Math 79SI Notes

Mason Rogers

2 Direct Proof

2.1 Background

Many mathematical questions cannot be answered by explicitly constructing an object or finding a counterexample, and we would like to develop strategies to handle them. To do so, we split statements into what they assume as input (their *hypotheses* or *assumptions*) and what they state as a consequence (their *conclusions*). For instance, the statement 'if $x \geq 0$ is a real number, then $x = y^2$ for some real number y' has the hypothesis 'x is a real number greater than or equal to 0' and the conclusion '$x = y^2$ for some real number y'.

A statement of form 'If P then Q', where P is a list of hypotheses and Q a list of conclusions, is called an *implication*. However, not all implications are immediately written in if/then format. For instance, the statement 'for all p in the set of prime numbers, there exists a larger prime number q' is an implication; we can put it in if/then format by writing the equivalent[1] statement 'if p is a prime number, then there exists a larger prime number q'. It is now clear that the hypothesis is that p is a prime, and the conclusion is that there exists a larger prime q.

To prove such a statement, we must show that whenever the hypotheses hold, the conclusions must hold. The first strategy to do so is called *direct proof.* To prove something directly, we begin with the the hypotheses, then state something relevant that they imply, then something relevant that *that* implies, and so on, until we imply the conclusions themselves. Of course, we do not want to plug forward blindly; we let our intuition tell us what exactly is relevant at each step. Shortly, we will do some examples to get comfortable with this strategy.

Before we start, we make life easier by introducing a few bits of notation. Mathematicians use double-bar letters to denote a few famous sets as follows:

- $\mathbb{N}$: the natural numbers, $\{1, 2, 3, \ldots\}$
- $\mathbb{Z}$: the integers, $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$
- $\mathbb{Q}$: the rational numbers, $\{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\}$

[1] Many proof-writing courses and textbooks begin with long discussions of implications, equivalences, and other formal logic concepts in terms of truth tables. We exclude this discussion because it is neither particularly interesting nor important moving forward, but if you are curious, then the further reading section 2.4 offers references.

- $\mathbb{R}$: the real numbers, for example π, 3, $0.57721566\ldots$, etc.[2]
- $\mathbb{C}$: the complex numbers, $\{a + bi : a, b \in \mathbb{R}, i^2 = -1\}$

These, and phrases such as 'choose $x \in \mathbb{R}$', are often seen in formal writing. One more shorthand that is more commonly used for taking notes or discussing math is the implication arrow $\implies$. For instance, we could rewrite 'all rational numbers have eventually repeating decimal expansions' first as 'if q is a rational number, then q has an eventually repeating decimal expansion', then as '$q \in \mathbb{Q} \implies q = 10^{-k} \left(n + \sum_{i=1}^{\infty} d_i 10^{-i}\right)$ for some $n, k \in \mathbb{Z}$ and $d_i \in \{0, \ldots, 9\}$ satisfying $d_i = d_{i+N}$ for some $N \in \mathbb{N}$'.

2.2 Examples

We start with a simple example to indicate the utility, potential pitfalls, and shortcomings of direct proof.

Proposition 2.1. Choose odd numbers a and b. Prove that their product ab is odd.

We provide two proofs, one cumbersome and one straightforward. In both proofs, we avoid invoking facts (to be proven later) about prime factorizations of integers, which are *overkill* for this statement.

Proof 1 (cumbersome). Let x be the units digit of a in base 10 and y be the units digit of b in base 10. Then because a and b are odd, $x, y \in \{1, 3, 5, 7, 9\}$. Let z be the units digit of ab in base 10. We can compute z in each of the 25 cases of (x, y), as summarized in table 1.

Table 1: Values of z for each case of (x, y)

$x \setminus y$	1	3	5	7	9
1	1	3	5	7	9
3	3	9	5	1	7
5	5	5	5	5	5
7	7	1	5	9	3
9	9	7	5	3	1

In every case, $z \in \{1, 3, 5, 7, 9\}$, so we conclude that ab is odd. □

Proof 2 (straightforward). Since a and b are odd, we write $a = 2k + 1$ and $b = 2\ell + 1$ for integers k and ℓ. Then $ab = (2k + 1)(2\ell + 1) = 2(2k\ell + k + \ell) + 1$. Let $n = 2k\ell + k + \ell$, which is an integer. Then $ab = 2n + 1$, which is odd. □

The proof by checking cases is indeed a valid proof, but it is inefficient. Starting with a slightly more conceptual characterization of odd numbers as in the second argument above leads to a quick and lucid proof. In general, case checking should only be used if no better

[2] It is difficult to define the real numbers using set builder notation, so we stick to giving a few examples. A common way is to define the reals as a type of *completion* of the rational numbers, which is a topic in real analysis (Math 115/171).

or faster argument is apparent. This may be the case when there are a small number of *functionally distinct* cases (e.g. even versus odd, but not different units digits within those categories), or when a computer can check a large (but finite) space of possibilities. However, a proof without checking cases is often more efficient and enlightening. The second proof of Proposition 2.1 argues that the product of two numbers that are 1 more than a multiple of 2 is itself 1 more than a multiple of 2. This is a more useful than a collection of 25 different statements about products of units digits.

In light of Proposition 2.1, we might want to know if the following related statement is true.

Proposition 2.2. If the product ab of two natural numbers a and b is odd, then a and b are both odd.

This is known as the *converse* of the statement in Proposition 2.1, which we will define in the next section. To prove it directly would be difficult, as starting with $ab = 2n + 1$ for some integer n does not seem to provide information about the parity of the factors a and b of $2n + 1$. You are welcome to try to think of another approach to proving this statement, which we will revisit in the next section with a new technique.

We have now seen a good direct proof, a bad direct proof, and a related true statement that is not easy to prove directly. In the previous section, we saw counterexamples generated with varying degrees of efficiency and statements that were not amenable to proof by construction or disproof by counterexample. This pattern will continue: every proof technique that we will encounter is better adapted to some situations than others. For the remainder of this section, we will gain practice with a few more challenging examples of direct proof with the broad theme of finding roots to polynomials.

Before diving in, we need a few preliminary notions that will come up in the examples we study. For integers r and r not both 0, we define the *greatest common divisor* $\gcd(r, s)$ to be the largest positive integer which divides[3] both r and s. If $s \neq 0$, we say that the fraction $\frac{r}{s}$ is in *lowest terms* if $\gcd(r, s) = 1$.

You may recognize these words and think of several associated facts, but it is important to remember when facts do not follow immediately from the definitions at hand. For instance, it is true (and not too difficult to prove) that any rational number q can be expressed as a fraction in lowest terms, but this fact is not included in the definition of 'lowest terms'. It is also true that $\gcd(r, s)$ is a multiple of any common divisor of r and s; however, this is more challenging to prove. If you try to think of a proof, you may think in terms of the prime factorizations of r and s. However, the fundamental theorem of arithmetic, which states that any natural number greater than 1 has a *unique* prime factorization, is itself somewhat tricky to prove.[4]

To maintain focus on our discussion of polynomials, we will assume the fundamental theorem of arithmetic for now, as well as the two non-obvious facts stated above. We also

[3] We say that m *divides* n if there exists an integer k such that $n = km$, i.e. m is a factor of n. Observe that with this definition, any positive integer divides 0, so for $s \neq 0$, $\gcd(0, s) = |s|$.

[4] The fundamental theorem of arithmetic asserts both the *existence* and *uniqueness* of prime factorizations of natural numbers $n \geq 2$. The existence part can be proven somewhat easily from the definition of a prime number, but the uniqueness part takes work. If you are skeptical that the uniqueness of prime factorizations is difficult to prove, try proving it!

need one preliminary claim that we will take the time to prove via the fundamental theorem of arithmetic.

Claim. Let a and b be integers such that $b > 0$ and $\gcd(a, b) = 1$. Then for any $n \in \mathbb{N}$, $\gcd(a^n, b) = 1$.

Proof of claim. We will prove the claim separately for the cases $|a| \leq 1$ and $|a| \geq 2$, which clearly exhaust all integers.

We begin with the case $a \in \{-1, 0, 1\}$. Observe that $\gcd(r, s) = \gcd(-r, s)$, since the positive divisors of r and $-r$ are the same. Since for all $a \in \{-1, 0, 1\}$, we have $a^n = \pm a$. For such a, we have that $\gcd(a^n, b) = \gcd(a, b)$. By assumption, $\gcd(a, b) = 1$, so we conclude that $\gcd(a^n, b) = 1$.

We now turn our attention to the case in which $|a| \geq 2$, for which we will invoke the fundamental theorem of arithmetic to write a uniquely as a product of primes $\pm p_1^{e_1} \cdot p_2^{e_2} \cdots p_\ell^{e_\ell}$. Since $\gcd(a, b) = 1$, we see that none of the p_i can divide b, as otherwise p_i would be a common divisor of a and b greater than 1. Now consider the unique prime factorization for a^n, which is given by $\pm p_1^{ne_1} \cdot p_2^{ne_2} \cdots p_\ell^{ne_\ell}$. Still, none of the p_i divides b, but any divisor $d > 1$ of a^n *must* be some product of the p_i since upon setting $a^n = dd'$ for an integer d', the *uniqueness* of the prime factorization of a^n forces any prime factor q of d to be some p_i. Thus, no such q can be a factor of b. We therefore have that $\gcd(a^n, b) = 1$, proving the claim. □

Claim in hand, we are ready to begin our discussion about polynomials. Polynomial equations are ubiquitous in math, physics, economics, and more, so it is important to understand when and how they can be solved. Let us take a look at a few simple polynomials with integer coefficients and their *real* roots (i.e. roots in $\mathbb{R}$).

- $x^2 - 1 = 0$ has the solutions $x = \pm 1$.
- $x^3 - 1 = 0$ has the solution $x = 1$.
- $x^2 - 2 = 0$ has the solutions $x = \pm\sqrt{2} \approx \pm 1.41421356237\ldots$
- $x^3 - 2 = 0$ has the solution $x = \sqrt[3]{2} \approx 1.25992104989\ldots$
- $x^2 - 3 = 0$ has the solutions $x = \pm\sqrt{3} \approx \pm 1.73205080757\ldots$
- $x^2 - 4 = 0$ has the solutions $x = \pm 2$.
- $x^3 - 4 = 0$ has the solution $x = \sqrt[3]{4} \approx 1.58740105197\ldots$

Already, a pattern is emerging: the square and cube roots of integers seem either to be integers themselves or to have very long, apparently non-repeating decimal expansions, i.e. are irrational. More precisely, it looks like the n-th roots of integers are either integers or irrational numbers. To decide whether or not this is true, we need to formulate a precise statement, then *prove* it. We thus propose the following.

Proposition 2.3. If q is a rational n-th root of an integer k, then q is an integer.

Proof. Suppose q is a rational n-th root of some integer k, so $q = \frac{a}{b}$ for some $a, b \in \mathbb{Z}$ with $b > 0$, and $q^n - k = 0$. Without loss of generality,[5] we assume that $\frac{a}{b}$ is a fraction in lowest terms.

To show that q is an integer, we must show that $b = 1$, which we begin to do by substituting $q = \frac{a}{b}$ into $x^n - k = 0$. This gives $\left(\frac{a}{b}\right)^n = k$, or after multiplying through by b^n,

$$a^n = kb^n. \tag{1}$$

The right side is a multiple of b, so by the definition of greatest common divisor, we know that $\gcd(kb^n, b) \geq b$. Using equation (1), we then have $\gcd(a^n, b) \geq b$. We know that $\gcd(a^n, b) = 1$ regardless of a as proven in the above claim, so we have

$$1 = \gcd(a^n, b) = \gcd(kb^n, b) \geq b > 0.$$

From this, it follows that $b = 1$, so $q = a$. We conclude that if q is a rational n-th root of an integer k, then q is necessarily an integer. □

Let us recap our above argument. Our hypothesis was that q is a rational n-th root of an integer k. This allowed us to substitute in a reduced fraction $\frac{a}{b}$ for q in the equation $q^n - k = 0$. We arrived at equation (1), then argued that for both sides of the equation to be divisible by b, we must have $b = 1$. From this, we concluded that q must be an integer, as we wanted.

The same argument may apply for a larger class of polynomials than $x^n - k$, and looking at a few more examples further supports the idea:

- $x^3 - 3x^2 + x + 1 = 0$ has the solutions $x = 1$, $x = 1 - \sqrt{2} \approx -.41421356\ldots$, and $x = 1 + \sqrt{2} \approx 2.41421356\ldots$

- $x^4 - 4x^3 - 6x^2 + 28x - 16 = 0$ has the solutions $x = 2$, $x = 4$, $x = -1 + \sqrt{3} \approx .73205080\ldots$, and $x = -1 - \sqrt{3} \approx -2.73205080\ldots$

These polynomials seem to match the pattern we witnessed for n-th roots even though they have more nonzero terms. In fact, the key commonality of these integer-coefficient polynomials is actually that the leading coefficient is 1. We formulate this in a statement as follows:

Proposition 2.4. If q is a rational root of a *monic* polynomial, i.e. a polynomial with leading coefficient 1, with integer coefficients, then q is an integer.

Proof. Our argument will be almost identical to before. Let q be a rational root of $p(x) = x^n + c_{n-1}x^{n-1} + \ldots + c_1x + c_0$, where $c_0, c_1, \ldots, c_{n-1} \in \mathbb{Z}$. We can write $q = \frac{a}{b}$, where $a, b \in \mathbb{Z}$, $b > 0$, and $\gcd(a, b) = 1$. Substituting $q = \frac{a}{b}$ into $p(q)$, we have

$$\left(\frac{a}{b}\right)^n + c_{n-1}\left(\frac{a}{b}\right)^{n-1} + \ldots + c_1\frac{a}{b} + c_0 = 0.$$

[5] Mathematicians often use the phrase 'without loss of generality' or the informal abbreviation 'WLOG' to indicate a simplifying assumption that can always be made. In this proof, we are only intersted in the fact that q is rational, not that q is expressed as a particular fraction $\frac{c}{d}$. Thus, we can assume that $\frac{a}{b}$ is in lowest terms without affecting the generality of the proof.

Ultimately, we want to use properties of integers such as divisibility, so we multiply through by $b^n \neq 0$ to obtain the equivalent expression

$$a^n + c_{n-1}a^{n-1}b + \ldots + c_1ab^{n-1} + c_0b^n = 0. \tag{2}$$

Subtracting a^n from both sides and factoring out a b from the left side gives

$$b\left(c_{n-1}a^{n-1} + \ldots + c_1ab^{n-2} + c_0b^{n-1}\right) = -a^n.$$

Invoking the claim from above to obtain $\gcd(a^n, b) = 1$, we can see that

$$1 = \gcd(a^n, b) = \gcd(-a^n, b) = \gcd\left(b(c_{n-1}a^{n-1} + \ldots + c_0b^{n-1}), b\right) \geq b > 0.$$

It is clear then that $b = 1$, so q is an integer. $\square$

We applied the same method used to reach the same conclusion as in the previous proposition, but under less restrictive hypotheses. Instead of assuming q to be a rational root of a polynomial of form $x^n - k$, we simply assumed q to be a rational root of *some* monic polynomial with integer coefficients. We will use the above style of argument (albeit with a slight twist) one more time to prove a significant theorem.

Theorem 2.5. *(Rational Root Theorem)* Let $p(x) = c_nx^n + c_{n-1}x^{n-1} + \ldots + c_1x + c_0$ be a polynomial with integer coefficients such that neither c_0 nor c_n is zero, and let $\frac{a}{b}$ be a rational root of $p(x)$ expressed in lowest terms. Then b divides c_n and a divides c_0.

Proof. Using the same manipulations as used in the previous proposition to arrive at equation (2), we have

$$c_na^n + c_{n-1}a^{n-1}b + \ldots + c_1ab^{n-1} + c_0b^n = 0. \tag{3}$$

We will begin by proving that b divides c_n by showing that $\frac{c_n}{b}$ is an integer. Subtracting c_na^n from both sides and factoring out a b from the left side gives

$$b\left(c_{n-1}a^{n-1} + \ldots + c_1ab^{n-2} + c_0b^{n-1}\right) = -c_na^n.$$

Since $b \neq 0$, we can divide both sides by b. Since the left side is a multiple of b, dividing either side by b must still give an integer. We thus have that $\frac{-c_na^n}{b}$ is an integer, and it remains to show that $\frac{c_n}{b}$ is an integer.

We will do so separately for the cases $b = 1$ and $b \geq 2$. If $b = 1$, then clearly $\frac{c_n}{b} = c_n$ is an integer. If b is greater than 1, then we can write b as a *unique* product of primes $q_1^{e_1} \cdot q_2^{e_2} \cdots q_\ell^{e_\ell}$ by the fundamental theorem of arithmetic.[6] Since $\gcd(-a^n, b) = 1$ by the same logic as previously used, we know that the unique prime factorization of $-a^n$ cannot contain any of the q_i. However, we know that the *unique* prime factorization of the product

[6] We are proving a form of Euclid's lemma, which states that if a prime p divides a product ab, then p divides at least one of a and b. This is a key step in the conventional proof of the fundamental theorem of arithmetic. It therefore may seem circular to use the fundamental theorem of arithmetic to prove it. However, we will see proofs later in the course of the fundamental theorem of arithmetic that do *not* require proving Euclid's lemma. We use the 'fundamental theorem' approach both to take advantage of our intuition and to minimize the number of facts we are assuming throughout this section, but a more sophisticated approach could avoid introducing the fundamental theorem of arithmetic here.

$-c_n a^n$ *must* contain $b = q_1^{e_1} \cdot q_2^{e_2} \cdots q_\ell^{e_\ell}$. Therefore, we can conclude that the unique prime factorization of c_n contains $q_1^{e_1} \cdot q_2^{e_2} \cdots q_\ell^{e_\ell}$, and accordingly that c_n is divisible by b as desired. This exhausts all possibilities of b, so we conclude that b must divide c_n always.

It remains to prove the second conclusion that a divides c_0. The proof will go similarly to that of the first conclusion. We return to equation (3) and subtract off $c_0 b^n$ from both sides and factor out an a from the left side to obtain

$$a\left(c_n a^{n-1} + c_{n-1} a^{n-2} b + \ldots + c_1 b^{n-1}\right) = -c_0 b^n.$$

Since neither c_0 nor b is equal to zero, we know that neither side of the above expression is zero, and consequently $a \neq 0$. We thus have that $-c_0 b^n$ is a nonzero multiple of a, and accordingly that $\frac{-c_0 b^n}{a}$ is an integer. It remains to show that $\frac{c_0}{a}$ is an integer. We would like to use the same argument that we used to show that $\frac{c_n}{b}$ is an integer, which we can do if $\gcd(a, -b^n) = 1$. Indeed, proving that if $\gcd(a, b) = 1$ then $\gcd(a, -b^n) = 1$ proceeds almost identically to the proof that if $\gcd(a, b) = 1$ then $\gcd(-a^n, b) = 1$, which we have already proven. Thus, we have that $\gcd(a, -b^n) = 1$, and can argue as before that $\frac{c_0}{a}$ is an integer.

We conclude that b divides c_n and that a divides c_0, which concludes the proof of the theorem. □

The rational root theorem is quite powerful because it narrows down the search for an integer-coefficient polynomial's rational roots to a finite list. As an application of the theorem, consider the following example.

Example 2.6. Find all rational roots of the polynomial $p(x) = 4x^3 - 3x^2 + 8x + 6$.

Proof. The rational root theorem tells us that any rational root of $p(x)$ must be in the set $\{\pm 6, \pm 3, \pm 2, \pm 1, \pm\frac{3}{2}, \pm\frac{1}{2}, \pm\frac{3}{4}, \pm\frac{1}{4}\}$. Testing all possibilities yields the fact that $p(x)$ has exactly one rational root, which is precisely $x = \frac{3}{4}$. □

In addition to answering questions about solvability of polynomial equations over the rationals, this also allows us to determine that some familiar roots of polynomials are irrational. For instance, we have the following corollary:

Corollary 2.7. The golden ratio ϕ and $\sqrt{2}$ are irrational.

Proof. Since $\phi = \frac{1+\sqrt{5}}{2}$, it is the positive root of $f(x) = x^2 - x - 1 = 0$. By the rational root theorem, the only possible roots of $f(x)$ are ± 1. Direct substitution shows that neither is a root, so any root of $f(x)$ must be irrational. Since ϕ is a root, ϕ is irrational.

We likewise have that $\sqrt{2}$ is the positive root of $g(x) = x^2 - 2 = 0$. By the rational root theorem, the only possible roots of $g(x)$ are ± 1 and ± 2. Direct substitution again shows that none of these is a root, so we must have that $\sqrt{2}$ is irrational. □

2.3 Exercises

1. A function $f : X \to Y$ between two sets X and Y is called *injective* if for any $x_1, x_2 \in X$ such that $f(x_1) = f(x_2)$, we necessarily have $x_1 = x_2$. Determine whether or not the following functions are injective. If so, prove your answer, and if not, give an example of distinct x_1 and x_2 such that $f(x_1) = f(x_2)$. *Note: Injective functions are sometimes referred to as 'one-to-one'.*

(a) $f : \mathbb{R} \to \mathbb{R}, f(x) = x$

(b) $f : \mathbb{R} \to \mathbb{R}, f(x) = \sin(x)$

(c) $f : \mathbb{C} \to \mathbb{R}, f(z) = |z|^2$

(d) $f : \mathbb{R} \to \mathbb{R}, f(x) = x^3$

(e) $f : \mathbb{C} \to \mathbb{C}, f(z) = z^3$

(f) $f : \mathbb{Z} \to \mathbb{R}, f(n) = \cos(n)$

(g) $f : \mathbb{N} \to \mathbb{R}, f(n) = \cos(n)$

2. For $f : \mathbb{R} \to \mathbb{R}$, suppose you are given a graph of f. Explain how you can intuitively tell (but not prove) if f is injective just by looking at its graph. How does the definition capture this intuition?

3. A function $f : X \to Y$ is called *surjective* if for every $y \in Y$, there exists some $x \in X$ such that $f(x) = y$. Determine whether or not each of the following functions is surjective. If so, prove your answer, and if not, give an example of a $y \in Y$ such that there is no $x \in X$ satisfying $f(x) = y$. *Note: Surjective functions are sometimes referred to as 'onto'.*

   (a) $f : \mathbb{R} \to \mathbb{R}, f(x) = x$

   (b) $f : (-\infty, 0) \cup (0, \infty) \to \mathbb{R}, f(x) = \frac{1}{x}$

   (c) $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$

   (d) $f : \mathbb{C} \to \mathbb{C}, f(z) = z^2$

   (e) $f : \mathbb{R} \to \mathbb{R}, f(x) = 2x$

   (f) $f : \mathbb{N} \to \mathbb{N}, f(n) = 2n$

   (g) $f : \mathbb{N} \to [-1, 1], f(n) = \sin(n)$

4. For $f : \mathbb{R} \to \mathbb{R}$, suppose you are given a graph of f. Explain how you can intuitively tell (though not prove) if f is surjective just by looking at its graph. How does the definition capture this intuition?

5. A function $f : X \to Y$ is called *bijective* if it is injective and surjective. Prove that if $f : X \to Y$ is a bijective function, then there exists a function $g : Y \to X$ such that for all $x \in X$ and $y \in Y$, $g(f(x)) = x$ and $f(g(y)) = y$.

6. Prove that the function g in the previous problem is *unique*, so if $f : X \to Y$ is bijective and g_1 and g_2 satisfy $g_i(f(x)) = x$ and $f(g_i(y)) = y$ for all $x \in X$, $y \in Y$, then $g_1(y) = g_2(y)$ for all $y \in Y$. Hence, we call g *the* inverse of f, often denoted f^{-1}.

7. For a bijective function $f : \mathbb{R} \to \mathbb{R}$, suppose you are given a graph of f. What will the graph of f^{-1} look like?

8. Prove that compositions of injective functions are injective.[7] In other words, if $f : X \to Y$ and $g : Y \to Z$ are injective, prove that $h : X \to Z$ given by $h = g \circ f$[8] is

[7] Note that although one can generally tell if certain types of functions are injective by looking at their graphs (cf. exercise 4), the *definition* of injectivity is necessary to prove this useful statement.

[8] This is standard notation for composition of functions, where $(f \circ g)(x) = f(g(x))$.

injective.

9. Prove that compositions of surjective functions are surjective. Conclude that compositions of bijective functions are also bijective.

2.4 Further Reading

CPZ Chapter 2.4-2.5	More on implications from a logic standpoint.
CPZ Chapter 2.8-2.9	What logical equivalence actually means.
Wikipedia entry for the Fundamental Theorem of Algebra	The proofs may be difficult to follow closely, but one can appreciate just how diverse the proofs are as well as the intuitions that they confirm. en.wikipedia.org/wiki/Fundamental_theorem_of_algebra
Wikipedia entry for Weierstrass Functions	Examples of everywhere continuous, nowhere differentiable functions with animations to supply intuition. en.wikipedia.org/wiki/Weierstrass_function

2.5 Further Examples

We will now broaden our attention from polynomials with integer coefficients and their *real* roots to polynomials with real coefficients and their *complex* roots. Let us take a look at a few examples.

- $z^2 - 1 = 0$ has solutions $z = \pm 1$.[9]
- $z^2 + 1 = 0$ has solutions $z = \pm i$.
- $z^2 + z + 1 = 0$ has solutions $z = -\frac{1}{2} \pm \frac{\sqrt{3}}{2} i$.
- $z^3 + z^2 + z - 3 = 0$ has solutions $z = 1$ and $z = -1 \pm \sqrt{2} i$.
- $z^4 + 5z^2 + 1 = 0$ has solutions $z = \pm i$ and $z = \pm 2i$.

Another pattern is emerging: It looks like whenever $z = a + bi$ is a root, so is its complex conjugate $\bar{z} = a - bi$. We would like to know whether or not this is always the case. As before, we need to formulate a statement capturing precisely what we mean, then prove it or disprove it. Such a statement could be 'if a is a root of a polynomial $p(z)$ with real coefficients, then so is its complex conjugate $\bar{a}$'. We will prove this statement, first beginning with a few facts about complex conjugates that will help us.

Proposition 2.8. Suppose z_1 and z_2 are complex numbers. Then $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$ and $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$.

Proof. Both of these facts will follow from straightforward computations. Since z_1 and z_2 are complex numbers, we can write $z_1 = a + bi$ and $z_2 = c + di$ for some $a, b, c, d \in \mathbb{R}$. We have the following computations:

$$\overline{z_1 + z_2} = \overline{a + bi + c + di} = \overline{(a + c) + (b + d)i} = (a + c) - (b + d)i$$

[9] It is conventional to let x denote a real variable and z denote a complex variable.

$$\bar{z_1} + \bar{z_2} = \overline{a+bi} + \overline{c+di} = a - bi + c - di = (a+c) - (b+d)i$$

It is now apparent that both expressions are equal. We thus have that for any z_1 and z_2, $\overline{z_1 + z_2} = \bar{z_1} + \bar{z_2}$. We now turn to multiplication:

$$\begin{aligned} \overline{z_1 z_2} &= \overline{(a+bi)(c+di)} = \overline{(ac-bd) + (ad+bc)i} = (ac-bd) - (ad+bc)i \\ \bar{z_1}\bar{z_2} &= \overline{a+bi}\ \overline{c+di} = (a-bi)(c-di) = (ac-bd) - (ad+bc)i \end{aligned}$$

Again, the two expressions are equal, concluding the proof. □

The proposition also holds for more than two complex numbers. To prove this, we would need a technique called *induction*, which we will develop a bit later in the course. Essentially, we would want to formalize the ideas that

$$\overline{z_1 + z_2 + \ldots + z_{n-1} + z_n} = \overline{z_1 + z_2 + \ldots + z_{n-1}} + \bar{z_n} = \ldots = \bar{z_1} + \bar{z_2} + \ldots + \overline{z_{n-1}} + \bar{z_n}$$

and

$$\overline{z_1 z_2 \cdots z_{n-1} z_n} = \overline{z_1 z_2 \cdots z_{n-1}}\bar{z_n} = \ldots = \bar{z_1}\bar{z_1}\cdots\overline{z_{n-1}}\bar{z_n},$$

but for now, we will simply accept these facts as true.

Proposition 2.9. Let x be a real number. Then $\bar{x} = x$.

Proof. As a complex number, we have $x = x + 0i$, so $\bar{x} = x - 0i = x$. □

We are now ready to prove the main statement that we developed before.

Theorem 2.10. If a is a root of a polynomial $p(z) = c_0 + c_1 z + \ldots + c_n z^n$ with real coefficients, then so is $\bar{a}$.

Proof. We assume $p(a) = 0$ and want to show that $p(\bar{a}) = 0$. There are a number of ways to proceed, but we will start by taking the complex conjugate of both sides of the expression $p(a) = 0$ to obtain $\overline{p(a)} = \bar{0}$. But $\bar{0} = 0$, so $\overline{p(a)} = 0$. We now substitute the coefficients of p to obtain

$$\overline{c_0 + c_1 a + \ldots + c_n a^n} = 0.$$

Using the fact (essentially) proven above that the complex conjugate of any finite sum equals the corresponding sum of the complex conjugates, we see that

$$\bar{c_0} + \overline{c_1 a} + \ldots + \overline{c_n a^n} = 0.$$

Next, we apply the fact that the complex conjugate of the product of two complex numbers is the same as the product of their complex conjugates to obtain that

$$\bar{c_0} + \bar{c_1}\bar{a} + \ldots + \bar{c_n}\overline{a^n} = 0.$$

Similarly, we use the fact that the complex conjugate of the product of k complex numbers equals the product of their complex conjugates to obtain $\overline{a^k} = \bar{a}^k$ for each $k \in \{1, \ldots, n\}$. Hence,

$$\bar{c_0} + \bar{c_1}\bar{a} + \ldots + \bar{c_n}\bar{a}^n = 0.$$

Lastly, since the coefficients c_i are assumed to be real, we know that for each $j \in \{0, 1, \ldots, n\}$, $\bar{c_j} = c_j$ by the above proposition. Applying this fact yields

$$c_0 + c_1\bar{a} + \ldots + c_n\bar{a}^n = 0.$$

This is precisely the statement that $p(\bar{a}) = 0$, concluding the proof. □

Combined with the fundamental theorem of algebra[10] which states that any n-th degree polynomial $p(z)$ with coefficients in $\mathbb{C}$ has (counted with multiplicity) n complex roots in $\mathbb{C}$, we will now show that Theorem 2.10 guarantees the existence of *real* roots to polynomials of *odd* degree with *real* coefficients. We will prove in two ways the slightly easier version where $p(z)$ has n distinct roots, each time relying on different facts beyond the scope of this course.

Corollary 2.11. Let $p(z)$ be a polynomial of degree $2n + 1$ with *real* coefficients such that all roots have multiplicity 1. Then $p(z)$ has a real root.

Proof 1. The fundamental theorem of algebra states that, counted with multiplicity, $p(z)$ must have $2n + 1$ roots. By assumption that each root has multiplicity 1, $p(z)$ has $2n + 1$ distinct roots in $\mathbb{C}$. Because the coefficients of $p(z)$ are assumed to be real, by Theorem 2.10, for every nonreal root a, $\bar{a}$ is also a root. Moreover, it can be checked easily from the definition of a complex conjugate that $\bar{\bar{a}} = a$. Thus, nonreal roots *come in pairs*, so $p(z)$ must have an even number of nonreal roots. However, $p(z)$ has an odd number of roots, so at least one of the roots of $p(z)$ must be real. □

Proof 2. In this proof, we will invoke concepts from calculus without formally defining them here, as the purpose is to communicate a proof based off a different idea.

We may assume that p is monic without loss of generality, for we may divide through by the leading coefficient of p without changing its roots. Let x denote a real variable. As x tends to $\pm\infty$, $p(x)$ is dominated by its leading term x^{2n+1} in the sense that $\lim_{x\to\pm\infty} p(x)/x^{2n+1} = 1$. Because $2n+1$ is odd, $\lim_{x\to+\infty} x^{2n+1} = +\infty$ and $\lim_{x\to-\infty} x^{2n+1} = -\infty$. We therefore have that $\lim_{x\to+\infty} p(x) = +\infty$ and $\lim_{x\to-\infty} p(x) = -\infty$. By the intermediate value theorem, $p(x)$ must hit 0 for some $x \in \mathbb{R}$, which is exactly the desired conclusion. □

As we are often interested in solving polynomial equations in which the variable represents a physical quantity like dollars, seconds, or meters, it is helpful to have such a tool to assure the existence of a real solution in many cases. A classic example is that of cubic polynomials with real coefficients. This is part of what makes mathematics exciting: it provides a framework for approaching meaningful questions, as well as one for settling them conclusively, often with the help of new concepts.

[10] The fundamental theorem of algebra requires surprisingly advanced machinery to prove despite the fact that the theorem carries somewhat simple geometric intuition via the polar form of complex numbers $re^{i\theta}$. There are *many* proofs of the theorem which take vastly different forms, as can be seen on the theorem's Wikipedia page.