The Fundamental Theorem of Algebra: A Visual Approach

Daniel J. Velleman
Department of Mathematics and Statistics
Amherst College
Amherst, MA 01002

In the real numbers, a linear equation $ax + b = 0$ (with $a \neq 0$) always has a solution. Quadratic equations don't always have real solutions, but the quadratic formula gives solutions in the complex numbers. What if we move on to polynomial equations of higher degree? Might we need to extend the complex numbers to some larger number system in order to find solutions to these equations? The answer is no: solutions to all polynomial equations can be found in the complex numbers. This is the content of the Fundamental Theorem of Algebra:

Fundamental Theorem of Algebra. *Every nonconstant polynomial with complex coefficients has a root in the complex numbers.*

Some version of the statement of the Fundamental Theorem of Algebra first appeared early in the 17th century in the writings of several mathematicians, including Peter Roth, Albert Girard, and René Descartes. The first proof of the Fundamental Theorem was published by Jean Le Rond d'Alembert in 1746 [2], but his proof was not very rigorous. Carl Friedrich Gauss is often credited with producing the first correct proof in his doctoral dissertation of 1799 [15], although this proof also had gaps. (For a comparison of these two proofs, see [26, pp. 195–200].) Today there are many known proofs of the Fundamental Theorem of Algebra, including proofs using methods of algebra, analysis, and topology. (The references include many papers and books containing proofs of the Fundamental Theorem; [14] alone contains 11 proofs.) Our focus in this paper will be on the use of pictures to see why the theorem is true.

Of course, if we want to use pictures to display the behavior of polynomials defined on the complex numbers, we are immediately faced with a difficulty: the complex numbers are two-dimensional, so it appears that a graph of a complex-valued function on the complex numbers will require four dimensions. Our solution to this problem will be to use color to represent some dimensions.

We begin by assigning a color to every number in the complex plane. Figure 1 is a picture of the complex plane in which every point has been assigned a different color. The origin is colored black. Traveling counterclockwise around a circle centered at the origin, we go through the colors of a standard color wheel: red, yellow, green, cyan, blue, magenta, and back to red. Points near the origin have dark colors, with the color assigned to a complex number z approaching black as z approaches 0. Points far from the origin are light, with the color of z approaching white as $|z|$ approaches infinity. Every complex number has a different color in this picture, so a complex number can be uniquely specified by giving its color.

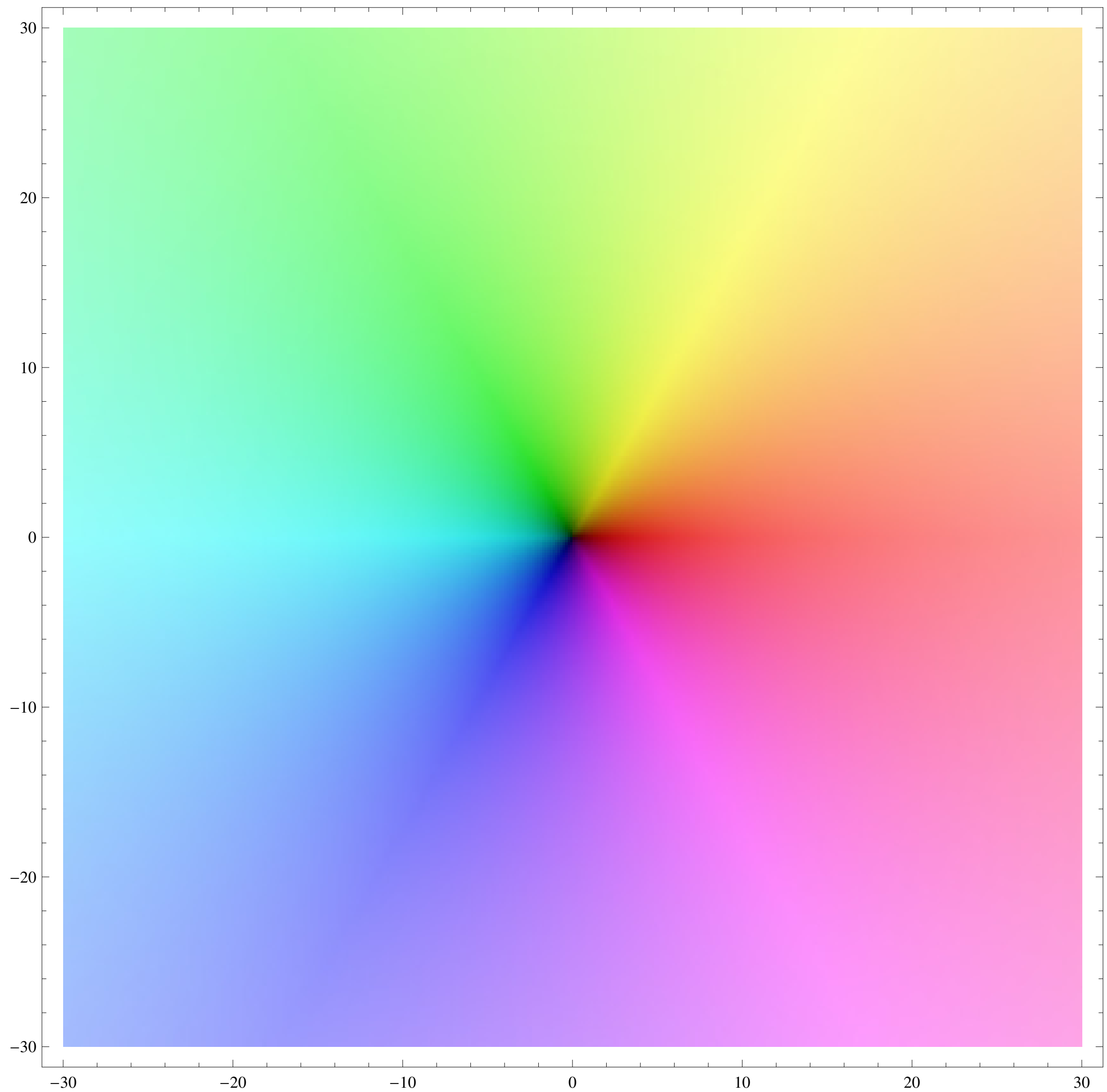

Figure 1: Assigning a color to each point in the complex plane.

We can now use this color scheme to draw a picture of a function $f : \mathbb{C} \to \mathbb{C}$ as follows: we simply color each point z in the complex plane with the color corresponding to the value of $f(z)$. From such a picture, we can read off the value of $f(z)$, for any complex number z, by determining the color of the point z in the picture, and then consulting Figure 1 to see what complex number is represented by that color.

For example, Figure 2 is a picture of the function $f(z) = z^3$. Three things are immediately evident in this picture. First, we see that the center of the picture is very dark. This is because when z is small, z^3 is very small, and therefore the color assigned to z^3 is very dark. Second, the colors fade out quickly when we move toward the outside of the picture. This is because when z is large, z^3 is very large, and therefore its color is very light. But what is most striking about the picture is that when we go counterclockwise around a circle centered at the origin, we go through the colors of the color wheel three times. This illustrates the fact that the argument of z^3 is three times the argument of z, and therefore the image of a circle centered at the origin under the cubing function wraps around the origin three times.

As an illustration of how such a picture can help us understand a function, note that it is immediately evident from Figure 2 that every nonzero complex number has three cube roots. For example, the color assigned to the number 1 in Figure 1 is a deep red. Therefore, the three cube roots of 1 are the three points in Figure 2 that are colored this particular shade of red.

Let us consider now a more complicated function. Figure 3 is a picture of the polynomial $f(z) = z^8 - 2z^7 + 2z^6 - 4z^5 + 2z^4 - 2z^3 - 5z^2 + 4z - 4$. Perhaps the first thing one notices in this picture is that the Fundamental Theorem of Algebra does, indeed, hold for f. Since the color assigned to the number 0 is black, the roots of f appear in this picture as six black dots. The fact that the polynomial has degree eight also shows up in the picture. For large z, the z^8 term in $f(z)$ dominates the other terms, and therefore the outer parts of the picture look similar to a picture of the function z^8: the colors begin to fade toward white as we move toward the edges of the picture, but before they fade out we can see that, as we go around the picture counterclockwise, the colors of the color wheel are repeated eight times.

Why does f, a polynomial of degree eight, have only six roots? The reason is that two of the roots are double roots, and this fact is also evident in the picture. The single roots occur at the points -1, 2, and $(-1 \pm i\sqrt{7})/2$, and the double roots are at $(1 \pm i\sqrt{3})/2$. The regions around the double roots are somewhat darker than those around the single roots, and at the double roots the colors of the color wheel wrap around the root twice, whereas at the single roots they wrap around only once.

In general, if a polynomial f has a root of multiplicity k at a point z_0, then when $f(z)$ is expanded in powers of $z - z_0$ it will have the form

$$f(z) = c(z - z_0)^k + (\text{higher degree terms}).$$

For z close to z_0, the first term in this expansion will dominate the higher degree terms, and therefore we have $f(z) \approx c(z - z_0)^k$. Thus, near the point z_0, the picture of f will be similar to the picture of the function cz^k near 0. In particular, the colors of the color wheel will wrap around the point z_0 k times, and the larger k is, the darker the picture will be near z_0. (It is not hard to see that the effect of the coefficient c on the picture is to alter the darkness of the picture near z_0, and also to rotate the arrangement of colors around z_0. For example, near the root at -1 in Figure 3, the colors have been rotated 180 degrees, so that red is to the left of the root rather than to the right.

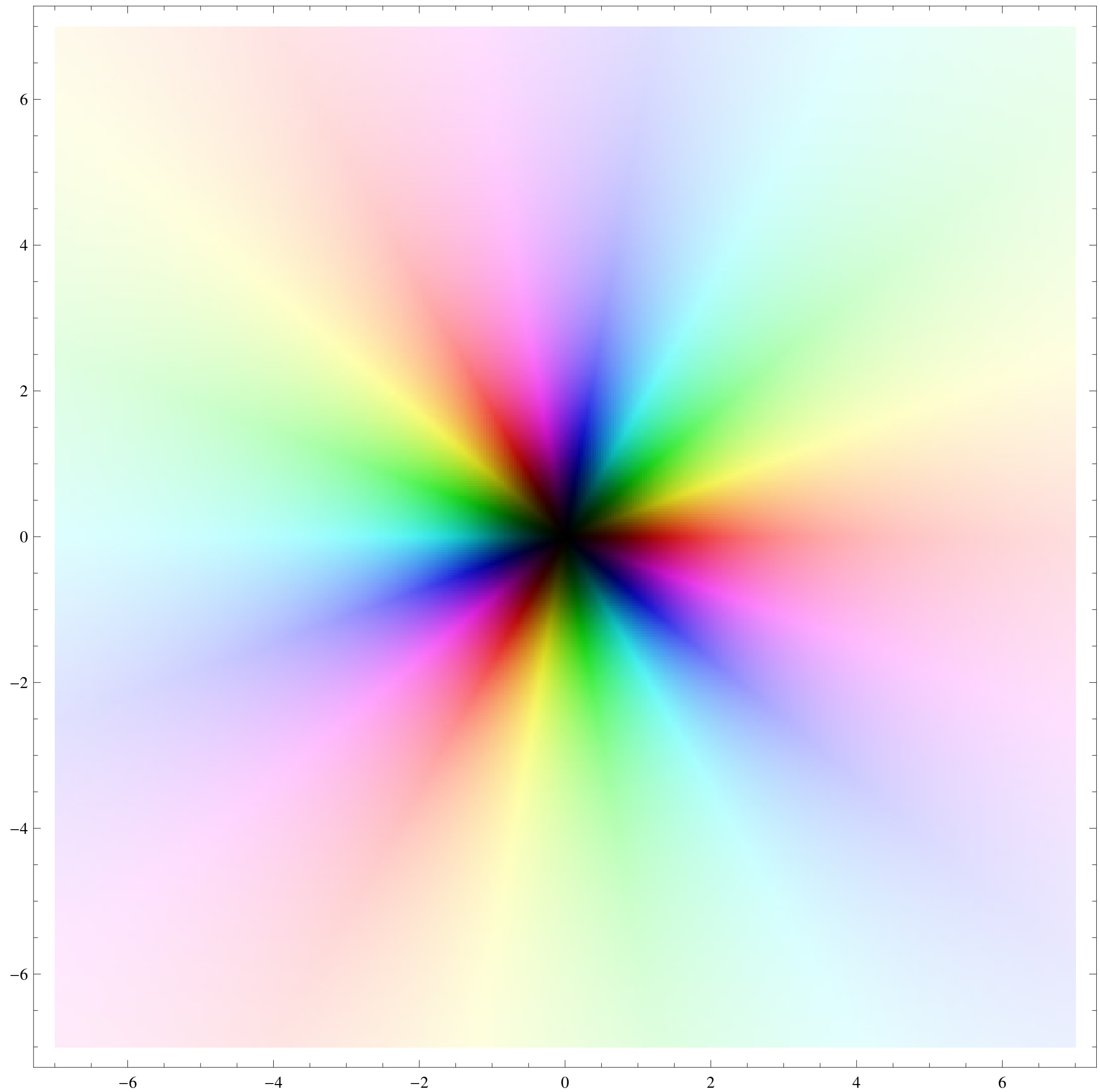

Figure 2: $f(z) = z^3$.

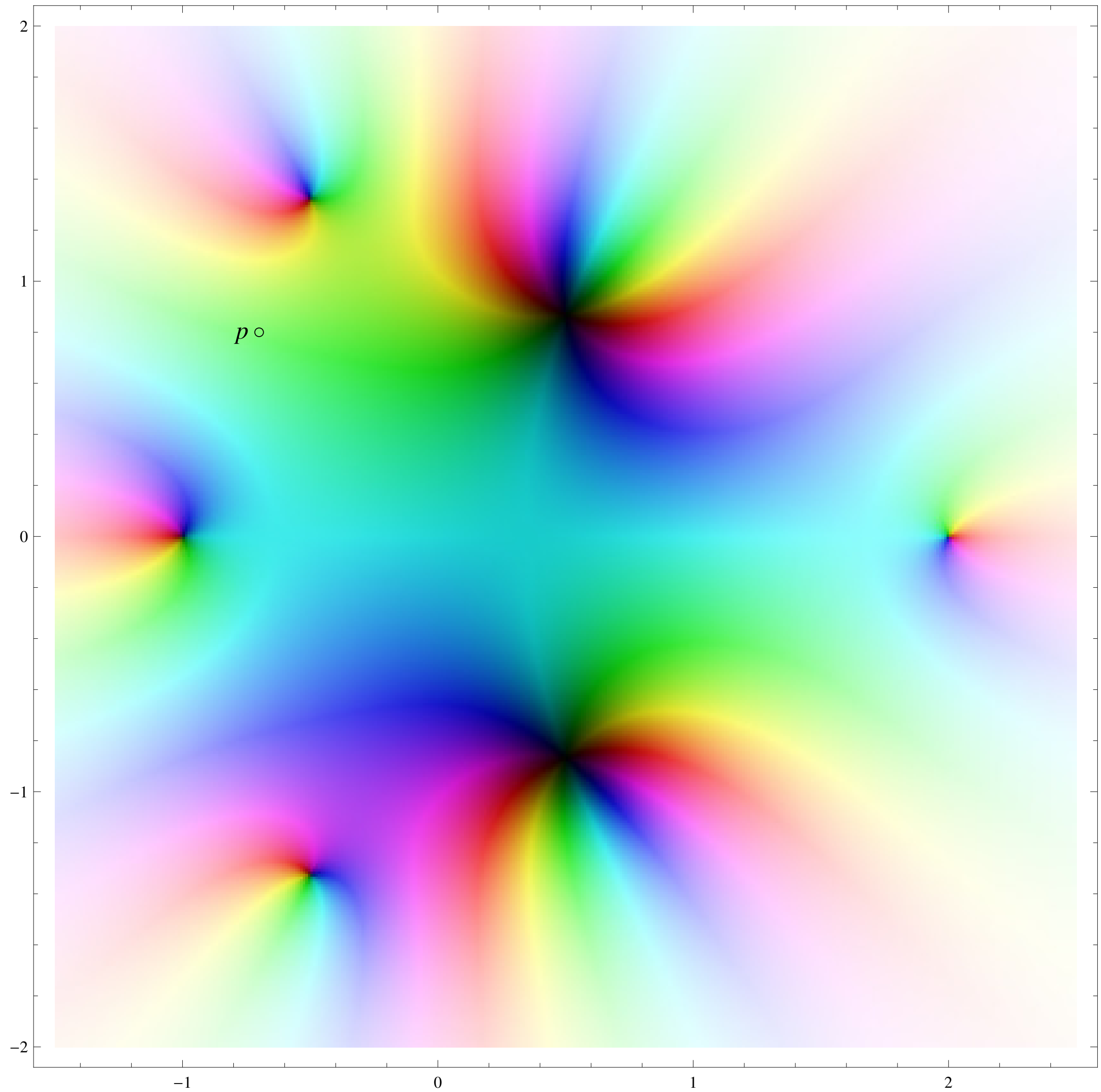

Figure 3: $f(z) = z^8 - 2z^7 + 2z^6 - 4z^5 + 2z^4 - 2z^3 - 5z^2 + 4z - 4.$

This is because when the polynomial $f(z)$ of Figure 3 is expanded in powers of $z+1$, the coefficient of $z+1$ is a negative real number.)

We might describe this situation by saying that at a single root, f is "locally linear," at a double root it is "locally quadratic," etc. In fact, a similar principle applies even at points that are not roots, although this is a little harder to see in our pictures. For any complex number z_0, by expanding $f(z)$ in powers of $z - z_0$ we can find complex numbers b and c and a positive integer k such that

$$f(z) = b + c(z - z_0)^k + (\text{higher degree terms}).$$

It follows that for z near z_0 we will have $f(z) \approx b + c(z - z_0)^k$. We have already seen that the picture of the function $c(z - z_0)^k$ has a black dot at z_0, with the colors that surround 0 in Figure 1 wrapping around z_0 k times. The effect of adding b to this function is to "shift" the colors in the color space of Figure 1 from 0 to b. The result is that z_0 will be colored with the color assigned to b, and it is the colors surrounding b in Figure 1 that will wrap around z_0 k times. (The fact that all of the colors surrounding b appear near z_0 illustrates that the Open Mapping Theorem holds for f.)

For example, in Figure 3 we have drawn a small circle at the point representing the complex number $-.7 + .8i$, and labeled this point p. The point p is colored light green, which is the color representing the complex number $f(-.7+.8i) \approx -7.5+15.8i$. Moving up from p, the colors become greenish-yellow and then yellow; moving down, they shift toward cyan. To the left there is a lighter shade of green, and to the right the shade of green gets darker. Referring to Figure 1, we see that these are the colors that surround light green in our color scheme. Thus, the colors surrounding light green wrap around p once; the polynomial is locally linear at p. (There are five points, other than the two double roots, at which the polynomial in Figure 3 is locally quadratic. It is an interesting exercise to try and locate them. Hint: There is one just below and to the right of the root at $(-1 + i\sqrt{7})/2$.)

Notice that one of the colors near p in Figure 3 is a darker shade of green. The reason, again, is that one of the colors near light green in Figure 1 is a darker green, and all of the colors surrounding light green are wrapped one or more times around every light green point in Figure 3. More generally, for every color in Figure 1 other than black, one of the nearby colors is a darker shade of the same color, and therefore in any picture of a nonconstant polynomial, any point that is not black will have a nearby point that is darker. It will be convenient to have a name for this principle:

Darker Neighbor Principle. *In any picture of a nonconstant polynomial, for any point that is not black, there is a nearby point that is darker.*

Using the Darker Neighbor Principle, we can now see why the Fundamental Theorem of Algebra is true. Suppose f is a nonconstant polynomial. Draw a picture of f on the square

$$S = \{x + iy : -R \leq x \leq R, -R \leq y \leq R\},$$

for some R. Since S is compact and $|f(z)|$ is continuous, there is a point in S at which $|f(z)|$ achieves its minimum value. This point will be the darkest point in the picture. We have already observed that, since the highest degree term of $f(z)$ will dominate the others when z is large, the colors in the picture will fade out toward white around the outside of the picture, if R is sufficiently large. It follows that the darkest point in the picture cannot be on the boundary of S, so this darkest point will be in the interior of S. But then this point must be black, because if it were not, then, by the Darker Neighbor Principle, some nearby point would be darker. This black point is a root of f.

The argument we have just given might be called a "colorized" version of d'Alembert's proof of 1746. The Darker Neighbor Principle is a colorized version of the key lemma of d'Alembert's proof:

D'Alembert's Lemma. *Suppose f is a nonconstant polynomial, and $f(z_0) \neq 0$. Then for every $\epsilon > 0$ there is some z such that $|z - z_0| < \epsilon$ and $|f(z)| < |f(z_0)|$.*

D'Alembert's proof of this lemma was not very rigorous, and it was unnecessarily complicated. (A simpler proof of the lemma was given by Jean-Robert Argand in 1806 [3].) Furthermore, the proof of the Fundamental Theorem of Algebra from d'Alembert's Lemma relies on the fact that a continuous real-valued function on a compact set achieves a minimum value, a fact that had not yet been rigorously proven in d'Alembert's time. Thus, d'Alembert's proof, while fairly easy to make rigorous using modern methods (see [13], [14, Section 3.5, pp. 31–33], [20, Problem 5.3, p. 44], [24], [27]), was not entirely convincing when d'Alembert published it.

Shortly after d'Alembert's proof, Leonhard Euler published an algebraic proof of the Fundamental Theorem of Algebra [11]. Euler's proof had a number of gaps in it, most of which were filled by Joseph-Louis Lagrange [19]. However, one significant gap remained: Euler and Lagrange both assumed that a polynomial of degree n would have n roots, and that the only thing that had to be proven was that these roots were complex numbers. (Today such reasoning could be justified by passing to an extension of $\mathbb{C}$ over which the polynomial splits, but in the 18th century the concepts needed to justify this reasoning had not yet been developed. See [30, Chapter 9].)

The first person to notice this gap was Gauss. In his doctoral dissertation in 1799, Gauss criticized Euler's proof:

> Since we cannot imagine forms of magnitudes other than real and imaginary magnitudes $a + b\sqrt{-1}$, it is not entirely clear how what is to be proved differs from what is assumed as [an axiom]; but granted one could think of other forms of magnitudes, say F, F', F'', ..., even then one could not assume without proof that every equation is satisfied either by a real value of x, or by a value of the form $a + b\sqrt{-1}$, or by a value of the form F, or of the form F', and so on. Therefore the [aforementioned axiom] can have only the following sense: every equation can be satisfied either by a real value of the unknown, or by an imaginary of the form $a + b\sqrt{-1}$, or, possibly, by a value of some as yet unknown form, or by a value not representable in any form. How these magnitudes, of which we can form no representation whatever—these shadows of shadows—are to be added or multiplied, this cannot be stated with the kind of clarity required in mathematics.[1]

Gauss also wrote:

> ... if one carries out operations with these impossible roots, as though they really existed, and says for example, the sum of all roots of the equation $x^m + ax^{m-1} + bx^{m-2} + \cdots = 0$ is equal to $-a$ even though some of them may be impossible (which really means: even if some are non-existent and therefore missing), then I can only say that I thoroughly disapprove of this type of argument.[2]

[1]The original text for this quotation, in Latin, can be found in [15, p. 14]. The translation is from [4, p. 98], but minor changes, indicated by brackets, have been made in the translation to clarify the meaning of the second sentence. My thanks to Cynthia Damon of the Amherst College Classics Department for help with the translation.

[2]The original Latin text can be found in [15, p. 5], and the translation is from [21].

Gauss then went on to give his own proof of the Fundamental Theorem of Algebra. We can illustrate the idea behind Gauss's proof in Figure 3. Gauss suggested that we consider separately the points where the real part of $f(z)$ is 0 and the points where the imaginary part is 0. Now, a complex number whose imaginary part is 0 is just a real number, and in Figure 1 we can see that the color assigned to a real number is either some shade of red (if the number is positive) or some shade of cyan (if it is negative). Similarly, complex numbers whose real part is 0 are those whose color is some shade of either yellow-green or magenta-blue. Thus, we can locate points in Figure 3 where the real or imaginary part of $f(z)$ is 0 by looking for these particular colors. Figure 4 is a copy of Figure 3 in which all of these points have been marked. The red curves in Figure 4 are the points where the real part of $f(z)$ is 0, and the green curves are the points where the imaginary part is 0. Notice that the red curves pass through points whose color is either yellow-green or magenta-blue, and the green curves pass through points whose color is either red or cyan.

As we observed earlier, going around the border of Figure 3, the color wheel cycle of colors is repeated eight times. Each cycle includes all four of the colors red, yellow-green, cyan, and magenta-blue, in order, and so along the border of Figure 4 there are 32 ends of curves, alternating red and green. Gauss asserted, without proof, that if we start at any one of these curve ends and follow the curve into the picture, we will emerge at another curve end of the same color. For example, starting at the red curve end just above the middle of the right side of Figure 4, we emerge at the red curve end just below the middle of the right side. Assuming that Gauss's assertion is correct, one can then use the fact that the colors of the curve ends around the border of the picture alternate between red and green to show that somewhere in the picture a red and green curve must intersect. This intersection point will be a point where the real and imaginary parts of $f(z)$ are both 0; in other words, it will be a root of f. Indeed, in Figure 4 we see that the red and green curves intersect at all six of the roots of f.

Although Gauss was critical of earlier attempts at proving the Fundamental Theorem, as we have seen his proof also included a step that was not rigorously justified. The first rigorous justification for this step was given in 1920 by Alexander Ostrowski [22]. Fortunately, Gauss eventually gave three more proofs of the theorem. His second proof, published in 1816 [16], was similar to Euler's proof, but did not assume the existence of the roots of the polynomial. This proof was perhaps the first essentially complete and correct published proof of the Fundamental Theorem of Algebra.

There is one more proof of the Fundamental Theorem that can be illustrated by reference to Figure 3. Imagine drawing a circle of some radius r, centered at the origin, on the picture in Figure 3. If r is small, then this circle will stay entirely in a region of the picture in which all points have a color that is close to some shade of cyan. It follows that the image of this circle under the function f will be a small closed curve in the complex plane that stays near some negative real number. On the other hand, if r is large then the circle will pass through eight cycles of the colors of the color wheel, and it follows that the image of the circle will be a closed curve that wraps around the origin eight times. This is confirmed by Figures 5 and 6, which show the images of circles of radius 0.1 and 3 respectively. If we now continuously increase r from 0.1 to 3, the first curve will be transformed continuously into the second. It seems clear that at some point in this transformation the curve must pass through the origin, which means that there must be some z such that $f(z) = 0$. Figure 7 shows what happens for a sequence of values of r ranging from 0.1 to 1.2 in steps of 0.1. (A similar argument can be found in [8].)

This intuitive argument can be turned into a rigorous proof by using the concept of the winding number of a closed curve. Suppose f is a polynomial of degree $n > 0$, and $f(0) \neq 0$. Then for

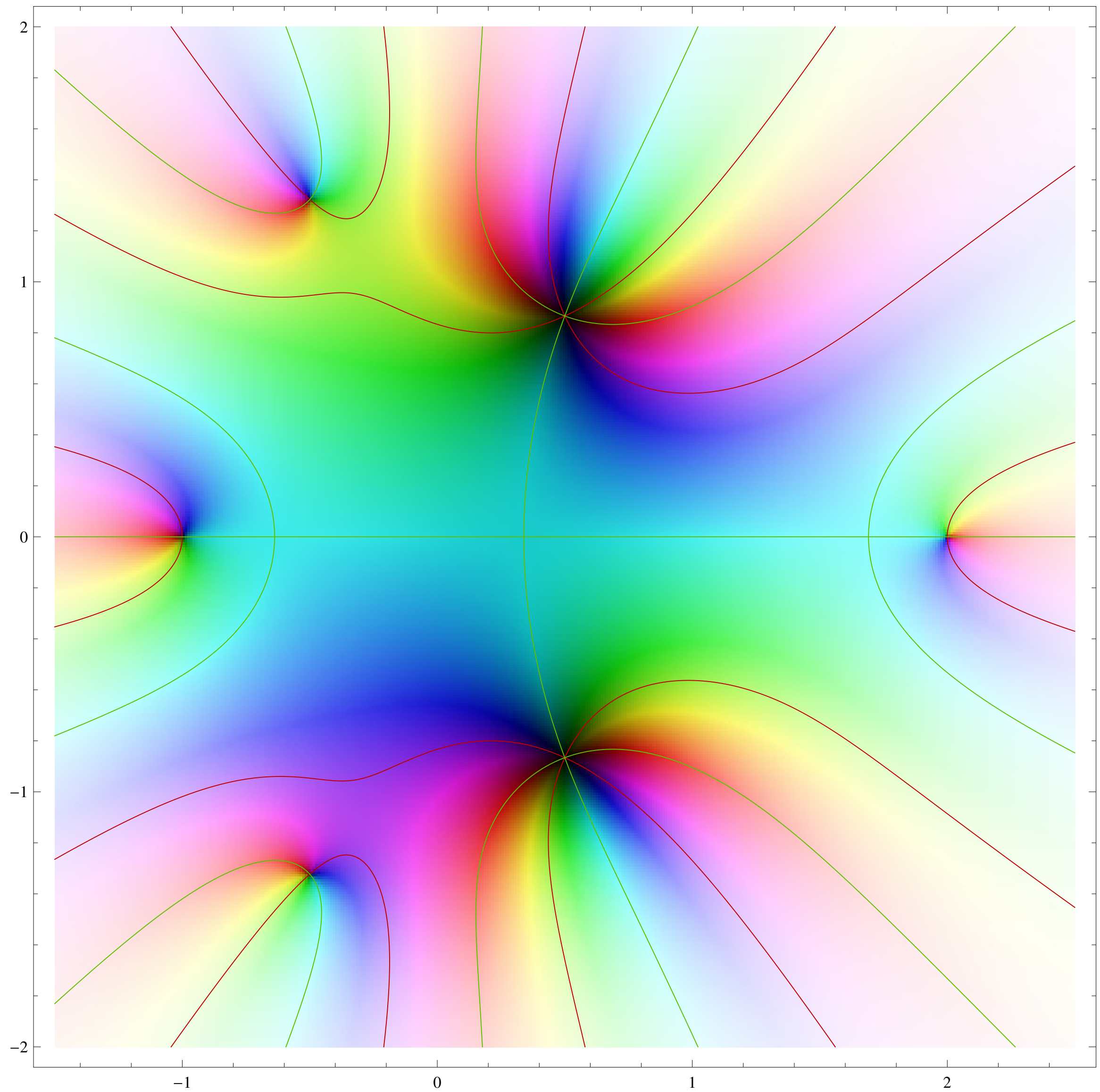

Figure 4: Gauss's first proof of the Fundamental Theorem of Algebra.

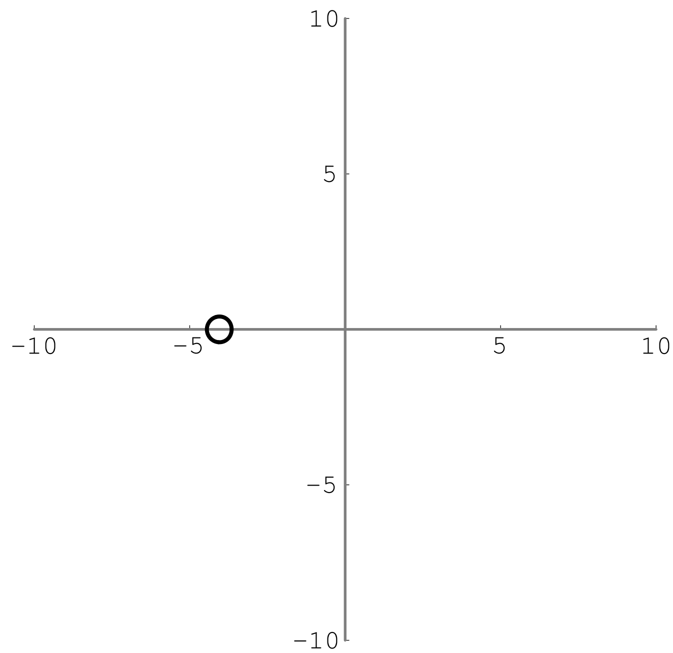

Figure 5: Image of a circle of radius 0.1 under f.

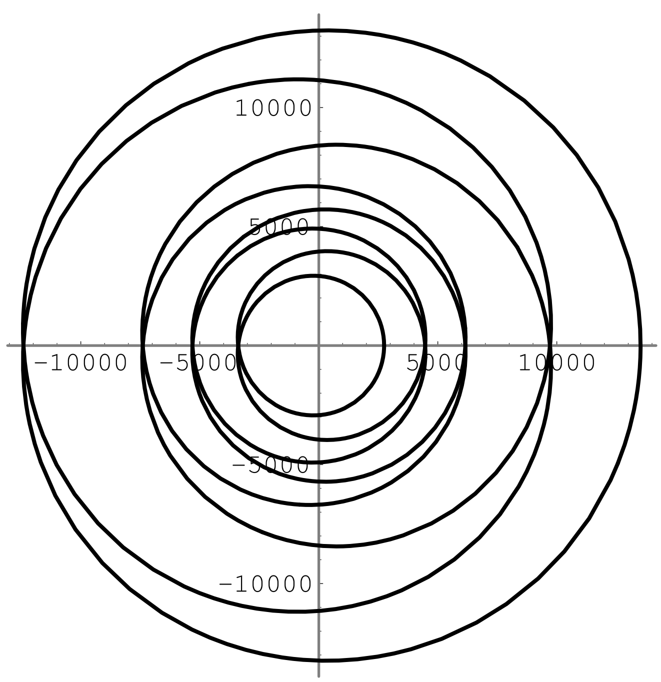

Figure 6: Image of a circle of radius 3 under f.

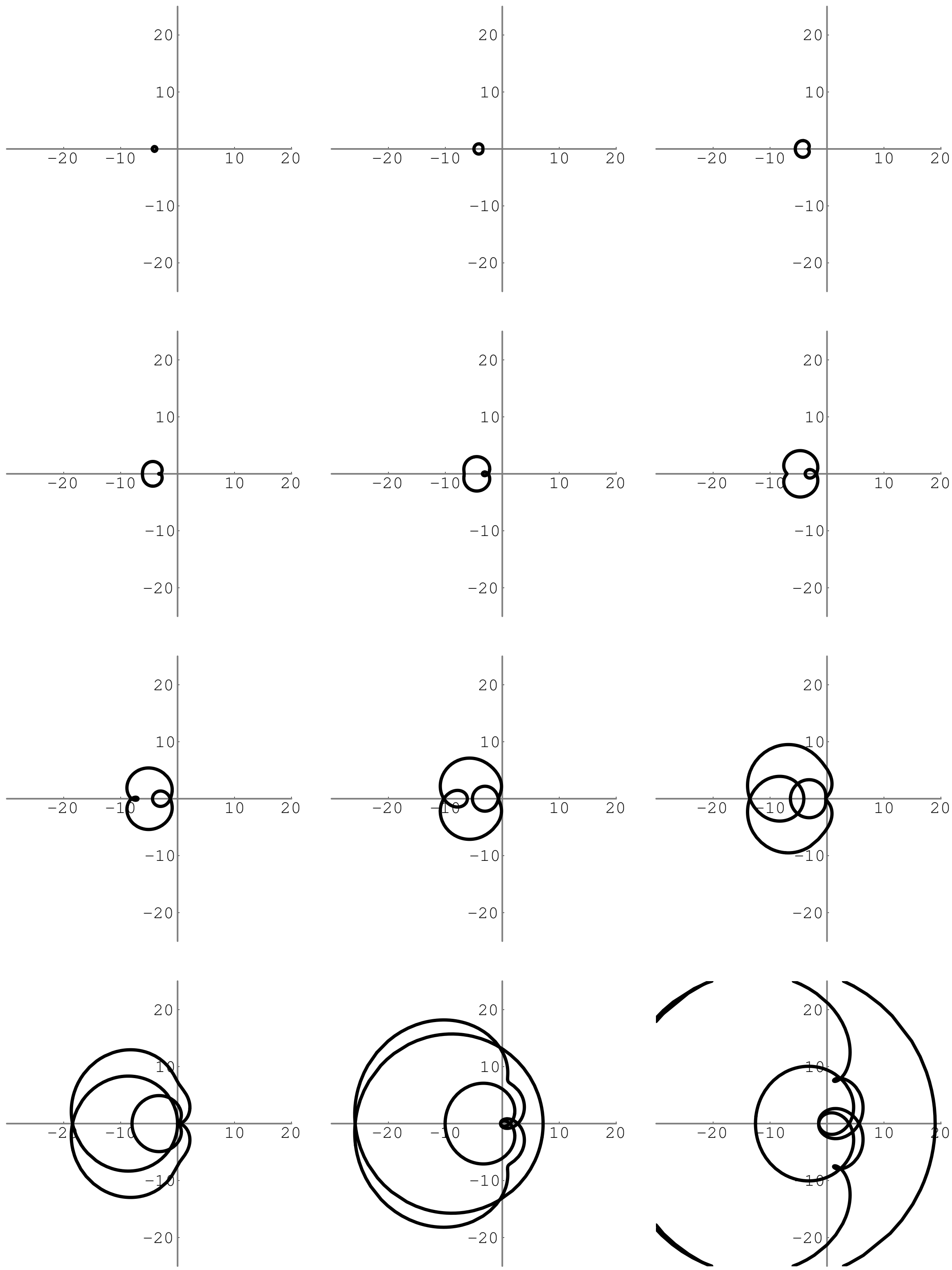

Figure 7: Images of circles with radii from 0.1 to 1.2 under f.

sufficiently small r, the image under f of a circle of radius r centered at the origin will be a closed curve whose winding number around the origin is 0. For large r, the image will be a curve with winding number n. But the winding number of a closed curve around the origin is unchanged if the curve is continuously transformed without passing through the origin. It follows that for some r, the image of the circle of radius r must pass through the origin. Details of this proof can be found in [14, Proof Five, pp. 134–136].

Although we have concentrated on pictures of polynomials in this paper, the scheme used in Figure 3 can be used to make pictures of any function $f : \mathbb{C} \to \mathbb{C}$. For the reader's amusement, we include in Figures 8 and 9 pictures of e^z and a branch of $\log(z)$. In Figure 8, it is evident that the magnitude and argument of e^z are determined by the real and imaginary parts of z, respectively. In Figure 9, the pole at 0 appears as a white dot, and a branch cut is visible along the negative real axis. Similar pictures can be found in [9], [12], [28], and [29].

For readers who would like to be able to create similar pictures themselves, we provide here the Mathematica code that was used to create our color figures:

```
brightness[x_] := If[x <= 1, Sqrt[x]/2, 1 - 8/((x + 3)^2)]
ComplexColor[z_] := If[z == 0, Hue[0, 1, 0],
  Hue[Arg[z]/(2 Pi), (1 - (b = brightness[Abs[z]])^4)^0.25, b]]
ComplexPlot[xmin_, ymin_, xmax_, ymax_, xsteps_, ysteps_] :=
  Block[{deltax = N[(xmax - xmin)/xsteps], deltay = N[(ymax - ymin)/ysteps]},
  Graphics[Raster[Table[f[x + I y],
  {y, ymin + deltay/2, ymax, deltay}, {x, xmin + deltax/2, xmax, deltax}],
  {{xmin, ymin}, {xmax, ymax}}, ColorFunction -> ComplexColor]]]
```

The function `ComplexColor` assigns a color to any complex number z, as in Figure 1. The color is given by specifying its hue, saturation, and brightness. The hue of the color assigned to z is determined by $\arg(z)$, and the saturation and brightness are expressed in terms of $|z|$, using formulas that were found by experimentation. The function `ComplexPlot` produces a color plot of the function f (which must be defined first).

References

[1] A. Abian, *A new proof of the fundamental theorem of algebra*, Caribbean J. Math. **5** (1986), no. 1, 9–12.

[2] Jean Le Rond d'Alembert, *Recherches sur le calcul integral*, Hist. Acad. Sci. Berlin **2** (1746), 182–224.

[3] Jean-Robert Argand, *Essai sur une manière de représenter les quantités imaginaires dans les constructions géométriques*, Paris, 1806.

[4] Isabella Bashmakova and Galina Smirnova, *The Beginnings and Evolution of Algebra*, Translated from the Russian by Abe Shenitzer, with the editorial assistance of David Cox, Mathematical Association of America, Washington, 2000.

[5] Joseph Bennish, *Another proof of the fundamental theorem of algebra*, Amer. Math. Monthly **99** (1992), 426.

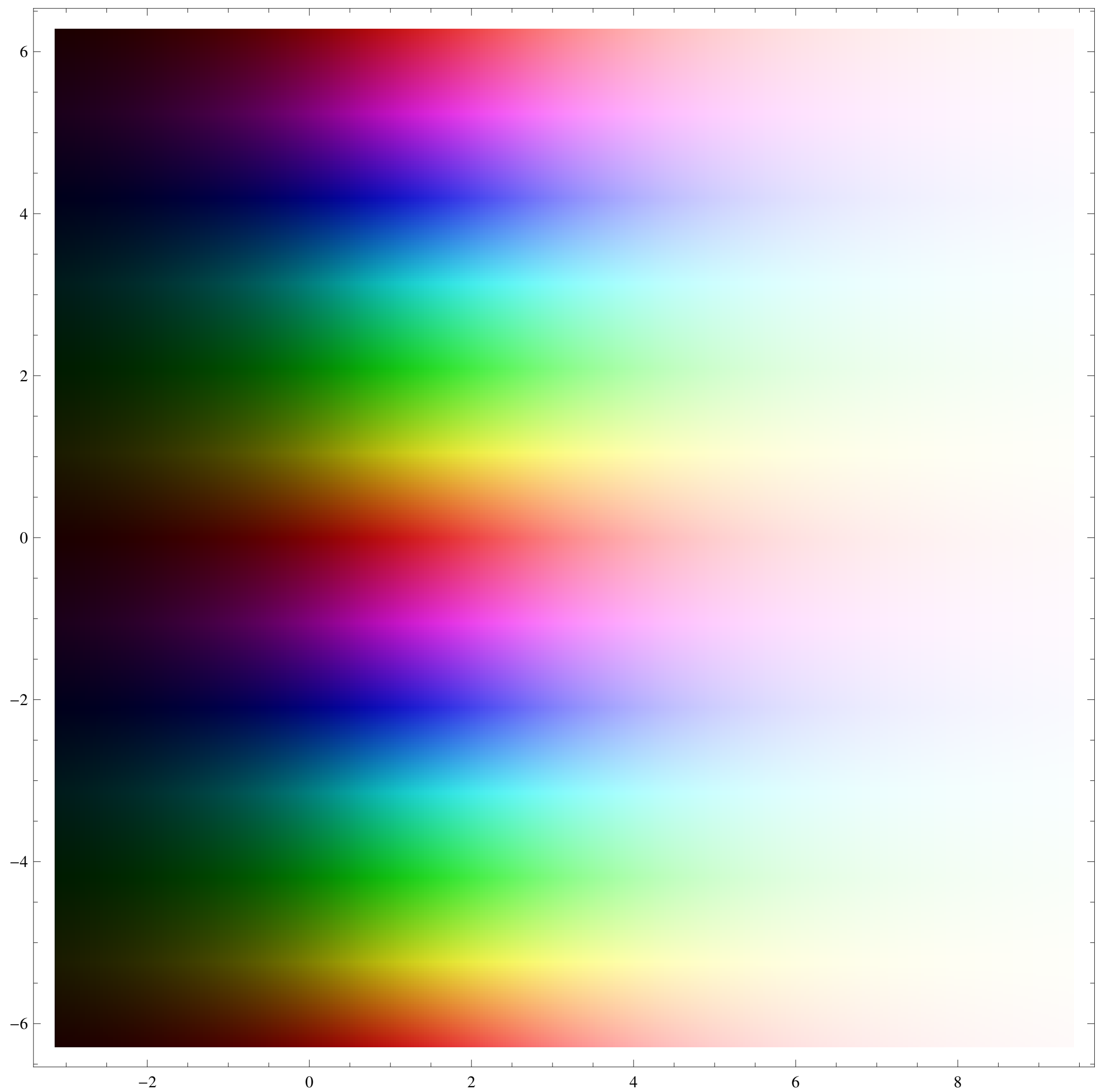

Figure 8: $f(z) = e^z$.

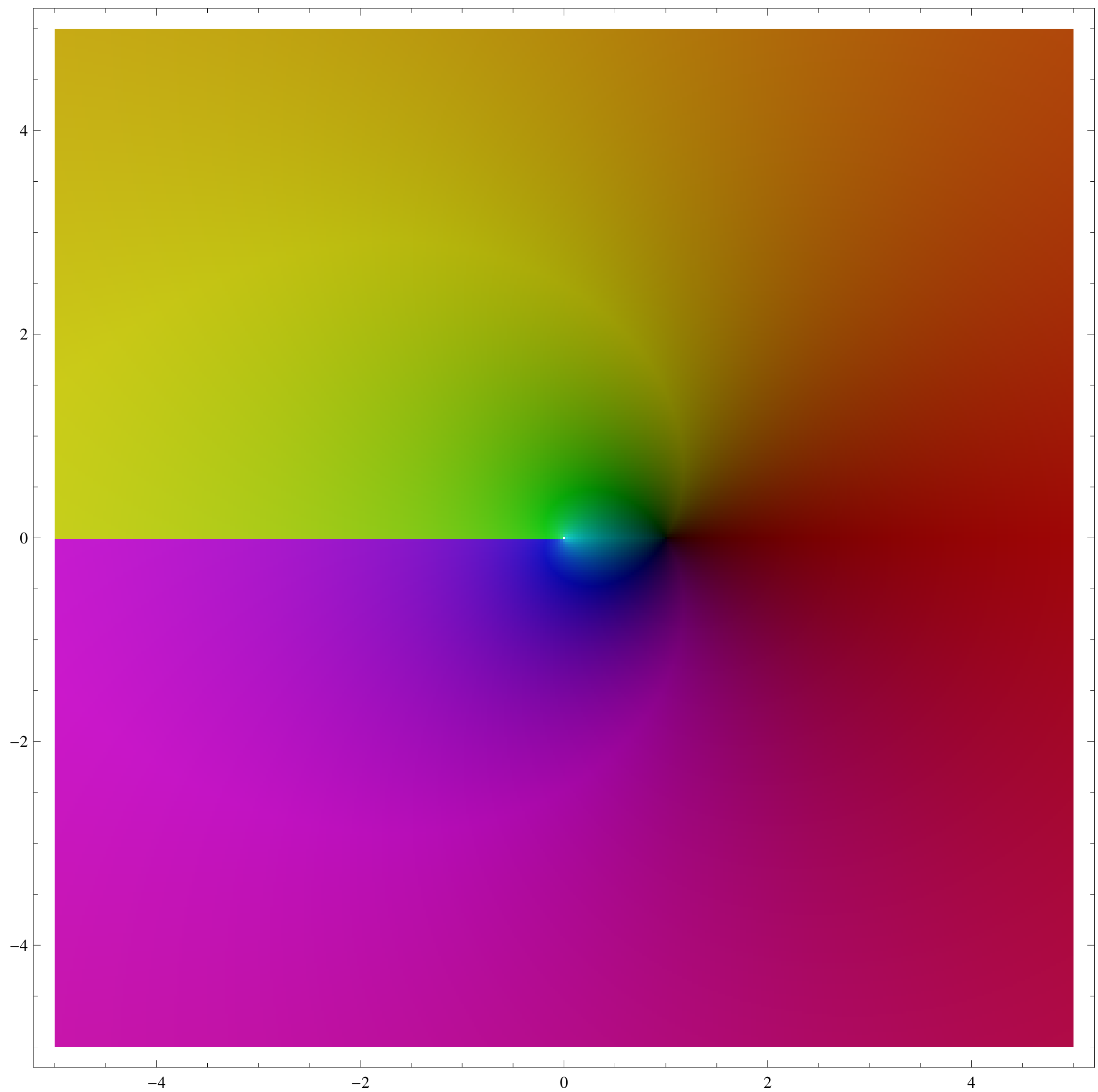

Figure 9: A branch of $f(z) = \log z$.

[6] R. P. Boas, Jr., *A proof of the fundamental theorem of algebra*, Amer. Math. Monthly **42** (1935), 501–502.

[7] R. P. Boas, Jr., *Yet another proof of the fundamental theorem of algebra*, Amer. Math. Monthly **71** (1964), 180.

[8] John Byl, *A simple proof of the fundamental theorem of algebra*, Internat. J. Math. Ed. Sci. Tech. **30** (1999), no. 4, 602–603.

[9] Lawrence Crone, *Color graphs of complex functions*, `http://nw08.american.edu/~lcrone/ComplexPlot.html`.

[10] William Dunham, *Euler and the fundamental theorem of algebra*, College Math. J. **22** (1991), 282–293.

[11] Leonhard Euler, *Recherches sur les racines imaginaires des équations*, Mém. Acad. Sci. Berlin **5** (1749), 222–288, reprinted in *Opera Omnia Series Prima*, vol. 6, 78–147.

[12] Frank A. Farris, *Visualizing complex-valued functions in the plane*, `http://www.maa.org/visualizing-complex-valued-functions-in-the-plane`.

[13] Charles Fefferman, *An easy proof of the fundamental theorem of algebra*, Amer. Math. Monthly **74** (1967), 854–855.

[14] Benjamin Fine and Gerhard Rosenberger, *The Fundamental Theorem of Algebra*, Springer-Verlag, New York, 1997.

[15] Carl Friedrich Gauss, *Demonstratio nova theorematis omnem functionem algebraicum rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse*, Helmstedt dissertation, 1799, reprinted in *Werke*, Vol. 3, 1–30.

[16] Carl Friedrich Gauss, *Demonstratio nova altera theorematis omnem functionem algebraicum rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse*, Comm. Recentiores (Gottingae) **3** (1816), 107–142, reprinted in *Werke*, Vol. 3, 31–56.

[17] Javier Gomez-Calderon and David M. Wells, *Why polynomials have roots*, College Math. J. **27** (1996), 90–91.

[18] Michael D. Hirschhorn, *The fundamental theorem of algebra*, College Math. J. **29** (1998), 276–277.

[19] Joseph-Louis Lagrange, *Sur la forme des racines imaginaires des équations*, Nouv. Mém. Acad. Berlin, 1772, 479–516, reprinted in *Œuvres*, Vol. 3, 479–516.

[20] Norman Levinson and Raymond M. Redheffer, *Complex Variables*, Holden-Day, San Francisco, 1970.

[21] MacTutor History of Mathematics Archive, *The fundamental theorem of algebra*, `http://www-groups.dcs.st-and.ac.uk/~history/HistTopics/Fund_theorem_of_algebra.html`.

[22] Alexander Ostrowski, *Über den ersten und vierten Gausschen Beweis des Fundamentalsatzes der Algebra*, in *Gauss Werke*, Vol. 10, Part 2, Abh. 3.

[23] Raymond Redheffer, *The fundamental theorem of algebra*, Amer. Math. Monthly **64** (1957), 582–585.

[24] R. M. Redheffer, *What! Another note just on the fundamental theorem of algebra?*, Amer. Math. Monthly **71** (1964), 180–185.

[25] Anindya Sen, *Fundamental theorem of algebra—yet another proof*, Amer. Math. Monthly **107** (2000), 842–843.

[26] John Stillwell, *Mathematics and its History*, Springer-Verlag, New York, 1989.

[27] Frode Terkelsen, *The fundamental theorem of algebra*, Amer. Math. Monthly **83** (1976), 647.

[28] Bernd Thaller, *Visualization of complex functions*, Mathematica J., Vol. 7, issue 2 (1998), `http://www.mathematica-journal.com/issue/v7i2/features/thaller/`.

[29] Bernd Thaller, *Visual Quantum Mechanics: Selected Topics with Computer-Generated Animations of Quantum-Mechanical Phenomena*, Springer/TELOS, New York, 2000.

[30] Jean-Pierre Tignol, *Galois' Theory of Algebraic Equations*, World Scientific, Singapore, 2001.

[31] Daniel J. Velleman, *Another proof of the fundamental theorem of algebra*, Math. Mag. **70** (1997), 216–217.