CHAPEL HILL MATH CIRCLE EXIT TICKET:
January 18, 2025: Continued Fractions, Part 1 of 2

Please remove this sheet, complete it, and return it before the end of our session.

- Did you find today's topic interesting?
- Was the this topic appropriately challenging relative to your background? That is, was the topic neither too elementary nor inaccessibly advanced?
- How could we improve this worksheet for future sessions?
- What did you enjoy about today's topic?
- What did you find particularly challenging?
- Was there anything you thought was too difficult?
- Was there anything you thought was too easy?
- Are there any topics you would be interested in seeing us cover in the future?

An Introduction to Continued Fractions, Part 1 of 2

Abstract

In this session, we shall explore *continued fractions* and their applications. Our approach is modeled on that of The Ross Mathematics Program (formerly The Ross Young Scholars Program), as well as more traditional texts such as [1] and [2].

Background needed: Prerequisites include basic algebra and some experience with inequalities. Prior experience with *mathematical induction* would be useful, but to the heretofore uninitiated, we will be presenting a self-contained introduction to induction. Other novel topics will be introduced as needed.

Note: Like most Chapel Hill Math Circle worksheets, this document shall be archived at chapelhillmathcircle.org. You can presently find CHMC's archives—and for *all* groups' worksheets, not just those for the advanced group—by navigating to the "Calendar" tab or page, selecting the relevant academic term, then finding the clickable links for each worksheet (where available). Blue text in this document, including in the References, typically provides a clickable link to an external website.

Contents

0 Warmup **2**

1 Continued Fractions: Basic Concepts and Notation **3**
1.1 Discussion 3
1.2 Exercises 6

2 Convergents, the Magic Table, and Mathematical Induction **7**
2.1 Discussion 7
2.2 Exercises 8

3 Infinite Continued Fractions **11**
3.1 Discussion 11
3.2 Exercises 12

4 Linear Diophantine Equations **13**
4.1 Discussion 13
4.2 Exercises 13

5 To Be Continued... **15**

0 Warmup

As prerequisites for this session, it will help to answer the following first. You do *not* need to know these answers already, and many questions will be revisited later in the worksheet.

0.1 What is *mathematical induction*? Can you given an example of how to prove something using mathematical induction?

0.2 What is the *greatest integer function* or *floor function*?

0.3 Let

$$M := \begin{bmatrix} a & c \\ b & d \end{bmatrix} \tag{0.1}$$

be a 2 × 2 matrix. What is the *determinant* of M, denoted $\det M$? Can you extend this definition for 2 × 2 matrices to 3 × 3 and 4 × 4 matrices?

0.4 What is a *quadratic equation*? What is the *quadratic formula*?

0.5 Let k be a positive integer that is not a perfect square. If a, b, c, d are integers, what does it mean to *rationalize* an expression of the form

$$\frac{a + b\sqrt{k}}{c + d\sqrt{k}}? \tag{0.2}$$

For example, how can we rewrite the value

$$\frac{2 + \sqrt{3}}{1 - 4\sqrt{3}}$$

in rationalized form?

1 Continued Fractions: Basic Concepts and Notation

1.1 Discussion

This week we explore the concept of *continued fractions*. These may first seem like artificial constructions, but they are incredibly useful in obtaining "good" rational approximations to real numbers, as well as solutions to at least two classes of Diophantine equations.

Definition 1.1.1. A *finite simple continued fraction* is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_{n-1} + \cfrac{1}{a_n}}}}}, \tag{1.1.1}$$

where a_0 is an integer, and $a_1, a_2, \ldots, a_n$ are all positive integers. To simplify this notation we let

$$[a_0; a_1, a_2, \ldots, a_n] \tag{1.1.2}$$

denote the above expression.

Remark. If $n \geq 1$, then we typically prefer to choose a_n such that $a_n \geq 2$. Otherwise, $a_{n-1} + \frac{1}{a_n} = a_{n-1} + \frac{1}{1} = a_{n-1} + 1$, in which case we could replace $a_{n-1} + \frac{1}{a_n}$ by $a_{n-1} + 1$, another positive integer.

Further, the expression in (1.1.1) is called a *simple* continued fraction because all the numerators are 1. There are more general continued fractions of the form

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{\ddots + \cfrac{b_{n-1}}{a_{n-1} + \cfrac{b_n}{a_n}}}}},$$

where the a_j and b_j are numbers, typically integers. *For our purposes, "continued fraction" here will by default mean simple continued fraction.*

Example 1.1.2. Let us express the continued fraction

$$[4; 1, 2, 3] := 4 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{3}}} \tag{1.1.3}$$

as a rational number of the form P/Q, where P, Q are integers, Q is positive, and $\gcd(P,Q) = 1$. In other words, we want to express the continued fraction $[4;1,2,3]$ as a fraction in lowest terms.

Working from the bottom up, we have

$$4+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{3}}} = 4+\cfrac{1}{1+\cfrac{1}{\frac{7}{3}}} = 4+\cfrac{1}{1+\frac{3}{7}} = 4+\cfrac{1}{\frac{10}{7}} = 4+\frac{7}{10}$$

$$\Rightarrow 4+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{3}}} = \boxed{\frac{47}{10}},$$

so we may take $\underline{P := 47,\ Q := 10}$.

The following, an algorithm for computing the *greatest common divisor ("gcd")* of two integers, shall later prove useful in computing continued fractions:

Theorem 1.1.3 (The Euclidean Algorithm)**.** *Let a, b be integers with $b \neq 0$. Let q_j, r_j be the unique quotients and remainders under the Division Algorithm such that*

$$\begin{aligned}
a &= bq_1 + r_1, && \text{with } 0 < r_1 < |b| && (1.1.4)\\
b &= r_1 q_2 + r_2, && \text{with } 0 < r_2 < r_1 && (1.1.5)\\
r_1 &= r_2 q_3 + r_3, && \text{with } 0 < r_3 < r_2 && (1.1.6)\\
\vdots &\quad\ \vdots && && \\
r_j &= r_{j+1} q_{j+2} + r_{j+2}, && \text{with } 0 < r_{j+2} < r_{j+1} && (1.1.7)\\
\vdots &\quad\ \vdots && && \\
r_{n-2} &= r_{n-1} q_n + r_n, && \text{with } 0 < r_n < r_{n-1} && (1.1.8)\\
r_{n-1} &= r_n q_{n+1}; && && (1.1.9)
\end{aligned}$$

that is, n and r_n are defined so that r_n last nonzero remainder from successive applications of the Division Algorithm. Then

$$\gcd(a,b) = r_n. \qquad (1.1.10)$$

That is, we divide a by b, then take the integer quotient q_1 and remainder r_1. Next, we divide b by r_1, taking integer quotient q_2 and remainder r_2. If we continue dividing each remainder r_j by the next remainder r_{j+1} to determine the next remainder r_{j+2}, then the final nonzero remainder r_n is equal to $\gcd(a,b)$, the greatest common divisor of a and b. The Euclidean Algorithm is explored in more detail in our session of October 7, 2023.

Example 1.1.4. Consider the case $a := 47$, $b := 10$. Then we have

$$\begin{aligned} 47 &= 10 \cdot \underline{4} + 7 \\ 10 &= \ \ 7 \cdot \underline{1} + 3 \\ 7 &= \ \ 3 \cdot \underline{2} + \boxed{1} \\ 3 &= \ \ 1 \cdot \underline{3} + 0. \end{aligned}$$

and $\gcd(a, b) = \gcd(47, 10) = \underline{1}$, the last nonzero remainder above.

Remark. Compare the underlined quotients in Example 1.1.4 to our solution to Example 1.1.2. As we shall see below, this similarity is not mere coincidence!

Corollary 1.1.4(a). *Let a, b be integers with $b > 0$. Then if q_i and r_i are integers as in (1.1.4)–(1.1.9), we have*

$$q_1 = \left\lfloor \frac{a}{b} \right\rfloor, q_2 = \left\lfloor \frac{b}{r_1} \right\rfloor, \ldots, q_n = \left\lfloor \frac{r_{n-2}}{r_{n-1}} \right\rfloor, \text{ and } q_{n+1} = \left\lfloor \frac{r_{n-1}}{r_n} \right\rfloor = \frac{r_{n-1}}{r_n},$$

and in general,

$$q_j = \left\lfloor \frac{r_{j-2}}{r_{j-1}} \right\rfloor \tag{1.1.11}$$

where $\lfloor x \rfloor$ denotes the greatest integer *or* floor function *of the real number x.*

In general, $\lfloor x \rfloor$ is the greatest integer n less than or equal to x. Equivalently, $\lfloor x \rfloor$ is the unique integer n such that

$$n \le x < n + 1.$$

Remark. Compare the result of Corollary 1.1.4(a) to Exercise 0.2. Further, this connection between the Euclidean Algorithm and greatest integers shall be useful later to compute the continued fraction for

In Example 1.1.2, we sought to simplify a continued fraction to express it as its lowest-terms quotient of integers. One application of Corollary 1.1.4(a) is allow us to compute the continued fraction for a rational number.

Example 1.1.5. Express the rational number $-\frac{49}{23}$ as a continued fraction.

The natural starting point is to compute a_0, the first coefficient in the continued fraction $[a_0; a_1, \ldots, a_n]$ for $\frac{49}{23}$. (After all, we don't *a priori* know how many terms are in the continued fraction $[a_0; a_1, \ldots, a_n]$ for $-\frac{49}{23}$, so it seems premature to start with the "bottom" entry a_n.) Note that $a_1, a_2, \ldots, a_n$ are all *positive* integers, so the expression $[a_1; a_2, a_3, \ldots, a_n]$ will be positive and satisfy the inequality $[a_1; a_2, a_3, \ldots, a_n] < 1$. Combining this, we conclude that

$$a_0 = \left\lfloor -\frac{49}{23} \right\rfloor, \tag{1.1.12}$$

whence $a_0 = -3$.

Beginning with a_0, we have

$$-\frac{49}{23} = -3 + \frac{20}{23} = -3 + \cfrac{1}{\cfrac{23}{20}} = -3 + \cfrac{1}{1 + \cfrac{3}{20}} = -3 + \cfrac{1}{1 + \cfrac{1}{\cfrac{20}{3}}}$$

$$= -3 + \cfrac{1}{1 + \cfrac{1}{6 + \cfrac{2}{3}}} = -3 + \cfrac{1}{1 + \cfrac{1}{6 + \cfrac{1}{\cfrac{3}{2}}}} = -3 + \cfrac{1}{1 + \cfrac{1}{6 + \cfrac{1}{1 + \cfrac{1}{2}}}},$$

and therefore

$$-\frac{49}{23} = [-3; 1, 6, 1, 2]. \tag{1.1.13}$$

1.2 Exercises

Let's begin with some simple computational exercises about computations with continued fractions.

1.1 Simplify $[-1; 7, 2]$ as a fraction of the form P/Q, where P, Q are integers, $Q > 0$, and P/Q is in lowest terms.

1.2 Compute, as above, $[0; 9, 1, 3]$.

1.3 Compute the continued fraction expansion for $\frac{29}{11}$.

1.4 Let $\alpha := [4; 1, 2, 3]$. We computed α in Example 1.1.2 above. Compute each of the continued fractions $[4]$, $[4; 1]$, $[4; 1, 2]$. (As we shall see in Section 2, these are called the *convergents* to α.)

1.5 Consider $\alpha := \frac{47}{10}$ as in Example 1.1.2 and Exercise #1.4. What is the best approximation to α with denominator 1? That is, what rational number of the form $\frac{P_1}{1}$, where P is an integer, is such that the distance from $\frac{P_1}{1}$ to $\frac{47}{10}$ is as small as possible?

Repeat this for denominators 2, 3, 4, all the way up to 10.

1.6 Let P, Q be integers, with $Q > 0$. Can you explain how to use the Euclidean Algorithm to compute the continued fraction for $\frac{P}{Q}$? What happens if $\frac{P}{Q}$ is not in lowest terms?

2 Convergents, the Magic Table, and Mathematical Induction

2.1 Discussion

Next, we consider some exercises exploring the properties of the convergents and the magic table for a given continued fraction. *Note:* proofs for many of these exercises can be obtained via *mathematical induction*; see Exercise #0.1 and the accompanying supplement on mathematical induction for background. If you are unfamiliar with induction, then find a volunteer or fellow student to help explain it.

Definition 2.1.1. Let α be either a simple finite continued fraction or an infinite simple continued fraction as in Definitions 1.1.1 and 3.1.1, respectively. Further, let k be an integer with $0 \leq k \leq n$. Then the kth *convergent* to α is the finite continued fraction

$$\alpha_k := [a_0; a_1, \ldots, a_k]. \tag{2.1.1}$$

In other words, a convergent to α will be a truncation of the continued fraction for α.

Definition 2.1.2. Let α be either a finite or infinite simple continued fraction as above. Then the *magic table* for α is an array of the form

$$\begin{array}{cc||c|c|c|c|c|c|c} & & a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & \cdots \\ \hline 0 & 1 & P_0 & P_1 & P_2 & P_3 & P_4 & P_5 & \cdots \\ \hline 1 & 0 & Q_0 & Q_1 & Q_2 & Q_3 & Q_4 & Q_5 & \cdots \end{array} \tag{2.1.2}$$

The P_j and Q_j are integers defined by the following recurrence relations:

$$P_{-2} := 0 \qquad Q_{-2} := 1 \tag{2.1.3}$$
$$P_{-1} := 1 \qquad Q_{-1} := 0 \tag{2.1.4}$$
$$P_0 := a_0 \qquad Q_0 := 1 \tag{2.1.5}$$
$$P_k := a_k P_{k-1} + P_{k-2} \qquad Q_k := a_k Q_{k-1} + Q_{k-2}, \tag{2.1.6}$$

and therefore (2.1.2) begins, equivalently, as

		a_0	a_1	a_2	a_3	$\cdots$
0	1	a_0	$a_1 a_0 + 1$	$a_2 a_1 a_0 + a_2 + a_0$	$a_3 a_2 a_1 a_0 + a_3 a_2 + a_3 a_0 + a_1 a_0 + 1$	$\cdots$
1	0	1	a_1	$a_2 a_1 + 1$	$a_3 a_2 a_1 + a_3 + a_1$	$\cdots$

(2.1.7)

Example 2.1.3. Let $\alpha := [3; 1, 2, 4]$. Then the magic table for α is

		3	1	2	4
0	1	3	4	11	48
1	0	1	1	3	13

In the top row, we have

$$\begin{aligned} 3 &= 3 \cdot 1 + 0 \\ 4 &= 1 \cdot 3 + 1 \\ 11 &= 2 \cdot 4 + 3 \\ 48 &= 4 \cdot 11 + 4; \end{aligned}$$

in the bottom row, we have

$$\begin{aligned} 1 &= 3 \cdot 0 + 1 \\ 1 &= 1 \cdot 1 + 0 \\ 3 &= 2 \cdot 1 + 1 \\ 13 &= 4 \cdot 3 + 1. \end{aligned}$$

For many of the exercises in this and subsequent sections, it will be useful to use *mathematical induction*. We present here a brief overview of the method. For a more thorough introduction to induction, see the accompanying supplement "Mathematical Induction" provided separately.

2.2 Exercises

2.1 Prove that for any continued fraction $[a_0; a_1, a_2, \ldots, a_n]$, we have $Q_0 \geq 1$, $Q_1 \geq 1$, $Q_2 \geq 2$, $Q_3 \geq 3$, $Q_4 \geq 5$, and in general, $Q_n > Q_{n-1}$ and $Q_n > n+1$ for all $n \geq 3$. (Can you provide an even better lower bound for Q_k?)

2.2 Let $[a_0; a_1, a_2, \ldots, a_n]$ be a continued fraction. Prove that for each integer k with $1 \le k \le n$, we have

$$\det \begin{bmatrix} P_{k-1} & P_k \\ Q_{k-1} & Q_k \end{bmatrix} := P_{k-1}Q_k - Q_{k-1}P_k = (-1)^k. \tag{2.2.1}$$

2.3 Let $[a_0; a_1, a_2, \ldots, a_n]$ be a continued fraction. Prove that for each integer k with $1 \le k \le n$, we have

$$\det \begin{bmatrix} P_{k-2} & P_k \\ Q_{k-2} & Q_k \end{bmatrix} := P_{k-2}Q_k - Q_{k-2}P_k = (-1)^{k-1}a_k. \tag{2.2.2}$$

2.4 Let $\alpha := [a_0; a_1, \ldots, a_n]$ be a continued fraction. Prove that $\alpha = \frac{P_n}{Q_n}$; that is, prove that the continued fraction α is recovered as the quotient entries under index n in the magic table. Moreover, prove that $\frac{P_n}{Q_n}$ is already in lowest terms.

2.5 Let $\alpha := [a_0; a_1, \ldots, a_n]$, where each $a_k > 0$. If $\alpha = \frac{P_n}{Q_n}$, prove that

$$\alpha' := [a_n; a_{n-1}, \ldots, a_1, a_0] = \frac{P_n}{P_{n-1}}. \tag{2.2.3}$$

That is, if α is a rational number with $\alpha > 1$, and $\alpha = [a_0; a_1, \ldots, a_n]$ as a continued fraction, then by taking the continued fraction $\alpha' = [a_n; a_{n-1}, \ldots, a_0]$ formed by reversing the order of the a_j, the result is the quotient P_n/P_{n-1} of the final two entries in the top row of the magic table for α.

2.6 Prove that for indices k for which all quantities makes sense,

$$\left|\frac{P_k}{Q_k} - \frac{P_{k+1}}{Q_{k+1}}\right| = \frac{1}{Q_k Q_{k+1}} \leq \frac{1}{Q_k^2}. \tag{2.2.4}$$

2.7 Let $\alpha := [a_0; a_1, \ldots, a_n]$ be a finite continued fraction. Prove that for each k with $0 \leq k \leq n$ and $\alpha_k := [a_0; a_1, \ldots, a_k]$, we have

$$\alpha_0 < \alpha_2 < \alpha_4 < \cdots < \alpha < \cdots < \alpha_5 < \alpha_3 < \alpha_1. \tag{2.2.5}$$

2.8 Let $[a_0; a_1, a_2, \ldots, a_n, \ldots]$ be an infinite continued fraction, as in Definition 3.1.1. Explain why this infinite continued fraction must represent an actual real number. For those of you with some basic understanding of the relevant concepts, this means that you should give some argument why

$$\lim_{n\to\infty} [a_0; a_1, a_2, \ldots, a_n] \tag{2.2.6}$$

exists.

2.9 Let α be any real number. Explain how to obtain a continued fraction representation for α. Must this continued fraction for α be unique?

2.10 **Challenging:** Let α be a rational number. Say that you know that in lowest terms, $\alpha = P/Q$, and $Q < 1000$, for example. If I give you some initial portion of the decimal expansion for α, explain how you might be able to recover α as a quotient of integers $\alpha = P/Q$.

For example, if I gave you the rational number $\alpha \approx 0.2064220183486\ldots$, then knowing that $Q < 1000$, how might you express α as a quotient of integers?

Hint: Given the decimal expansion for α, how would you compute the continued fraction for α? How does knowing that $Q < 1000$ help?

3 Infinite Continued Fractions

3.1 Discussion

We begin by generalizing our definition of finite simple continued fractions (Definition 1.1.1) to continued fractions that are infinitely "deep":

Definition 3.1.1. An *infinite simple continued fraction* is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_{n-1} + \cfrac{1}{a_n + \cfrac{1}{\ddots}}}}}}, \tag{3.1.1}$$

where a_0 is an integer, and $a_1, a_2, \ldots, a_n, \ldots$ are all positive integers. As in Definition 1.1.1, we let

$$[a_0; a_1, a_2, \ldots, a_n, \ldots] \tag{3.1.2}$$

denote the above infinite simple continued fraction.

As above with finite continued fractions, since we shall be considering *only* simple continued fractions, we consider "simple" to be implicit whenever we use the term "continued fraction", including for infinite continued fractions.

Definition 3.1.2. Let $\alpha := [a_0; a_1, a_2, \ldots, a_n, \ldots]$ be an infinite simple continued fraction. We say that α is *periodic* if the sequence eventually repeats. That is, α is periodic if and only if it is of the form

$$[a_0; a_1, \ldots, a_k, a_{k+1}, \ldots, a_m, a_{k+1}, \ldots, a_m, \ldots]. \tag{3.1.3}$$

We denote a periodic infinite continued fraction as above by

$$[a_0; a_1, \ldots, a_k, \dot{a}_{k+1}, \ldots, \dot{a}_m] \tag{3.1.4}$$

or

$$[a_0; a_1, \ldots, a_k, \overline{a_{k+1}, \ldots, a_m}]. \tag{3.1.5}$$

3.2 Exercises

3.2.1 Let $\alpha := \sqrt{3}$. What is the infinite continued fraction expansion for α?

3.2.2 Compute the infinite continued fraction for $\sqrt{2}$, and compute the first few entries of its magic table. Further, compute values of $P_k^2 - 2Q_k^2$ for the first few columns of the magic table.

3.2.3 Assuming convergence, compute the value for the infinite continued fraction $\alpha := [1; 1, 1, \ldots] = [1; \dot{1}]$. Construct the magic table, and compute the first few entries. Can you recognize the pattern?

3.2.4 Verify that $\varphi := \frac{1+\sqrt{5}}{2}$ is a solution to the equation $x^2 - x - 1 = 0$. Use this to find an alternate derivation for the continued fraction expansion for φ.

Hint: Note that since $x^2 - x - 1 = 0$, we have $x^2 = x + 1$ and therefore $x = 1 + \frac{1}{x}$. What happens when we repeatedly substitute the right-hand side of this equation into itself?

3.2.5 Compute the infinite continued fraction expansion for $\sqrt{41}$. For the first few columns of its magic table, compute the values $P_k^2 - 41Q_k^2$.

4 Linear Diophantine Equations

4.1 Discussion

Given integers a, b, and c, an important subject of study in number theory is the *linear Diophantine equation*

$$ax + by = c. \tag{4.1.1}$$

It should be familiar that if a, b, and c are given, then if (4.1.1) has any solution, then the set of all points (x, y) satisfying this equation forms a line in the plane. In the case of a *Diophantine* equation, though, we are interested in only those solutions (x, y) to a given equation such that both coordinates are *integers*.

One of the most useful notions in number theory is that of a *modular multiplicative inverse* modulo m. Namely, if a and m are integers, we wish to compute—if possible—an integer x such that $ax \equiv 1 \pmod{m}$. Using continued fractions and convergents, we shall explain how to compute such multiplicative inverse efficiently and explicitly, provided they exist. The relevant definitions and concepts, many of which may be familiar from past sessions, shall be presented in the exercises of Subsection 4.2.

4.2 Exercises

4.1 Let a, b be positive integers. If $\gcd(a, b) = 1$, describe a method guaranteed to produce integers x, y such that $ax + by = 1$. More generally, if $\gcd(a, b) = d$, how can we produce integers x, y such that $ax + by = d$?

4.2 Let a, b be positive integers. If $d := \gcd(a, b)$, then describe a method to find integers x and y such that $ax + by = d$.

4.3 Let a, b, and c be arbitrarily given integers. Provide, with justification, a set of criteria for which the equation $ax + by = c$, Equation 4.1.1 above, has a solution such that x and y are both integers.

4.4 Let a and m be integers, with $m > 1$. We say that x is the *multiplicative inverse of a modulo m* if and only if x is an integer and $ax \equiv 1 \pmod{m}$. (For those unfamiliar with modular arithmetic: this means that $ax - 1$ is divisible by m.) *Important:* note that x must itself be an integer!

Fix an integer $m > 1$. Give a complete characterization of all integers a such that a has a multiplicative inverse modulo m. For those a admitting a multiplicative inverse modulo m, provide a method for computing such a multiplicative inverse.

4.5 Compute, if possible, the following multiplicative inverses modulo m for a:

(a) $a := 4$, $m := 17$

(b) $a := 43$, $m := 257$

(c) $a := 17$, $m := 119$

4.6 Let $a_1, \dots, a_k, c$ be integers. When can we solve the Diophantine equation

$$a_1 x_1 + a_2 x_2 + \cdots a_k x_k = c$$

for integers $x_1, \dots, x_k$? For those cases where a solution exists, describe an algorithm for producing at least one solution.

5 To Be Continued...

In our next session, we will further explore continued fraction methods. In particular, we shall further explore rational approximations to real numbers using convergents, and we consider in more generality *Pell's Equation*, the Diophantine equation of the form $x^2 - dy^2 = 1$, where d is a positive integer.

References

[1] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers.* Oxford University Press, Walton Street, Oxford OX2 6DP, fifth edition, 1979.

[2] Ivan Niven, Herbert S. Zuckerman, and Hugh L. Montgomery. *An Introduction to the Theory of Numbers.* John Wiley & Sons, Inc., New York, fifth edition, 1991.

Mathematical Induction

David Yavenditti
Chapel Hill Math Circle

January 18 and February 1, 2025

Abstract

This supplementary document provides an introduction to the proof technique known as *mathematical induction*. In addition to an explanation of the principle itself, we include several proofs illustrating the technique. Examples will include both traditional and strong (or complete) mathematical induction.

Note: Like most Chapel Hill Math Circle worksheets, this document shall be archived at chapelhillmathcircle.org. You can presently find CHMC's archives—and for *all* groups' worksheets, not just those for the advanced group—by navigating to the "Calendar" tab or page, selecting the relevant academic term, then finding the clickable links for each worksheet (where available). Blue text in this document, including in the References, typically provides a clickable link to an external website.

Contents

1 Preliminaries 2

2 Mathematical Induction 2

3 Proof Examples Using Mathematical Induction 3

4 Variants of Mathematical Induction 6

5 Additional Proof Examples 7

6 Potential Pitfalls 13

7 An Especially Bonkers Use of Mathematical Induction 14

8 Closing Remarks 17

1 Preliminaries

Let $\{P(n)\}$ be a collection of statements associated with every positive integer n. These statements should, in principle, be either true or false.

Example 1.1. For all positive integers n, let $P(n)$ denote the statement

$$3n+1 \text{ is even.}$$

Then $\underline{P(1) \text{ is true}}$, since $3 \cdot 1 + 1 = 4$ is even. Conversely, $\underline{P(2) \text{ is false}}$, since $3 \cdot 2 + 1 = 7$ is odd.

In general, one can show that for positive integers n, $P(n)$ is true—that is, $3n+1$ is even—if and only if n is odd.

Example 1.2. For all positive integers n, let $P(n)$ be the statement

$$19^n + 4^{n+1} \text{ is divisible by 5.}$$

For $P(1)$, we have $19^1 + 4^{1+1} = 19 + 16 = 35$, which is divisible by 5. Therefore, $\underline{P(1) \text{ is true}}$. Similarly, $19^2 + 4^{2+1} = 361 + 64 = 425$, also divisible by 5, so $\underline{P(2) \text{ is true}}$.

Later, we shall show that $P(n)$ is true for *all* positive integers n. (Those of you familiar with modular arithmetic may already see how to prove this directly.)

Example 1.3.

For all positive integers n, let $P(n)$ be the statement

$$2^{2^n} + 1 \text{ is prime.}$$

Since $2^{2^1} + 1 = 2^2 + 1 = 5$ is prime, $\underline{P(1) \text{ is true}}$. Similarly, $2^{2^2} + 1 = 17$ is prime, $2^{2^3} + 1 = 257$, and $2^{2^4} + 1 = 65{,}537$ is prime. Therefore, $\underline{P(2), P(3), \text{ and } P(4) \text{ are all true}}$. Next, $2^{2^5} + 1 = 6{,}700{,}417$. One can show, however, that $641 \mid 6{,}700{,}471$; that is, 641 is a divisor of $2^{2^5} + 1$ with no remainder. Therefore, $2^{2^5} + 1$ is *not* prime, so $\underline{P(5) \text{ is false}}$.

Note: The numbers $2^{2^n} + 1$ are called *Fermat numbers*. Fermat conjectured that all Fermat numbers are prime, but Euler discovered the above nontrivial divisor of $2^{2^5} + 1$. Since then, no larger Fermat number has been shown to be prime. There are open conjectures whether all larger Fermat numbers are composite, finitely many are prime, or finitely many are composite.

2 Mathematical Induction

A typical problem in mathematics is to show that for a given statement P, we want to prove that $P(n)$ is true for *every* positive integer n. In the next section, we introduce a general technique to prove such statements.

Theorem 2.1 (Mathematical Induction)**.** *Let* $\{P(n)\}$ *be a collection of statements for every positive integer* n*. Then* $P(n)$ *is true for every positive integer* n *if and only if*

(a) $P(1)$ *is true, and*

(b) for every positive integer, if $P(n)$ *is true,* then $P(n+1)$ *is true.*

The condition in Theorem 2.1(a) is called the *base case*, and the condition in Theorem 2.1(b) is the *inductive step*. Note, in particular, that the inductive step is a conditional statement where the goal is to show that *if* $P(n)$ is true, *then as a consequence* it follows that $P(n+1)$ is also true. In trying to verify the inductive step holds, then, we assume $P(n)$ is true by hypothesis, and we attempt to deduce the truth of $P(n+1)$ as a corollary of $P(n)$ being true.

Theorem 2.1 is sometimes described as *weak* induction to distinguish it from *strong induction* as described in Section 4. Here, we shall characterize mathematical induction as a theorem, though it can alternately be taken as an axiom for the positive integers.

A common way to explain the intuition behind induction is to imagine each of the statements $P(1)$, $P(2)$, $P(3)$, ... are all arranged as an infinite sequence dominoes, and the goal is to establish that every domino falls over. In this analogy, the base case tells us the first domino falls over. The inductive step tells us that *if* a given domino falls over, *then* its successor domino likewise topples. Since $P(1)$ is true by the base case, $P(1+1) = P(2)$ is true by the inductive step. Similarly, since $P(2)$ was just established to be true, $P(2+1) = P(3)$ must be true by the inductive step. Continuing in this way, we see that $P(4)$, $P(5)$, $P(6)$, and so on must all be true.

3 Proof Examples Using Mathematical Induction

The best way to understand how mathematical induction works is from examples—and your own practice.

Proposition 3.1. *Let* n *be a positive integer. Then*

$$1+2+\cdots+(n-1)+n = \frac{n(n+1)}{2}. \tag{3.1}$$

Proof of Proposition 3.1 by mathematical induction. Let $P(n)$ be the statement in (3.1). We prove that $P(n)$ is true for all positive integers n by mathematical induction.
Base Case: We have

$$1 = \frac{1(1+1)}{2}$$

by inspection, so $P(1)$ is true.
Inductive Step: Assume that $P(n)$ is true; that is, by hypothesis,

$$1+2+\cdots+(n-1)+n = \frac{n(n+1)}{2}. \tag{3.2}$$

Our goal is to show that $P(n+1)$ is true; that is,

$$1+2+\cdots+(n-1)+n+(n+1)=\frac{(n+1)[(n+1)+1]}{2}. \tag{3.3}$$

Adding $n+1$ to both sides of (3.2), we have

$$\begin{aligned}
1+2+\cdots+(n-1)+n &= \frac{n(n+1)}{2}\\
\Longrightarrow 1+2+\cdots+(n-1)+n+(n+1) &= \frac{n(n+1)}{2}+(n+1)\\
\Longrightarrow 1+2+\cdots+(n-1)+n+(n+1) &= (n+1)\left(\frac{n}{2}+1\right), \quad \text{since } n+1 \text{ is a common factor}\\
\Longrightarrow 1+2+\cdots+(n-1)+n+(n+1) &= (n+1)\left(\frac{n+2}{2}\right)\\
\Longrightarrow 1+2+\cdots+(n-1)+n+(n+1) &= \frac{(n+1)(n+2)}{2}\\
\Longrightarrow 1+2+\cdots+(n-1)+n+(n+1) &= \frac{(n+1)[(n+1)+1]}{2},
\end{aligned}$$

and this final statement is precisely $P(n+1)$ from (3.3).

Since both the base case and inductive step are true, by mathematical induction, we conclude that $P(n)$ is true for every positive integer, completing the proof. □

Proposition 3.2. *For every positive integer n, 19^n+4^{n+1} is divisible by* 5.

(Compare to Example 1.2.)

Proof of Proposition 3.2 by mathematical induction. Let $P(n)$ be the statement

$$19^n+4^{n+1} \text{ is divisible by } 5;$$

equivalently, there exists some integer k (depending on n) such that

$$19^n+4^{n+1}=5k.$$

We prove that $P(n)$ is true for all positive integers n by mathematical induction.
Base Case: For $P(1)$, we have

$$\begin{aligned}
19^1+4^{1+1} &= 19+4^2\\
&= 19+16\\
&= 35,
\end{aligned}$$

which is divisible by 5. Therefore, $P(1)$ is true.
Inductive Step: Assume that $P(n)$ is true; that is, by hypothesis, 19^n+4^{n+1} is divisible by 5. Concretely, let k be the integer such that

$$19^n+4^{n+1}=5k.$$

To show that $P(n+1)$ is also true, we must show that $19^{n+1}+4^{n+2}$ is also divisible by 5.

We have

$$\begin{aligned}
19^{n+1}+4^{n+2} &= 19\cdot 19^{n}+4\cdot 4^{n+1}\\
&= 19\left[(19^{n}+4^{n+1})-4^{n+1}\right]+4\cdot 4^{n+1}\\
&= 19(19^{n}+4^{n+1})-19\cdot 4^{n+1}+4\cdot 4^{n+1}\\
&= 19(19^{n}+4^{n+1}+(-19+4)4^{n+1}\\
&= 19(19^{n}+4^{n+1})-15\cdot 4^{n+1}\\
&= 19\cdot 5k-5(3\cdot 4^{n+1})\\
&= 5\left(19k-3\cdot 4^{n+1}\right).
\end{aligned}$$

Since $19k-3\cdot 4^{n+1}$ is an integer, this entire expression is therefore a multiple of 5. Therefore, if $P(n)$ is true, so is $P(n+1)$.

Because both the base case and inductive steps are true, by mathematical induction it follows that $P(n)$ is true for every positive integer n, as desired. □

Proposition 3.3. *For every positive integer n, $2n<3^{n}$.*

Proof of Proposition 3.3 by mathematical induction. Let $P(n)$ be the statement $2n<3^{n}$.

Base Case: For $P(1)$, we have

$$2\cdot 1=2<3=3^{1},$$

so $2\cdot 1<3^{1}$, and the base case $P(1)$ is therefore true.

Inductive Step: Assume that $P(n)$ is true for some positive integer n; that is, assume that $3^{n}>2n$ for some positive integer n. We wish to prove $P(n+1)$ is true; that is, that $3^{n+1}>2(n+1)$ follows as a consequence of $P(n)$.

As a preliminary matter, note that for any positive integer n,

$$\begin{aligned}
\frac{n+1}{n} &= 1+\frac{1}{n}\\
&\le 1+\frac{1}{1},\ \text{since } n\in\mathbb{N}\\
&= 1+1\\
&= 2\\
&< 3,
\end{aligned}$$

so

$$\frac{n+1}{n}<3. \tag{3.4}$$

for every positive integer n.

Now, from the hypothesis that $P(n)$ is true, we have

$$\begin{aligned} 2n < 3^n &\implies 2n \cdot \frac{n+1}{n} < 3^n \cdot 3, \text{ multiplying by (3.4)} \\ &\implies 2(n+1) < 3^{n+1}, \end{aligned}$$

so $P(n+1)$ is true as well.

Since both the base case and inductive steps hold, the proposition is true by mathematical induction. □

4 Variants of Mathematical Induction

Theorem 2.1 is the most familiar version of mathematical induction. Alternatives to weak induction are often more versatile and powerful than this elementary version:

Theorem 4.1 (Strong Mathematical Induction)**.** *Let $\{P(n)\}$ be a collection of statements for every positive integer n. Then $P(n)$ is true for every positive integer n if and only if*

(a) $P(1)$ is true, and

(b) for every positive integer, if *$P(k)$ is true* for every *positive integer $k \le n$,* then *$P(n+1)$ is true.*

Theorem 4.2 (Mathematical Induction, Multiple Base Cases)**.** *Assume k_0 is a positive integer, and let $\{P(n)\}$ be a collection of statements for every positive integer n. Then $P(n)$ is true for every positive integer n if and only if*

(a) $P(1), \ldots, P(k_0)$ are each true, and

(b) for every integer $n \ge k_0$, if *$P(n)$ is true,* then *$P(n+1)$ is true.*

Theorem 4.3 (Mathematical Induction, Alternate Base Cases)**.** *Let k_0 be an integer, and let $\{P(n)\}$ be a collection of statements for every integer $n \ge k_0$. Then $P(n)$ is true for every integer $n \ge k_0$ if and only if*

(a) $P(k_0)$ is true, and

(b) for every integer $n \ge k_0$, if *$P(n)$ is true,* then *$P(n+1)$ is true.*

The following powerful principle—alternately taken as an axiom for the integers or a consequence of other axiomatic descriptions of the integers such as the Peano Axioms—has all the above versions of induction as corollaries:

Axiom 4.4 (The Well-Ordering Principle)**.** Let $\mathbb{N} := \{1, 2, 3, 4, \ldots\}$ denote the set of natural numbers. If S is any nonempty subset of $\mathbb{N}$, then S contains a minimal element. That is, there exists an element $\ell \in S$ such that for every $s \in S$, $\ell \le s$.

Corollary 4.4(a) (Corollary to Well-Ordering Principle)**.** *Let X be a subset of the integers that is bounded below. If S is a nonempty subset of X, then S contains a minimal element.*

Another corollary of Axiom 4.4 is a proof method known as *proof by descent*, which we shall use implicitly to prove Proposition 5.4 below.

5 Additional Proof Examples

In the next example, we shall require multiple base cases and strong induction:

Proposition 5.1. *Let $\{F_n\}$ denote the sequence of Fibonacci numbers, defined recursively by*

$$\begin{aligned} F_1 &= 1 \\ F_2 &= 1 \\ F_n &= F_{n-1} + F_{n-2}, \textit{ for all } n \geq 3. \end{aligned}$$

Then for every positive integer n, we have Binet's Formula*:*

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]. \tag{5.1}$$

Since $\varphi := \frac{1+\sqrt{5}}{2}$ means $-\frac{1}{\varphi} = \frac{1-\sqrt{5}}{2}$, equivalently,

$$F_n = \frac{1}{\sqrt{5}}\left(\frac{\varphi^{2n} - (-1)^n}{\varphi^n}\right) \tag{5.2}$$

for every positive integer n.

Remark. The value $\varphi = \frac{1+\sqrt{5}}{2}$ is the *golden ratio*, the topic for the advanced group's session of October 12, 2024.

Proof. We prove Proposition 5.1 using both Theorems 4.1 and 4.2 with $k_0 := 2$. Let $P(n)$ be the statement in (5.2).
Base cases $P(1)$ and $P(2)$: We have that

$$\begin{aligned} \frac{1}{\sqrt{5}}\left(\frac{\varphi^2 - (-1)^1}{\varphi^1}\right) &= \frac{1}{\sqrt{5}}\left(\frac{\varphi^2 + 1}{\varphi}\right) \\ &= \frac{1}{\sqrt{5}}\left(\varphi + \frac{1}{\varphi}\right) \\ &= \frac{1}{\sqrt{5}} \cdot \sqrt{5}, \qquad \text{since } \varphi = \frac{1+\sqrt{5}}{2}, \frac{1}{\varphi} = \frac{-1+\sqrt{5}}{2} \\ &= 1 \\ &= F_1, \end{aligned}$$

so $P(1)$ is true. Similarly,

$$\begin{aligned}
\frac{1}{\sqrt{5}}\left(\frac{\varphi^4-(-1)^2}{\varphi^2}\right) &= \frac{1}{\sqrt{5}}\left(\frac{\varphi^4-1}{\varphi^2}\right)\\
&= \frac{1}{\sqrt{5}}\left[\frac{(\varphi^2-1)(\varphi^2+1)}{\varphi^2}\right]\\
&= \frac{1}{\sqrt{5}}\left(\frac{\varphi(\varphi^2+1)}{\varphi^2}\right), \qquad \text{since } \varphi^2-1=\varphi\\
&= \frac{1}{\sqrt{5}}\left(\frac{\varphi^2+1}{\varphi}\right)\\
&= \frac{1}{\sqrt{5}}\left(\varphi+\frac{1}{\varphi}\right)\\
&= 1, \qquad \text{as above}\\
&= F_2.
\end{aligned}$$

Therefore, $P(2)$ is also true, so both base cases hold.
Inductive Step: Assume that $n \geq 2$ and both $P(n-1)$ and $P(n)$ are true. We wish to show that $P(n+1)$ is true; that is if $P(n-1)$ and $P(n)$ are true, then

$$F_{n+1} = \frac{1}{\sqrt{5}}\left(\frac{\varphi^{2n+2}-(-1)^{n+1}}{\varphi^{n+1}}\right). \tag{5.3}$$

First, note that

$$\varphi^2 = \varphi + 1, \tag{5.4}$$

which can be verified by inspection.

Now, by our inductive hypothesis, assume $n \geq 2$, and $P(n-1)$ and $P(n)$ are both true. We therefore have

$$\begin{aligned}
F_{n+1} &= F_n + F_{n-1}\\
&= \frac{1}{\sqrt{5}}\left(\frac{\varphi^{2n}-(-1)^n}{\varphi^n}\right) + \frac{1}{\sqrt{5}}\left(\frac{\varphi^{2n-2}-(-1)^{n-1}}{\varphi^{n-1}}\right)\\
&= \frac{1}{\sqrt{5}}\left(\frac{\varphi^{2n+1}-(-1)^n\varphi}{\varphi^{n+1}}\right) + \frac{1}{\sqrt{5}}\left(\frac{\varphi^{2n}-(-1)^{n-1}\varphi^2}{\varphi^{n+1}}\right)\\
&= \frac{1}{\sqrt{5}}\left(\frac{\varphi^{2n+1}+\varphi^{2n}-(-1)^n\varphi-(-1)^{n-1}\varphi^2}{\varphi^{n+1}}\right)\\
&= \frac{1}{\sqrt{5}}\left(\frac{\varphi^{2n+1}+\varphi^{2n}-(-1)^{n-1}\left[\varphi^2-\varphi\right]}{\varphi^{n+1}}\right)\\
&= \frac{1}{\sqrt{5}}\left(\frac{\varphi^{2n}(\varphi+1)-(-1)^{n-1}\cdot 1}{\varphi^{n+1}}\right)\\
&= \frac{1}{\sqrt{5}}\left(\frac{\varphi^{2n}\cdot\varphi^2-(-1)^{n+1}}{\varphi^{n+1}}\right),
\end{aligned}$$

by (5.4), and since $n-1 \equiv n+1 \pmod 2$ means $(-1)^{n-1} = (-1)^{n+1}$

$$\implies F_{n+1} = \frac{1}{\sqrt{5}} \left(\frac{\varphi^{2n+2} - (-1)^{n+1}}{\varphi^{n+1}} \right),$$

so (5.2) holds for F_{n+1}, too. The (strong) inductive step therefore holds, so we have proven Proposition 5.1 by strong induction. □

Next, we consider a proof that uses the strong inductive step in a more essential way:

Proposition 5.2. *Let n be a positive integer with $n \geq 2$. Then n is expressible as a product of primes. (Explicitly, there exist finitely many primes $p_1, p_2, \ldots, p_m$ such that $n = p_1 p_2 \cdots p_m$.)*

Proof. For all positive integers n with $n \geq 2$, let $P(n)$ be the statement that n is expressible as a product of primes.
Base Case: For $n := 2$, the result is immediate, since 2 is itself a prime. Therefore, the base case holds.
Inductive Step: Assume that $P(2), P(3), \ldots, P(n)$ are all true, and we consider $P(n+1)$.

Case 1: If $n+1$ is prime, then $n+1$ is immediately expressible as a product of primes.

Case 2: If $n+1$ is not prime, then since $n \geq 2$, $n+1$ must be composite. (I.e., because $n \geq 2$, we rule out the cases $n+1 = 0$ and $n+1 = 1$. Since $n+1$ is *not* prime, it must therefore be composite.) Therefore, there exist positive integers r, s such that $1 < r, s < n+1$ and $n+1 = rs$. By our strong inductive hypothesis, each of r and s is expressible as a product of primes. Since $n+1 = rs$, $n+1$ is therefore a product of products of primes, whence $n+1$ is itself a product of primes.

Since Cases 1–2 are both true, the (strong) inductive step has been verified. Therefore, by the principle of strong induction, Proposition 5.2 is true. □

Our next example is a proposition involving two different variables, n and k. First, we recall the definition of factorial:

Notation. Let n be a nonnegative integer. Then *n factorial*, denoted $n!$, is defined by

$$n! = \begin{cases} 1, & \text{if } n = 0 \text{ or } n = 1 \\ 1 \cdot 2 \cdot \cdots \cdot (n-1) \cdot n, & \text{otherwise.} \end{cases} \tag{5.5}$$

Equivalently,[1]

$$n! := \prod_{j=1}^{n} j.$$

[1] The equivalence is immediate if $n \geq 1$. For $n := 0$, the product notation yields the *empty product*, which is 1 by definition.

Proposition 5.3. *Let n and k be nonnegative integers. Define the* binomial coefficient $\binom{n}{k}$, *read as "n choose k", by*

$$\binom{n}{k} := \begin{cases} \dfrac{n!}{k!(n-k)!}, & \text{if } 0 \le k \le n \\ 0, & \text{otherwise.} \end{cases}$$

Prove that for all positive integers n and every integer k with $0 \le k \le n$*, we have*

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}. \tag{5.6}$$

In particular, Proposition 5.3 implies $\binom{n}{k}$ is itself an integer for every positive integer n and every nonnegative integer k with $k \le n$.

Proof. First, let us fix a positive integer n. Let $P(n)$ be the statement that

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}, \textit{ for every } \text{integer } k \text{ with } 0 \le k \le n,$$

Our strategy shall be to induct on n, letting k vary.
<u>*Base Case:*</u> To verify that $P(1)$ is true, we must show

$$\binom{1}{0} = \binom{0}{-1} + \binom{0}{0}$$

and

$$\binom{1}{1} = \binom{0}{1} + \binom{0}{0}.$$

Since $\binom{1}{0} = \binom{1}{1} = \binom{0}{0} = 1$ and $\binom{0}{-1} = \binom{0}{1} = 0$, the base case $P(1)$ is true.
<u>*Inductive Step:*</u> For a positive integer n, assume that $P(n)$ is true, so that (5.6) holds for <u>*every*</u> integer k such that $0 \le k \le n$. By the base case, we may assume $n \ge 1$.

Case 1: $k = 0$.

Then $k - 1 = -1$, so $\binom{n+}{k-1} = 0$. Further, $\binom{n+1}{k} = \binom{n+1}{0} = 1 = 1 + 0 = \binom{n}{k} + \binom{n}{k-1}$, as desired.

Case 2: $k = n + 1$.

Then $\binom{n+1}{n+1} = 1$. Further, $\binom{n+1}{k} = \binom{n+1}{n+1} = 1 = 0 + 1 = \binom{n}{n+1} + \binom{n}{n}$, as desired.

Case 3: $1 \le k \le n$.

In this case, $\binom{n}{k}$ and $\binom{n}{k-1}$ will both be defined and nonzero. Then

$$\binom{n}{k} + \binom{n}{k-1} = \frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)![n-(k-1)]!}$$

$$
\begin{aligned}
&= \frac{(n-k+1)\cdot n!}{k!\cdot(n-k+1)\cdot(n-k)!} + \frac{k\cdot n!}{k\cdot(k-1)!(n-k+1)!}\\
&= \frac{(n-k+1)\cdot n!}{k!(n-k+1)!} + \frac{k\cdot n!}{k!(n-k+1)!}\\
&= \frac{(n-k+1)\cdot n! + k\cdot n!}{k!(n-k+1)!}\\
&= \frac{[(n-k+1)+k]\cdot n!}{k![(n+1)-k)!}\\
&= \frac{(n+1)\cdot n!}{k![(n+1)-k)!}\\
&= \frac{(n+1)!}{k![(n+1)-k)!}
\end{aligned}
$$

$$\implies \binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k},$$

as desired.

Therefore, the inductive step also holds, so Proposition 5.3 is true by induction. □

Next, we revisit Proposition 3.1, this time proving it using the Well-Ordering Principle, Axiom 4.4:

Proof of Proposition 3.1 by the Well-Ordering Principle. We use proof-by-contradiction and the Well-Ordering Principle to prove $1+2+\cdots+n = \frac{n(n+1)}{2}$ for all positive integers n.

Set

$$S := \left\{ n \in \mathbb{N} : 1+2+\cdots+n \neq \frac{n(n+1)}{2} \right\}; \tag{5.7}$$

that is, S is the set of positive integers for which the assertion in Proposition 3.1 is *false*. Our goal is therefore to prove that $S = \varnothing$.

ASSUME instead that $S \neq \varnothing$. Then by well-ordering, there exists a minimal element $\ell \in S$. That is,

$$1+2+\cdots+\ell \neq \frac{\ell(\ell+1)}{2},$$

and

$$1+2+\cdots+k = \frac{k(k+1)}{2}$$

for every positive integer $k < \ell$.

First, observe that by inspection, $1 = \frac{1(1+1)}{2}$, so $1 \notin S$. That is, the equation *is* true for $k = 1$, it is not false for $k = 1$. Since S is the set of all k for which the equation is false, it follows that $1 \notin S$.

Since ℓ is the minimal element of S, we must have $\ell > 1$. It follows that $\ell - 1$ is a *positive* integer. Since $\ell - 1 < \ell$ and ℓ is the *smallest* element in S, $\ell - 1 \notin S$. (Otherwise, we would have an even smaller element than ℓ lying in S.) Therefore,

$$1+2+\cdots+(\ell-1)=\frac{(\ell-1)\ell}{2}.$$

Adding ℓ to both sides of the previous equation, we obtain

$$\begin{aligned}
1+2+\cdots+(\ell-1)&=\frac{(\ell-1)\ell}{2}\\
\implies 1+2+\cdots+(\ell-1)+\ell&=\frac{(\ell-1)\ell}{2}+\ell\\
\implies 1+2+\cdots+(\ell-1)+\ell&=\ell\left(\frac{\ell-1}{2}+1\right)\\
\implies 1+2+\cdots+(\ell-1)+\ell&=\ell\left(\frac{\ell+1}{2}\right)\\
\implies 1+2+\cdots+(\ell-1)+\ell&=\frac{\ell(\ell+1)}{2}\\
&\implies \ell\notin S.
\end{aligned}$$

This is a contradiction, though: ℓ was defined to be the minimal element of S, but we cannot simultaneously have $\ell\in S$ and $\ell\notin S$.

Our original assumption is therefore false, so $S=\varnothing$, as desired. Since the set of all positive integers where the equation is *fales* is empty, $1+2+\cdots+n=\frac{n(n+1)}{2}$ for *every* positive integer n, completing the proof. □

Compare the above proof of Proposition 3.1 using Axiom 4.4 to the initial proof in Section 3. In general, for an induction-like proof using well-ordering, you consider the set S of all n for which the claim is *false*. The goal is to prove $S=\varnothing$. Employing a proof-by-contradiction strategy, we assume instead that $S\neq\varnothing$, so by well-ordering, a nonempty S has a minimal element ℓ. Using techniques similar to a direct induction proof (or otherwise), show that $\ell\notin S$ or that ℓ is not the *minimal* element of S. Either would contradict the definition of ℓ, so our assumption $S\neq\varnothing$ is false, completing the proof.

The following application of well-ordering shows how it can be even more versatile than standard formulations of induction:

Proposition 5.4. *The smallest positive integer is* 1. *That is, if n is a positive integer, then* $1\le n$. *In particular, there is no positive integer n such that* $0<n<1$.

Proof of Proposition 5.4. Let $S:=\mathbb{N}$, the set of all positive integers. In particular, $S\subseteq\mathbb{N}$. Clearly $S\neq\varnothing$ since, in particular, $1\in S$. Since S is a nonempty subset of $\mathbb{N}$, by Axiom 4.4 it therefore contains a minimal element $\ell\in S$. I claim that $\ell\geq 1$

Assume otherwise; that is, assume that $\ell<1$. Since $\ell>0$, we combine this as the chain of inequalities

$$0<\ell<1.$$

Then since ℓ is a positive integer, multiplying all sides of the previous by ℓ preserves the inequalities, yielding

$$0 < \ell^2 < \ell.$$

This means that ℓ^2 is also a positive integer, but it is strictly smaller than ℓ. That is a contradiction, since ℓ is the smallest positive integer by definition. Our assumption is therefore false, so $\ell \geq 1$ as claimed.

Since $1 \leq \ell$ from above, and since $\ell \leq n$ for every positive integer n by the minimality of ℓ, this implies $1 \leq n$, completing the proof. □

Note that Proposition 5.4 is not a statement asking us to prove that $P(n)$ is true for all positive integers n. Axiom 4.4 is more flexible than mere induction, and we can use it to prove, for example, that $\sqrt{2}$ is irrational.

6 Potential Pitfalls

For a proof by mathematical induction to be valid, it's essential to establish both the base case (or base cases) and the inductive step, as the following examples shall illustrate.

Very Important Note: Each of the numbered Claims in Section 6 is false. These purported proofs are examples of invalid applications of mathematical induction.

Claim 6.1 (FALSE). For every positive integer n, $n^2 + n + 1$ is even.

"Proof" of Claim 6.1. Assume that $n^2 + n + 1$ is even for some positive integer n. Then we have

$$\begin{aligned}(n+1)^2 + (n+1) + 1 &= n^2 + 2n + 1 + n + 1 + 1 \\ &= (n^2 + n + 1) + (2n + 2) \\ &= (n^2 + n + 1) + 2(n + 1).\end{aligned}$$

By hypothesis, $n^2 + n + 1$ is even, and clearly $2(n + 1)$ is likewise even. Therefore, we conclude that $(n + 1)^2 + n + 1$ is also even. □

This purported, *incorrect* "proof", though, is unsound. To see why, note that for $n := 1$, $n^2 + n + 1 = 1^1 + 1 + 1 = 3$, which is odd. The inductive step—establishing the truth of the *conditional statement* that if $n^2 + n + 1$ even, $(n + 1)^2 + (n + 1) + 1$ is also even—is therefore valid. However, the base case is false, so we have not met all the criteria for a proof by induction.

This is typical of faulty proofs by induction: the argument is incomplete by failing to establish the base case. For a more subtle example of this mistake, we consider the following:

Claim 6.2. Let $S_1, S_2, \ldots, S_n$ be finite sets. Then each of these n sets has the same size;[2] that is, $|S_1| = |S_2| = \cdots = |S_n|$.

"Proof" of Claim 6.2. We proceed by induction. For the base case $n = 1$, we have a single set S_1, and clearly S_1 has the same size as itself. Therefore, the base case holds.

For the inductive step, assume that for any collection of n finite sets, they all have the same size. Consider a collection of $n+1$ finite sets $S_1, S_2, \ldots, S_n, S_{n+1}$. Note that the collections

$$\{S_1, S_2, \ldots, S_{n-1}, S_n\} \qquad \text{and} \qquad \{S_2, S_3, \ldots, S_n, S_{n+1}\}$$

are each collections of n finite sets. By our inductive hypothesis, then,

$$|S_1| = |S_2| = \cdots = |S_{n-1}| = |S_n| \qquad \text{and} \qquad |S_2| = |S_3| = \cdots = |S_n| = |S_{n+1}|.$$

Since $|S_2|$ is common to both collections, these respective common sizes must be equal. Therefore, $|S_1| = |S_2| = \cdots = |S_n| = |S_{n+1}|$, so the inductive step is true. By mathematical induction, then, every set has the same size. □

As in the "proof" for Claim 6.1, we again have a problem with the base case. Before, we had simply ignored establishing the base case altogether. The "proof" for Claim 6.2, while it nominally considers the case $n = 1$, doesn't establish its base case completely.

To see why, think about the case $n := 2$. We would then have partitioned the collection $\{S_1, S_2\}$ into

$$\{S_1\} \qquad \text{and} \qquad \{S_2\},$$

and there is no common element in these two subcollections.

This issue is a common danger in induction proof, where we might consider an index like $n-1$ or $n-2$. To proceed, we must first establish such indices are themselves *positive* integers, though, or else manipulation of objects with those indices is invalid.

7 An Especially Bonkers Use of Mathematical Induction

Our final example is a proof of one of the classical inequalities, the Arithmetic Mean-Geometric Mean Inequality ("AM-GM"). We begin with some preliminary definitions:

Definition 7.1. Let $a_1, a_1, \ldots, a_n$ be a collection of n numbers. Then the *arithmetic mean* of $a_1, a_2, \ldots, a_n$ is

$$\frac{a_1 + a_2 + \cdots + a_n}{n}.$$

Definition 7.2. Let $a_1, a_1, \ldots, a_n$ be a collection of n nonnegative numbers. Then the *geometric mean* of $a_1, a_2, \ldots, a_n$ is

$$(a_1 a_2 \cdots a_n)^{\frac{1}{n}} = \sqrt[n]{a_1 a_2 \cdots a_n}.$$

[2] This is often formulated in slightly different ways: every car has the same color, every person has the same name, or some similar variant.

Theorem 7.3 (Arithmetic Mean-Geometric Mean Inequality)**.** *Let n be a positive integer, and $a_1, a_2, \dots, a_n$ be nonnegative real numbers. Then the geometric mean of $\{a_1, a_2, \dots, a_n\}$ is no greater than the arithmetic mean of $\{a_1, a_2, \dots, a_n\}$; that is,*

$$(a_1 a_2 \cdots a_n)^{\frac{1}{n}} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}. \tag{7.1}$$

Further, equality in (7.1) holds if and only if $a_1 = a_2 = \cdots = a_n$.

We first need the special case of Theorem 7.3, with $n := 2$ as a preliminary step:

Lemma 7.4. *Let x, y be nonnegative real numbers. Then*

$$\sqrt{xy} \leq \frac{x+y}{2}, \tag{7.2}$$

with equality if and only if $x = y$.

Proof of Lemma 7.4 (Cauchy). Let x, y be nonnegative real numbers. Then

$$0 \leq (\sqrt{x} - \sqrt{y})^2,$$

with equality if and only if $x = y$. Therefore,

$$\begin{aligned} 0 \leq (\sqrt{x} - \sqrt{y})^2 &\Longrightarrow 0 \leq x - 2\sqrt{xy} + y \\ &\Longrightarrow 2\sqrt{xy} \leq x + y, \end{aligned}$$

so

$$\sqrt{xy} \leq \frac{x+y}{2}$$

with equality if and only if $x = y$, as claimed. □

Using Lemma 7.4, we can now proceed with a proof of Theorem 7.3. The following proof of the general AM-GM Inequality, attributed to Cauchy, is an especially creative application of mathematical induction.

Cauchy's Induction Proof of the AM-GM Inequality. We prove the AM-GM Inequality by induction.

Base Cases: Clearly Theorem 7.3 holds for $n = 1$ trivially, and the theorem holds for $n = 2$ by Lemma 7.4. These provide our base cases for the proof.

"Forward" Inductive Step: I claim that for all positive integers k, if (7.1) holds for all sequences of length $n := 2^k$, then (7.1) also holds for all sequences of length $n := 2^{k+1}$.

Let k be a positive integer, and let $a_1, a_2, \cdots, a_{2^{k+1}}$ be a set of nonnegative real numbers. Consider the related sequence $b_1, b_2, \dots, b_{2^k}$ of length 2^k defined by $b_1 := \frac{a_1+a_2}{2}$, $b_2 := \frac{a_3+a_4}{2}$, and so on up to $b_{2^k} := \frac{a_{2^{k+1}-1}+a_{2^{k+1}}}{2}$; in general,

$$b_j := \frac{a_{2j-1} + a_{2j}}{2}.$$

Since $b_1, b_2, \ldots, b_{2^k}$ is a sequence of length 2^k, by hypothesis the AM-GM Inequality holds, with equality if and only if $b_1 = b_2 = \cdots = b_{2^k}$. Therefore,

$$\begin{aligned}
\left(b_1 b_2 \cdots b_{2^k}\right)^{\frac{1}{2^k}} &\le \frac{b_1 + b_2 + \cdots + b_{2^k}}{2^k} \\
\implies \left[\left(\frac{a_1 + a_2}{2}\right)\left(\frac{a_3 + a_4}{2}\right)\cdots\left(\frac{a_{2^{k+1}-1} + a_{2^{k+1}}}{2}\right)\right]^{\frac{1}{2^k}} &\le \frac{\left(\frac{a_1+a_2}{2}\right) + \left(\frac{a_3+a_4}{2}\right) + \cdots + \left(\frac{a_{2^{k+1}-1}+a_{2^{k+1}}}{2}\right)}{2^k} \\
\implies \left[\left(\frac{a_1 + a_2}{2}\right)\left(\frac{a_3 + a_4}{2}\right)\cdots\left(\frac{a_{2^{k+1}-1} + a_{2^{k+1}}}{2}\right)\right]^{\frac{1}{2^k}} &\le \frac{a_1 + a_2 + \cdots + a_{2^{k+1}-1} + a_{2^{k+1}}}{2^{k+1}} \\
\implies \left(\sqrt{a_1 a_2} \cdot \sqrt{a_3 a_4} \cdot \cdots \cdot \sqrt{a_{2^{k+1}-1} a_{2^{k+1}}}\right)^{\frac{1}{2^k}} &\le \frac{a_1 + a_2 + \cdots + a_{2^{k+1}-1} + a_{2^{k+1}}}{2^{k+1}},
\end{aligned}$$

by Lemma 7.4, whence

$$\implies \left(a_1 a_2 a_3 a_4 \cdot \cdots \cdot a_{2^{k+1}-1} a_{2^{k+1}}\right)^{\frac{1}{2^{k+1}}} \le \frac{a_1 + a_2 + \cdots + a_{2^{k+1}-1} + a_{2^{k+1}}}{2^{k+1}}.$$

This establishes the inequality (7.1) holds for $a_1, a_2, \ldots, a_{2^{k+1}}$. Further, equality holds if and only if $b_1 = b_2 = \cdots = b_{2^k}$ *and* $a_1 = a_2$, $a_3 = a_4$, etc., and $a_{2^{k+1}-1} = a_{2^{k+1}}$. Together, this implies equality holds if and only if $a_1 = a_2 = \cdots = a_{2^{k+1}}$. Therefore, as claimed, the "forward" induction step holds.

“Backward” Inductive Step: I claim that if $n \ge 2$ is a positive integer, and if (7.1) holds for all sequences of length n, then (7.1) also holds for all sequences of length $n - 1$.

Let n be any positive integer with $n \ge 2$, and let $a_1, a_2, \ldots, a_{n-1}$ be nonnegative real numbers. We build the associated sequence $b_1, b_2, \ldots, b_{n-1}, b_n$ of length n by

$$\begin{aligned}
b_1 &:= a_1 \\
b_2 &:= a_2 \\
\vdots\ &\quad \vdots \\
b_{n-1} &:= a_{n-1} \\
b_n &:= \frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1}.
\end{aligned}$$

In particular, b_n has been selected in this way so that

$$\frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1} = \frac{b_1 + b_2 + \cdots + b_{n-1} + b_n}{n}; \tag{7.3}$$

that is, the arithmetic mean of $a_1, a_2, \ldots, a_{n-1}$ equals the arithmetic mean of $b_1, b_2, \ldots, b_n$. In particular, if each $b_j \ge 0$, then $a_n \ge 0$, as required, too.

Since $b_1, b_2, \ldots, b_n$ is a sequence of length n, by our inductive hypothesis, the AM-GM Inequality holds for $b_1, \ldots, b_n$, with equality if and only if $b_1 = b_2 = \cdots b_n$. Therefore, we have

$$\begin{aligned}
(b_1 b_2 \cdots b_{n-1} b_n)^{\frac{1}{n}} &\leq \frac{b_1 + b_2 + \cdots + b_{n-1} + b_n}{n} \\
\implies \left[a_1 a_2 \cdots a_{n-1} \cdot \left(\frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1} \right) \right]^{\frac{1}{n}} &\leq \frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1}, && \text{by (7.3)} \\
\implies a_1 a_2 \cdots a_{n-1} \cdot \left(\frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1} \right) &\leq \left(\frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1} \right)^n \\
\implies a_1 a_2 \cdots a_{n-1} &\leq \left(\frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1} \right)^{n-1} \\
\implies (a_1 a_2 \cdots a_{n-1})^{\frac{1}{n-1}} &\leq \frac{a_1 + a_2 + \cdots + a_{n-1}}{n-1},
\end{aligned}$$

Therefore, if the AM-GM Inequality holds for sequences of length n, it also holds for sequences of length $n-1$ as well.

Summary: First, we proved that (7.1) holds for all collections of nonnegative real numbers where $n = 1$ or 2. Next, we proved that (7.1) holds for arbitrarily large positive integers n, namely those of the form $n = 2^k$, where k is a positive integer. Finally, we proved that if $n > 1$ and (7.1) holds for n, then it also holds for $n-1$. In each case, we also showed that equality obtains if and only if all elements are equal.

For any positive integer n, then, choose a positive integer k such that $n \leq 2^k$. If $n = 2^k$, then our "forward induction" establishes the AM-GM Inequality for n. Otherwise, we can use the "backward induction" to show that AM-GM holds for $2^k - 1$, $2^k - 2$, and so on until after sufficiently many decrements, we show that AM-GM holds for n itself. Therefore, by this circuitous variant of induction, we have shown that the AM-GM Inequality holds for every positive integer n. □

Whew!

8 Closing Remarks

Mathematical induction, as well as its siblings, are powerful tools to prove certain kinds of statements. That said, it is worth noting some limitations of mathematical induction, too.

8.1 Mathematical induction helps us *prove* certain assertions, but it give no insight how to *derive* them.

For example, using induction, we can prove Binet's Formula, Proposition 5.1, is a closed-form expression for the Fibonacci numbers. By itself, though, induction gives us no way to discover that formula in the first place.

8.2 Induction can be powerful where it applies, but its range of applicability may be narrow.

The types of assertions eligible to be proven via induction are those of the form "prove that for all positive integers n, $P(n)$ is true". For statements of this form, induction can be effective. But for nearly any other type of statement, induction can't apply. (That said, Axiom 4.4 is more flexible than induction alone, and it can be applied to an even wider class of mathematical propositions.)

8.3 Induction is a valid method of proof, but *how* to use induction may not be obvious.

In particular, the AM-GM Inequality proof shows how we may need to modify the structure of induction itself before arriving at a valid proof.

None of these observations, though, are dismissals of induction. As you'll discover with further experience, induction is a powerful, versatile proof strategy, often indispensable for many types of exercises you'll be asked to prove.