Chapter 8
The Fundamental Theorem of Algebra

The Fundamental Theorem of Algebra (stated below) provides an ideal case study for illustrating the roles of alternative proofs in mathematical practice. Like the Pythagorean Theorem, the Fundamental Theorem of Algebra has been proved in many different ways since its enunciation by Euler in 1739. Unlike the Pythagorean Theorem, however, early attempts to prove the Fundamental Theorem of Algebra are not shrouded in the mists of antiquity, so we know how the adequacy of those attempts was evaluated by mathematicians of the time. We can see how criticisms of earlier efforts to prove the theorem led to alternative proof strategies, and we can analyze why the proof given by Gauss in his 1799 inaugural dissertation was the first to be accorded general acceptance, though it too would later be deemed not fully rigorous.

As with the theorems considered in earlier chapters, besides questions of rigor there have been other impetuses for devising alternative proofs of the Fundamental Theorem of Algebra: issues of perspicuity, simplicity, generality, purity of method and constructivity have also been matters of concern; and in a pedagogical context, different proofs of the Fundamental Theorem have been employed as a vehicle for introducing a variety of topics in higher-level mathematics (complex line integrals, field extensions, Galois theory, and notions from algebraic topology) in a text designed for a capstone course for senior mathematics majors (Fine and Rosenberger 1997).

8.1 Alternative formulations of the theorem

In its earliest and simplest form, the Fundamental Theorem of Algebra was the conjecture that every polynomial with real coefficients can be expressed as a product of linear and quadratic polynomials with real coefficients. The question whether that is so arose in connection with Leibniz's attempts to integrate functions by the method of partial fractions, and Leibniz himself believed the conjecture to be false.

J.W. Dawson, Jr., *Why Prove it Again?*, DOI 10.1007/978-3-319-17368-9_8

Euler, however, showed that putative counterexamples put forward by Leibniz and by Nikolaus Bernoulli did, in fact, possess factorizations of the stated form, and it was he, in a letter to Bernoulli of 1 October 1742, who first asserted the truth of the statement. Two months later, however, in a letter of 15 December to his friend Goldbach, he confessed that he was unable to produce a fully satisfactory proof of the theorem.[1]

Today the Fundamental Theorem of Algebra is more often stated in the form "Every polynomial $p(X)$ of degree n with complex coefficients possesses exactly n complex roots, counting multiplicities." The equivalence of that statement with the one given above rests not only upon recognizing complex numbers as meaningful entities, but upon the quadratic formula (which shows how to express any quadratic polynomial as a product of two complex linear factors), upon the factor theorem of Descartes (that a is a root of a polynomial $p(X)$ if and only if $X - a$ is a factor of $p(X)$), and upon the observation (made by Bombelli around 1560, and again by Euler in his 1742 letter to Goldbach) that the complex roots of any polynomial with real coefficients always occur in conjugate pairs, so that the product of the corresponding linear factors guaranteed by Descartes's theorem is a *real* quadratic polynomial.

Consideration of the properties of complex conjugates shows that if $p(X)$ is a polynomial with complex coefficients and $\overline{p}(X)$ is the polynomial whose coefficients are the conjugates of those of $p(X)$, then the product $p(X)\overline{p}(X)$ is a polynomial with *real* coefficients. If z_0 is a complex root of $p(X)\overline{p}(X)$, then it must either be a root of $p(X)$ or of $\overline{p}(X)$; and in the latter case, again using the overbar to denote complex conjugation, $\overline{\overline{p}(z_0)} = \overline{\overline{p}}(\overline{z_0}) = p(\overline{z_0}) = 0$, so $\overline{z_0}$ is a root of $p(X)$. To prove the Fundamental Theorem it therefore suffices to establish it for polynomials $p(X)$ whose coefficients are real numbers.

8.2 Early attempts to prove the theorem

The task of showing that every polynomial with real coefficients possesses at least one complex root involves two separate aspects: showing (1) that a root *of some definite sort* exists, and (2) that any such root must in fact be of the form $a + bi$ (in modern terms, proving the existence of a splitting field K over $\mathbb{R}$ for $p(X)$, and then showing that K must be isomorphic to $\mathbb{C}$). Before Gauss, however, those who endeavored to prove the existence of complex roots either explicitly *assumed* (1) to be true (in part, perhaps, because it was believed that formulas for the roots of polynomials of degree ≥ 5 similar to those obtained by Ferro, Tartaglia, Ferrari, and

[1]The dates given here for Euler's letters are based on the account in Kline (1972), pp. 597–598. They disagree with those given in Remmert (1990), which are inconsistent with one another.

Bombelli for cubic and quartic polynomials would eventually be found[2]), or else unwittingly employed arguments that smuggled in that assumption. Their efforts focused instead on establishing (2). But, as Gauss trenchantly observed in the critique of prior proof attempts that he gave in his dissertation, there was need not only to *justify* the existence of roots, but, if algebraic operations were to be applied to them, to characterize their structure; for it made no sense to attempt to manipulate hypothetical quantities that were mere "shadows of shadows."

The remainder of this section is devoted to outlining the strategies employed by D'Alembert, Euler, Lagrange, and Laplace in their attempted proofs of the Fundamental Theorem (published in 1746, 1749, 1772 and 1795, respectively) and to analyzing the deficiencies in their arguments.

d'Alembert's 'proof': In his memoir 'Recherches sur le calcul intégral' (d'Alembert 1746), Jean Le Rond d'Alembert is generally credited with having made the first serious attempt to prove the Fundamental Theorem of Algebra. The memoir was apparently hastily written, however, and is not notable for its clarity. Indeed, there is marked disparity among the descriptions given by modern commentators both of the mechanics of d'Alembert's argument and of the extent of its deficiencies. (My own reading of d'Alembert's text is in accord with the descriptions of it in Gilain (1991) and Baltus (2004), but at variance with that in Remmert (1990).) d'Alembert began by noting that if $p(X) = X^m + c_{m-1}X^{m-1} + \cdots + c_1X + c_0$ is a monic polynomial with real coefficients of degree $m \geq 1$, then $p(0) = 0$ if $c_0 = 0$. He then replaced the constant term c_0 by the *parameter* z and set the resulting function $F(X, z)$ equal to 0, so that the Fundamental Theorem became the statement that for any real value of z, there is a (possibly complex) value x for which $F(x, z) = 0$. To establish that, d'Alembert first claimed that for any real number z_0, if (x_0, z_0) is a point for which $F(x_0, z_0) = 0$ — in particular, if $x_0 = z_0 = 0$ — then for all real z sufficiently close to z_0, there is a complex value x for which $F(x, z) = 0$. He then went on to claim that for any real value z^* of z, an overlapping chain of discs can be found, starting at $(0, 0)$, yielding a sequence of points (x_n, z_n) such that $F(x_n, z_n) = 0$ for each n, the values x_n converge to a complex number x^* and the (real) values z_n to z^*, with $F(x^*, z^*) = 0$.

To establish the first claim d'Alembert alleged, without proof, that if

$$F(x_0, z_0) = 0,$$

then for all z sufficiently close to z_0, there is a natural number q and a convergent series of fractional powers of $z - z_0$ such that

[2]The impossibility of expressing the roots of arbitrary polynomials of degree ≥ 5 in terms of radicals was finally established by Abel in 1826.

$$x = x_0 + \sum_{i=1}^{\infty} c_i (z - z_0)^{i/q}$$

satisfies $F(x, z_0) = 0$. More than a century later that fact was finally proved by Victor Puiseaux, as a *consequence* of the Fundamental Theorem (which by then had been rigorously proved by other means); so d'Alembert's argument was circular. There were difficulties, too, with his second claim: What ensures that the radii of the discs are such that the z_n converge to z^* ? And even if they do, why does the fact that $F(x_n, z_n) = 0$ for each n entail that $F(x^*, z^*) = 0$? Those and other criticisms were lodged against d'Alembert's argument by Gauss, who nevertheless thought that it might be possible to repair its defects. But he and others chose instead to seek different ways to establish the Fundamental Theorem.[3]

Euler's attack on the theorem: Three years after the appearance of d'Alembert's memoir, Euler attempted to prove the Fundamental Theorem in its original formulation. By invoking the fact that a real polynomial of odd degree must have a real root (a consequence of the intermediate-value theorem, a principle generally accepted at the time, but first rigorously proved by Bolzano around 1816), he argued that a real quintic polynomial must have at least one real linear factor, and then went on to show how any real quartic polynomial could be expressed as a product of two real quadratic factors. Having thus established the truth of the Fundamental Theorem for polynomials of degree ≤ 5, he attempted to extend the proof to polynomials of higher degree, but was unable to do so.

At first glance it might appear that Euler had made but a minor advance beyond the work of Ferrari and Bombelli two centuries earlier, since their explicit formulas for the roots of real quartic polynomials, in which the complex roots occur in conjugate pairs, immediately entail that all such quartics can be factored into a product of real polynomials of degree at most 2. But in the formulas obtained by the Italians for the roots of cubic and quartic polynomials, complex numbers play an essential role. In its original form, however, the Fundamental Theorem makes no reference to complex numbers, so, as noted in Remmert (1990), p. 117, their employment in proofs thereof appears to invoke a *deus ex machina*. Euler's method of factoring quartics, however, made no use of complex numbers, so from the standpoint of purity of method it was superior.

A detailed and very readable discussion of what Euler did in his paper Euler (1749) is given in Dunham (1991). Following the lead of the Italian school, Euler noted that any monic polynomial of degree 4 in the variable x with real coefficients can be converted into an equivalent quartic in the variable y that lacks a cubic term (via the substitution $x = y - c_3/4$, where c_3 is the coefficient of x^3 in the original polynomial). The factorization of the resulting quartic $y^4 + By^2 + Cy + D$ then

[3] An attempt to repair d'Alembert's proof is given in Baltus (2004), but that effort, too, appears to be flawed.

depends on the values of the coefficients B, C and D. If $C = 0$, the quartic is a quadratic in y^2, which, if $B^2 - 4D \geq 0$, factors as

$$\left[y^2 + \frac{B + \sqrt{B^2 - 4D}}{2}\right]\left[y^2 + \frac{B - \sqrt{B^2 - 4D}}{2}\right].$$

On the other hand, if $B^2 - 4D < 0$, then both $\sqrt{D} > 0$ and $\sqrt{2\sqrt{D} - B} > 0$, so

$$y^4 + By^2 + D = \left[y^2 + \sqrt{D}\right]^2 - \left[y\sqrt{2\sqrt{D} - B}\right]^2,$$

a difference of two squares that factors once again into the product of two quadratics. If $C \neq 0$, the absence of the y^3 term in $y^4 + By^2 + Cy + D$ implies that any factorization of that quartic into quadratic factors must be of the form $(y^2 + uy + \alpha)(y^2 - uy + \beta)$ for some constants u, α and β. Expanding that product, setting it equal to $y^4 + By^2 + Cy + D$ and equating coefficients of like powers of y, Euler obtained three equations in the unknowns u, α and β, from which after further algebra he deduced the equation $u^6 + 2Bu^4 + (B^2 - 4D)u^2 - C^2 = 0$, in which all the powers of u are even. The graph of $Y = u^6 + 2Bu^4 + (B^2 - 4D)u^2 - C^2$ is therefore symmetric about the Y-axis, Y approaches $+\infty$ as u approaches $\pm\infty$, and $Y(0) < 0$ (see Figure 8.1), so by the intermediate-value principle, $Y = u^6 + 2Bu^4 + (B^2 - 4D)u^2 - C^2$ must have real roots $\pm u_0$, either of which can be substituted back into the earlier equations to find real values for α and β.

To extend to polynomials of arbitrary degree, Euler noted that by multiplying, if necessary, by some positive integral power of X, any polynomial $p(X)$ of degree

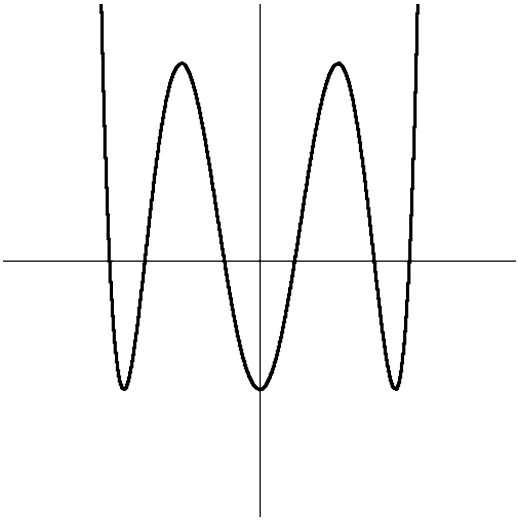

Fig. 8.1 Graph of a sixth-degree polynomial with no odd powers

d could be converted into a polynomial $q(X)$ of degree 2^m, for some $m > 0$. He then attempted to mimic the procedure he had employed for factoring quartics. A direct approach led to systems of equations too complex to allow derivation of an equation for u, but he showed that an alternative approach, avoiding the need to find an explicit equation for u_0, was also possible in the quartic case. However, to extend that approach to polynomials of degree 2^m for $m > 2$ it was necessary to *assume* that $2m$ roots of some sort existed; and without specifying the nature of those roots, Euler's attempt to show that algebraic combinations of them would yield *real* coefficients for the putative factors of $q(X)$ was doomed to failure (as Gauss was to point out).

Lagrange's improvement of Euler's argument: In a long and important paper that appeared in 1770/1771,[4] Joseph Louis Lagrange investigated the properties of symmetric polynomials and established the result now known as the fundamental theorem about them: that any polynomial $S(X_1, \ldots, X_n)$ symmetric in $X_1, \ldots, X_n$ has a unique representation as a polynomial $P(s_1, \ldots, s_n)$, where $s_1, \ldots, s_n$ are the elementary symmetric polynomials in $X_1, \ldots, X_n$. Using that result, and assuming that a real polynomial $p(X)$ of degree 2^m had 2^m roots that could be manipulated like ordinary real numbers (in modern terms, that $p(X)$ had roots in some field extending $\mathbb{R}$), Lagrange was able to establish (even to Gauss's satisfaction) that the factors Euler had sought for $p(X)$ would indeed have real coefficients. Only the justification for the existence of such roots remained to be proven.

Laplace's proof: Under the same basic assumptions that Euler and Lagrange had made (the existence of a splitting field and the intermediate-value principle), together with DeMoivre's theorem on roots of complex numbers, proved earlier in the eighteenth century, Pierre Simon de Laplace employed Lagrange's theorem on symmetric polynomials to prove the Fundamental Theorem of Algebra in its second formulation. A version of his proof in modern terminology, as given in Remmert (1990), pp. 120–122, goes as follows.[5]

Suppose a monic polynomial $p(X)$ of degree $n \geq 1$ with real coefficients has roots $r_1, \ldots, r_n$ in some splitting field F over $\mathbb{R}$, and rewrite n as $2^m q$, where q is odd. The proof proceeds by induction on m. If $m = 0$, $p(X)$ has a real root by the intermediate-value principle. For $m \geq 1$, suppose that every polynomial of degree $n = 2^k q$ with $k < m$ has a complex root. Laplace then considered the symmetric polynomials over F given by

$$L_t(X) = \prod_{1 \leq i < j \leq n} (X - r_i - r_j - t r_i r_j),$$

[4]"Réflexions sur la résolution algébrique des équations," reprinted in *Oeuvres de Lagrange* III, 205–421)

[5]Full background details can be found in Chapter 6 of Fine and Rosenberger (1997), where, however, the strategy underlying the proof is not credited to Laplace.

for each positive integer t. By the Fundamental Theorem on Symmetric Polynomials, each L_t, when written as a polynomial in powers of X, has coefficients that are elementary symmetric polynomials in the roots of the *real* polynomial $p(X)$. But the coefficient of each power X^j in $p(X)$ *is* just $(-1)^j s_j(r_1, \ldots, r_n)$, where $s_j(r_1, \ldots, r_n)$ is the jth elementary symmetric polynomial in $r_1, \ldots, r_n$; so each coefficient of $L_t(X)$ is a real number. Moreover, each $L_t(X)$ has degree

$$\binom{n}{2} = \binom{2^m q}{2} = 2^{m-1} q[2^m q - 1],$$

where $q[2^m q - 1]$ is odd. By the induction hypothesis, each $L_t(X)$ thus has a complex root c_t, so for some pair (r_i, r_j) with $1 \le i < j \le n$, $c_t = r_i + r_j + t r_i r_j$; and since there are infinitely many integers t but only finitely many pairs (r_i, r_j) with $1 \le i < j \le n$, there must be *distinct* integers t_1 and t_2 such that for the *same* i and j, $c_{t_1} = r_i + r_j + t_1 r_i r_j$ and $c_{t_2} = r_i + r_j + t_2 r_i r_j$ are both complex numbers. The difference $c_{t_1} - c_{t_2} = (t_1 - t_2) r_i r_j$ is then also a complex number, whence so is $r_i r_j$. Therefore $c_{t_1} - t_1 r_i r_j = r_i + r_j$ is a complex number as well. So by DeMoivre's theorem and the quadratic formula, the roots of the polynomial $X^2 - (r_i + r_j)X + r_i r_j = (X - r_i)(X - r_j)$, that is, the roots r_i and r_j of the original polynomial $p(X)$, must be complex numbers. q.e.d.

8.3 Gauss's first proof

Gauss's doctoral dissertation, submitted to the University of Helmstedt in 1799 and written in Latin, was entitled *Demonstratio nova theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse*[6] — that is, "New proof of the theorem that every rational integral algebraic function [i.e, polynomial] of one variable can be resolved into real factors of first or second degree." Gauss thus stated the Fundamental Theorem in its original formulation, and declared his aim to be that of giving "a new and stronger proof" of that result. On the second page of the dissertation he noted the equivalent formulation in terms of complex roots, but stated that he would eschew the use of complex numbers in his demonstration. He went on to criticize earlier proofs, all of which he faulted for presuming without justification that a polynomial of degree m must possess m roots of some (unspecified) sort. To avoid that presumption, he gave a geometric argument to establish the desired factorization.

[6]Reprinted in Gauss's *Werke* III, 1-30. The discussion here is based on the German translation by E. Netto in Gauss (1890).

Outline of proof: Gauss begins by proving two lemmas.

Lemma 1: If m is any positive integer, then $x^2 - 2r\cos\phi x + r^2$ is a factor of $\sin\phi x^m - r^{m-1}\sin(m\phi)x + r^m\sin(m-1)\phi$.

(The latter expression is 0 if $m = 1$. If $m = 2$, the other factor is $\sin\phi$, and if $m > 2$, $\sum_{i=1}^{m-1}\sin(i\phi)r^{i-1}x^{m-i-1}$ is the other factor.)

Lemma 2: If r and ϕ satisfy the equations

$$(8)\quad r^m\cos(m\phi) + Ar^{m-1}\cos(m-1)\phi + Br^{m-2}\cos(m-2)\phi + \ldots + Kr^2\cos(2\phi) + Lr\cos\phi + M = 0$$

and

$$(9)\quad r^m\sin(m\phi) + Ar^{m-1}\sin(m-1)\phi + Br^{m-2}\sin(m-2)\phi + \ldots + Kr^2\sin(2\phi) + Lr\sin\phi = 0,$$

then the expression $x^m + Ax^{m-1} + Bx^{m-2} + \cdots + Kx^2 + Lx + M$ has the factor $x - r\cos\phi$ if $r\sin\phi = 0$ and the factor $x^2 - (2r\cos\phi)x + r^2$ if $r\sin\phi \neq 0$. (Gauss notes that complex numbers are usually invoked to prove Lemma 2, but he gives an alternative proof that avoids them, based on Lemma 1.)

To prove the Fundamental Theorem it therefore suffices to show that r and ϕ can be found that satisfy the two equations of Lemma 2.[7]

Toward that end, Gauss considers the surfaces generated by the functions

$$T = r^m\sin(m\phi) + Ar^{m-1}\sin(m-1)\phi + Br^{m-2}\sin(m-2)\phi + \ldots + Kr^2\sin(2\phi) + Lr\sin\phi$$

and

$$U = r^m\cos(m\phi) + Ar^{m-1}\cos(m-1)\phi + Br^{m-2}\cos(m-2)\phi + \ldots + Kr^2\cos(2\phi) + Lr\cos\phi + M$$

[7]The connection with the complex formulation of the theorem is readily seen, since by DeMoivre's Theorem, if the variable X is written in polar form as $X = r(\cos\phi + i\sin\phi)$, the left members of those equations are just the real and imaginary parts of the expression $X^m + Ax^{m-1} + Bx^{m-2} + \cdots + Kx^2 + Lx + M$.

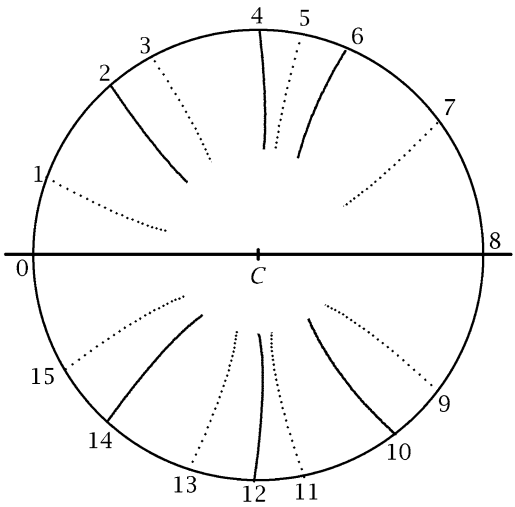

Fig. 8.2 Alternating T- and U- arcs

above and below the (r, ϕ)-plane and the traces in that plane where those surfaces intersect it.[8] The problem then becomes that of showing that there is at least one point in the (r, ϕ)-plane where the T-trace and the U-trace themselves intersect.

Further analysis shows that the T- and U-traces each contain $2m$ arcs that extend to infinity. Two arcs of the T-trace join to form the horizontal axis. The other arcs are each asymptotic to lines where $\sin(m\phi) = 0$, that is, to lines through the origin that are inclined to the axis at one of the angles $k\pi/m$, for $0 < k < m$. The arcs of the U-trace are likewise asymptotic to lines where $\cos(m\phi) = 0$, that is, to lines through the origin that are inclined to the axis at one of the angles $(2k-1)\pi/2m$, for $0 < k \leq m$. Accordingly, those arcs will intersect a circle of sufficiently large radius at $2m$ points, which divide its circumference into $2m$ intervals in which T is alternately positive and negative. Moreover, those points are alternately one where a T-arc intersects the circle and one where a U-arc does. (See Figure 8.2, based on Gauss's own illustration for the quartic polynomial $X^4 - 2X^2 + 3X + 10$. The solid curves there represent the T-arcs and the dotted ones the U-arcs.)

Assuming that for at least one k the arc of the T-trace that intersects the circle at point k and the arc of the T-trace that intersects the circle at point $k + 2$ are both part of the *same* continuous T-branch, and likewise that the arc of the U-trace that intersects the circle at point $k + 1$ and the arc of the U-trace that intersects the circle at point $k + 3$ are both part of the same continuous U-branch, then that U-branch,

[8] Since the leading terms of T and U dominate the others and can be made positive or negative by appropriate choice of the angle ϕ, it is clear by continuity that both surfaces do intersect the (r, ϕ)-plane.

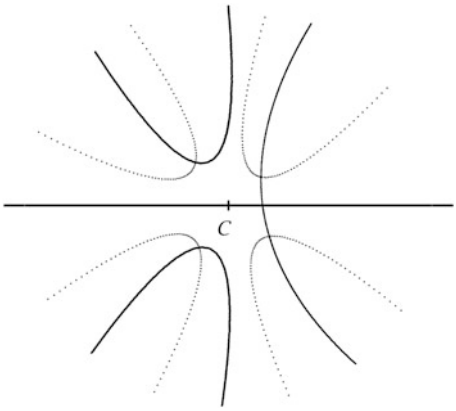

Fig. 8.3 Intersecting T- and U-branches

in passing from one intersection point where T is positive to another where it is negative, must at some point within the circle cross that T-branch. (See Figure 8.3, also taken from Gauss's original text.) The theorem would thereby be proved.

Gauss declared that the assumption involved could be justified in "many different ways," one of which he endeavored to outline. But ultimately he had to rely on his geometric intuition that "an algebraic curve can neither suddenly end abruptly ... nor lose itself, so to speak, ... after an infinity of circuits (as in the case of a logarithmic spiral)" — a fact that he believed could "be taken as [having been] sufficiently securely established" (Gauss 1890, footnote to p. 33).

Gauss's contemporaries evidently agreed, for they found no fault with his proof. Only much later — long after Bolzano's proof of the intermediate-value theorem, and after Kronecker, Dedekind and others, by relatively straightforward means, had shown how to construct splitting fields and thereby justified the earlier proofs of the Fundamental Theorem that had relied on those facts — did mathematicians come to regard the principle that Gauss had relied on (that a non-compact branch of an algebraic curve that enters a bounded space must eventually emerge from it) as a statement (like the Jordan Curve Theorem) that required more rigorous demonstration. The principle was finally proved rigorously by Alexander Ostrowski in 1920, using sophisticated topological notions. (See Ostrowski 1983.)

Fifty years after his receipt of the doctorate, Gauss gave another proof of the Fundamental Theorem (his fourth) that was a minor variant of the one given in his dissertation.[9] Since (as he remarked) complex numbers had by then come to be generally accepted by the mathematical community (in large part due to the Fundamental Theorem itself), he felt free to employ them in the revised version of his proof; and there he also allowed the polynomials to have complex coefficients.

[9] A detailed exposition of a modernized version of Gauss's fourth proof is given in Fine and Rosenberger (1997), pp. 182–186.

8.4 Argand's proof

The assertion that every nonconstant polynomial with complex coefficients must have a complex root was first made not by Gauss, but by Jean Robert Argand, a Paris bookkeeper who introduced the planar representation of complex numbers now named after him;[10] and in 1814 Argand gave his own, very different proof of that result (Argand 1814), for whose understanding nothing beyond some basic knowledge of advanced calculus is required.

Specifically, Argand's proof depends on knowing (1) that polynomials are continuous functions; (2) that every continuous function defined on a closed disc $|z| \leq R$ assumes a minimum in that disc; and (3) that every complex number has a kth root for each integer $k > 1$ (an immediate consequence of DeMoivre's Theorem).[11] The proof then proceeds as follows:

Given a polynomial $p(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ with $n \geq 1$, $|p(z)|$ approaches ∞ as $|z|$ does, so for any positive constant C, there is an $R > 0$ such that $|p(z)| > C$ for $|z| > R$. Taking $C = \inf_{z\in\mathbb{C}} |p(z)|$, it follows that $\inf_{z\in\mathbb{C}} |p(z)| = \inf_{z\leq R} |p(z)|$, so by (2), $|p((z)|$ assumes a minimum for some z_0 with $|z_0| \leq R$. That $p(z_0) = 0$ then follows directly from

Argand's inequality: For any polynomial $p(z)$ of degree $n \geq 1$, if $p(z_0) \neq 0$ then there is a $z_1 \in \mathbb{C}$ for which $|p(z_1)| < |p(z_0)|$.

Sketch of proof: Since $p(z_0) \neq 0$, we can divide $p(z)$ by $|p(z_0)|$ to obtain a polynomial $q(z)$ of the same degree for which $q(z_0) = 1$; and if we define $h(z) = q(z + z_0)$, then $h(0) = 1$, so $h(z)$ may be written as

$$\begin{aligned} 1 + b_1 z + b_2 z^2 + \cdots + b_n z^n &= 1 + b_k z^k + b_{k+1} z^{k+1} + \cdots + b_n z^n \\ &= 1 + b_k z^k + z^k [b_{k+1} z + \cdots + b_n z^{n-k}], \end{aligned}$$

where k is the least index i for which $b_i \neq 0$. The expression in brackets is a polynomial that has $z = 0$ as a root, so it is continuous at $z = 0$ and therefore can be made arbitrarily small for z sufficiently close to 0. If r is any kth root of $-1/b_k$, the triangle inequality then shows that $|h(rt)| < 1$ for sufficiently small positive real numbers t. For any such t, setting $z_1 = z_0 + rt$ yields $|p(z_1)| < |p(z_0)|$.

[10] See the entry on Argand by Phillip Jones in the *Dictionary of Scientific Biography*, vol. 1, pp. 237–240.

[11] Argand assumed fact (2), which was not proved rigorously until later.

8.5 Gauss's second proof

Despite the acceptance of his 1799 proof by other mathematicians of his time, Gauss published a second proof in 1815[12] that was based on algebraic rather than geometric principles. In his opening remarks, he maintained that his first proof "probably [*sic*] leaves nothing more to be desired with respect to rigor or simplicity".[13] He deemed his new proof to be "no less rigorous" than the first, but he did not claim it was simpler (as it certainly was not). Why then did he offer it?

Perhaps, as the qualifying *wohl* might be taken to suggest, he did after all harbor some doubts about whether the geometric principles he had invoked in his first proof had been rigorously established; but at least ostensibly, he was concerned about purity of method.

Like the proofs of Laplace and Lagrange, Gauss's 1815 proof was by induction on the highest power of 2 dividing the degree of the polynomial $Y(x)$. But Gauss avoided having to assume that $Y(x)$ had any roots by working in terms of indeterminates $a_1, a_2, \ldots, a_n$. He defined various auxiliary polynomials symmetric in those indeterminates, and applied the Fundamental Theorem on Symmetric Polynomials to each.[14]

The very long proof is divided into twenty numbered sections. To make the argument self-contained, Gauss first established a number of preliminary results: Given two polynomials $Y_1(x)$ and $Y_2(x)$, whose coefficients might include other indeterminates in addition to x, he defined their greatest common divisor and used the Euclidean algorithm to show that the g.c.d. must be a linear combination of $Y_1(x)$ and $Y_2(x)$, so that, in particular, if $Y_1(x)$ and $Y_2(x)$ have no common divisor of positive degree, there must be polynomials $Z_1(x)$ and $Z_2(x)$ such that $Z_1(x)Y_1(x) + Z_2(x)Y_2(x) = 1$, and conversely. In section 3, for any positive integer m he defined the elementary symmetric polynomials in the indeterminates $a_1, a_2, \ldots, a_m$ and noted that any polynomial function of *them* (again, possibly containing other indeterminates as well) must also be symmetric in $a_1, a_2, \ldots, a_m$. Section 4 was devoted to proving the converse (the fundamental theorem), and section 5 to establishing the uniqueness of that representation.

Then, for any integer $m \geq 1$, he defined π_m to be the symmetric polynomial in the indeterminates $a_1, a_2, \ldots, a_m$ given by

$$\pi_m = \prod_{\substack{1 \leq i,j \leq m \\ i \neq j}} (a_i - a_j) = (-1)^{\binom{m}{2}} \prod_{\substack{1 \leq i,j \leq m \\ i < j}} (a_i - a_j)^2. \tag{10}$$

[12] Translated into German on pp. 37–60 of Gauss (1890).

[13] "*in Anbetracht der Strenge wie der Einfachheit wohl nichts zu wünschen übrig lässt*"

[14] In the description of Gauss's proof given below we depart from his notation in some respects. In particular, we use subscripts rather than primes or distinct letters to distinguish among different objects of the same type. For an overview in English that retains Gauss's notation, see Baltus (2000).

By the fundamental theorem, π_m can be represented as a polynomial in the elementary symmetric polynomials $s_1, s_2, \ldots s_m$ of the indeterminates $a_1, a_2, \ldots, a_m$. Let p_m be the polynomial in the indeterminates $i_1, i_2, \ldots, i_m$ obtained from the latter polynomial by replacing each occurrence of s_j therein by i_j, for $1 \leq j \leq m$.[15] Then, given a monic polynomial $Y(x) = x^m - C_1 x^{m-1} + C_2 x^{m-2} - \cdots \pm C_m$ of degree m with real coefficients $C_1, -C_2, \ldots, \pm C_m$, let P_Y be the real number obtained by substituting for each i_j in p_m the value C_j from Y (that is, P_Y is the value of $p_m(i_1, i_2, \ldots, i_m)$ at $(C_1, C_2, \ldots, C_m)$).

Sections 6–9 of Gauss's paper are devoted to proving

Lemma 1: If $Y'(x)$ is the (formal) derivative of $Y(x)$, then $Y(x)$ And $Y'(x)$ have a nonconstant common factor if and only if $P_Y = 0$.

Gauss began by noting that *if* $Y(x)$ could be decomposed into (not necessarily distinct) linear factors, say

$$Y(x) = (x - A_1)(x - A_2) \cdots (x - A_m),$$

then

$$Y(x) = x^m - (\sum_{1 \leq i \leq m} A_i) x^{m-1} + (\sum_{\substack{1 \leq i,j \leq m \\ i \neq j}} A_i A_j) x^{m-2} - \cdots \pm \prod_{1 \leq i \leq m} A_i ;$$

that is, for each $1 \leq i \leq m$, $C_i = s_i(A_1, A_2, \ldots, A_m)$. Hence

$$\begin{aligned} P_Y &= p_m\left(s_1(A_1, A_2, \ldots, A_m), s_2(A_1, A_2, \ldots, A_m), \ldots, s_m(A_1, A_2, \ldots, A_m)\right) \\ &= \prod_{\substack{1 \leq i,j \leq m \\ i \neq j}} (A_i - A_j), \end{aligned}$$

so $P_Y = 0$ if and only if for some distinct i, j, $A_i = A_j$. It then follows directly from the product rule, applied to the factored form of $Y(x)$, that the latter condition holds if and only if $Y(x)$ and $Y'(x)$ have some factor $(x - A_i)$ in common.

To avoid circularity in proving the Fundamental Theorem, however, it was necessary to prove the Lemma without presuming that $Y(x)$ could be decomposed into linear factors. To do so, Gauss noted that any monic polynomial

$$Y(x) = x^m - C_1 x^{m-1} + C_2 x^{m-2} - \cdots \pm C_m$$

[15] Gauss called p_m the *determinant* of the polynomial $y(x) = x^m - i_1 x^{m-1} + i_2 x^{m-2} - \cdots \pm i_m$; today, it is called the *discriminant*. He described it as "that function of the indeterminates $i_1, \ldots, i_m$ that is transformed into the product of every pair of differences of distinct indeterminates $a_1, \ldots, a_m$ when each i_k is replaced by s_k." Note that p_m depends only on m.

in the indeterminate x could be regarded as a substitution instance of the polynomial

$$y(x, i_1, \ldots, i_m) = x^m - i_1 x^{m-1} + i_2 x^{m-2} - \cdots \pm i_m, \tag{11}$$

in which the indeterminates $i_1, \ldots, i_m$ had been replaced by the real numbers $C_1, \ldots, C_m$. If, on the other hand, those indeterminates were replaced by the elementary functions $s_1, \ldots, s_m$ of the indeterminates $a_1, \ldots, a_m$, the resulting polymonial v could be written as

$$v = \prod_{i=1}^{m} (x - a_i). \tag{12}$$

As a tool for proving the 'only if' direction of the Lemma, Gauss considered the polynomial in the indeterminates $x, a_1, a_2, \ldots, a_m$, symmetric in $a_1, a_2, \ldots, a_m$, given by

$$\rho = \pi_m \sum_{i=1}^{m} \prod_{\substack{1 \le j \le m \\ j \ne i}} \frac{(x - a_j)}{(a_i - a_j)^2}$$

(That ρ *is* a polynomial in $x, a_1, a_2, \ldots, a_m$ follows from the rightmost member of equation (10), which shows that the denominator in each summand of ρ divides π_m.) If $1 \le k \le m$ and $x = a_k$, then only one summand of ρ and one of the derivatives v' is non-zero (in each case, that for which $i = k$), and $\rho v' = \pi_m$. Thus for each integer k between 1 and m, $x - a_k$ is a factor of $\pi_m - \rho v'$, so v divides $\pi_m - \rho v'$.

The quotient, σ, is then another polynomial in $x, a_1, a_2, \ldots, a_m$ symmetric in all the a_i. Applying the Fundamental Theorem on Symmetric Polynomials to each term in the equation $\pi_m = \sigma v + \rho v'$ and replacing each elementary symmetric polynomial s_j therein by the indeterminate i_j produces the equation

$$p_m = s(x)y(x) + r(x)y'(x), \tag{13}$$

where $r(x)$ and $s(x)$ are polynomials in $x, i_1, i_2, \ldots, i_m$, and $y(x)$ is the polynomial defined in (11). Replacing each i_j in (13) by the real number C_j then yields $P_Y = S(x)Y(x) + R(x)Y'(x)$, whose left member is a real number and whose right member is a polynomial in x. If $P_Y \ne 0$, division by P_Y yields $1 = \frac{S_Y(x)}{P_Y} Y(x) + \frac{R_Y(x)}{P_Y} Y'(x)$; so $Y(x)$ and $Y'(x)$ have no nonconstant common factor unless $P_Y = 0$.

Conversely, if $Y(x)$ and $Y'(x)$ have a nonconstant common factor, there are functions $f(x)$ and $\phi(x)$ such that $f(x)Y(x) + \phi(x)Y'(x) = 1$. Moving the left member of that equation to the right and adding $f(x)v + \phi(x)v'$ to both sides then gives

$$f(x)v + \phi(x)v' = 1 + f(x)(v - Y(x)) + \phi(x)(v - Y(x))'. \tag{14}$$

Gauss abbreviates the expression $f(x)(y(x) - Y(x)) + \phi(x)(y(x) - Y(x))'$, a polynomial in the indeterminates $x, i_1, \ldots, i_m$, by $F(x, i_1, \ldots, i_m)$, and the right member of (14), regarded as a polynomial in x and the elementary symmetric polynomials $s_1, \ldots, s_m$, by $1 + F(x, s_1, \ldots, s_m)$. Similarly, he uses $F(x, C_1, \ldots, C_m)$ to denote the result of replacing each i_k in $F(x, i_1, \ldots, i_m)$ by C_k; and since $y(x)$ is thereby transformed into $Y(x)$, it follows that for any value of x,

$$F(x, C_1, \ldots, C_m) = 0. \tag{15}$$

Gauss next applies the product rule to v, which, for any j between 1 and m, yields

$$v' = \prod_{\substack{1 \le i \le m \\ i \ne j}} (x - a_i) + (x - a_j)\left(\prod_{\substack{1 \le i \le m \\ i \ne j}} (x - a_i) \right)'. \tag{16}$$

Replacing v' in (14) by the expression on the right of (16) and setting $x = a_j$ successively for each j between 1 and m then gives the sequence of equations

$$\phi(a_j) \prod_{\substack{1 \le i \le m \\ i \ne j}} (x - a_i) = 1 + F(a_j, s_1, \ldots, s_m) \qquad (1 \le j \le m).$$

Since the expressions on either side of each equation in that sequence are polynomials symmetric in the indeterminates $a_1, \ldots, a_m$, the same is true of the product of those equations:

$$\pi_m \phi(a_1)\phi(a_2) \ldots \phi(a_m) = \prod_{i=1}^{m} (1 + F(a_i, s_1, \ldots, s_m)). \tag{17}$$

The Fundamental Theorem on Symmetric Polynomials thus ensures that there are polynomials in the indeterminates $i_1, \ldots, i_m$, say t and ψ, such that t is transformed into $\phi(a_1)\phi(a_2) \ldots \phi(a_m)$ and ψ into $\prod_{i=1}^{m}(1 + F(a_i, s_1, \ldots, s_m))$ when each i_j is replaced by s_j. From (17) it then follows that $p_m t = \psi$.

Replacing each indeterminate i_j in this last equation by the real number C_j, and writing T for the resulting value of t, we conclude from (15) that $P_Y T = 1$. Hence $P_Y \ne 0$.

Lemma 1, just proved, implies that if $P_Y = 0$, then $Y(x)$ must have a nontrivial factor; so repeating that argument, if need be, shows that it must in fact have a factor Q for which $P_Q \ne 0$. Hence without loss of generality we may assume that $P_Y \ne 0$. Moreover, if $Y(x)$ has degree $m = k2^{\mu}$, where k is odd, then at least one factor of $Y(x)$ must be of degree $l2^{\tau}$, where l is odd and $\tau \le \mu$, for otherwise the power of 2 in m would exceed μ.

Gauss went on in section 12 to consider the polynomial

$$\zeta(u,x) = \prod[u - (a_i + a_j)x + a_i a_j], \tag{18}$$

where the product is taken over the $m(m-1)/2$ unordered pairs $\{a_i, a_j\}, i \neq j$ of the indeterminates $a_1, \ldots, a_m$. Once again, ζ is symmetric in those indeterminates, so there is a function $z(u,x)$ in the indeterminates $i_1, \ldots, i_m$ that transforms to ζ when the latter indeterminates are replaced by $s_1, \ldots, s_m$, and a function $Z(u,x)$ that results from $z(u,x)$ when each s_i is replaced by C_i (the ith coefficient of Y). Regarded as functions of u alone, ζ and z are monic polynomials of degree $n = m(m-1)/2$, with coefficients that are polynomials in $x, a_1, \ldots, a_m$ and in $x, i_1, \ldots, i_m$, respectively, and Z is a monic polynomial of degree n in u whose coefficients are polynomials in x, say $c_1(x), \ldots, c_n(x)$. We may then consider the discriminant $P_Z(x)$ of Z, that, is, the function $p_n(c_1(x), \ldots, c_n(x))$. (See the last footnote above.)

Gauss's next task was to prove

Lemma 2: If $P_Y \neq 0$, then $P_Z(x)$ cannot be identically zero.

He noted that, once again, that would be straightforward if $Y(x)$ were a product of linear factors. To establish the result without that assumption, he observed that the discriminant of ζ is the product of all the $n(n-1)$ non-zero differences of distinct pairs of the n expressions $(a_i + a_j)x - a_i a_j$. Hence the discriminants of ζ and of z, regarded as polynomials in x, each have degree $d = n(n-1) = \frac{1}{4}m(m-1)(m+1)(m-2)$, while the discriminant $P_Z(x)$ of Z may have lesser degree if the particular values of the C_i cause the coefficient of x^d in $P_Z(x)$ to vanish. The problem is to show that not all the coefficients of $P_Z(x)$ will vanish.

Closer examination of the discriminant of ζ reveals that that product may be split into two groups of factors, the first consisting of those differences of the form

$$[(a_i + a_j)x - a_i a_j] - [(a_i + a_k)x - a_i a_k] = (a_j - a_k)(x - a_i),$$

for distinct i, j, k, and the second of those differences of the form

$$[(a_i + a_j)x - a_i a_j] - [(a_k + a_l)x - a_k a_l] = (a_i + a_j - a_k - a_l)x - a_i a_j + a_k a_l, \tag{19}$$

for distinct i, j, k, l. In the first group, each factor $(a_j - a_k)$ will occur $m-2$ times (once for each value of i distinct from j and k), whereas each factor $(x - a_i)$ will occur $(m-1)(m-2)$ times (once for every ordered pair of distinct values j, k different from i). If the product of the second group of factors (a polynomial symmetric in $a_1, \ldots, a_m$) be denoted by κ, then from (10) and (12), the discriminant of ζ is ${\pi_m}^{m-2}{\nu}^{(m-1)(m-2)}\kappa$.

Likewise, if $k(x, i_1, \ldots, i_m)$ is the function that transforms into κ when each i_j in it is replaced by s_j, and $K(x)$ is the result of replacing each such i_j by C_j, then the discriminant of z is ${p_m}^{m-2}y^{(m-1)(m-2)}k$ and that of Z, P_Z, is ${P_Y}^{m-2}Y^{(m-1)(m-2)}K$. Since $P_Y \neq 0$ by assumption, it remains to show that K is not identically zero.

Toward that end Gauss introduced the function of the new indeterminate w given by

$$\prod[(a_i + a_j - a_k - a_l)w + (a_i - a_k)(a_i - a_l)], \tag{20}$$

where i, j, k, l are distinct integers between 1 and m and no factors are repeated. (Note that each factor is invariant under the interchange of a_k and a_l, and so would appear twice in the product if repeated factors were allowed.) That product is symmetric in the indeterminates $a_1, \ldots, a_m$, so it is uniquely expressible as a polynomial function $f(w, s_1, \ldots, s_m)$ of the elementary symmetric polynomials and w. Since the number of factors in the product is $\frac{1}{2}m(m-1)(m-2)(m-3)$, the degree of each substitution instance of $f(w, s_1, \ldots, s_m)$ is at most that. Also,

$$\begin{aligned} f(0, s_1, \ldots, s_m) &= {\pi_m}^{(m-2)(m-3)}, \\ f(0, i_1, \ldots, i_m) &= {p_m}^{(m-2)(m-3)}, \quad \text{and} \\ f(0, C_1, \ldots, C_m) &= {P_Y}^{(m-2)(m-3)}. \end{aligned}$$

In particular, the last equation shows that the constant term of

$$f(w, C_1, \ldots, C_m)$$

does not vanish.

Let the non-zero term of highest degree in $f(w, C_1, \ldots, C_m)$ be Nw^γ. Then for each j between 1 and m, if w be replaced by $x - a_j$, $f(x - a_j, C_1, \ldots, C_m)$ may be regarded as a polynomial in x whose leading term is Nx^γ and whose other coefficients depend upon a_j. Consequently

$$\prod_{j=1}^{m} f(x - a_j, C_1, \ldots, C_m) \tag{21}$$

is a polynomial in x with leading term $N^m x^{m\gamma}$, in which the coefficients of the remaining terms are functions of $a_1, \ldots, a_m$.

Similarly,

$$\prod_{j=1}^{m} f(x - a_j, i_1, \ldots, i_m)$$

is a polynomial function of $x, a_1, \ldots, a_m, i_1, \ldots, i_m$ symmetric in $a_1, \ldots, a_m$, which by the Fundamental Theorem on Symmetric Polynomials may be rewritten as a polynomial $\varphi(x, s_1, \ldots, s_m, i_1, \ldots, i_m)$. Replacing each i_j by s_j then yields

$$\varphi(x, s_1, \ldots, s_m, s_1, \ldots, s_m) = \prod_{j=1}^{m} f(x - a_j, s_1, \ldots, s_m).$$

Now, for any fixed value of i, when w is replaced by $x - a_i$, the factor

$$(a_i + a_j - a_k - a_l)w + (a_i - a_k)(a_i - a_l)$$

in (20) reduces after cancellation of like terms to

$$(a_i + a_j - a_k - a_l)x - a_i a_j + a_k a_l,$$

which is the same as the right member of (19). So each factor of κ is also a factor of $\varphi(x, s_1, \ldots, s_m, s_1, \ldots, s_m)$; that is, κ divides $\varphi(x, s_1, \ldots, s_m, s_1, \ldots, s_m)$, say $\varphi(x, s_1, \ldots, s_m, s_1, \ldots, s_m) = \kappa\chi(x, s_1, \ldots, s_m)$. Therefore also

$$\varphi(x, C_1, \ldots, C_m, C_1, \ldots, C_m) = K\chi(x, C_1, \ldots, C_m).$$

But $\varphi(x, s_1, \ldots, s_m, C_1, \ldots, C_m)$ is the product in (21), which has leading term $N^m x^{m\gamma}$, not involving any of $s_1, \ldots, s_m$; so $N^m x^{m\gamma}$ must also be the leading term of $\varphi(x, C_1, \ldots, C_m, C_1, \ldots, C_m)$. In particular, $\varphi(x, C_1, \ldots, C_m, C_1, \ldots, C_m)$ does not vanish identically. Therefore neither does K, as was to be shown.

Before beginning the induction that lay at the heart of his second proof, Gauss stated and proved one final lemma.

Lemma 3: Let $\Phi(u, x)$ denote the product $\prod_{i=1}^{k}(\alpha_i + \beta_i u + \gamma_i x)$ of any number of factors linear in the indeterminates u and x, and let v be another indeterminate. Then the function

$$\Omega = \Phi(u + v\frac{\partial\Phi}{\partial x}, x - v\frac{\partial\Phi}{\partial u})$$

is divisible by $\Phi(u, x)$.

Proof: For each i between 1 and k, we have

$$\Phi(u, x) = (\alpha_i + \beta_i u + \gamma_i x)Q_i,$$

where Q_i denotes the product of all the factors $\alpha_j + \beta_j u + \gamma_j x$ with $j \neq i$. (Each Q_i is thus a polynomial in $u, x, \alpha_j, \beta_j, \gamma_j$, for $1 \leq j \leq k, j \neq i$.) So

$$\frac{\partial\Phi}{\partial x} = \gamma_i Q_i + (\alpha_i + \beta_i u + \gamma_i x)\frac{\partial Q_i}{\partial x} \quad \text{and} \quad \frac{\partial\Phi}{\partial u} = \beta_i Q_i + (\alpha_i + \beta_i u + \gamma_i x)\frac{\partial Q_i}{\partial u}.$$

Substituting the expressions on the right sides of those equations into the corresponding factor

$$\alpha_i + \beta_i u + \gamma_i x + \beta_i v\frac{\partial\Phi}{\partial x} - \gamma_i v\frac{\partial\Phi}{\partial u}$$

of Φ yields

$$(\alpha_i + \beta_i u + \gamma_i x)(1 + \beta_i v\frac{\partial Q}{\partial x} - \gamma_i v\frac{\partial Q}{\partial u}),$$

and consequently,

$$\Omega = \Phi(u, x)\prod_{i=1}^{k}(1 + \beta_i v\frac{\partial Q}{\partial x} - \gamma_i v\frac{\partial Q}{\partial u}). \qquad \text{q.e.d.}$$

When applied to $\zeta(u, x) = z(u, x, s_1, \ldots, s_m)$, Lemma 3 shows that ζ divides

$$z(u + v\frac{\partial \zeta}{\partial x}, x - v\frac{\partial \zeta}{\partial u}, s_1, \ldots, s_m),$$

say with quotient $\Psi(u, x, v, s_1, \ldots, s_m)$. Likewise,

$$z(u + v\frac{\partial z}{\partial x}, x - v\frac{\partial z}{\partial u}, i_1, \ldots, i_m) = z(u, x)\Psi(u, x, v, i_1, \ldots, i_m)$$

and

$$Z(u + v\frac{\partial Z}{\partial x}, x - v\frac{\partial Z}{\partial u}, C_1, \ldots, C_m) = Z(u, x)\Psi(u, x, v, C_1, \ldots, C_m).$$

Now, given definite values U and X for the indeterminates u and x, let U' and X' denote

$$\left.\frac{\partial Z}{\partial u}\right|_{(U,X)} \quad \text{and} \quad \left.\frac{\partial Z}{\partial x}\right|_{(U,X)},$$

respectively. Then

$$Z(U + vX', X - vU') = Z(U, X)\Psi(U, X, v, C_1, \ldots, C_m).$$

If $U' \neq 0$, replacing v by $\dfrac{X - x}{U'}$ yields

$$Z(U + \frac{XX'}{U'} - \frac{xX'}{U'}, x) = Z(U, X)\Psi(U, X, \frac{X - x}{U'}, C_1, \ldots, C_m). \tag{22}$$

In other words, setting $u = U + \dfrac{X - x}{U'}X'$ transforms $Z(u, X)$ into

$$Z(U, X)\Psi(U, X, \frac{X - x}{U'}, C_1, \ldots, C_m).$$

By Lemma 2, the assumption that $P_Y \neq 0$ implies that the polynomial P_Z does not identically vanish. Therefore $P_Z(x) = 0$ for only finitely many values of x, so a real number X may be chosen for which $P_Z(X) \neq 0$; that is, the discriminant of the function $Z(u, X)$ is non-zero. By Lemma 1, the polynomial $Z(u, X)$ and its derivative $\dfrac{dZ}{du}$ thus have no common factor. Also, as noted earlier, $Z(u, X)$ has degree $n = m(m-1)/2$ in u, where $m = k2^{\mu}$, k odd, is the degree of $Y(x)$. Hence $n = (k2^{\mu})(k2^{\mu} - 1)/2 = (k^2)2^{2\mu-1} - k2^{\mu-1} = [k(k2^{\mu} - 1)]2^{\mu-1}$. The quantity in brackets in the last member of that equation is odd, so the power of 2 in n is less than the power of 2 in m. Therefore we may assume by induction that there is a real or complex value U for which $Z(U, X) = 0$.[16] By the factor theorem, $u - U$ must then be a factor of $Z(u, X)$, but, *ipso facto*, not a factor of $\dfrac{dZ}{du}$. So $U' \neq 0$, again by the factor theorem.

For those particular values of X and U, the right-hand member of (22) is identically zero, independent of the value of x. Thus $Z(u, x)$, regarded as a polynomial in u with coefficients that are polynomials in x, vanishes when $u = U + \dfrac{X - x}{U'}X'$, and so has $u - U - \dfrac{X - x}{U'}X'$ as a factor. If we then let $u = x^2$, the polynomial $Z(x^2, x)$ must have the quadratic polynomial

$$x^2 - U - \frac{X - x}{U'}X' = x^2 + \frac{X'}{U'}x - (U + \frac{XX'}{U'})$$

as a factor. The quadratic formula then provides a real or complex root of $Z(x^2, x)$.

Finally, recall that $z(x^2, x, i_1, \ldots, i_m)$ is the unique polynomial that transforms into $\zeta(x^2, x)$ when each i_j is replaced by the elementary symmetric polynomial s_j, and that $Z(x^2, x)$ is obtained from $z(x^2, x, i_1, \ldots, i_m)$ by replacing each i_j by the coefficient C_j of $Y(x)$. But

$$\zeta(x^2, x) = \prod[x^2 - (a_i + a_j)x + a_i a_j] = \prod(x - a_i)(x - a_j)$$
$$= \prod_{i=1}^{m}(x - a_i)^{m-1} = v^{m-1}$$

(where the first two products are taken over all unordered pairs $\{a_i, a_j\}, i \neq j$; cf. (18) and (12) above), and the unique polynomial in $x, i_1, \ldots, i_m$ that transforms into v^{m-1} when each i_j is replaced by s_j is $y(x)^{m-1}$ (cf. (11)). Therefore $z(x^2, x) = (y(x))^{m-1}$ and $Z(x^2, x) = (Y(x))^{m-1}$, so the root found for $Z(x^2, x)$ is also a root of $Y(x)$, completing the proof of the theorem.

[16] Gauss observed that the coefficients of $Z(u, X)$ will be real numbers if X is real and all coefficients of $Y(x)$ are real — a fact needed for the base case of the induction (that a real polynomial of odd degree must have a real root).

The proof just given is remarkable not only for purity of method but for its economy of means. The principal tool invoked is the Fundamental Theorem on Symmetric Polynomials, applied over and over again to a sequence of carefully chosen polynomials, and the only non-algebraic principle used is the intermediate-value theorem. As such it is a *tour de force*. Its length, however, is a pedagogical deterrent, and the justifications for some of the definitions and substitutions employed become apparent only with hindsight; it is thus less perspicuous than Argand's nearly contemporaneous argument.

8.6 Proofs based on integration

The proofs of the Fundamental Theorem of Algebra discussed above are all direct proofs. There are indirect proofs as well, several of which are based on the theory of integration. Among them is another by Gauss, published just one year after his second.[17]

Gauss's third proof: As in his first proof, given a monic polynomial

$$Y = x^m + A_1 x^{m-1} + \cdots + A_{m-1}x + A_m$$

with real coefficients, Gauss began by writing the variable x in polar form as $x = r(\cos\phi + i\sin\phi)$ and considered the real and imaginary parts of Y, which he denoted by t and u. Thus (replacing Gauss's $A, B, \ldots, M$ by $A_1, \ldots, A_m$)

$$t = \sum_{j=0}^{m} A_j r^{m-j} \cos(m-j)\phi \quad \text{and} \quad u = \sum_{j=0}^{m} A_j r^{m-j} \sin(m-j)\phi,$$

where $A_0 = 1$ and, for $1 \le j \le m$, the coefficients A_j are arbitrary real numbers. He further defined

$$t' = \sum_{j=0}^{m} (m-j) A_j r^{m-j} \cos(m-j)\phi,$$

$$u' = \sum_{j=0}^{m} (m-j) A_j r^{m-j} \sin(m-j)\phi,$$

$$t'' = \sum_{j=0}^{m} (m-j)^2 A_j r^{m-j} \cos(m-j)\phi,$$

[17] Translated into German on pp. 61–67 of Gauss (1890).

$$u'' = \sum_{j=0}^{m} (m-j)^2 A_j r^{m-j} \sin(m-j)\phi, \quad \text{and}$$

$$y = \frac{(t^2+u^2)(tt''+uu'') + (tu'-ut')^2 - (tt'+uu')^2}{r(t^2+u^2)^2},$$

and stipulated that R should be a real number greater than the largest of the numbers $(m|A_j|\sqrt{2})^{1/j}$, for $1 \le j \le m$. He then claimed that setting $r = R$ would ensure that $tt' + uu'$ was positive, for any angle ϕ.

Proof of claim: Corresponding to the definitions of t, u, t' and u', let

$$T = \sum_{j=0}^{m} A_j R^{m-j} \cos(\frac{\pi}{4} + j\phi),$$

$$U = \sum_{j=0}^{m} A_j R^{m-j} \sin(\frac{\pi}{4} + j\phi),$$

$$T' = \sum_{j=0}^{m} (m-j) A_j R^{m-j} \cos(\frac{\pi}{4} + j\phi), \quad \text{and}$$

$$U' = \sum_{j=0}^{m} (m-j) A_j R^{m-j} \sin(\frac{\pi}{4} + j\phi).$$

Gauss observed that T could be rewritten as

$$\sum_{j=1}^{m} \frac{R^{m-j}}{m\sqrt{2}} (R^j + mA_j \sqrt{2} \cos(\frac{\pi}{4} + j\phi)),$$

and similarly for U, T' and U'; so, by the stipulation on R, those four quantities, and hence $TT' + UU'$ must all be positive. But when $r = R$, $tt' + uu' = TT' + UU'$.

To see that, note first that when $r = R$, the quantity t is equal to

$$T \cos(\frac{\pi}{4} + m\phi) + U \sin(\frac{\pi}{4} + m\phi). \tag{23}$$

For, by the definitions of T and U, each term of (23) is of the form

$$A_j R^{m-j} [\cos(\frac{\pi}{4} + j\phi) \cos(\frac{\pi}{4} + m\phi) + \sin(\frac{\pi}{4} + j\phi) \sin(\frac{\pi}{4} + m\phi)] \tag{24}$$

$$= \frac{A_j R^{m-j}}{2} [\cos((m-j)\phi) + \cos(\frac{\pi}{2} + (m+j)\phi) + \cos((m-j)\phi)$$

$$- \cos(\frac{\pi}{2} + (m+j)\phi)] = A_j R^{m-j} \cos((m-j)\phi),$$

the corresponding term of t. Likewise, when $r = R$, the quantities u, t' and u' are equal to

$$T\sin(\frac{\pi}{4} + m\phi) - U\cos(\frac{\pi}{4} + m\phi),$$
$$T'\cos(\frac{\pi}{4} + m\phi) + U'\sin(\frac{\pi}{4} + m\phi), \quad \text{and}$$
$$T'\sin(\frac{\pi}{4} + m\phi) - U'\cos(\frac{\pi}{4} + m\phi),$$

respectively. Then, letting $A = \dfrac{\pi}{4} + m\phi$, we have

$$tt' = TT'\cos^2 A + T'U\sin A\cos A + TU'\sin A\cos A + UU'\sin^2 A \quad \text{and}$$
$$uu' = TT'\sin^2 A - T'U\sin A\cos A - TU'\sin A\cos A + UU'\cos^2 A,$$

so $tt' + uu' = TT' + UU' > 0$ when $r = R$, as claimed.

In addition, when $r = R$,

$$t^2 = T^2\cos^2 A + 2TU\sin A\cos A + U^2\sin^2 A \quad \text{and}$$
$$u^2 = T^2\sin^2 A - 2TU\sin A\cos A + U^2\cos^2 A,$$

so $t^2 + u^2 = T^2 + U^2$. Consequently, for any r satisfying the stipulations on R, $t^2 + u^2$ must be positive, whence t and u cannot simultaneously equal 0.

On the other hand, within the circle C of radius R centered at the origin there must be a point (r, ϕ) where both $t = 0$ and $u = 0$ (and thus a point $x = r(\cos\phi + i\,\sin\phi)$ where $Y = 0$, proving the theorem). For suppose not. Then let

$$\Omega = \int\int_C y\,dA = \int_0^R \int_0^{2\pi} y\,d\phi\,dr = \int_0^{2\pi} \int_0^R y\,dr\,d\phi.$$

Note that $\dfrac{\partial t}{\partial \phi} = -u'$, $\dfrac{\partial u}{\partial \phi} = t'$, $\dfrac{\partial t'}{\partial \phi} = -u''$ and $\dfrac{\partial u'}{\partial \phi} = t''$. Using those relations, one computes that

$$\frac{\partial}{\partial \phi}\left[\frac{tu' - ut'}{r(t^2 + u^2)}\right] = y, \quad \text{that is,} \quad \int y\,d\phi = \frac{tu' - ut'}{r(t^2 + u^2)}. \tag{25}$$

Since u and u' both equal 0 when $\phi = 0$ or $\phi = 2\pi$, the last expression above is also zero for those values of ϕ, whence

$$\int_0^{2\pi} y\,d\phi = 0, \quad \text{and so} \quad \Omega = \int_0^R \int_0^{2\pi} y\,d\phi\,dr = 0. \tag{26}$$

Similarly, from $r\dfrac{\partial t}{\partial r} = t'$, $r\dfrac{\partial u}{\partial r} = u'$, $r\dfrac{\partial t'}{\partial r} = t''$, and $r\dfrac{\partial u'}{\partial r} = u''$, one computes that

$$\frac{\partial}{\partial r}\left[\frac{tt' + uu'}{t^2 + u^2}\right] = y, \quad \text{that is,} \quad \int y\, dr = \frac{tt' + uu'}{t^2 + u^2}. \tag{27}$$

Consequently,

$$\int_0^R y\, dr = \left.\frac{tt' + uu'}{t^2 + u^2}\right|_0^R = \frac{TT' + UU'}{T^2 + U^2} > 0 \quad \text{by the claim proved earlier.}$$

But then

$$\Omega = \int_0^{2\pi}\int_0^R y\, dr\, d\phi = \int_0^{2\pi} \frac{TT' + UU'}{T^2 + U^2} > 0, \quad \text{contrary to (26).}$$

In his prefatory remarks, Gauss said merely that continued reflection on the Fundamental Theorem had led him to this third proof, which, like the second, was "purely analytic," but was based on entirely different principles and far surpassed the second in simplicity. And indeed, like Argand's proof, nothing beyond advanced calculus is needed for understanding the argument just given. However, several of the functions used in the proof, especially y, are introduced seemingly out of the blue, and it seems almost miraculous that the partial derivatives in (25) and (27) turn out to equal y. Thus, though succinct and requiring minimal prerequisites, the proof is not perspicuous: It provides convincing *verification* that the Fundamental Theorem is true, but it is not explanatory, since it does not convey understanding of *why* it is.

Other indirect proofs of the Fundamental Theorem are based on Cauchy's theory of complex contour integration. The best known is perhaps that based on **Liouville's Theorem** (that a bounded entire function must be constant): For if the polynomial $p(z)$ had no zero in the complex plane, then $\dfrac{1}{p(z)}$ would be an entire function; and as in Argand's proof, for any positive constant C there is an $R > 0$ such that $|p(z)| > C$ whenever $|z| > R$. Thus $\dfrac{1}{p(z)} < \dfrac{1}{C}$ for $|z| > R$, and as a continuous function, $\dfrac{1}{p(z)}$ would also be bounded within the disc $|z| \le R$. Thus $\dfrac{1}{p(z)}$ would be bounded throughout the complex plane, and hence a constant by Liouville's Theorem — a contradiction for any $p(z)$ of positive degree.

Liouville's Theorem is itself a consequence of **Cauchy's integral formula**, which asserts that if $f(z)$ is any function analytic in a simply connected domain containing the simple closed curve γ, then for any point z_0 inside γ,

$$f(z_0) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z - z_0}\, dz;$$

and, as first noted in Zalcman (1978) (see also Lax and Zalcman 2012), Cauchy's formula may be applied directly to yield an even simpler proof of the Fundamental Theorem of Algebra. For if $|p(z)| \neq 0$ throughout the complex plane, then $q(z) = \frac{1}{p(z)}$ is entire and $\frac{1}{p(0)} = q(0) \neq 0$. Hence

$$q(0) = \frac{1}{2\pi i} \int_{|z|=R} \frac{q(z)}{z}\, dz = \frac{1}{2\pi} \int_0^{2\pi} q(Re^{i\theta})\, d\theta,$$

for any $R > 0$. But as R approaches ∞, the last integral approaches zero, contrary to $q(0) \neq 0$.

Alternatively, a proof of the Fundamental Theorem may be couched in terms of **winding numbers**, where the winding number of a continuously differentiable closed curve γ about the origin is given by

$$\frac{1}{2\pi i} \int_\gamma \frac{dz}{z},$$

if γ does not pass through the origin. That notion can, however, also be defined without reference to line integrals: Less formally, and more generally, if $f(z)$ is a continuous function that is never zero, the winding number of $f(z)$ around the origin as z traces out a continuously differentiable closed curve γ may be defined, as in Courant and Robbins (1941), as "the net number of complete revolutions made by a vector joining the origin to $f(\gamma(z))$ as z traces out γ." Courant and Robbins then offer the following indirect proof of the Fundamental Theorem.

Suppose that the monic polynomial $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$ of degree $n > 0$ is never zero. Let C_t be the circle about the origin of radius t, given by the equation $z = te^{i\theta}$, and let $\phi(t)$ be the winding number of $p(z)$ around the origin as z traces out C_t. Then ϕ is a continuous, integer-valued function of t, and so must be a *constant*; and since $\phi(0) = 0$, we must have $\phi(t) = 0$ for all t.

But

$$|z^n - p(z)| \leq |a_{n-1}||z|^{n-1} + \cdots + |a_0| = |z|^{n-1}\left[|a_{n-1} + \cdots + \frac{|a_0|}{|z|^{n-1}}\right],$$

so for values of t greater than $|a_0| + |a_1| + \cdots + |a_{n-1}| + 1$, the length $|z^n - p(z)|$ of the vector from $p(z)$ to z^n will be less than or equal to

$$t^{n-1}[|a_{n-1}| + \cdots + \frac{|a_0|}{t^{n-1}}] < t^n = |z^n|, \text{ the distance from } z^n \text{ to the origin.}$$

(See Figure 8.4.)

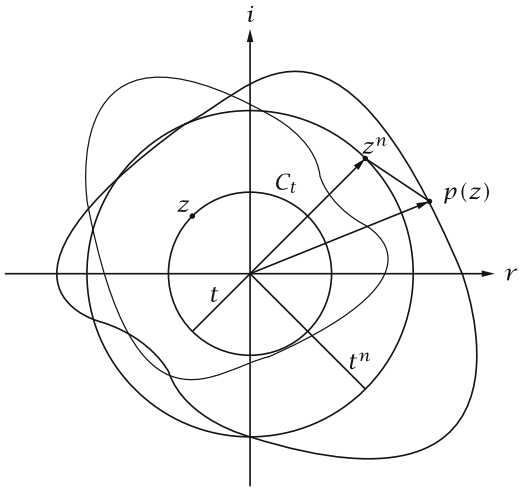

Fig. 8.4 A winding-number proof. Adapted from Figure 150, p. 270 in *What is Mathematics?*, 2nd. ed. (1996), by Richard Courant and Herbert Robbins. By permission of Oxford University Press, U.S.A.

The segment joining $p(z)$ to z^n thus cannot pass through the origin when z is on C_t. Deforming the curve traced by $p(z)$ to the circle traced by z^n by shrinking each such line segment to zero will thus not alter the value of $\phi(t)$, which must be the same as the winding number of z^n around the origin as z traces out C_t.[18] But that number is n, which is greater than zero, contrary to what was found above.

The preface to the first edition of *What is Mathematics?* states that that book "presupposes only knowledge that a good high school course could impart," and the proof just given is an excellent example of a perspicuous informal proof. By dispensing with reference to line integrals, Courant and Robbins succeeded in offering a proof of the Fundamental Theorem that should be both convincing and understandable to mathematically inclined high school students — a remarkable achievement from a pedagogical standpoint, since it is at that level that students first encounter the Fundamental Theorem, and all other proofs known to this author presume at least knowledge of advanced calculus.[19]

[18]The formal statement of that fact is Rouché's Theorem.

[19]I vividly recall my own frustration on repeatedly reading, not only in high school, but throughout my undergraduate courses at M.I.T., that a proof of the Fundamental Theorem was "beyond the scope of this text." Only when I took a graduate complex analysis course did I finally see the theorem proved as a corollary to Liouville's theorem — an experience I found distinctly anticlimactic, given my long years of expectant waiting. (William Dunham, in his article Dunham (1991), describes his own very similar experience.)

8.7 Other modern proofs

As part of the development of field theory in the nineteenth century, Kronecker proved the basic result needed to establish the existence of splitting fields: that if $p(x)$ is an irreducible polynomial with coefficients in a field F, then there is a field F' extending F, of finite degree over F, in which $p(x)$ has a root. The earlier arguments put forward as proofs of the Fundamental Theorem of Algebra by Lagrange and Laplace were thereby validated.

Later in the nineteenth century the work of Évariste Galois was belatedly published, and in 1872 Peter Ludvig Sylow proved the theorem in group theory now bearing his name, according to which for any prime number p, if p^m is the largest power of p dividing the order of a finite group G, then G must possess subgroups of order p^k for each $k \leq m$. In addition, the number of subgroups of G of order p^m divides the order of G and is congruent to 1 modulo p. Emil Artin then gave a proof of the Fundamental Theorem in terms of those concepts.

Proof via Galois theory:[20] Let K be a splitting field for the polynomial $p(x)$, of degree $n > 1$ with real coefficients. K is a finite extension of the real field $\mathbb{R}$, say of degree $d = 2^m q$ over $\mathbb{R}$, where q is odd. Since $\mathbb{R}$ has characteristic 0, K is a simple extension $\mathbb{R}(\alpha)$ of $\mathbb{R}$, where α is a root of a unique monic irreducible polynomial $r(x)$ of degree d with real coefficients. If $m = 0$, $r(x)$ must have degree $q = 1$, since every real polynomial of odd degree has a real root. So in that case $K = \mathbb{R}$. If $m = 1$ and $q = 0$, $r(x)$ is a quadratic polynomial, whose roots must lie in $\mathbb{C}$, whence $K = \mathbb{C}$. So suppose $m > 0$ and $q \neq 0$, and let G be the Galois group of K over R. By the fundamental theorem of Galois theory, G must have order d, and by Sylow's theorem, G therefore has a subgroup of order 2^m and index q. Again by the fundamental theorem of Galois theory, there must then be an extension field E of $\mathbb{R}$ intermediate between $\mathbb{R}$ and K, having degree q over $\mathbb{R}$. Then as above, since q is odd, $q = 1$ and $E = \mathbb{R}$. Accordingly, K must have degree 2^m over $\mathbb{R}$, so the order of G is 2^m. By Sylow's Theorem, G has a subgroup H_1 of order 2^{m-1}. Let L_1 be the fixed field of H_1. Then L_1 is an extension of $\mathbb{R}$ of degree 2, so L_1 is isomorphic to $\mathbb{C}$ (since L_1 is obtained from $\mathbb{R}$ by adjoining a root of a quadratic polynomial), and K is an extension of L_1 of degree 2^{m-1}. If $m > 1$, let H_2 be a subgroup of H_1 of order 2^{m-2}, and let L_2 be the fixed field of H_2. Then L_2 is an extension of L_1 (and so is isomorphic to an extension of $\mathbb{C}$) of degree 2. But that is impossible, since quadratic polynomials with complex coefficients always have complex roots. Thus $m = 1$ and K has degree $2^0 = 1$ over L_1; that is, $K \simeq L_1 \simeq \mathbb{C}$, where $\simeq$ denotes isomorphism. q.e.d.

The prerequisites for understanding the proof just given are substantially greater than those for the other proofs so far considered. In particular, few undergraduates would be able to comprehend it. Nonetheless, since Galois theory was developed to

[20] For further background details, see chapter 7 of Fine and Rosenberger (1997).

answer questions about the solvability of polynomials by radicals, it is appropriate to use it to prove the Fundamental Theorem of Algebra; and like Gauss's second proof, it is as methodologically pure as possible, invoking nothing non-algebraic except the intermediate-value theorem.

There are also proofs of the Fundamental Theorem based on sophisticated notions from algebraic topology (Brouwer degree, homotopy theory, simplicial complexes, homology theory), three of which are sketched in chapter 9 and appendix D of Fine and Rosenberger (1997). They are not discussed further here, since the means they employ were developed to address very different concerns and go far beyond what is required to establish the Fundamental Theorem. Nevertheless, they demonstrate the power of topological techniques (another example of benchmarking) and illustrate the coherence of disparate mathematical theories.

A much simpler proof, based on notions from point-set topology, establishes the Fundamental Theorem in the equivalent form: Every polynomial of non-zero degree with complex coefficients, regarded as a mapping with domain $\mathbb{C}$, has range all of $\mathbb{C}$.[21]

A topological proof:[22] Let $p(z)$ be a polynomial of degree $n > 1$ with complex coefficients. Since p is continuous and $|p(z)|$ approaches infinity as $|z|$ does, the range R of p must be closed (by Weierstrass's theorem that every bounded sequence has a convergent subsequence). Consider then the set T of all points $p(z)$ where $p'(z) = 0$. Since $\mathbb{C} - R$ is open, it suffices to show that $R - T$ is open as well. For if so, then $(\mathbb{C} - R) \cup (R - T) = \mathbb{C} - T$. Since T is finite, $\mathbb{C} - T$ is a connected set, which cannot be the union of two disjoint nonempty open sets; and since $p^{-1}(T)$ is finite and $n > 1$, $R - T$ is nonempty. Thus $\mathbb{C} - R$ must be empty.

The proof is completed by noting that for every $p(z_0)$ in $R - T$, $p'(z_0) \neq 0$, so the inverse function theorem implies that there are neighborhoods U of z_0 and V of $p(z_0)$ such that p maps U one-to-one onto V. Hence every point of $R - T$ is an interior point.

There remains the purity question broached earlier in connection with Euler's failed proof attempt: Can the original statement of the Fundamental Theorem, that any non-constant polynomial with real coefficients can be expressed as a product of real linear and quadratic factors, be proved without reference to complex numbers or splitting fields?

Such a proof has been given, based on concepts from linear algebra. The proof is by induction on the degree of the polynomial and uses properties of the so-called Bezoutian resultant (the determinant of an $n \times n$ symmetric matrix defined for any pair of polynomials p, q, where n is the maximum of the degrees of p and q). Specifically, writing p as a function of the variable x and q as a function of the variable w, the factor theorem implies that $x - w$ must be a factor of the

[21] The intermediate-value theorem may similarly be cast as the assertion that the range of any polynomial of *odd* degree with *real* coefficients, regarded as a mapping from $\mathbb{R}$ into $\mathbb{R}$, is all of $\mathbb{R}$.

[22] This is essentially the proof given in Sen (2000).

polynomial $P(x,w) = p(x)q(w) - p(w)q(x)$, say $P(x,w) = (x-w)b(x,w)$, where b is a polynomial each of whose terms is of the form $b_{ij}x^{i-1}w^{j-1}$, for $1 \le i, j \le n$. The coefficients b_{ij} are the entries of the Bezoutian matrix $B(p,q)$, which is nonsingular if and only if p and q have no common root.[23]

In outline, the proof then proceeds as follows: Given a real polynomial $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ with $a_0 \neq 0$ and $a_n \neq 0$, if n is odd, then $p(x)$ has a real root x_0 by the intermediate-value theorem, so $p(x) = (x - x_0)q(x)$, where $q(x)$ has degree $n-1$. By the induction hypothesis, $q(x)$ then is a product of real linear and quadratic factors, whence so is $p(x)$. If n is even, say $n = 2m$, then let r be a real parameter and define auxiliary polynomials $p_1(x)$ and $p_2(x)$ by

$$p_1(x) = p(rx) = a_n r^n x + a_{n-1} r^{n-1} x^{n-1} = \cdots + a_0 \quad \text{and}$$

$$p_2(x) = a_0 x^n + a_1 r x^{n-1} + \cdots + a_n r^n .$$

The intermediate-value theorem, together with properties of principal minors, can then be invoked to show that there must be a value r_0 of r for which the determinant of $B(p_1, p_2) = 0$; so when $r = r_0$, $p_1(x)$ and $p_2(x)$ will have some maximal common factor $f(x)$. Writing $p_1(x) = f(x)Q_1(x)$ and $p_2(x) = f(x)Q_2(x)$ yields $p_1(x)Q_2(x) = p_2(x)Q_1(x)$, where Q_1 and Q_2 are relatively prime polynomials of degree less than n. By the induction hypothesis, if Q_2 is not constant, it must be a product of linear and quadratic factors, all of which divide p_1. The quotient $p_1(x)/Q_2(x)$ then also has degree less than n, so it too must be a product of linear and quadratic factors. Hence $p_1(x)$ is a product of linear and quadratic factors, and replacing x in each of those factors by x/r_0 produces a similar factorization of $p(x)$.

On the other hand, if $Q_2(x)$ is a constant, then $f(x)$ has degree n, so $Q_1(x)$ is also a constant. Hence $p_1(x) = cp_2(x)$ for some constant c, so for each i between 0 and n, the coefficient of x^i in $p_1(x)$ must equal c times the coefficient of x^i in $p_2(x)$; that is, $a_i r_0^i = ca_{n-i} r_0^{n-i}$ for each such i. In particular, for $i = m$, $a_m r_0^m = ca_m r_0^m$, so $c = 1$ and $a_j r_0^j = a_{n-j} r_0^{n-j}$ for $0 \le j \le n$, which means that $p_1(x)$ is a palindromic (self-reciprocal) polynomial. But then either $x+1$ is a factor of $p_1(x)$, or $p_1(x) = x^m q(X)$, where q is a polynomial of degree m and $X = x + x^{-1}$.

By the induction hypothesis, q is a product of real linear and quadratic polynomials in X, so $p_1(x)$ is a product of real linear, quadratic, and quartic polynomials in x. Since quartics can always be factored (e.g., by Euler's technique) into products of real linear and quadratic expressions, so can $p_1(x)$. The desired factorization of $p(x)$ is then obtained once again by replacing x everywhere by x/r_0.

The proof just given (published in Eaton 1960) employs very different techniques than the others considered here. However, the idea of using the resultant of two polynomials as a tool in proving the Fundamental Theorem of Algebra is not new.

[23] Indeed, the nullity of $B(p,q)$ equals the number of common zeroes of p and q, counting multiplicities.

Indeed, the British mathematician James Wood gave an incomplete proof of that sort in 1798, the year before Gauss's first proof.[24] A full proof using resultants was later published by Paul Gordan, who presented it as a simpler alternative to Gauss's second proof (Gordan 1876).

8.8 Constructive proofs

Although the foregoing proofs establish the *existence* of a root for any nonconstant polynomial, they provide no means for actually *finding* one. One may therefore wonder, as Weierstrass did in 1891, whether it is possible "for any given polynomial f in $\mathbb{C}(Z)$, to produce a sequence z_n of complex numbers by an effectively defined procedure, so that $|f(z_n)|$ is sufficiently small in relation to $|f(z_{n-1})|$ that it converges to a zero of f?"[25] Beyond that, one might ask whether the proof that such a procedure converges can be done constructively, and how computationally *efficient* the procedure is. The importance of such questions to mathematical practice is indicated by the appearance in 1969 of a volume entitled *Constructive Aspects of the Fundamental Theorem of Algebra* (Dejon and Henrici 1969), containing the proceedings of a symposium on that topic held two years before at the IBM Research Laboratory in Zürich.

In 1924 Hermann Weyl gave an intuitionistic proof of the Fundamental Theorem of Algebra that invoked winding numbers (Weyl 1924, pp. 142–146). Later Hellmuth Kneser presented a modification of Argand's proof in which, given a nonconstant polynomial $p(x)$, he defined a sequence of complex numbers and proved, also by means acceptable to intuitionists, that it converged to a root of p (H. Kneser 1940). Specifically, given a monic polynomial $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ of degree $n > 1$ with complex coefficients, Kneser defined a sequence of complex numbers x_m designed to make the ratio $|p(x_{m+1})|/|p(x_m)|$ as small as possible. Toward that end he expressed $p(x_m + y)$ as $p(x_m) + \sum_{i=1}^{n} b_i y^i$, chose y so that one of the terms in the sum would have the same argument as $-p(x_m)$ and would strongly dominate the other terms, and set $x_{m+1} = x_m + y$. To find such a y he employed a lemma whose proof was lengthy and delicate. Forty-one years later his son Martin published a simplified version of the proof (M. Kneser 1981) based on the following much simpler lemma:

Given a monic polynomial $p(x)$ of degree $n > 1$ with constant term a_0, there is a positive number $q < 1$, depending only on n, such that if $|a_0| \leq c$, a complex number y can be specified for which $|y| \leq c^{1/n}$ and $|p(y)| \leq qc$.

[24] Wood (1798), his only published mathematical paper. An analysis of Wood's argument, as well as a completion of it, has recently been given by Frank Smithies (Smithies 2000). I am indebted to Peter Goddard for bringing both those papers to my attention.

[25] Quoted on p. 115 of Remmert (1990) from the third volume of Weierstrass's *Mathematische Werke*, pp. 251–269.

The sequence $\{x_m\}$ may then be defined inductively, as follows, starting with $x_0 = 0$. Suppose that x_m has already been defined and satisfies $|p(x_m)| \leq q^m c$. Apply the lemma to $f(x) = p(x_m + x)$, which has constant term $p(x_m)$, and set $x_{m+1} = x_m + y$. Then $|x_{m+1} - x_m| = |y| \leq (q_m c)^{1/n}$ and $|p(x_{m+1})| \leq q(q^m c) = q^{m+1}c$. The first of those inequalities shows that the sequence $\{x_m\}$ converges to some value x and the second that the inductive hypothesis is satisfied by x_{m+1}, so that $p(x_m)$ converges to 0. Since p is continuous, it follows that $p(x) = 0$. The proof of the lemma is constructive, and can be made to satisfy intuitionistic demands as well.

Two years before Martin Kneser's paper appeared, Steve Smale and Morris Hirsch also defined a sequence of values guaranteed to converge to a root of any given polynomial of degree $n > 0$ (Hirsch and Smale 1979, pp. 303–309). Their procedure was based on a modification of Newton's method, and again involved an auxiliary proposition, namely:

For any positive integer n there exist real numbers $\sigma_1, \ldots, \sigma_n$ *with* $0 < \sigma_k \leq 1$ *for* $k = 1, \ldots, n$ *and* K_n *satisfying* $0 < K_n < 1$, *such that if* $h_k = \sigma_k e^{i\pi/k}$, *then for any n-tuple* $(\nu_1, \ldots, \nu_n)$ *in* $\mathbb{C}^n - 0$ *there is an m for which*

$$\left|1 + \sum_{k=1}^{n} \left(\frac{\nu_k}{\nu_m} h_m\right)^k\right| < K_n.$$

Suppose then that $p(x)$ is a polynomial of degree $n > 0$ and z is any complex number. If $p(z) = 0$, let $z' = z$. Otherwise, for each positive integer $k \leq n$, let ν_k be a k^{th} root of $p^{(k)}(z)/k!p(z)$. Since $n > 0$, $p^{(n)}(z) \neq 0$, so $(\nu_1, \ldots, \nu_n)$ is in $\mathbb{C}^n - 0$. The auxiliary proposition then yields a $\nu_m \neq 0$, whence z' can be taken to be $z + h_m/\nu_m$. Then by Taylor's theorem,

$$\begin{aligned}
|p(z')| &= \left|p(z) + \sum_{k=1}^{n} \frac{p^{(k)}(z)}{k!} \left(\frac{h_m}{\nu_m}\right)^k\right| \\
&= \left|p(z)\left[1 + \sum_{k=1}^{n} \frac{p^{(k)}(z)}{p(z)k!} \left(\frac{h_m}{\nu_m}\right)^k\right]\right| \\
&= |p(z)|\left|1 + \sum_{k=1}^{n} \left(\frac{\nu_k}{\nu_m} h_m\right)^k\right| < K_n |p(z)|.
\end{aligned}$$

To prove the Fundamental Theorem of Algebra, let z_0 be any complex number, and assume by induction that $z_0, \ldots, z_j$ have already been defined. If $p(z_j) = 0$, then z_j is a root of p, so put $z_{j+1} = z_j$. Otherwise define z_{j+1} to be $z_j + h_{m_0}/\nu_{m_0}$, where m_0 is the least value of m that satisfies the inequality given in the auxiliary proposition. Then the equations and inequality displayed above show that $|p(z_{j+1}| < K_n |p(z_j)|$. So $|p(z_j)| < (K_n)^j |p(z_0)|$ for every j. Since $K_n < 1$, $p(z_j)$ therefore converges

to 0 as j approaches ∞. The convergence may be very slow, however, since the authors note that for large values of n, K_n is very close to 1.[26]

That a subsequence of $\{z_j\}$ must converge to a root of p follows by a compactness argument:[27] For since $|p(z)|$ approaches ∞ as z does, and since $|p(z_j)| < K_n|p(z_0)|$ for all j, all the values z_j must lie within some disc $|z| \leq R$. If there are infinitely many distinct values z_j, a subsequence of them must approach some point within that disc, since otherwise for each point z with $|z| \leq R$ some disc D_z centered at z would contain at most one of the z_j (namely z_j if and only if $z = z_j$). But since the disc $|z| \leq R$ is compact, a finite number of those D_z would cover it. Thus either the sequence $\{z_j\}$ is eventually constant, say $z_k = z_j$ for all $k \geq j$ (which by construction implies that $p(z_j) = 0$), or else every neighborhood of some z^* in $\mathbb{C}$ contains infinitely many of the z_j. In the latter case, the continuity of p implies that $p(z^*) = 0$, so z^* is a root of p.

Further algorithms and constructive proofs of the Fundamental Theorem of Algebra may be found in the volume (Dejon and Henrici 1969) cited earlier. See in particular the article "A never failing fast convergent root-finding algorithm," by Bruno Dejon and Karl Nickel, pp. 1–35.

References

Argand, J.R.: Réflexions sur la nouvelle théorie d'analyse. Annales Math. **5**, 197–209 (1814)

Baltus, C.: Gauss's second proof of the fundamental theorem of algebra, 1815. Proc. Canad. Soc. Hist. Phil. Math. **13**, 97–106 (2000)

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

Courant, R., Robbins, H.: What is Mathematics? Oxford U., Oxford (1941)

D'Alembert, J.L.R.: Recherches sur le calcul intégral. Histoire de l'Academie Royale Berlin, 182–224 (1746)

Dejon, B., Henrici, P.: Constructive Aspects of the Fundamental Theorem of Algebra. Proceedings of a Symposium Conducted at the IBM Research Laboratory, Zürich-Rüschlkon, Switzerland, June 5–7, 1967. Wiley, New York (1969)

Dunham, W.: Euler and the fundamental theorem of algebra. College Math. J. **22**(4), 282–293 (1991)

Eaton, J.E.: The fundamental theorem of algebra. Amer. Math. Monthly **67**, 578–579 (1960)

Euler, L.: Recherches sur les racines imaginaires des equations. Mem. l'acad. sci. Berlin **5**, 222–288 (1749)

Fine, B., Rosenberger, G.: The Fundamental Theorem of Algebra. Springer, New York (1997)

Gauss, C.F.: Die vier Gauss'schen Beweise für die Zerlegung ganzer algebraischer Funktionen in reelle Factoren ersten oder zweiten Grades (1799–1849), trans. E. Netto. Ostwalds Klassiker der exakten Wissenschaften 14. Wilhelm Engelmann, Leipzig (1890)

[26] In his later paper (Smale 1981) Smale analyzed algorithms for finding roots of polynomials from the perspective of computational complexity theory, where questions of speed of convergence are paramount.

[27] The authors assert without proof that "in fact it is easy to see that the [full] sequence $\{z_j\}$ converges to a root" of p.

Gilain, C.: Sur l'histoire du theéorème d'algèbre: theorie des équations et calcul intégral. Arch. Hist. Exact Sci. **42**(2), 91–136 (1991)
Gordan, P.: Ueber den Fundamentalsatz der Algebra. Math. Ann. **10**, 572–575 (1876)
Hirsch, M., Smale, S.: On algorithms for solving $f(x) = 0$. Comm. Pure App. Math. **32**, 281–312 (1979)
Kline, M.: Mathematical Thought from Ancient to Modern Times. Oxford U., Oxford (1972)
Kneser, H.: Der Fundamentalsatz der Algebra und Intuitionismus. Math. Z. **46**, 287–302 (1940)
Kneser, M.: Ergänzung zu einer Arbeit von Hellmuth Kneser über den Fundamentalsatz der Algebra. Math. Z. **177**, 285–287 (1981)
Lax, P.D., Zalcman, L.: Complex Proofs of Real Theorems. University Lecture Series 58. Amer. Math. Soc., Providence (2012)
Ostrowski, A.: Über den ersten und vierten Gaussschen Beweise des Fundamentalsatzes der Algebra. In Alexander Ostrowski, Collected Mathematical Papers, vol. 1, pp. 538–553. Birkhäuser, Basel et al. (1983)
Remmert, R.: The fundamental theorem of algebra. In H-D. Ebbinghaus (ed.), Numbers, pp. 97–122. Springer, New York (1990)
Sen, A.: Fundamental theorem of algebra—yet another proof. Amer. Math. Monthly **107**(9), 842–843 (2000)
Smale, S.: The fundamental theorem of algebra and complexity theory. Bull. Amer. Math. Soc. **4**(1), 1–36 (1981)
Smithies, F.: A forgotten paper on the fundamental theorem of algebra. Notes Rec. Royal Soc. London **54**(3), 333–341 (2000)
Weyl, H.: Randbemerkungen zu Hauptproblemen der Mathematik. Math. Z. **20**, 131–150 (1924)
Wood, J.: On the roots of equations. Philos. Trans. Royal Soc. London **88**, 368–377 (1798)
Zalcman, L.: Picard's theorem without tears. Amer. Math. Monthly **85**, 265–268 (1978)