorem, it also postpones the proof of such an important theorem until one takes an upper level course in complex analysis. However, Gauss himself proved the theorem without any appeal to complex numbers at all! Instead, he used ideas from geometry, trigonometry, and calculus. The ideas present in his dissertation are accessible to any interested student who has completed the calculus sequence. While Gauss himself avoided using complex numbers, he sometimes did so only with great effort. Indeed, at several points Gauss took theorems that are easily proved using basic properties of complex numbers and reproved

*Department of Mathematics, The Davidson Academy Online, Reno, Nevada, 89503; shagen@davidsononline.org.

[1]This is quite a claim! We point the interested reader to [May, 1972], [O'Connor and Robertson, 1996], or [Dunnington et al., 2004] to learn more about the life and work of this remarkable mathematician.

[2]English translations in this project were predominantly taken from a translation by Ernest Fandreyer, which can be found at https://quantresearch.org/Gauss_PhD_Dissertation.pdf, and secondarily from [Struik, 1969], with occasional modification by these authors while referencing the original Latin text [Gauss, 1799]. The figures that are cited are due to Gauss himself. They were copied from the Fandreyer translation, with slight modification by these authors to emphasize the dotted nature of certain lines.

them using somewhat convoluted trigonometric calculations. While he succeeded in avoiding complex numbers in doing so, it is the opinion of these authors that he muddied the water as a result and made it mysterious how or why anyone would have come up with such theorems. Thus, we have made the decision to use complex numbers (though sparingly) in our discussion for the sake of clarity and efficiency. However, the presentation is still intended to be understandable to students who have completed the calculus sequence and have only a passing familiarity with complex numbers.

In this project we will work through the essential aspects of Gauss's proof. Along the way, we will point out what might be considered to be gaps in Gauss's own proof, and discuss whether Gauss deserves the credit that he is given.

1 The Statement of the Theorem

Gauss wrote his doctoral dissertation *Demonstratio nova theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse* (in Latin) in 1799. It is not surprising, given that his paper was written hundreds of years ago in a different language, that Gauss's notation differs somewhat from our modern notation. In particular, whereas nowadays might mathematicians refer to a general m^{th} degree monic polynomial by $f(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_2x^2 + a_1x + a_0$, Gauss instead wrote such a function as $X = x^m + Ax^{m-1} + Bx^{m-2} + \text{etc.} + M$. For consistency's sake, we will adopt Gauss's notation for polynomials throughout this project.

Gauss began his dissertation by laying out several theorems regarding roots of polynomials whose proofs were accepted at his time. One theorem in particular is important for understanding his argument. Following Gauss, we will use this theorem (stated below in his notation) without proof:

> **Factor Theorem:** Given a polynomial X of degree m, if $X = 0$ at $x = a$, then X may be factored into a product of $x - a$ and a polynomial of degree $m - 1$.

With his brief discussion of elementary facts about polynomials out of the way, Gauss wrote:

> What has so far been set forth is sufficiently proved in algebraic books and does not in any way offend mathematical rigor. But analysts seem to have adopted far too quickly and without previous solid proof a theorem upon which almost all of the teaching of equations is built: *That any such function X can always be resolved into m simple factors or, which agrees with that entirely, that every equation of degree m has indeed m roots.*[3] But already with equations of second degree such cases are quite often encountered which disagree with this theorem. In order to make these cases conform to it, the algebraists were forced to invent an imaginary quantity whose square is -1.[4]

[3] Gauss noted earlier that by "simple factor" he meant a factor of the form $x - a$. In addition, it is clear that when discussing this theorem that he meant such polynomials have m roots with repetition.

[4] Note that Gauss actually got the history wrong here! Early mathematicians considered such irreducible quadratic equations to simply be unsolvable. Famously, it is due to the work of Girolamo Cardano (1501–1576) and Rafael Bombelli (1526–1572) on solving the general cubic equation that mathematicians "were forced to invent" complex numbers. For a wonderful discussion of the historical development of complex numbers see [Nahin, 2010] (in particular, the first chapter).

Task 1 Give an example of a degree 2 polynomial that cannot be factored using real numbers, but can be factored if one allows complex numbers.

Gauss continued:

I observe that it suffices to show only this: Every equation of whatever degree $x^m + Ax^{m-1} + Bx^{m-2} + \text{etc.} + M = 0$, or $X = 0$ (where the coefficients A, B, etc. are real numbers) will be satisfied at least once by a value of x of the form $a + b\sqrt{-1}$.

REMEMBER THIS! This is the statement of the theorem that Gauss set out to prove. Namely, that every non-constant polynomial with real coefficients has a complex root.

Task 2 Prove Gauss's claim here that this statement is sufficient for proving his broader claim. That is, prove that if a degree m polynomial has at least one complex root, then such a polynomial has m complex roots (with repetition). Hint: you will want to use the Factor Theorem.

Gauss, however, wanted to state his theorem using only real quantities. Thus, in the next line he wrote:

For it is well known that then X is divisible by a real factor $x^2 - 2ax + a^2 + b^2$, if b is not $= 0$, and by a simple factor $x - a$ if $b = 0$.

The claim that if X has a complex factor $x - (a + bi)$ then X has a real quadratic factor $x^2 - 2ax + a^2 + b^2$ follows from what mathematicians now refer to as the Complex Conjugate Root Theorem. This theorem states that if a polynomial with real coefficients has a complex root $a + bi$, then it also has a complex root $a - bi$.[5]

Task 3 (a) Use the Complex Conjugate Root Theorem to show that if a polynomial with real coefficients has a complex factor $x - (a + bi)$ with $b \neq 0$ then it has a real quadratic factor $x^2 - 2ax + a^2 + b^2$.

(b) OPTIONAL: Prove the conjugate root theorem. Hint: Let $p(x)$ be a polynomial with real coefficients and suppose $p(a+bi) = 0$. Note that this implies $\overline{p(a + bi)} = \overline{0}$, and use this to show that $p(\overline{a + bi}) = 0$.

To sum up, Gauss's task is still to prove that every polynomial of degree m has a complex factor $x - (a + bi)$. However, due to the fact complex numbers were still viewed with some suspicion at the time of his writing, he noted that this is equivalent to finding either a real simple factor or an irreducible real quadratic factor of the polynomial.

[5] Recall that the complex conjugate of $a + bi$, denoted $\overline{a + bi}$, is defined to be $a - bi$, hence the name of the theorem.

2 A Lemma and a Restatement of the Theorem

Before proceeding to his main argument, Gauss established the following lemma:[6]

Lemma: If the quantity r and the angle ϕ are so determined that the equations

$$r^m \cos m\phi + Ar^{m-1}\cos(m-1)\phi + Br^{m-2}\cos(m-2)\phi + \text{etc.} \\ + Krr\cos 2\phi + Lr\cos\phi + M = 0, \quad (1)$$

$$r^m \sin m\phi + Ar^{m-1}\sin(m-1)\phi + Br^{m-2}\sin(m-2)\phi + \text{etc.} \\ + Krr\sin 2\phi + Lr\sin\phi = 0 \quad (2)$$

hold, then the function

$$x^m + Ax^{m-1} + Bx^{m-2} + \text{etc.} + Kxx + Lx + M = X$$

will be divisible by the second-degree factor $xx - 2\cos\phi \cdot rx + rr$, unless $r\sin\phi = 0$. If $r\sin\phi = 0$, then the function will be divisible by the simple factor $x - r\cos\phi$.

Gauss noted that this lemma is standardly proved using complex numbers. Let's see how such a proof would go.

Task 4 Consider the polynomial $X = x^m + Ax^{m-1} + Bx^{m-2} + \text{etc.} + Kxx + Lx + M$. Note that we are letting x take complex values, not just real values, so we may write $x = re^{i\phi}$ or $x = r(\cos(\phi) + i\sin(\phi))$. Show that the left-hand-side of Gauss's equation (1) is just the real part of X and that the left-hand-side of Gauss's equation (2) is the imaginary part of X. Hint: You may find it useful to first write x in exponential form ($x = re^{i\phi}$) and then switch at some point to writing $e^{i\phi}$ in polar form ($e^{i\phi} = \cos\phi + i\sin\phi$).

So we see that if r and ϕ satisfy equations (1) and (2), then this makes the real and imaginary parts of X each equal to 0. In this case, X is thus equal to 0 at $x = re^{i\phi} = r(\cos\phi + i\sin\phi)$. So we have that $x - r(\cos\phi + i\sin\phi)$ is the desired factor of X when both equations (1) and (2) hold!

Task 5 OPTIONAL: For completeness's sake, let's confirm the final part of the claims of the lemma. We already have just seen that if equations (1) and (2) hold then $x - r(\cos\phi + i\sin\phi)$ is the desired factor of X. What is left is to show that in this case, (a) if $r\sin\phi = 0$ then $x - r\cos\phi$ is a real factor of X, and (b) if $r\sin\phi \neq 0$ then $x^2 - 2xr\cos\phi + r^2$ is an irreducible second-degree factor of X.

(a) Explain why the $r\sin\phi = 0$ really implies that $\sin\phi = 0$. Then argue for the fact that if $r\sin\phi = 0$ then $x - r\cos\phi$ must be a factor of X.

[6] Note that xx was often used in mathematics before the 20th century to represent x^2. We see this crop up several times in what follows.

(b) Show that if $r \sin\phi \neq 0$ then $x^2 - 2xr\cos\phi + r^2$ is a factor of X. Hint: Use the complex conjugate root theorem.

With this lemma in hand, Gauss wrote:

> It is quite manifest that for the proof of our theorem nothing more is required than to show: *When any function X of the form $x^m + Ax^{m-1} + Bx^{m-2} +$ etc. $+ Lx + M$ is given, then r and ϕ can be determined in such a way that the equations (1) and (2) hold.*

REMEMBER THIS![7] This is the statement that Gauss went on to prove. Before moving on, make sure you understand why this, as Gauss claimed, is all he needs to show.

3 The Geometric Set-Up

Gauss's next step was to recast the project in geometric terms. Helpfully, in addition to his general arguments, Gauss also provided a concrete example to illustrate his ideas. This section explores Gauss's treatment of each.

3.1 Establishing a Geometric Perspective

Thinking of the left-hand-sides of equations (1) and (2) as functions of two variables, r and ϕ, Gauss geometrically interpreted these functions as surfaces using cylindrical coordinates. That is, the inputs r and ϕ, are interpreted in polar coordinates in the plane, and the output is the height of the surface at that point. He then detailed how he planned to refer to these surfaces as follows:

> ...the expression $r^m \sin m\phi + Ar^{m-1}\sin(m-1)\phi + Br^{m-2}\sin(m-2)\phi +$etc.$+ Krr\sin 2\phi + Lr\sin\phi$, which for brevity I shall designate T from here on... will lie above the plane for positive values of T, below it for negative values, and on the plane itself if T vanishes; and they will be on a curved surface, continuous and in all directions infinite, which for brevity I shall call *the first surface* from here on. Again, in a similar manner a second surface may be referred to the same plane and center and axis whose altitude above any point of the plane shall be $r^m \cos m\phi + Ar^{m-1}\cos(m-1)\phi + Br^{m-2}\cos(m-2)\phi +$ etc. $+ Krr\cos 2\phi + Lr\cos\phi + M$, which expression I shall always denote by U, for brevity. This surface, which will also be continuous and infinite in every direction, I shall distinguish from the former surface by the term *second surface*.[8]

[7] We know we said this earlier, but this time we really mean it.

[8] We are not sure why Gauss decided to refer to the surface from Equation (2) as the first surface and the surface from Equation (1) as the second surface. We'll just have to deal with it.

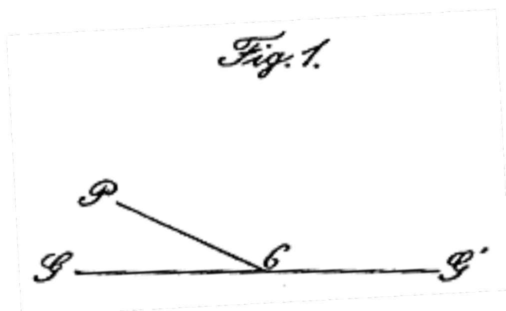
Note to the reader: Since we will frequently be referring back to these surfaces, make sure you have them written down in your notes. That is, in your notes, write out the following and fill in the blanks:

- *The first surface* is given by $T =$ ____________________
- *The second surface* is given by $U =$ ____________________

Task 6 Given the polynomial $X = x^4 - 2x^2 + 3x + 10$, find T and U.

Task 7 OPTIONAL: How do we know that these curved surfaces given by T and U are indeed continuous? Gauss needed this to be true, but he did not say how he knew it.

Interestingly, Gauss's orientation of the plane is different from our own orientation. For Gauss, when viewing the plane in polar coordinates, the point (r, ϕ) is the point in the plane that is a distance r from the center and with angle ϕ rotating **clockwise** from the **left** horizontal axis. In Gauss's figure below, the point C is the origin, the point G lies on the positive x-axis (G$'$ lies on the negative x-axis), and the point P has polar coordinates (r, ϕ) where ϕ is an acute angle.



Gauss then continued:

One can easily see that the first surface lies partly above and partly below the plane; for it is clear that the distance from the center can be taken so large that the remaining terms in T contribute nothing compared to the first $r^m \sin m\phi$; but that one can be made positive as well as negative accordingly as the angle ϕ is determined.

Task 8 DISCUSSION: Suppose that ϕ is fixed and that $\sin(m\phi) \neq 0$. Discuss Gauss's claim that if r is large enough then "the remaining terms in T contribute nothing compared to the first $r^m \sin m\phi$." Why is this true? Why did we include the caveats that ϕ is fixed and $\sin(m\phi) \neq 0$?

Task 9 Let us now verify the details of Gauss's claims in this last excerpt. Recall that

$$T = r^m \sin m\phi + Ar^{m-1}\sin(m-1)\phi + Br^{m-2}\sin(m-2)\phi + \text{etc.} + Krr\sin 2\phi + Lr\sin\phi.$$

(a) Let $r > |A| + |B| + \cdots + |L|$, and suppose further that $r > 1$. Let $\phi = (90/m)^\circ$. Explain why the values of r and ϕ just given guarantee that $T(r, \phi)$ is positive. Hint: Show that these values guarantee that

$$|r^m \sin m\phi| > |Ar^{m-1}\sin(m-1)\phi + Br^{m-2}\sin(m-2)\phi + \text{etc.} + Krr\sin 2\phi + Lr\sin\phi|.$$

(b) Given $r > |A| + |B| + \cdots + |L|$ (and greater than 1), come up with a value of ϕ that guarantees that $T(r, \phi)$ is negative.

(c) Let $r > (|A| + |B| + \cdots + |L|)\sqrt{2}$, and suppose further that $r > 1$. Let $\phi = (45/m)^\circ$. Explain why the values of r and ϕ just given guarantee that $T(r, \phi)$ is positive.

(d) OPTIONAL: Given $r > (|A| + |B| + \cdots + |L|)\sqrt{2}$ (and greater than 1), come up with *another* value of ϕ that guarantees that $T(r, \phi)$ is positive. Come up with *two* values of ϕ that guarantees that $T(r, \phi)$ is negative.

Having argued that the first surface lies partly above and partly below the fixed plane, Gauss could now state the conclusion he needed.

Therefore the fixed plane will necessarily be intersected by the first surface. This intersection of the plane with the first surface I shall call *the first curve*, which will therefore be determined by the equation $T = 0$.

Task 10 The claim is that since $T > 0$ at some point, $T < 0$ at another point, and T is continuous, it must be the case that $T = 0$ at some point.

(a) What standard theorem from calculus guarantees this result?

(b) Do a quick internet search to find out when that theorem was first rigorously proved by modern standards. Was it proved at the time Gauss wrote his dissertation in 1799?

(c) DISCUSSION: What do you think now about the claim that Gauss's dissertation provides the first "rigorous" proof of the Fundamental Theorem of Algebra?

Gauss next stated the same conclusion about the second surface.

For the same reason, the plane will be intersected by the second surface. The intersection constitutes a curve determined by the equation $U = 0$, which I will call *the second curve*.

Task 11 OPTIONAL: Find values of r and ϕ that make $U = r^m \cos m\phi + Ar^{m-1}\cos(m-1)\phi + Br^{m-2}\cos(m-2)\phi + \text{etc.} + Krr\cos 2\phi + Lr\cos\phi + M$ positive and then find values that make it negative.

Gauss next noted that each of these curves $T = 0$ and $U = 0$ will generally consist of several disconnected curves, which he called "branches." Indeed, as Gauss himself noted, one branch of $T = 0$ will always be the x-axis (do you see why?). He then wrote:

> Obviously, the task has now been reduced to that of demonstrating that there exists in our plane at least one point where some branch of the first curve is intersected by a branch of the second curve. For this, it is imperative to survey the nature of these curves more closely.

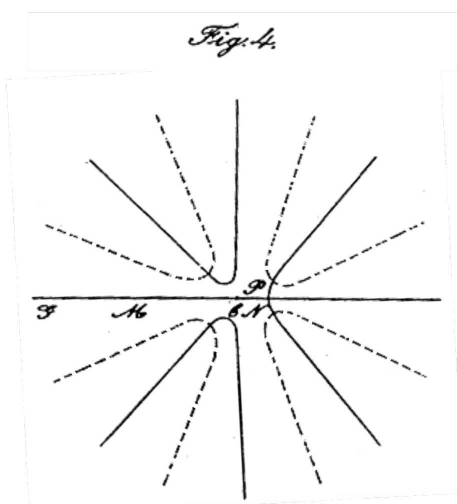
Task 12 In your own words, describe why the task now amounts to showing that there exists a point where the first curve intersects the second curve.

REMEMBER THIS![9] This is the final formulation of the task. What is left is to prove this claim.

3.2 Gauss's Concrete Example

One final thing before moving on. To understand Gauss's proof, it will be useful to have a specific example to work with. Gauss himself recognized this and gave the following.

> The fourth figure has been constructed assuming $X = x^4 - 2x^2 + 3x + 10$, by which readers less familiar with general and abstract investigations can then look at the position of both curves in a concrete case.



[9] We know, we know. We've now said this twice before. But this time we really, really mean it. We promise, this is the last time we'll say it.

In Gauss's figure the branches of the curve $T = 0$ are given by solid lines (with the x-axis one of these branches) while the branches for $U = 0$ are given by dotted lines.

Task 13 Make at least three observations and write at least two questions about Figure 4.

Let's confirm for ourselves that Gauss's figure is correct. This will take a few steps. You may have noticed that the example that Gauss gives is the example we used in Task 6. Review your work for that task and confirm that you calculated:

$$T = r^4 \sin(4\phi) - 2r^2 \sin(2\phi) + 3r \sin(\phi),$$
$$U = r^4 \cos(4\phi) - 2r^2 \cos(2\phi) + 3r \cos(\phi) + 10.$$

Now $T = 0$ and $U = 0$ are implicitly defined polar equations. Most graphing programs will not graph these equations as they are written. Thus, they need to be converted into Cartesian coordinates (since most graphing programs can handle graphing implicitly defined Cartesian equations). Gauss himself supplied the required conversion formulas:

If now the origin is assumed at C, the abscissa x taken in the direction of G, the corresponding y toward P, then $x = r\cos\phi$, $y = r\sin\phi$ and therefore generally for any n

$$r^n \sin n\phi = nx^{n-1}y - \frac{n \cdot n-1 \cdot n-2}{1 \cdot 2 \cdot 3}x^{n-3}y^3 + \frac{n \cdots n-4}{1 \cdots 5}x^{n-5}y^5 - \text{etc.},$$

$$r^n \cos n\phi = x^n - \frac{n \cdot n-1}{1 \cdot 2}x^{n-2}y^2 + \frac{n \cdot n-1 \cdot n-2 \cdot n-3}{1 \cdot 2 \cdot 3 \cdot 4}x^{n-4}y^4 - \text{etc.}$$

Task 14

(a) Use Gauss's formulas to convert $r^4 \sin(4\phi) - 2r^2 \sin(2\phi) + 3r \sin(\phi) = 0$ and $r^4 \cos(4\phi) - 2r^2 \cos(2\phi) + 3r \cos(\phi) + 10 = 0$ (that is, $T = 0$ and $U = 0$ from the example) into equations in terms of x and y.

(b) Plot these new (Cartesian coordinate) equations for $T = 0$ and $U = 0$ in Desmos or whatever your preferred graphing program is (make sure you plot them on the same graph). Does your graph resemble Gauss's graph, just flipped across the y axis? (Remember, Gauss oriented his axes differently from our modern orientation.)

(c) DISCUSSION: How remarkable is it that Gauss was able to produce his Figure 4 without the aid of calculators or computers? Can you think of any ways Gauss might have achieved this result?

Task 15 OPTIONAL: It is clear that the branches for both $T = 0$ and $U = 0$ have asymptotes as they approach infinity. What are the equations for these asymptotes?

Task 16 OPTIONAL: Prove that Gauss's conversions in the last excerpt are correct using de Moivre's formula together with the binomial theorem. Recall that de Moivre's formula states that for any real number ϕ and integer n we have $(\cos(\phi)+i\sin(\phi))^n = \cos(n\phi)+i\sin(n\phi)$.

Now that we have confirmed that the graphs in Gauss's fourth figure are correct, we will see how they play a critical role in his proof in the next section.

4 The Proof

In this section, we see how Gauss analyzed the branches of the curves T and U for a general polynomial to prove that the two curves intersect.[10] To make his ideas easier to follow, we divide this proof up into three separate subsections.

4.1 Finding the Right Circle

Gauss stated the theorem at the crux of his proof as follows:

Theorem. With everything the same as above, a circle can be described around C on whose periphery are $2m$ points where $T=0$ and as many where $U=0$, and in such a way that the latter ones lie singly between pairs of the former points.

Having stated this theorem, Gauss set out to prove it by first defining a sufficient radius for such a circle.

Let the sum of all coefficients (taken positive) A, B, etc. K, L, M be $= S$; let furthermore R be at the same time $> S\sqrt{2}$ and > 1. Then I assert that what has been stated in the theorem will necessarily take place in a circle of radius R.

Let us note right away that surely when Gauss said "the sum of all coefficients (taken positive)" he meant to consider the sum of the *absolute values* of the coefficients. That is, he intended that $S = |A|+|B|+\cdots+|K|+|L|+|M|$.

[10] Did you remember that we told you to remember all this back in footnote 8? This is all Gauss needs to complete his proof!

Task 17 Recall that in Tasks 6 and 14 you were given the polynomial $X = x^4 - 2x^2 + 3x + 10$, you calculated $T = 0$ and $U = 0$ in Cartesian coordinates, and then plotted them on the same graph. For this task, complete the following:

(a) Given the polynomial $X = x^4 - 2x^2 + 3x + 10$, how large must R be (according to Gauss's formula) in order to guarantee that a circle of radius R will each intersect $T = 0$ and $U = 0$ a total of $2m = 8$ times each?

(b) On your plot from Task 14, plot a circle with a radius at least as large as you calculated in part (a). Does it indeed intersect $T = 0$ and $U = 0$ a total of 8 times each?

The conditions on R that Gauss gave above should be familiar from Task 9. You may want to review that task before continuing on.

Gauss then elaborated:

In other words; Let us for the sake of brevity designate by (1) that point of this circle's circumference which is $\frac{1}{m}45$ degrees away from the intersection of the circle and the left part of the axis, or for which $\phi = \frac{1}{m}45^\circ$; similarly by (3) the point which is $\frac{3}{m}45^\circ$ away from that intersection, or for which $\phi = \frac{3}{m}45^\circ$; next by (5) the point where $\phi = \frac{5}{m}45^\circ$ etc. up to ($8m$–1) which is $\frac{8m-1}{m}45^\circ$ away from that intersection if you proceed always in the direction of that part (or $\frac{1}{m}45^\circ$ towards the opposite part). Whereupon a total of $4m$ points will be obtained on the periphery spaced at equal intervals.

Task 18 In Task 17 you added a circle to your plot of $T = 0$ and $U = 0$. On that circle, label the points (1), (3), (5), ... (31).

Then one point will fall between $(8m - 1)$ and (1) for which $T = 0$, and certainly similar single points will be between (3) and (5), between (7) and (9), between (11) and (13) etc. Thus there are $2m$ of them. In the same way, single points for which $U = 0$ will fall between (1) and (3), (5) and (7), (9) and (11), etc., wherefore their number is also $2m$. Finally, there will be no other points, on the whole periphery, beyond these $4m$ points for which T or $U = 0$.

Task 19 Let us continue working with the example polynomial $X = x^4 - 2x^2 + 3x + 10$. In Task 6 you calculated $T = r^4 \sin(4\phi) - 2r^2 \sin(2\phi) + 3r\sin(\phi)$ and $U = r^4\cos(4\phi) - 2r^2\cos(2\phi) + 3r\cos(\phi) + 10$.

(a) Find a value of R such that $R > 1$ and $R > (|A| + |B| + \cdots + |K| + |L| + |M|)\sqrt{2}$.

(b) In Task 6 you calculated $T(r, \phi) = r^4 \sin(4\phi) - 2r^2 \sin(2\phi) + 3r\sin(\phi)$. Use WolframAlpha or your favorite calculator to calculate $T(R, \phi)$ for the value of R from part (a) and the values of ϕ at the following points: $\phi = \frac{1}{m}45°, \frac{3}{m}45°, \phi = \frac{5}{m}45°, \ldots, \frac{8m-1}{m}45°$.

(c) For which values of ϕ is T positive? For which values of ϕ is T negative? Using the Intermediate Value Theorem again (see Task 10), does your work agree with Gauss's claim that "one point will fall between $(8m-1)$ and (1) for which $T = 0$, and certainly similar single points will be between (3) and (5), between (7) and (9), between (11) and (13) etc."?

(d) OPTIONAL: In Task 6 you calculated $U = r^4\cos(4\phi) - 2r^2\cos(2\phi) + 3r\cos(\phi) + 10$. Calculate $U(R, \phi)$ for the value of R from part (a) and the values of ϕ at the following points: $\phi = \frac{1}{m}45°, \frac{3}{m}45°, \phi = \frac{5}{m}45°, \ldots, \frac{8m-1}{m}45°$.

(e) OPTIONAL: For which values of ϕ is U positive? For which values of ϕ is U negative? Does your work agree with Gauss's claim that "points for which $U = 0$ will fall between (1) and (3), (5) and (7), (9) and (11), etc."?

Task 20 OPTIONAL: Why do you think Gauss chose the angles he did for determining the reference points on the circle? That is, why did Gauss choose odd multiples of $\frac{1}{m}45°$ as opposed to, say, multiples of $\frac{1}{m}90°$ or $\frac{1}{m}180°$?

Gauss then proceeded to prove his claim that T will equal 0 between points (3) and (5), (7) and (9), (11) and (13), up through (8m-1) and (1):

Proof: I. At point (1), $m\phi = 45°$ and therefore

$$T = R^{m-1}\left(R\sqrt{\frac{1}{2}} + A\sin(m-1)\phi + \frac{B}{R}\sin(m-2)\phi + \text{etc.} + \frac{L}{R^{m-2}}\sin\phi\right).$$

However, the sum $A\sin(m-1)\phi + \frac{B}{R}\sin(m-2)\phi$ etc. can certainly not be greater than S and is therefore necessarily less than $R\sqrt{\frac{1}{2}}$; whence it follows that the value of T at this point is certainly positive.

Task 21 Let's confirm each of the claims in the proof so far.

(a) Confirm that at point (1) we have

$$T = R^{m-1}\left(R\sqrt{\frac{1}{2}} + A\sin(m-1)\phi + \frac{B}{R}\sin(m-2)\phi + \text{etc.} + \frac{L}{R^{m-2}}\sin\phi\right).$$

(Recall that $T = r^m \sin m\phi + Ar^{m-1}\sin(m-1)\phi + Br^{m-2}\sin(m-2)\phi + \text{etc.} + Krr\sin 2\phi + Lr\sin\phi$.)

(b) Confirm that the sum $A\sin(m-1)\phi + \frac{B}{R}\sin(m-2)\phi + \text{etc.} + \frac{L}{R^{m-2}}\sin\phi$ cannot be greater (in absolute value) than S, and so is less than $R\sqrt{\frac{1}{2}}$.
(Recall that $S = |A| + |B| + \cdots + |K| + |L| + |M|$.)

(c) Confirm that T is in fact positive at this point.

More importantly, T will consequently have a positive value as long as $m\phi$ falls between $45°$ and $135°$, i.e. from point (1) up to (3) the value of T will be positive. For the same reason will T have a positive value from point (9) and up to (11), and in general from any point ($8k$ + 1) up to ($8k$ + 3) where k denotes any integer. Similarly, T will have negative values everywhere between (5) and (7), between (13) and (15) etc. and generally between ($8k$ + 5) and ($8k$ + 7), and certainly will nowhere in all these intervals $= 0$. But because at (3) this value is positive, in (5) negative, it will necessarily be $= 0$ somewhere between (3) and (5); and certainly also between (7) and (9), between (11) and (13) etc. up to and including the interval between ($8m$ – 1) and (1), so that altogether $T = 0$ at $2m$ points.

Task 22 (a) Why will T be positive as long as $m\phi$ falls between $45°$ and $135°$?

(b) Why will T be positive between points of the form ($8k$+1) and ($8k$+3)?

(c) Why will T be negative between points of the form ($8k$+5) and ($8k$+7)?

You just showed that $T \neq 0$ between points of the form ($8k$+1) and ($8k$+3) since it is always positive then. You also showed that $T \neq 0$ between points of the form ($8k$+5) and ($8k$+7) since it is always negative then. The Intermediate Value Theorem then tells us that $T = 0$ between points of the form ($8k$+3) and ($8k$+5) and between points of the form ($8k$+7) and ($8k$+1). Since there are $2m$ such intervals, there must be at least $2m$ points on the circle where $T = 0$. The next part of Gauss's proof showed that those are in fact exactly $2m$ points on the circle where $T = 0$. That is, he went on to prove that between points of the form ($8k$+3) and ($8k$+5) and between points of the form ($8k$+7) and ($8k$+1) there is exactly one point where $T = 0$. As we see, he used a proof by contradiction:

II. Beyond these $2m$ points there are no others having this property, as can be seen thus: Because there are none between (1) and (3), (5) and (7) etc., such points could exist in no other way than that in some interval between (3) and (5), or between (7) and (9) etc. would lie at least two. Then T would necessarily be a maximum or a minimum in the same interval and therefore $\frac{dT}{d\phi} = 0$. But $\frac{dT}{d\phi} = mR^{m-2}\left(R\cos m\phi + \frac{m-1}{m}A\cos(m-1)\phi + \text{etc.}\right)$, and $\cos m\phi$ is always negative between (3) and (5) and $> \sqrt{\frac{1}{2}}$.[11] Whence it is easily seen that in this whole interval $\frac{dT}{d\phi}$ is a negative quantity; likewise it is positive everywhere between (7) and (9), negative between (11) and (13) etc., so that it can be 0 in none of these intervals, and therefore that supposition cannot be maintained.

Task 23 (a) How did Gauss know that $\cos m\phi$ is always less than $-\sqrt{\frac{1}{2}}$ between points (3) and (5)?

(b) Why must $\frac{dT}{d\phi} < 0$ between points (3) and (5)? Hint: Recall that $R > (|A| + |B| + \cdots + |K| + |L| + |M|)\sqrt{2}$.

Task 24 (a) Gauss claimed that if $T = 0$ twice in an interval, then T would necessarily attain a maximum or a minimum in the closed interval between the points where $T = 0$. What standard theorem from calculus guarantees this result?

(b) Do a quick internet search to find out when that theorem was first rigorously proved by today's standards. Was it proved by the time Gauss wrote his dissertation in 1799?

Without going into as much detail, Gauss gave the similar claims for the values of U:

III. Again, in a similar way it is proved that U has a negative value everywhere between (3) and (5), (11) and (13) etc. and generally between $(8k + 3)$ and $(8k + 5)$, but a positive value between (7) and (9), (15) and (17) etc. and generally between $(8k + 7)$ and $(8k + 9)$. It follows from this immediately that U must become $= 0$ somewhere between (1) and (3), between (5) and (7) etc., i.e. at $2m$ points. Again, in none of these intervals can $\frac{dU}{d\phi} = 0$ (which is easily proved in the same manner as above); wherefore more than those $2m$ points where $U = 0$ will not be possible on the periphery of the circle.

So Gauss had finally proven that his circle was sufficiently large to satisfy his constraints, which he helpfully summarized below:

These conclusions can also be expressed in the following manner: When a circle of the required magnitude is drawn around center C then $2m$ branches of the first curve and as many of the second curve will enter it, and indeed in such a way that any two successive branches of the first curve will be separated by some one branch of the second curve alternately.

[11] Gauss clearly meant here that $\cos m\phi$ in *absolute value* is greater than $\sqrt{\frac{1}{2}}$.

4.2 Entering and Exiting the Circle

With these claims now established, Gauss turned to the next piece of his argument, exploring the nature of algebraic curves to ensure that they would cooperate in his final step.

Already from this relative position of the branches entering the circle it can be deduced in many ways that there must necessarily be an intersection of some branch of the first curve with a branch of the second curve within the circle. And I hardly know which method should preferably be chosen above the others.

Gauss must have felt like he was nearly there. At this point it seemed obvious to him that since the branches of $T = 0$ and $U = 0$ enter and leave at alternating points, somewhere they must intersect. And of course, finding an intersection of $T = 0$ and $U = 0$ is exactly what he wanted! Indeed, it seemed so obvious that Gauss felt there were several tacks he could take to prove it.

Task 25

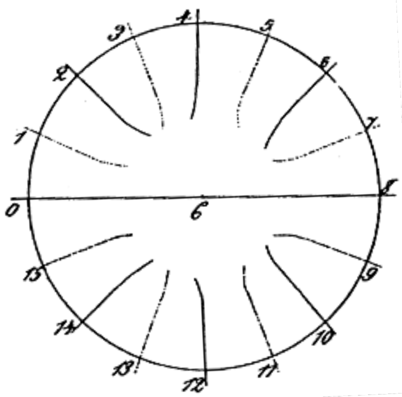
(a) How obvious to you is Gauss's claim that "there must necessarily be an intersection of some branch of the first curve with a branch of the second curve within the circle" given the branches enter and leave at alternating points?

(b) OPTIONAL: How might you try to convince someone who had doubts about this claim that it is indeed true? No need to write a full proof, but perhaps give an idea of how a proof might go, or give a reason or two in favor of the claim.

We will see, though, that nailing down this final aspect of the proof was not without its difficulty.

Let us designate by 0 (Figure 2) the point on the periphery of the circle where it is intersected by the left part of the axis (which itself is one of the $2m$ branches of the first curve); by 1, the nearest point where a branch of the second curve enters; the point nearest to this one where the second branch of the first line enters, by 2, and so on up to $4m$–1 . In this way a branch of the first line enters the circle at every point marked by an even number, whereas a branch of the second line enters at all points represented by an odd number. But according to higher mathematics, any algebraic curve (or the individual parts of such an algebraic curve if it perhaps consists of several parts) either turns back into itself or extends to infinity. Consequently, a branch of any algebraic curve which enters a limited space, must necessarily exit from this space somewhere.

Fig. 2.



By "algebraic curve" here Gauss likely meant an implicitly defined curve of the form $p(x, y) = 0$ where $p(x, y)$ is a polynomial in x and y (note that these are the types of curves T and U are when converted to Cartesian coordinates). Gauss's claim is that any branch of such a curve that enters an enclosed space must also exit that enclosed space.

Task 26 OPTIONAL: Do you agree with Gauss's claim that "a branch of any algebraic curve which enters a limited space, must necessarily exit from this space somewhere"? Does this claim seem like something that requires proof? Explain why or why not.

Interestingly, though Gauss claimed that this result had already established in "higher mathematics," he walked back from this claim in a footnote, saying:

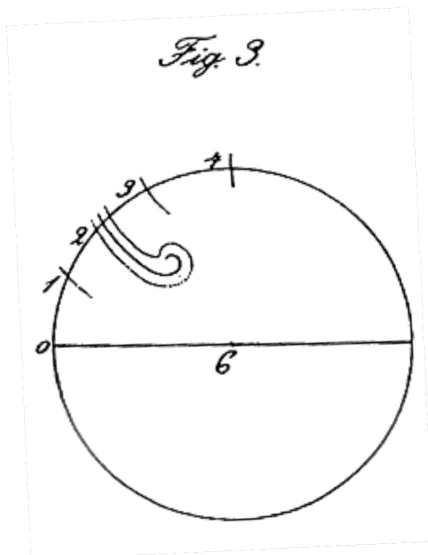
> It seems to have been proved with sufficient certainty that an algebraic curve can neither be broken off suddenly anywhere (as happens e.g. with the transcendental curve whose equation is $y = \frac{1}{\ln(x)}$) nor lose itself, so to say, in some point after infinitely many coils (like the logarithmic spiral). As far as I know, nobody has raised any doubts about this. However, should someone demand it then I will undertake to give a proof that is not subject to any doubt, on some other occasion.

Task 27 Check the graphs for $y = \frac{1}{\ln(x)}$ and the logarithmic spiral (which has polar equation $r = e^{\frac{\theta}{2\pi}}$). Either plot them yourself or look them up to get a sense of what Gauss was worried about. Explain how this illustrates what Gauss meant when he discussed a graph that is "broken off suddenly" or one that may "lose itself, so to say, in some point after infinitely many coils."

It turns out that Gauss never did provide a proof that is not subject to any doubt. Indeed, no such proof was given until 1920, and even then, the proof was nontrivial. See [Basu and Velleman, 2017] for an excellent discussion of this topic.

However, Gauss did not stop there. He continued in his footnote to argue for the claim at least in the specific case we are considering. He wrote:

In the present case this is really manifest: Suppose some branch, e.g. 2, does not exit the circle anywhere (Figure 3). Then you could enter the circle between 0 and 2, thereafter could go around this whole branch (which would have to lose itself within the space of the circle), and finally could exit again from the circle between 2 and 4 so that in the whole path you shall not have intersected the first curve. That is this is really absurd is clear from the fact that at the point where you enter the circle you had the first surface above you, at the exit below you. Therefore you must necessarily have somewhere come upon the first surface itself, that is to say: Upon a point of the first curve.



Note that "had the first surface above you" implies that T is positive, while having the first surface below you implies that T is negative.

Task 28

(a) Where does this explanation make use of the Intermediate Value Theorem?

(b) Do you agree with Gauss's claim "That is this is really absurd is clear from the fact that at the point where you enter the circle you had the first surface above you, at the exit below you"?

We see that Gauss actually went further than simply arguing that the curves in question that enter the circle must also exit the circle. He argued that in fact any branch that enters the circle must exit at a *different* point. Let us proceed, then, as Gauss did, accepting that any branch that enters the circle must exit at a different point.

From this it is easily concluded that any point marked by an even number (or in short, any even point) must be connected within the circle to another even point by a branch of the first curve, and likewise any point marked by an odd number to a similar point by a branch of the second curve. Now although this connection of two such points can be very varied according to the nature of the function X so that in general it cannot be clearly described, it is yet easy to prove *that an intersection of the first curve with the second always occurs whatever that connection may be eventually.*

4.3 Branch Intersections and the Jordan Curve Theorem

Gauss's final task was to show that, given this knowledge of connections between points on the outside circle, the curves must indeed cross. He proceeded using proof by contradiction:

The proof that this happens necessarily seems to be most suitably given indirectly. That is to say, let us assume the junction of any two of the even points and any two of the odd points can be arranged in such a way that no intersection of a branch of the first curve with a branch of the second curve results. Because the axis is a part of the first curve the point 0 is obviously connected with the point $2m$. Point 1 can therefore not be connected with any point beyond the axis, i.e. with no point marked by a number greater than $2m$, for otherwise the connecting curve would necessarily intersect the axis. If it is therefore assumed that 1 is connected with point n, then n will be $< 2m$.

Task 29

(a) How did Gauss know that for the first curve the point 0 is connected to the point $2m$?

(b) How obvious is Gauss's claim "Point 1 can therefore not be connected with any point beyond the axis, i.e. with no point marked by a number greater than $2m$, for otherwise the connecting curve would necessarily intersect the axis"? Is this something that you think requires proof? Why or why not?

Mathematicians now accept that Gauss here (and again later in his proof) implicitly used a result known as the Jordan Curve Theorem. This theorem can be stated as follows:

Jordan Curve Theorem: Any closed curve on a plane divides that plane into two parts, an inside and an outside, such that any continuous line connecting a point on the inside to a point on the outside must intersect the curve.

We see this theorem being used above in the following way. The line connecting point 0 to point $2m$ and the upper half of the circle form a simple closed curve. Consider the branch that enters the interior of this enclosed region through point 1. As Gauss argued earlier, this branch must also exit this region. By the Jordan Curve Theorem, it must intersect the closed curve. By hypothesis, it cannot intersect the curve along the line connecting point 0 to point $2m$. Therefore, it must exit at some point on the upper half of the circle. Such an exit point will be labeled point n where $n < 2m$.

So the argument is sound, with appeal to the Jordan Curve Theorem. First stated and partially proved by Camille Jordan (1838–1922) in the late 1800s, the proof of the general form of this theorem is also highly nontrivial.[12]

Now we come to the final argument of Gauss's proof, where the Jordan Curve Theorem is again implicitly used.

> For a similar reason, if 2 is established to be connected to n', then n' will be $< n$, because otherwise the branch $2 \cdots n'$ would necessarily intersect the branch $1 \cdots n$. For the same reason point 3 will be connected with one of the points falling between 4 and n'; and obviously when 3, 4, 5 etc. are taken as connected with n'', n''', n'''' etc., then n''' falls between 5 and n'', n'''' between 6 and n''' etc. Whence it is evident that at last there will be reached some point h connected with the point $h+2$, and then the branch that enters the circle at point $h+1$ will necessarily intersect the branch connecting points h and $h+2$. But because one of these two branches belongs to the first curve, the other to the second curve, it is thus manifest that the assumption is contradicted and that indeed an intersection of the first curve with the second curve necessarily exists somewhere.

Task 30 Explain how the Jordan Curve Theorem is used to show that "if 2 is established to be connected to n', then n' will be $< n$, because otherwise the branch $2 \cdots n'$ would necessarily intersect the branch $1 \cdots n$."
Hint: Follow the argument given earlier in the project, where Gauss's first implicit use of the Jordan Curve Theorem is explained.

Task 31

(a) How evident do you find it that "at last there will be reached some point h connected with the point $h+2$"? Explain why this must be the case in your own words, using a diagram to illustrate the idea if you like.

(b) At the end of this part of Gauss's proof, there are two branches: "the branch that enters the circle at point $h+1$" and "the branch connecting points h and $h+2$." Explain why it must be the case that these two branches intersect.

[12]Things get tricky with the Jordan Curve Theorem, for example, when dealing with nondifferentiable curves or fractal curves with lots of twists to them. One "intuitive" way to define the interior and exterior of a closed curve is the following: for a given point, draw a line from that point to a point at infinity. If there is an odd number of intersections between the line and our curve, the point is on the inside; if there is an even number of intersections, it is on the outside. Look up images of the Cesàro fractal and the Osgood Curve to get a sense for some curves where this intuitive definition breaks down.

And we are done! As Gauss stated,

When this is combined with the preceding, it will be concluded from all the investigations set forth that the theorem has been proved with all rigor *that any algebraic rational whole function in one variable can be resolved into real factors of the first or the second degree.*

5 Conclusion

To summarize, we see that Gauss took the following steps to prove that every nonconstant polynomial of degree m with real coefficients has m (complex) roots when counted with multiplicity. First, he showed that it sufficed to prove that such a polynomial contains at least one complex root. Allowing the domain of the polynomial to be complex numbers he split the polynomial into its real and complex parts. He then showed that to find a root $x = re^{i\theta} = r(\cos(\theta) + i\sin(\theta))$ of the original polynomial, it sufficed to find values of r and θ that simultaneously made the real and imaginary parts of the polynomial equal to 0. He found such a pair by interpreting the real and imaginary parts of the polynomial as surfaces in cylindrical coordinates. He noted that the points where these surfaces equal 0 form continuous curves in the plane. Finally, he showed that these curves must intersect. That intersection point then provided the necessary values of r and θ to determine the complex root.

We have also seen that at several points along the way Gauss's proof lacked the rigor that we expect of modern day proofs. Indeed, he made use of several theorems that were only proved later. Some of these results he relied on were not proved until *much* later, and some were extremely difficult or impossible to prove using elementary methods available at Gauss's time. Clearly, the standards for what counts as an air-tight proof have shifted over time.

Indeed, Gauss wrote this proof in the middle of one of the larger shifts in how mathematics is perceived. 1799 was only a few decades before mathematicians began to introduce formal definitions for concepts in calculus, independent of intuitive notions of time, motion, and continuity. As such, Gauss mentioned that certain functions (such as T and U) were continuous, but he did not provide justifications for those claims, likely because he did not think he needed to do so. He also wrote at a time just before standards concerning the use of geometry in proofs underwent a monumental change. Here, for example, is what Bernard Bolzano (1781–1848) wrote about Gauss's work on the Fundamental Theorem of Algebra in his monograph [Bolzano, 1817]:[13]

Indeed, this outstanding scholar [Gauss] had already presented us with a proof of this proposition [the Fundamental Theorem of Algebra] in 1799 but it had, as he admitted, the defect that it proved a purely analytic truth on the basis of a geometrical consideration. But

[13]Bolzano wrote these comments in the preface of the monograph in which he gave the first geometry-free proof of the Intermediate Value Theorem, based on the first geometry-free definition of a continuous function. In fact, Bolzano may have become interested in providing such a proof of the Intermediate Value Theorem in part because of the role that it played in Gauss's proofs of the Fundamental Theorem of Algebra [Barnett, 2016].

his two most recent proofs [both in 1816] are quite free of this defect; the trigonometric functions which occur in them can, and must, be in a purely understood analytical sense.

Task 32 DISCUSSION: How concerned are you about the ways in which Gauss used geometry in his 1799 proof? In what sense, if any, do you see this as a defect, and why (or why not)?

At the same time, Gauss maintained many conventions that soon came to be considered outdated. He deliberately avoided the use of complex numbers, to an extent that seems irrational to modern eyes. For example, despite having full knowledge of the (relatively) easy method outlined in Tasks 4 and 5, Gauss chose instead to use a monstrously complicated series of trigonometric identities to achieve the same result, due to a prevailing impression at the time that complex numbers were less rigorous than real numbers. Interestingly, Gauss presented yet a fourth proof of the Fundamental Theorem of Algebra on the occasion of the 50th anniversary of his original 1799 proof. Published as [Gauss, 1850], this fourth proof followed a strategy similar to the one that we have just read, but made free use of complex numbers which had come to be accepted as legitimate mathematical objects by that time (see [Bottazzini and Gray, 2013], p.76, for greater context).

Regardless, of what Bolzano and others may have regarded (or still regard) to be gaps, Gauss's dissertation remains a remarkable achievement. It is widely accepted to be the first proof related to the Fundamental Theorem of Algebra that is free of any major flaws. The strategy works. The theorems implicitly used were all later made rigorous. Indeed, when compared with previous attempts to prove the theorem, Gauss's proof comes across as a masterpiece of rigor.

We end by noting that the theorem proved in Gauss's dissertation is not exactly the same as what we now refer to as the Fundamental Theorem of Algebra. Gauss proved that every nonconstant polynomial of degree m with *real* coefficients has m complex roots when counted with multiplicity. The Fundamental Theorem of Algebra states that every nonconstant polynomial of degree m with *complex* coefficients has m complex roots when counted with multiplicity.[14] Let us see how Gauss's theorem implies the Fundamental Theorem of Algebra, thus justifying the credit Gauss is given as the first to prove this important theorem.

Task 33 OPTIONAL: Let $p(z) = a_m z^m + a_{m-1} z^{m-1} + \cdots + a_0$ be a nonconstant polynomial with complex coefficients a_i. Now consider $\overline{p(\overline{z})} = \overline{a_m} z^m + \overline{a_{m-1}} z^{m-1} + \cdots + \overline{a_0}$.

(a) Explain why $f(z) = p(z)\overline{p(\overline{z})}$ is a nonconstant polynomial with real coefficients.

(b) By Gauss's theorem, there exists a complex number z_0 such that $f(z_0) = p(z_0)\overline{p(\overline{z_0})} = 0$. Explain why this means that either z_0 is a root of $p(z)$ or $\overline{z_0}$ is a root of $p(z)$.

References

C. Baltus. D'Alembert's proof of the fundamental theorem of algebra. *Historia Mathematica*, 31(4): 414–428, 2004.

[14] Gauss actually proved this more general statement of the theorem in his 1850 proof.

J. H. Barnett. Why be so Critical? Nineteenth Century Mathematics and the Origins of Analysis, 2016. URL https://digitalcommons.ursinus.edu/triumphs_analysis/1.

S. Basu and D. J. Velleman. On Gauss's First Proof of the Fundamental Theorem of Algebra. *The American Mathematical Monthly*, 124(8):688–694, 2017. doi: 10.4169/amer.math.monthly.124.8.688.

B. Bolzano. *Rein analytischer beweis des lehrsatzes, dass zwischen je zwey werthen, die ein entgegengesetzes resultat gewähren, wenigstens eine reelle wurzel der gleichung liege (A purely analytic proof of the theorem that between two values which give results of opposite sign there lies at least one real root of the equation)*. Gottlieb Haase, Prague, 1817. English translation by S. J. Russ in [Russ, 1980] and [Russ, 2004, pp. 251–278].

U. Bottazzini and J. Gray. *Hidden Harmony—Geometric Fantasies: The Rise of Complex Function Theory*. Sources and Studies in the History of Mathematics and Physical Sciences. Springer New York, 2013. ISBN 9781461457244.

G. W. Dunnington, J. Gray, and F. E. Dohse. *Carl Friedrich Gauss: Titan of Science*. MAA Spectrum. Mathematical Association of America, 2004. ISBN 9780883855478. Reprint of original 1955 publication.

C. F. Gauss. *Demonstratio Nova Theorematis Omnem Functionem Algebraicam Rationalem Integram unius Variabilis in Factores Reales Primi vel Secundi Gradus resolvi posse. (New Proof of the Theorem That Every Algebraic Rational Integral Function In One Variable can be Resolved into Real Factors of the First or the Second Degree)*. C. G. Fleckeisen, Helmstadt, 1799. Also in *Werke*, Volume 3, pages 1–30. Partial English translation by D. J. Struik in [Struik, 1969, pp. 115–122]. Full English translation by E. Fandreyer available at https://www.quantresearch.org/Gauss_PhD_Dissertation.pdf.

C. F. Gauss. Beitrage zur Theorie der algebraischen Glehungen. In *Abhandlungen der K. Gesellschaft der Wissenschaften zu Göttingen. Mathematisch–Physikalische Klasse*, volume 4 of *Mathematisch–Physikalische Klasse*, pages 3–35. 1850. Also in *Werke* 3, pages 71–103.

K. O. May. Carl Friedrich Gauss. In *Dictionary of Scientific Biography*, volume 5, pages 298–315. Scribner, 1972.

P. J. Nahin. *An Imaginary Tale*. Princeton Science Library. Princeton University Press, 2010.

J.J. O'Connor and E.F. Robertson. Johann Carl Friedrich Gauss, 1996. URL https://mathshistory.st-andrews.ac.uk/Biographies/Gauss/.

S. B. Russ. A Translation of Bolzano's Paper on the Intermediate Value Theorem. *Historia Mathematica*, 5:156–185, 1980.

S. B. Russ. *The Mathematical Works of Bernard Bolzano*. Oxford University Press, New York, NY, 2004.

J. Stillwell. *Mathematics and Its History*. Undergraduate Texts in Mathematics. Springer New York, 2010. ISBN 9781441960528.

D. J. Struik. *A Source Book in Mathematics, 1200–1800*. Harvard University Press, Cambridge MA, 1969.

Notes to Instructors

PSP Content: Topics and Goals

This Primary Source Project (PSP) most naturally fits into a first course in complex variables. However, the inspiration from this project came from teaching linear algebra. In that class the fundamental theorem of algebra is used to show that an $n \times n$ matrix has n complex eigenvalues counted with multiplicity. In Otto Bretcher's textbook *Linear Algebra with Applications*, he does not prove the theorem himself, but challenges the curious reader to "check Gauss's original proof." We did, and here we are! Since the proof only requires concepts from the calculus sequence and basic properties of complex numbers, this project could certainly fit into a linear algebra course. It could also profitably be used in an Introduction to Proofs course, a capstone seminar for mathematics or secondary education majors, or as an independent study.

Student Prerequisites

- Students should be familiar with standard theorems from calculus, in particular the Intermediate Value Theorem and Extreme Value Theorem.
- Students should be familiar with polar and (ideally) cylindrical coordinates. Note that students need not have seen multivariable calculus to work through this PSP. Students can get the gist of what Gauss did with cylindrical coordinates if they are familiar enough with polar coordinates.
- Students should be comfortable writing a complex number z both as $z = re^{i\theta}$ and as $z = r\cos(\theta) + ir\sin(\theta)$.
- Students should be familiar with the following properties of the complex conjugate: $z\overline{z} = |z|^2$, $\frac{z+\overline{z}}{2} = \mathrm{Re}(z)$, $\frac{z-\overline{z}}{2i} = \mathrm{Im}(z)$, $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$, and $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$. However, as these properties only come up in the optional tasks, it is not strictly necessary that students know these properties. Alternatively, proving these properties is an easy exercise, so the instructor could simply introduce these properties along the way.
- Ideally, students should be familiar with proof by induction. If they are not, the main consequence is that the argument in Task 2 will not be fully rigorous.
- Students should understand the basics of proof by contradiction.

PSP Design and Task Commentary

This PSP is broken up into 5 sections. What follows is a brief overview of each section and what to expect as an instructor.

- Section 1 lays down preliminaries needed to understand the rest of the project. The goal of this section is to establish that, in order to prove that every nonconstant polynomial of degree m has m complex roots, it suffices to show that each such polynomial has at least one complex root. This can be glossed over pretty quickly, and the tasks can even be skipped if students are ready to accept the conclusion right away.
- In Section 2 we interpret a general polynomial as having complex domain, replacing x with $re^{i\phi}$. We then write out its corresponding real and imaginary parts in polar coordinates. The goal of this section is to reduce the task of finding a complex factor $(x - re^{i\phi})$ of a polynomial

to that of finding values of r and ϕ that simultaneously make the real and imaginary parts equal to 0. This section is relatively short and not too challenging.

- In Section 3 the real and imaginary parts of the polynomial are interpreted geometrically as surfaces in cylindrical coordinates. We then consider the curves where these surfaces intersect the zero plane. The goal of this section is to reduce our task to finding a point of intersection of these curves. The reader is guided through this section using a concrete example to ground the interpretation and calculations. Tasks 8 and 9 are important for understanding the crux of Gauss's proof in the next section (in particular, Task 9 sets up Task 21). Task 9 could be challenging and/or time consuming and instructors may wish to provide their students with more guidance. Alternatively, instructors may choose to skip task 9 while having robust discussions of Task 8 and (in the following section) Task 21. Task 14(b) requires a graphing program that can graph implicitly defined equations. This is a somewhat challenging section (especially if Task 9 is included).

- Section 4 is the heart of the project. This section is longer and more challenging than the other sections. The goal of this section is to show that the implicitly defined curves (points where the real and imaginary parts of the polynomial are equal to zero) do indeed intersect. There are several subtleties along the way, and the arguments rely on geometric reasoning as well as calculus theory. Note that Task 17 and Task 18 require the use of a graphing program and Task 19 requires use of a good calculator.

- Section 5 summarizes the work of the PSP. The section is short and contains one discussion task and one optional task.

Note that all tasks labeled as "optional" are truly optional. They are not strictly necessary for moving on with the arguments. Some may be discussed in class (depending on instructor interest), while others could be good homework problems (see the following suggestions for classroom implementation).

Suggestions for Classroom Implementation

This PSP is designed with the following implementation strategy in mind: (1) Students will complete a portion of the reading and tasks ahead of time as preparation for class; (2) Students will work in groups in class on the more challenging tasks; (3) Students will complete a portion tasks as homework problems. See the next section for a detailed sample implementation schedule.

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

Sample Implementation Schedule (based on three 75-minute class periods)

Class Period 1:

- **Before Class:** Have students read up to (and including) Section 3.1. Ask students to complete Task 1, Task 3(a), Task 6, and Task 10(a)-(b). Ideally these will be turned in ahead of class and graded for completion. Ask students to think about Task 2, Task 4, Task 8, Task 9(a)–(c), Task 10(c), and Task 12.

- **During Class:** Have students in groups begin by comparing their answers to Task 1, Task 3(a), Task 6, and Task 10(a)–(b), which they completed before class. Once they are in agreement on the answers to these tasks, ask students to work together to complete Task 2, Task 4, Task 8, Task 9(a)–(c). Depending on the class, the instructor may want to tackle Tasks 8 and 9(a)–(c) together as a whole-class discussion. Task 9 can be challenging and time consuming. Its main purpose is to set up Gauss's claims in the fourth section (in particular, Task 21). However, it is not strictly necessary for proceeding with the project, and so can be skipped if the instructor is feeling rushed for time. In this case, the instructor should make sure the discussion in Task 8 is robust. Have students return to Task 12 in small groups (perhaps have them turn in their answers to this task before the next class). End class with the discussion prompt in Task 10(c).

Class Period 2:

- **Before Class:** Have students re-read the final primary source quote in Section 3.1 and review their answers to Task 12. Have students continue reading up through Section 4.1. Ask students to complete Task 13, Task 14(a)–(b), Task 17, Task 19(a), Task 21(a), Task 23(a), and Task 24. Ideally these will be turned in ahead of class and graded for completion. Ask students to think about Task 18, Task 19(b)–(c), Task 21(b)-(c), Task 22, and Task 23(b).

- **During Class:** Have students in groups begin by comparing their answers to Task 13, Task 14(a)–(b), Task 17, Task 19(a), Task 21(a), Task 23(a), and Task 24. Once they are in agreement on the answers to these tasks, as a class (on the projector) complete Task 18 - note that this graph can be time-consuming to produce on the fly, so you may want to prepare this ahead of time. You may want to ask the students to sketch the graph in their notes. Next, have students return to their groups to complete Task 19(b)–(c) (which will require a good calculator), Task 21(b)–(c) (which will be easier if Task 9 was completed in the previous class), Task 22, and Task 23(b). The discussion of rigor can be continued at the end of class, prompted by Task 24.

Class Period 3:

- **Before Class:** Have students re-read the final primary source quote in Section 4.1 and then read through the rest of the project. Ask students to complete Task 25(a), Task 26 (which is optional), Task 27, Task 28, Task 29, and Task 31(a). Ideally these will be turned in ahead of class and graded for completion. Have students think about Task 31(b) and Task 32.

- **During Class:** Note: this day could be used to catch up on any tasks that were not covered in the previous classes, as there are fewer tasks to complete. Have students in groups begin by comparing their answers to Task 25(a), Task 26 (which is optional), Task 27, Task 28, Task 29, and Task 31(a). Once they are in agreement on the answers to these tasks, ask students to work together to complete Task 31(b). Discuss Task 32 as a class, and end with a final discussion regarding the changing norms in mathematical proofs, emphasizing that mathematics is a human endeavor.

Homework: For weekly homework assign Task 5, Task 9(d), Task 11, Task 19(d)–(e), Task 30, and any of the other optional tasks.

The actual number of class periods spent on each section naturally depends on the instructor's goals and on how the PSP is actually implemented with students.

Connections to other Primary Source Projects

The following additional projects based on primary sources are also freely available for use in a complex variables course (with the project topic listed parenthetically if this is not evident from the PSP title). Shorter PSPs that can be completed in at most 2 class periods are designated with an asterisk (*). Classroom-ready versions of each can be downloaded from https://digitalcommons.ursinus.edu/triumphs_complex.

- An Introduction to the Algebra of Complex Numbers and the Geometry in the Complex Plane, Nicholas A. Scoville and Diana White
- Euler's Square Root Laws for Negative Numbers,* Dave Ruch (multivalued functions)
- The Logarithm of -1,* Dominic Klyve
- Gauss and Cauchy on Complex Integration, Dave Ruch
- Rieman's Development of the Cauchy-Riemann Equations, Dave Ruch

Additional PSPs that are appropriate for use in an Introduction to Proofs course include the following:

- A Look at Desargues' Theorem from Dual Perspectives, Carl Lienert
  https://digitalcommons.ursinus.edu/triumphs_geometry/3/
- Stitching Dedekind Cuts to Construct the Real Numbers, Michael P. Saclolo
  https://digitalcommons.ursinus.edu/triumphs_analysis/15/
- Greatest Common Divisor: Algorithm and Proof, Mary Flagg
  https://digitalcommons.ursinus.edu/triumphs_number/10/
- Wronskians and Linear Independence: A Theorem Misunderstood by Many, Adam E. Parker
  https://digitalcommons.ursinus.edu/triumphs_differ/2

Although Gauss published little during his lifetime, he was extraordinarily productive as a creative mathematician. The following PSPs are based primarily on excerpts from his unpublished works (e.g., his Mathematical Diary or correspondence). The first three share the title "Gaussian Guesswork" and are designed for use in a second-semester calculus course; each can be used either alone or in conjunction with any of the other three to introduce students to the power of numerical experimentation and the use of analogy in discovering mathematical ideas while consolidating their understanding of the content topic listed in the subtitle. The fourth examines the problem of estimating the number of primes up to a given number x and is designed for use in a course on number theory.

- Gaussian Guesswork: Elliptic Integrals and Integration by Substitution,* Janet Heine Barnett
  https://digitalcommons.ursinus.edu/triumphs_calculus/8/
- Gaussian Guesswork: Polar Coordinates, Arc Length and the Lemniscate Curve,* Janet Heine Barnett
  https://digitalcommons.ursinus.edu/triumphs_calculus/3/
- Gaussian Guesswork: Infinite Sequences and the Arithmetic-Geometric Mean,* Janet Heine Barnett
  https://digitalcommons.ursinus.edu/triumphs_calculus/2/
- The Origin of the Prime Number Theorem,* Dominic Klyve
  https://digitalcommons.ursinus.edu/triumphs_number/9/

Additional Historical Notes

D'Alembert is another mathematician who is sometimes credited with the first fixable proof of the fundamental theorem of algebra. For a nice discussion of D'Alembert's proof and attempts to fix it using elementary methods, see Christopher Baltus's paper [Baltus, 2004]. (For the interested reader, Gauss himself gave a brief overview of D'Alembert's proof and what he takes to be its flaws in the introduction to his dissertation.)

Some mathematicians seem to think that Gauss's proof retains a significant gap insofar as he never proves the claim that "a branch of any algebraic curve which enters a limited space, must necessarily exit from this space somewhere." We disagree, given Gauss's elaboration in his footnote. However, Basu and Velleman have a nice discussion of this claim and provide their own proof of it using elementary methods in their paper [Basu and Velleman, 2017].

For background on the reasons why Bolzano and other nineteenth-century mathematicians became concerned about the use of geometrical intuition to justify theorems in analysis and other motivating factors behind the changes that occurred in standards of proof rigor at that time, see the project "Why be so Critical? 19th Century Mathematics and the Origins of Analysis" (author Janet Heine Barnett), available at https://digitalcommons.ursinus.edu/triumphs_analysis/1/.

Recommendations for Further Reading

We suggest reading the first 12 sections of Gauss's dissertation, in which he outlines previous attempts at the proof and reasons why they are all inadequate.

We also recommend reading Sections 14.6 and 14.7 of Stillwell's *Mathematics and Its History, 3rd ed.* in which he discusses the Fundamental Theorem of Algebra and compares Gauss and D'Alembert's proofs.

Acknowledgments

We would like to heartily thank Samuel Goodman for his help in working through the mathematics in Gauss's dissertation.

The development of this student project has been partially supported by the TRansforming Instruction in Undergraduate Mathematics via Primary Historical Sources (TRIUMPHS) Program with funding from the National Science Foundation's Improving Undergraduate STEM Education Program under Grant Nos. 1523494, 1523561, 1523747, 1523753, 1523898, 1524065, and 1524098. Any opinions, findings, and conclusions or recommendations expressed in this project are those of the author and do not necessarily represent the views of the National Science Foundation.

For more information about TRIUMPHS, visit https://blogs.ursinus.edu/triumphs/.