Contents

6 Rational Approximation and Diophantine Equations **1**

6.1 Simple Examples of Diophantine Equations 2

6.1.1 Linear Diophantine Equations 2

6.1.2 The Frobenius Coin Problem 4

6.1.3 The Equation $x^2 + y^2 = z^2$: Pythagorean Triples 5

6.2 Rational Approximation and Transcendence 7

6.2.1 The Farey Sequences 7

6.2.2 Continued Fractions 11

6.2.3 Infinite Continued Fractions 14

6.2.4 Rational Approximation Via Continued Fractions 19

6.2.5 Irrationality and Transcendence 22

6.3 Pell's Equation 24

6.3.1 Motivation and Small Examples 24

6.3.2 Proofs of Main Results 27

6.3.3 The Super Magic Box 30

6.4 An Assortment of Other Diophantine Equations 31

6.4.1 Assorted Diophantine Equations 32

6.4.2 The Fermat Equation $x^n + y^n = z^n$ 34

6 Rational Approximation and Diophantine Equations

In this chapter, we discuss Diophantine equations, which are concerned with the problem of solving equations over the integers: one of the earliest nontrivial examples was posed by Diophantus, whence the general name of "Diophantine equation" for this class of problems. One of the most famous Diophantine equations is the Fermat equation $x^n + y^n = z^n$, which we will make a central focus of studying.

There is no general procedure for deciding whether a given Diophantine equation possesses any solutions, or (even if existence is known) for finding them all[1]. Thus, many of the methods for solving Diophantine equations are rather *ad hoc*, and so our goals in this chapter are primarily to provide a survey of elementary techniques. One recurring theme, however, will be to exploit the structure of the rings $\mathbb{Z}/m\mathbb{Z}$ and $\mathbb{Z}[\sqrt{D}]$, as we can illustrate in some of our early examples.

We then take a bit of a detour to develop some results regarding another important topic in classical number theory: continued fractions and rational approximation. We then apply our results to study Pell's equation $x^2 - Dy^2 = 1$, along with a smattering of other Diophantine equations.

[1]The question of whether there exists an algorithm that can solve an arbitrary Diophantine equation is Hilbert's tenth problem. A 1970 theorem of Matiyasevich, building off of work of Davis, Putnam, and Robinson, established that determining whether a given Diophantine equation has any solutions is undecidable.

6.1 Simple Examples of Diophantine Equations

- We begin with a short discussion of some simpler examples of Diophantine equations, which will serve as motivation for some of our later discussion.

6.1.1 Linear Diophantine Equations

- The simplest equations are linear equations in two variables, since solving a linear equation in one variable over the integers is trivial (the solution to $ax = b$ is $x = b/a$, assuming a is nonzero and divides b).
    - The general form of a linear equation in two variables is $ax + by = c$, for some fixed integers a, b, and c: our goal is to determine when this equation has an integral solution (x, y), and then to characterize all the solutions.
- In order to give the general solution of a linear equation in two variables we will use modular arithmetic to reduce the two-variable equation to a one-variable equation, which will require the following proposition about linear congruences modulo m:
- Proposition (Linear Equations): The equation $ax \equiv b \pmod{m}$ has a solution for x if and only if $d = \gcd(a, m)$ divides b. If $d|b$, then the set of all such x is given by the residue class $\overline{r}$ modulo m/d, where r is any solution to the equation.
    - Proof: If x is a solution to the congruence $ax \equiv b \pmod{m}$, then there exists an integer k with $ax - mk = b$. Since $d = \gcd(a, m)$ divides the left-hand side, it must divide b.
    - Now suppose $d = \gcd(a, m)$ divides b, and set $a' = a/d$, $b' = b/d$, and $m' = m/d$.
    - Then the original equation becomes $a'dx \equiv b'd \pmod{m'd}$, which is equivalent to $a'x \equiv b' \pmod{m'}$, by one of our properties of congruences.
    - But since a' and m' are relatively prime, a' is a unit modulo m', so we can simply multiply by its inverse to obtain $x \equiv b' \cdot (a')^{-1} \pmod{m'}$. This means that there is a unique solution to the congruence modulo $m' = m/d$, as claimed.
- Now we can solve linear Diophantine equations in two variables:
- Theorem (Linear Diophantine Equations): Let a, b, c be integers with $ab \neq 0$, and set $d = \gcd(a, b)$. If $d \nmid c$, then the equation $ax + by = c$ has no solutions in integers (x, y). If $d \mid c$, then the equation has infinitely many solutions, and if (x_0, y_0) is one solution, then all the others are $(x_0 - bt/d,\ y_0 + at/d)$, for some integer t.
    - Proof: If $a = b = 0$, then the equation $ax + by = c$ is either trivially true (if $c = 0$) or trivially false (if $c \neq 0$), so we can assume that the gcd d is nonzero. If one of a, b is zero, the equation is also trivial, so we may also deal only with the case where $ab \neq 0$.
    - In this case, observe that there is an integral solution to $ax + by = c$ if and only if there is a solution to the congruence $ax \equiv c \pmod{b}$, since then $y = \dfrac{c - ax}{b}$.
    - From our proposition above, we know that $ax \equiv c \pmod{b}$ has a solution only if $d = \gcd(a, b)$ divides c.
    - In this case, if we set $a' = a/d$, $b' = b/d$, and $c' = c/d$, the set of all such x is given by the residue class $\overline{x_0}$ modulo b', where $x_0 \equiv c' \cdot (a')^{-1} \pmod{b'}$.
    - Now if (x, y) is any solution, then by the above, we see that $x = x_0 - bt/d$ for some integer t, and then $y = y_0 + at/d$. This yields the full characterization of the solutions given above.
- Example: Find all solutions to $14x + 18y = 12$ in integers (x, y).
    - First, we compute $\gcd(14, 18) = 2$, and then divide through by the gcd to get $7x + 9y = 6$.
    - This is equivalent to solving $7x \equiv 6 \pmod{9}$.
    - We compute (via the Euclidean algorithm) that the inverse of 7 mod 9 is 4, so multiplying both sides by 4 yields $x \equiv 24 \equiv 6 \pmod{9}$.

- Hence one solution is $(x, y) = (6, -4)$. The set of all solutions is then $(x, y) = \boxed{(6 - 9t, -4 + 7t)}$ for $t \in \mathbb{Z}$.

- Example: Find all solutions to $354x + 936y = 34$ in integers (x, y).
  - Using the Euclidean algorithm we can quickly compute $\gcd(736, 354) = 6$. Since 6 does not divide 34, there are $\boxed{\text{no solutions}}$.

- We will remark that we could also solve linear Diophantine equations in two variables by making changes of variable. We illustrate the idea using an example:

- Example: Find all solutions to the equation $4x + 13y = 5$.
  - By the division algorithm, we have $13 = 3 \cdot 4 + 1$, so we can write the system in the form $4x + (3 \cdot 4 + 1)y = 5$, and rearrange this into the form $4(x + 3y) + 1y = 5$.
  - If we substitute $u = x + 3y$, this new system becomes $4u + y = 5$, which we can easily solve to get $y = 5 - 4u$.
  - Substituting back yields $x = u - 3y = u - 3(5 - 4u) = -15 + 13u$.
  - Thus, we obtain the general solution $(x, y) = \boxed{(-15 + 13u,\ 5 - 4u)}$.

- This latter method, using changes of variable, is the most efficient way to solve systems of linear Diophantine equations involving more variables or equations.
  - The approach is essentially the same as the standard linear algebra procedure of row-reducing a matrix to solve a system of equations.
  - The standard solution technique is to convert the system into matrix form, and then perform row and column operations on the matrix until it is in a sufficiently simple form that the solution to the original system is obvious.
  - The general procedure for solving a system of linear equations over $\mathbb{Z}$ is essentially the same, except for the added complication that all of the row and column operations need to be done over $\mathbb{Z}$. Specifically, the following operations are permissable:
    1. Swap two rows or negate a row (this does not change the system) or add/subtract an integer multiple of one row from another (this yields an equivalent system).
    2. Swap two columns or negate a column (this swaps / negates the underlying variables) or add/subtract an integer multiple of one column from another (this performs a change of variables $x' = x + ay$).
  - As with a system of equations over a field, the end result will be either that the system has no solution, a unique solution, or an infinite family of solutions with some number of free parameters.
  - We will not go into the technical details, since the procedure falls more properly into a course in linear algebra or abstract algebra[2]. Instead, we will just give an example.

- Example: Find all solutions to $3x + 5y + 7z = 11$ in integers (x, y, z).
  - Motivated by the division algorithm, we rewrite the equation as $3(x + y + 2z) + 2y + z = 11$, and then substitute $w = x + y + 2z$.
  - The new equation is $3w + 2y + z = 11$, which we can easily solve, obtaining $z = 11 - 2y - 3w$.
  - Solving for x yields $x = w - y - 2z = -22 + 7w + 3y$, so we obtain the general solution $(x, y, z) = \boxed{(-22 + 7w + 3y,\ y,\ 11 - 2y - 3w)}$, where w, y are arbitrary integers.

[2] In fact, it is equivalent to the procedure for converting a presentation of a finitely generated additive abelian group into a description of the abelian group as a direct product of cyclic groups, which is in turn a special case of the general classification theorem for finitely-generated modules over a principal ideal domain.

6.1.2 The Frobenius Coin Problem

- We just characterized when there exists a solution to the equation $ax + by = c$ in integers (x, y). In various settings (some of which are actually motivated by real-world concerns for once!), we can be interested in knowing for which values of c this equation has a solution in *nonnegative* integers (x, y).
  - If a and b are not relatively prime, clearly c must be divisible by their gcd, and (by dividing through by the gcd) we can reduce to the case where a and b are relatively prime.
- One version of this problem uses postage stamps, which often cost irregular amounts: if, for example, there are postage stamps worth 5 cents and stamps worth 13 cents, is it possible to use them to put exactly 79 cents' worth of postage on an envelope?
  - The most obvious method is simply to make a list of totals that are attainable: 0, 5, 10, 13, 15, 18, 20, 23, 25, 26, 28, 30, 31, 33, 35, 36, 38, 39, 40, 41, 43, 44, 45, 46, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, ...
  - Based on our list, it seems that every value above 47 is attainable. (Indeed, it is easy to see that since 48, 49, 50, 51, and 52 are all attainable, we can obtain any larger number by adding additional 5-cent stamps.)
  - Thus, for example, we get exactly 79 cents of postage by using three 13-cent stamps and eight 5-cent stamps.
- Another version occurs in sports: In American football, a team can score 3 points for a field goal, or 7 points for a touchdown. What possible scores can a team obtain? (Ignore safeties, missed extra points, and so forth.)
  - Like above, we can simply list the totals that are attainable: 0, 3, 6, 7, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, ...
  - Based on our list, only a few values are unattainable: 1, 2, 4, 5, 8, and 11. Indeed, it is easy to see that since 12, 13, and 14 are attainable, any larger number is also attainable by adding more field goals.
- The problem of describing the largest integer that cannot be written as a nonnegative linear combination of two integers (sometimes called the Frobenius coin problem) was first solved by Sylvester:
- Theorem (Sylvester): If a and b are relatively prime integers, then there are exactly $\frac{1}{2}(a-1)(b-1)$ integers that cannot be written in the form $ax + by$ with $x, y \geq 0$, and the largest such integer is $ab - a - b$.
  - Remark: In mathematics competition circles, this result is often known as the "Chicken McNuggets Theorem".
  - Proof: For brevity, we say an integer is "representable" if it can be written in the form $ax + by$ with $x, y \geq 0$.
  - Without loss of generality, assume $a < b$. Arrange the nonnegative integers in an array in the following manner:

$$\begin{array}{ccccc} 0 & 1 & 2 & \cdots & a-1 \\ a & a+1 & a+2 & \cdots & 2a-1 \\ 2a & 2a+1 & 2a+2 & \cdots & 3a-1 \\ \vdots & \vdots & \vdots & & \vdots \\ ab-a & ab-a+1 & ab-a+2 & \cdots & ab-1 \end{array}$$

  - Now we use the array to mark all of the representable integers. We first circle all of the multiples of b: then an integer is representable precisely if it appears in the same column as some multiple of b, lower down.
  - For illustration, here is the array with $a = 3$ and $b = 5$:

$$\begin{array}{ccc} \boxed{0} & 1 & 2 \\ \underline{3} & 4 & \boxed{5} \\ \underline{6} & 7 & \underline{8} \\ \underline{9} & \boxed{10} & \underline{11} \\ \underline{12} & \underline{11} & \underline{14} \end{array}$$

- Since a and b are relatively prime, the integers 0, b, $2b$, ... , $(a-1)b$ all lie in different columns. Thus, the largest element that is left unmarked is the element one row above $(a-1)b$, which is $ab-a-b$, so this is the largest integer not expressible as $ax+by$ with $x, y \geq 0$.
- For the other part, we simply count the number of unmarked integers in the array: the number of integers lying above kb is $\lfloor kb/a \rfloor$, so there are a total of $\sum_{k=0}^{a-1} \left\lfloor \frac{kb}{a} \right\rfloor$ unmarked integers in the array.
- We can interpret this sum geometrically as the number of lattice points lying under the line $y = \frac{b}{a}x$, with $1 \leq x \leq a-1$. Equivalently, this is the total number of lattice points lying strictly inside the rectangle with vertices $(0,0)$, $(a,0)$, (a,b), $(0,b)$ and below the diagonal.
- By symmetry, since there are no lattice points on the interior of the diagonal, the points below the diagonal represent exactly half of the lattice points lying strictly inside the $a \times b$ rectangle. Since this full set of points represents an $(a-1) \times (b-1)$ rectangle, there are $(a-1)(b-1)$ such lattice points. Therefore, the number of unmarked integers in the array is $\frac{1}{2}(a-1)(b-1)$, as claimed.
- Remark: Another way to obtain this count is to prove that if $0 \leq c \leq ab-a-b$, then c is representable if and only if $ab-a-b-c$ is not representable.

- Example: If there are postage stamps worth 5 cents and stamps worth 13 cents, then with $a=5$ and $b=13$ in the theorem above, the largest non-representable integer is $5 \cdot 13 - 13 - 5 = 47$, and there are in total $\frac{1}{2} \cdot 4 \cdot 12 = 24$ unattainable totals.
- We could of course generalize this problem, to ask: for given integers a_1, a_2, ... , a_k, what is the largest integer n that cannot be written as a nonnegative integer linear combination of the a_i?
  - It turns out that there is no known general formula when $k > 2$.
  - For a fixed number of denominations k, there does exist a polynomial-time algorithm (polynomial in $\log a_k$, specifically) for computing this maximum integer n, but it is not appreciably faster than merely attempting to list the possibilities!
  - For a variable number of denominations k, it is known that computing n is NP-hard.

6.1.3 The Equation $x^2 + y^2 = z^2$: Pythagorean Triples

- Now that we have discussed solving linear equations in integers, we turn our attention to the one of the simplest quadratic Diophantine equations: characterizing integer triples (x, y, z) such that $x^2 + y^2 = z^2$.
  - Such triples are naturally called Pythagorean triples because (by the Pythagorean theorem) they form the sides of a right triangle: a familiar example is the nearly-ubiquitous $(3, 4, 5)$ triangle.
  - Since the defining equation is homogeneous (i.e., all of the terms have the same degree), if we have one Pythagorean triple, we can create more by scaling x, y, and z: thus we obtain $(6, 8, 10)$, $(9, 12, 15)$, and so forth from $(3, 4, 5)$.
  - To exclude these essentially repetitious cases, we say a Pythagorean triple (x, y, z) is primitive if $\gcd(x, y, z) = 1$, and would like to characterize the primitive triples.
  - First we note that if (x, y, z) is a primitive Pythagorean triple, x and y cannot both be even, since then z would also be even. Also, x and y cannot both be odd, since then $x^2 + y^2 \equiv 2 \pmod 4$, but 2 is not a square modulo 4. So we conclude that z must be odd, and that exactly one of x and y is also odd.

- Theorem (Primitive Pythagorean Triples): Every primitive Pythagorean triple of the form (x, y, z) with x even is of the form $(x, y, z) = (2st,\ s^2 - t^2,\ s^2 + t^2)$, for some relatively prime integers $s > t$ of opposite parity, and (conversely) any such triple is Pythagorean and primitive. As a consequence, the positive-integer solutions (x, y, z) to $x^2 + y^2 = z^2$ can be uniquely written as $(x, y, z) = (2kst, k(s^2 - t^2), k(s^2 + t^2))$ for a unique positive integer k and relatively prime positive integers $s > t$ of opposite parity.

- For the reverse direction, it is easy to see that $(2st)^2 + (s^2 - t^2)^2 = (s^2 + t^2)^2$ simply by multiplying out. Furthermore, it is easy to check that if s and t are relatively prime and have opposite parity, that $\gcd(s^2 - t^2, s^2 + t^2) = 1$, so this triple is primitive. The characterization of all triples follows from our discussion of primitive triples above, since we may take $k = \gcd(x, y, z)$.
- We will give three proofs of the nontrivial direction: one using the arithmetic of $\mathbb{Z}$, one using the arithmetic of $\mathbb{Z}[i]$, and one using geometry.
- The central idea in the first proof is to rearrange the equation and use the arithmetic of $\mathbb{Z}$. The central idea in the second proof is to exploit the fact that $\mathbb{Z}[i]$ has unique factorization, while the central idea in the third proof is to use the geometry of the dehomogenized curve $x^2 + y^2 = 1$ to study the rational solutions.
- Proof 1: Suppose $x^2 + y^2 = z^2$ and x, y, z are relatively prime.
- Since y and z are both odd and x is even, we can rewrite the equation as $\frac{z-y}{2} \cdot \frac{z+y}{2} = \left(\frac{x}{2}\right)^2$.
- Now we claim that $\frac{z-y}{2}$ and $\frac{z+y}{2}$ are relatively prime: their gcd divides their sum z and their difference y, and since y and z are relatively prime, the gcd must be 1.
- Since $\frac{z-y}{2}$ and $\frac{z+y}{2}$ share no prime divisors and their product is a square, each of them must individually be a square, by the uniqueness of prime factorization.
- Hence there exist integers s and t such that $\frac{z-y}{2} = t^2$ and $\frac{z+y}{2} = s^2$.
- Then $z = s^2 + t^2$ and $y = s^2 - t^2$, and then we also obtain $x = 2st$, as claimed. Furthermore, s and t are necessarily relatively prime and have opposite parity, since (x, y, z) is primitive.
- Proof 2: Suppose $x^2 + y^2 = z^2$ and x, y, z are relatively prime.
- In $\mathbb{Z}[i]$, we factor the equation as $(x + iy)(x - iy) = z^2$.
- Now we claim that $x + iy$ and $x - iy$ are relatively prime as elements of $\mathbb{Z}[i]$: any greatest common divisor in $\mathbb{Z}[i]$ must divide $2x$ and $2y$, so since x and y are relatively prime integers, the gcd must divide 2. However, $x + iy$ is not divisible by the Gaussian prime $1 + i$, since x and y are of opposite parity.
- Hence, since $x + iy$ and $x - iy$ are relatively prime and have product equal to a square, by the uniqueness of prime factorization in $\mathbb{Z}[i]$, there exists some $s + it \in \mathbb{Z}[i]$ and some unit $u \in \{1, i, -1, -i\}$ such that $x + iy = u(s + it)^2$.
- Multiplying out yields $x + iy = u\left[(s^2 - t^2) + (2st)i\right]$. Since x is even and y is odd, we must have $u = \pm i$: then writing out the various possibilities yields the given parametrization.
- Proof 3: Suppose $x^2 + y^2 = z^2$ and x, y, z are relatively prime.
- Dividing by z^2 yields the equivalent equation $\left(\frac{x}{z}\right)^2 + \left(\frac{y}{z}\right)^2 = 1$, so it is sufficient to describe all points (a, b) on the unit circle $x^2 + y^2 = 1$ whose coordinates are both rational numbers.
- To do this, consider all non-vertical lines passing through the point $(-1, 0)$.
- Such a line will intersect the circle $x^2 + y^2 = 1$ in exactly one other point. If the coordinates of this point are rational, then the line will have rational slope.
- Conversely, if the line has rational slope $\frac{t}{s}$, its equation is $y = \frac{t}{s}(x + 1)$ so we can simply compute the other intersection point to see that it is $(x, y) = \left(\frac{s^2 - t^2}{s^2 + t^2}, \frac{2st}{s^2 + t^2}\right)$, which is rational.
- Thus, the rational points on the unit circle are those of the form $\left(\frac{s^2 - t^2}{s^2 + t^2}, \frac{2st}{s^2 + t^2}\right)$ for some integers s and t. Clearing the denominator yields the desired Pythagorean triples.
- Remark: The third proof is closely related to the Weierstrass substitution $u = \tan(\theta/2)$, which transforms an integral of any rational function of $\sin(\theta)$ and $\cos(\theta)$ into an integral of a rational function of u, which can then be evaluated using partial fraction decomposition. (With the notation above, $u = s/t$.)

- Using the characterization above, we can easily generate a list of Pythagorean triples with small hypotenuses.

- Here is a table of the Pythagorean right triangles with hypotenuse ≤ 100:

s	t	Primitive Triple	Non-Primitive Triples	s	t	Primitive Triple
2	1	$(3,4,5)$	$(6,8,10)$, $(9,12,15)$, ... , $(60,80,100)$	7	2	$(28,45,53)$
3	2	$(5,12,13)$	$(10,24,26)$, $(15,36,39)$, ... , $(35,84,91)$	7	4	$(33,56,65)$
4	1	$(8,15,17)$	$(16,30,34)$, $(24,45,51)$, ... , $(40,75,85)$	7	6	$(13,84,85)$
4	3	$(7,24,25)$	$(14,48,50)$, $(21,72,75)$	8	1	$(16,63,65)$
5	2	$(20,21,29)$	$(40,42,58)$, $(60,63,87)$	8	3	$(48,55,73)$
5	4	$(9,40,41)$	$(18,80,82)$	8	5	$(39,80,89)$
6	1	$(12,35,37)$	$(24,70,74)$	9	2	$(36,77,85)$
6	5	$(11,60,61)$		9	4	$(65,72,97)$

- Using our characterization, we can enumerate all of the Pythagorean right triangles having a side of a particular length.
- Example: Find all Pythagorean right triangles having one side of length 20.
  - From our result above, any such right triangle has legs of lengths $k(2st)$ and $k(s^2-t^2)$, with hypotenuse $k(s^2+t^2)$, where $s>t$ are positive integers of opposite parity and k is some positive integer.
  - If $20 = 2stk$, then $10 = stk$, so $(s,t,k) = (10,1,1)$ or $(5,2,1)$ or $(2,1,5)$, yielding 20-99-101, 20-21-29, and 15-20-25 triangles.
  - If $20 = k(s^2-t^2)$, then k must be divisible by 4. Since $k \neq 20$ we see $k=4$, and then $s^2-t^2=5$ requires $s=3$ and $t=2$. This yields a 20-48-52 triangle.
  - If $20 = k(s^2+t^2)$, then since $s^2+t^2 \geq 5$ the only possibilities are $k=4$ (yielding $s=2$ and $t=1$), $k=2$ (yielding $s=3$ and $t=1$ but these are not of opposite parity) or $k=1$ (yielding $s=4$ and $t=2$ but again these are not of opposite parity). This yields a 12-16-20 triangle.
  - Hence there are five such triangles: $\boxed{(20,99,101),\ (20,21,29),\ (15,20,25),\ (20,48,52),\ (12,16,20)}$.

6.2 Rational Approximation and Transcendence

- When describing real numbers, for convenience we often want to give a nearby rational number that is a good approximation.
  - Indeed, this idea is implicitly embedded in the notion of the decimal expansion of a real number.
  - For example, writing $\pi = 3.1415926535\ldots$ formally means that π is the limit of the sequence 3, 3.1, 3.14, 3.141, 3.1415, 3.14159, ..., and so truncating this sequence after some finite number of steps will provide a good approximation of π. More specifically, in the case of the decimal expansion to n digits, the approximation is accurate to within an error of 10^{-n}.
  - Decimal numbers are all well and good, but we can often get better approximations using arbitrary rational numbers, rather than just ones whose denominators are powers of 10.
  - We will now study some problems related to rational approximation of real numbers by rational numbers.

6.2.1 The Farey Sequences

- If we are seeking to approximate a real number α, one thing we might first look at is the set of rational numbers of small denominator. Since we want to understand distances between nearby numbers, we should arrange the rationals in increasing order. This yields the famous Farey sequences:
- Definition: The Farey sequence of level n is the set of rational numbers between 0 and 1 whose denominators (in lowest terms) are $\leq n$, arranged in increasing order.
  - Example: The Farey sequence of level 4 is $\frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1}$.
  - Example: To obtain the Farey sequence of level 5, we simply insert the terms with denominator 5 in the proper locations: $\frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{1}{1}$.

- There are several natural and immediate questions about the Farey sequence of level n: for example, how many terms does it have? Are consecutive terms related to each other? In going from the $(n-1)$st to the nth sequence, how many terms are added and where do they go?
  - Several of these can be immediately answered: of all rational numbers of the form $\frac{k}{n}$, the ones in lowest terms will be those with $\gcd(k,n) = 1$. Thus, the number of such terms added in going from the $(n-1)$st Farey sequence to the nth is simply $\varphi(n)$.
  - By a trivial induction, we see that the length of the Farey sequence of level n is $1 + \sum_{d=1}^{n} \varphi(d)$. It is a nontrivial problem to estimate the rate of growth of this function, but it turns out to be approximately $\frac{3}{\pi^2} n^2$.
  - Some brief numerical investigation quickly leads to a number of simple properties of the Farey sequences:

- Proposition (Properties of Farey Sequences): Let n be a positive integer.

  1. If $\frac{a}{b}$ and $\frac{c}{d}$ are consecutive terms in the Farey sequence of level n, then $bc - ad = 1$.
     - Proof: Suppose $\frac{a}{b}$ and $\frac{c}{d}$ are consecutive terms in the Farey sequence of level n.
     - In the plane, draw the triangle whose vertices are (0,0), (b,a), and (d,c).
     - By Pick's Theorem[3], the area of this lattice-point triangle is $\frac{1}{2}B + I - 1$, where B is the number of lattice points on the boundary and I is the number of points in the interior. We claim $B = 3$ and $I = 0$.
     - To see this, suppose there were a lattice point (x,y) in the interior, where (necessarily) $y \leq \max(b,d)$. Then the slope of the line joining $(0,0)$ to (x,y) would lie strictly between a/b and c/d: but then y/x would be between a/b and c/d in the Farey sequence, which by hypothesis it is not.
     - Now suppose there were a lattice point on the boundary not equal to one of the vertices. It cannot lie on the side joining (0,0) and (b,a), since a and b are relatively prime. Similarly, it cannot lie on the side joining (0,0) and (d,c). If it were on the side joining (b,a) and (d,c), then by the same argument given above, there would be a term between a/b and c/d in the Farey sequence.
     - Thus, $B = 3$ and $I = 0$, so the triangle has area $\frac{1}{2}$. By basic geometry (either by enclosing this triangle with larger right triangles, or by noting that the area of the triangle is half of the magnitude of the cross product $\langle b,a,0\rangle \times \langle d,c,0\rangle = \langle 0,0,bc-ad\rangle$), the area of this triangle is $\frac{1}{2}|bc-ad|$, so since $bc > ad$ we conclude immediately that $bc - ad = 1$.
  2. If $\frac{a}{b}$, $\frac{e}{f}$, and $\frac{c}{d}$ are three consecutive terms in a Farey sequence, then $\frac{e}{f} = \frac{a+c}{b+d}$.
     - Notation: This last expression is sometimes called the mediant of a/b and c/d. (It is also occasionally called the baseball average, since it is the expression, frequently computed in baseball, used when combining several hit percentages into a single statistic.)
     - Proof: Suppose $\frac{a}{b}$, $\frac{e}{f}$, and $\frac{c}{d}$ are consecutive. By (1), since a/b and e/f are consecutive we have $be - af = 1$, and by (2) since e/f and c/d are consecutive we have $cf - de = 1$.
     - This is a system of two linear equations in the two variables e and f, so solving it (e.g., by multiplying the first equation by d, the second by a, and adding) yields $e = (a+c)/(bc-ad)$ and $f = (b+d)/(bc-ad)$: thus, $\frac{e}{f} = \frac{a+c}{b+d}$, as claimed.
     - One can check directly that $\frac{a+c}{b+d}$ appears between $\frac{a}{b}$ and $\frac{c}{d}$ in the Farey sequence of level $b+d$, since $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$.

[3]Pick's theorem is a result from elementary geometry that says that the area of a plane lattice polygon is equal to $\frac{1}{2}B + I - 1$, where B is the number of lattice points on the boundary and I is the number of points in the interior. To prove this result, one may first establish that it holds for rectangles and is also consistent under gluing regions together or removing pieces of regions. Applying these results shows that it holds for right triangles, then arbitrary triangles, and finally arbitrary polygons.

3. If $\frac{a}{b}$ and $\frac{c}{d}$ are rational numbers between 0 and 1 with $bc-ad=1$, then these two terms are consecutive entries in the Farey sequence of level $\max(b,d)$. The first term that will appear between them (in a later sequence) is $\frac{a+c}{b+d}$, and this first occurs in the Farey sequence of level $b+d$.
   - Proof: For the first statement, suppose $\frac{e}{f}$ is the term immediately following $\frac{a}{b}$ in the Farey sequence of level $\max(b,d)$.
   - Then $be-af=1$ by (1). Subtracting $bc-ad=1$ yields $b(c-e)-a(d-f)=0$, so $b(c-e)=a(d-f)$. Since a and b are relatively prime, we conclude that b divides $d-f$. Since $f \leq \max(b,d) < b+d$, the only possibility is that $f=d$, and then $e=c$.
   - Alternatively, we could have observed that both (e,f) and (c,d) are solutions to the linear Diophantine equation $bx-ay=1$, and used the structure of the solutions to deduce this result.
   - For the second statement, we just showed that a/b and c/d are consecutive in the Farey sequence of level $\max(b,d)$.
   - Now increase the level of the sequence in increments of 1: if e/f is the first term to appear between a/b and c/d, then by (2), it would necessarily be the case that $e=(a+c)/(bc-ad)=a+c$ and $f=(b+d)/(bc-ad)=b+d$.
4. The rational numbers $\frac{a}{b}$ and $\frac{c}{d}$ are consecutive terms in the Farey sequence of level n if and only if $bc-ad=1$ and $b+d>n$.
   - Proof: We must have $bc-ad=1$ by (1). Also, if $b+d \leq n$, then $\frac{a+c}{b+d}$ is a term between a/b and c/d as noted in (2).
   - Thus, we must have $bc-ad=1$ and $b+d>n$. But if both conditions hold, then (3) immediately implies that there are no terms between a/b and c/d in the Farey sequence of level n.
5. If a/b and e/f are consecutive terms in the Farey sequence of level n, the term immediately following e/f is c/d, where $c=\left\lfloor \frac{n+b}{f} \right\rfloor e-a$ and $d=\left\lfloor \frac{n+b}{f} \right\rfloor f-b$.
   - Proof: By the mediant property (2), we know that $\frac{e}{f}=\frac{a+c}{b+d}$. Thus, there must exist some integer k such that $a+c=ke$ and $b+d=kf$, so that $c=ke-a$ and $d=kf-b$. Since the closest term to e/f will have k as large as possible, and since $d \leq n$, the largest possible value of k is $\left\lfloor \frac{n+b}{f} \right\rfloor$.

- Using the above results, we can construct the portion of any Farey sequence around any desired rational number, without needing to compute all of the terms in the sequence.
- Example: Find the first three terms after 11/202 in the Farey sequence of level 500.
   - By the above results, if 11/202 and c/d are consecutive terms, then $202c-11d=1$.
   - Solving this Diophantine equation using the Euclidean algorithm produces the solutions $(c,d)=(3+11k,\ 55+202k)$ for $k \in \mathbb{Z}$.
   - The larger the value of k is, the smaller the value of $\frac{c}{d}-\frac{11}{202}=\frac{1}{202d}$ will be. The largest possible value for k is $k=2$, so the first term is $\frac{25}{457}$.
   - Now we can apply the two-term recursion to 11/202 and 25/457 to quickly find the next terms: they are 14/255 and 17/308.
   - Thus, the three terms are $\boxed{\frac{25}{457},\ \frac{14}{255},\ \frac{17}{308}}$.
- In some cases, we can fill in the portion of any desired Farey sequence between two consecutive terms of some Farey sequence by taking mediants.
- Example: Find all terms between 7/33 and 14/65 in the Farey sequence of level 100.

- First, we notice that these terms are not consecutive (in any Farey sequence), because $14 \cdot 33 - 7 \cdot 65 = 7$, not 1.
- We start by finding terms between them: the mediant of these two terms is $21/98 = 3/14$.
- Now $7/33$ and $3/14$ are consecutive in the Farey sequence of level 33, since $33 \cdot 3 - 7 \cdot 14 = 1$.
- Also, $3/14$ and $14/65$ are consecutive in the Farey sequence of level 65, since $14 \cdot 14 - 3 \cdot 65 = 1$.
- At this point, we just need to fill in the missing terms. Because the terms we have identified are all adjacent, these are all given by computing mediants of the terms already found. We can stop computing mediants when the sum of two consecutive denominators exceeds 100 in each step.
- Filling in all of the remaining mediants yields the sequence $\boxed{\frac{7}{33}, \frac{17}{80}, \frac{10}{47}, \frac{13}{61}, \frac{16}{75}, \frac{19}{89}, \frac{3}{14}, \frac{20}{93}, \frac{17}{79}, \frac{14}{65}}$.

- Example: Find all terms between $6/77$ and $5/62$ in the Farey sequence of level 80.
  - First, we notice that these terms are not consecutive (in any Farey sequence), because $5 \cdot 77 - 6 \cdot 62 = 13$, not 1.
  - The mediant of these terms is $12/139$, which is not in the desired Farey sequence of level 80.
  - Instead, we can search for the term a/b immediately following $6/77$, which necessarily has $77a - 6b = 1$. Solving this linear Diophantine equation using the Euclidean algorithm yields $a = 5 + 6k$, $b = 64 + 77k$, so we may take $a/b = 5/64$.
  - Since the mediant of $6/77$ and $5/64$ is not in the Farey sequence of level 80, we may generate the remaining terms up to $5/62$ using the two-term recursion.
  - This yields the sequence $\boxed{\frac{6}{77}, \frac{5}{64}, \frac{4}{51}, \frac{3}{38}, \frac{5}{63}, \frac{2}{25}, \frac{5}{62}}$.

- We can use the Farey sequences to give some basic results about rational approximation:

- Proposition (Rational Approximation via Farey): Let n be a positive integer and α be a real number. Then the following hold:

1. There exists a rational number $\frac{p}{q}$ such that $0 < q \leq n$ and $\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q(n+1)}$.
   - Proof: By replacing α with $\alpha - \lfloor \alpha \rfloor$ as necessary, we can assume that α lies in $[0, 1]$.
   - Now consider the Farey sequence of level n, and let $\frac{a}{b}$ and $\frac{c}{d}$ be two consecutive terms such that $\frac{a}{b} \leq \alpha \leq \frac{c}{d}$. By our earlier results, we know that $bc - ad = 1$ and $b + d \geq n + 1$.
   - The number α either lies in the interval $\left[\frac{a}{b}, \frac{a+c}{b+d}\right]$ or in $\left[\frac{a+c}{b+d}, \frac{c}{d}\right]$.
   - In the first case, $\left|\alpha - \frac{a}{b}\right| \leq \left|\frac{a}{b} - \frac{a+c}{b+d}\right| = \frac{|ad-bc|}{b(b+d)} \leq \frac{1}{b(n+1)}$.
   - In the second case, $\left|\alpha - \frac{c}{d}\right| \leq \left|\frac{c}{d} - \frac{a+c}{b+d}\right| = \frac{|ad-bc|}{d(b+d)} \leq \frac{1}{d(n+1)}$.
   - Hence, in either case, we obtain a rational number $\frac{p}{q}$ such that $\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q(n+1)}$.
2. If α is irrational, then there are infinitely many distinct rational numbers p/q such that $|\alpha - p/q| < 1/q^2$.
   - Proof: Apply (1) to the Farey sequence of level n for each n: this yields a collection of rational numbers $\frac{p_n}{q_n}$ such that $\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n(n+1)} < \frac{1}{q_n^2}$, and with $q_n \leq n$.
   - Since α is irrational, none of these differences can be zero, and so there must be infinitely many different terms $\frac{p_n}{q_n}$, since the distances $\left|\alpha - \frac{p_n}{q_n}\right|$ become arbitrarily small, but remain nonzero.
3. If α is irrational, then there are infinitely many pairs of positive integers (m, n) such that $|m\alpha - n| < 1/m$.

- This is an approximation theorem first proven by Dirichlet and is sometimes known as Dirichlet's Diophantine approximation theorem.
- Proof: Clear denominators in (2).

- We can illustrate these results with a typical irrational $\alpha = \sqrt{2} \approx 1.4142136\ldots$ for various n.
  - For example, with $n = 5$ (in the Farey sequence of level 5) the two entries surrounding $\sqrt{2} - 1$ are $2/5$ and $1/2$. We can see that $\left|\sqrt{2} - 7/5\right| \approx 0.0142 < \dfrac{1}{5 \cdot 5}$, so $7/5$ has the desired property in (1) of the proposition. In fact, $3/2$ also has the desired property, since $\left|\sqrt{2} - 3/2\right| \approx 0.0858 < \dfrac{1}{5 \cdot 2}$.
  - Taking an increasing sequence of values of n up to $n = 100$ then yields various $\dfrac{p}{q}$ with $\left|\sqrt{2} - \dfrac{p}{q}\right| < \dfrac{1}{q^2}$ as indicated in (2): specifically, we obtain the sequence 1, 2, 3/2, 4/3, 7/5, 10/7, 17/12, 24/17, 41/29, 58/41, 99/70, 140/99, ...

6.2.2 Continued Fractions

- We now discuss another method for generating rational approximations of a given real number α.
  - If we want to give an approximation to α, it will be of the form $a_0 + x$ where $a_0 = \lfloor \alpha \rfloor$ is the greatest integer less than or equal to α and $0 \leq x < 1$.
  - In such a situation, we have $1/x > 1$, so we could approximate $1/x$ as an integer $a_1 = \lfloor 1/x \rfloor$, yielding an approximation to α of the form $a_0 + \dfrac{1}{a_1}$.
  - For example, if we wanted to approximate π, we would compute $\lfloor \pi \rfloor = 3$, and then note $x = \pi - 3 = 0.141592\ldots$ has $1/x \approx 7.06251\ldots$, and so we get the well-known approximation to π of $3 + 1/7 = 22/7$.
  - Alternatively, instead of stopping after one step, we could then approximate $1/x$ in the same way: it is of the form $a_1 + y$ where $a_1 = \lfloor 1/x \rfloor$ and $0 \leq y < 1$. We can continue this procedure as long as each of the rounded-off values are not exact integers.
  - For π, the next step would be noting that $y = 1/x - 7 \approx 0.06251$ has $1/y \approx 15.9966$, and so we get an approximation $1/x - 7 \approx 16$, which yields an approximation to π of $3 + 1/(7 + 1/16) = 355/113 \approx 3.14159292$, which is accurate to 6 decimal places.
  - It is clear that we can continue this procedure to generate increasingly accurate rational approximations of α.
  - The resulting expression has the form $a_0 + \dfrac{1}{a_1 + \dfrac{1}{a_2 + \cdots}}$, which is called a continued fraction:

- Definition: A finite continued fraction is an expression of the form $a_0 + \dfrac{1}{a_1 + \dfrac{1}{a_2 + \dfrac{1}{\ddots + \dfrac{1}{a_{k-1} + \dfrac{1}{a_k}}}}}$, where the a_i are positive real numbers. For brevity, we will denote this expression using the much more compact notation $[a_0, a_1, \cdots, a_k]$. If the a_i are integers, we term it a simple continued fraction.
  - Example: $[2, 3, 4] = 2 + \dfrac{1}{3 + \dfrac{1}{4}} = \dfrac{30}{13}$.
  - We note a few very basic properties: $[a_0, a_1, \ldots, a_k] = a_0 + \dfrac{1}{[a_1, \ldots, a_k]} = [a_0, a_1, \ldots, a_{k-1} + \dfrac{1}{a_k}]$.

- Clearly, every simple continued fraction is a rational number. Conversely, every rational number can be written as a simple continued fraction: if p/q is any positive rational number in lowest terms, then if we apply the Euclidean algorithm to write

$$\begin{aligned} p &= q_1 q + r_1 \\ q &= q_2 r_1 + r_2 \\ r_1 &= q_3 r_2 + r_3 \\ &\vdots \\ r_{k-1} &= q_k r_k + 1 \\ r_k &= q_{k+1} \end{aligned}$$

  where we know the last remainder will be 1 since p/q is in lowest terms, then it is an easy induction to verify that $\dfrac{p}{q} = [q_1, q_2, \ldots, q_k, q_{k+1}]$.

  - Furthermore, by the uniqueness of the Euclidean algorithm, all of the quotients are unique, so the expression is unique, except for the fact that we can write $[q_1, q_2, \ldots, q_k] = [q_1, q_2, \ldots, q_k - 1, 1]$.
  - If we exclude the case where the final term is equal to 1, then every positive rational number can be written uniquely as a continued fraction.

- Example: To convert $\dfrac{17}{7}$ into a continued fraction, we first compute

$$\begin{aligned} 17 &= 2 \cdot 7 + 3 \\ 7 &= 2 \cdot 3 + 1 \\ 3 &= 3 \cdot 1 \end{aligned}$$

  so that, by the above, $\dfrac{17}{7} = \boxed{[2, 2, 3]}$.

  - Another way of doing this is just to work it out explicitly, by writing $\dfrac{17}{7} = 2 + \dfrac{3}{7} = 2 + \dfrac{1}{7/3} = 2 + \dfrac{1}{2 + \dfrac{1}{3}}$.

- Example: To convert $\dfrac{67}{19}$ into a continued fraction, we compute

$$\begin{aligned} 67 &= 3 \cdot 19 + 10 \\ 19 &= 1 \cdot 10 + 9 \\ 10 &= 1 \cdot 9 + 1 \\ 9 &= 9 \cdot 1 \end{aligned}$$

  so that $\dfrac{67}{19} = \boxed{[3, 1, 1, 9]}$.

- If we truncate a continued fraction after some number of terms, we will obtain an approximation to the true value.

- Definition: If $C = [a_0, a_1, \ldots, a_k]$ is given, then the continued fraction $C_n = [a_0, a_1, \ldots, a_n]$ for $n < k$ is called the nth convergent to C.

  - Example: For $\dfrac{117}{101} = [1, 6, 3, 5]$, the successive convergents are $[1] = 1$, $[1, 6] = \dfrac{7}{6}$, $[1, 6, 3] = \dfrac{22}{19}$, and $[1, 6, 3, 5] = \dfrac{117}{101}$.
  - Observe that $\dfrac{117}{101} \approx 1.1584$, while $\dfrac{7}{6} \approx 1.1666$ and $\dfrac{22}{19} \approx 1.1580$.
  - Notice that the convergents are fairly close to the actual value of the continued fraction, and their accuracy improves as we take higher convergents. We will make this idea rigorous in a moment.

- Here are some simple properties of the convergents of continued fractions:
- Proposition (Properties of Convergents): Let $C = [a_0, a_1, \ldots, a_k]$ where the a_i are positive, and define $p_{-1} = 1$, $p_0 = a_0$, and $p_n = a_n p_{n-1} + p_{n-2}$, and also $q_{-1} = 0$, $q_0 = 1$, and $q_n = a_n q_{n-1} + q_{n-2}$. We then have the following:

  1. The convergent $C_n = p_n/q_n$.
     - Proof: We use induction on n. The base cases $n = 1$ and $n = 2$ are trivial, since $[a_0] = a_0/1$ and $[a_0, a_1] = a_0 + 1/a_1 = (a_0 a_1 + 1)/a_1$, as claimed.
     - For the inductive step, suppose we know that the result holds for $n \leq m$. By hypothesis, for any x, it is the case that $[a_0, a_1, \ldots, a_{m-1}, x] = \dfrac{p_{m-1}x + p_{m-2}}{q_{m-1}x + q_{m-2}}$.
     - Now observe that $[a_0, a_1, \ldots, a_{m-1}, a_m, a_{m+1}] = [a_0, a_1, \ldots, a_{m-1}, a_m + \dfrac{1}{a_{m+1}}]$ and apply the above result with $x = a_m + \dfrac{1}{a_{m+1}}$ to obtain

$$\begin{aligned}
[a_0, a_1, \ldots, a_{m-1}, a_m + \frac{1}{a_{m+1}}] &= \frac{\left(a_m + \dfrac{1}{a_{m+1}}\right) p_{m-1} + p_{m-2}}{\left(a_m + \dfrac{1}{a_{m+1}}\right) q_{m-1} + q_{m-2}} \\
&= \frac{(a_m p_{m-1} + p_{m-2}) + p_{m-1}/a_{m+1}}{(a_m q_{m-1} + q_{m-2}) + q_{m-1}/a_{m+1}} \\
&= \frac{p_m + p_{m-1}/a_{m+1}}{q_m + q_{m-1}/a_{m+1}} \\
&= \frac{a_{m+1}p_m + p_{m-1}}{a_{m+1}q_m + q_{m-1}} = \frac{p_{m+1}}{q_{m+1}},
\end{aligned}$$

     as desired.
  2. We have $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$ and $p_n q_{n-2} - p_{n-2} q_n = (-1)^{n-2} a_n$.
     - Proof: For the first statement, by the recursion we can write

$$\begin{aligned}
p_n q_{n-1} - p_{n-1} q_n &= (a_n p_{n-1} + p_{n-2}) q_{n-1} - p_{n-1}(a_n q_{n-1} - q_{n-2}) \\
&= -(p_{n-1} q_{n-2} - p_{n-2} q_{n-2})
\end{aligned}$$

     so since $p_1 q_0 - p_0 q_1 = 1$, by a trivial induction we see that $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$.
     - The second statement follows in the same way. (We skip the algebra.)
  3. We have $C_n - C_{n-1} = \dfrac{(-1)^{n-1}}{q_{n-1} q_n}$ and $C_n - C_{n-2} = \dfrac{(-1)^{n-2} a_n}{q_{n-2} q_n}$.
     - Proof: Divide the relations $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$ and $p_n q_{n-2} - p_{n-2} q_n = (-1)^{n-2} a_n$ from (3) by $q_n q_{n-1}$ and $q_n q_{n-2}$ respectively.
  4. We have $C_1 > C_3 > C_5 > \cdots > C_6 > C_4 > C_2$, and $|C - C_n| \leq \dfrac{1}{q_n q_{n+1}} < \dfrac{1}{q_n^2}$.
     - Proof: From $C_n - C_{n-2} = \dfrac{(-1)^{n-2} a_n}{q_{n-2} q_n}$ in (3), we see that $C_n < C_{n-2}$ if n is odd, and $C_n > C_{n-2}$ if n is even.
     - Hence, by a trivial induction, we see $C_1 > C_3 > C_5 > \cdots$ and $\cdots > C_6 > C_4 > C_2$.
     - Furthermore, since $C_{2n+1} > C_{2n}$ for every n, we can combine the two chains of inequalities to obtain the third statement.
     - For the last statement, we simply observe that the inequalities above imply that C is between C_n and C_{n+1} for every n, hence the triangle inequality implies $|C - C_n| \leq |C_{n+1} - C_n| = \dfrac{1}{q_n q_{n+1}} < \dfrac{1}{q_n^2}$.
- We will mention that property (1) gives a fairly efficient procedure for computing the numerators and denominators of the convergents $C_n = p_n/q_n$, using a computational procedure that is sometimes referred to as the "magic box". It works as follows:

- The rows in the table are the sequences a_n, p_n, and q_n.
- Starting with the terms a_n, which we assume are given to us, we can then evaluate $p_n = a_n p_{n-1} + p_{n-2}$ and $q_n = a_n q_{n-1} + q_{n-2}$ after starting with the initial conditions $p_{-2} = 0$, $q_{-2} = 1$, $p_{-1} = 1$, $q_{-1} = 0$.
- Explicitly, we calculate the next term in the row for p_n by evaluating a_n times the previous entry plus the entry before that, and similarly for the next term in the row for q_n.
- For example, here is the magic box calculation for the continued fraction $[2, 4, 5, 1, 3] = 215/96$:

n	-2	-1	0	1	2	3	4
a_n			2	4	5	1	3
p_n	0	1	2	9	47	56	215
q_n	1	0	1	4	21	25	96

To illustrate, the entries in the column for $n = 4$ are calculated as $3 \cdot 56 + 47 = 215$ and $3 \cdot 25 + 21 = 96$.

6.2.3 Infinite Continued Fractions

- We have discussed finite continued fraction expansions (of rational numbers) and shown that their convergents obey some nice relations, and we can compute the simple continued fraction expansion of any rational number using the Euclidean algorithm.
  - We would now like to extend our discussion to include irrational numbers: what, for example, does it mean to ask for the continued fraction expansion of $\sqrt{2}$, or of π, or $\ln(2)$?
  - None of these is a rational number, so any such expansion cannot be finite.
  - To handle this, we simply extend our definition of continued fraction to an infinite continued fraction by taking a limit.
- Definition: Given a sequence a_0, a_1, a_2, ... of positive integers, we define the infinite continued fraction $\alpha = [a_0, a_1, a_2, \dots]$ to be the limit $\lim_{n\to\infty} [a_0, a_1, \dots, a_n]$ of its finite continued fraction convergents.
  - It is worthwhile explaining why this limit exists. From our results on convergents, we know that if $C_n = [a_0, a_1, \dots, a_n]$, then $C_1 > C_3 > C_5 > \cdots > C_6 > C_4 > C_2$.
  - Thus, the sequence C_1, C_3, C_5, ... is monotone decreasing and bounded below (by C_2), hence it has a limit by the monotone convergence theorem[4].
  - Similarly, the sequence C_2, C_4, C_6, ... is monotone increasing and bounded above (by C_1), hence it also has a limit by the monotone convergence theorem.
  - These two limits must be equal because $|C_n - C_{n+1}| < 1/q_n^2$, which tends to zero as $n \to \infty$.
  - Alternatively (though essentially equivalently), we could observe that the intervals $[C_{2n}, C_{2n-1}]$ form a set of nested closed intervals of lengths tending to zero, so by the nested intervals theorem[5], their intersection is a single point C equal to the limit of the sequence C_i.
- We can now establish some of the basic properties of infinite continued fractions.
- Proposition (Properties of Infinite Continued Fractions): Let $\alpha = [a_0, a_1, a_2, \dots]$ be an infinite simple continued fraction with nth convergent $C_n = [a_0, a_1, \dots, a_n] = p_n/q_n$. Then the following hold:

1. We have $|\alpha - C_n| \leq \dfrac{1}{q_n q_{n+1}} < \dfrac{1}{q_n^2}$.
   - Proof: The proof follows identically from the finite case done earlier, since α lies between C_n and C_{n+1}.

[4]The monotone convergence theorem says that any monotone increasing sequence that is bounded above (i.e., any sequence $a_1 < a_2 < a_3 < \cdots$ such that all terms are less than some finite number M) has a limit. By negating everything, it equivalently says that any monotone decreasing sequence bounded below has a limit.

[5]The nested intervals theorem says that if I_1, I_2, I_3, ... is an infinite sequence of nested closed intervals (i.e., where $I_{n+1} \subseteq I_n$ for each n) that are bounded, then the intersection $\cap_{n=1}^{\infty} I_n$ is also a closed interval. Furthermore, if the lengths of the intervals tend to zero, then the intersection consists of a single point. This result is a special case of Cantor's intersection theorem in $\mathbb{R}^n$, which says that the intersection of a nested sequence of compact sets is a nonempty compact set.

2. Any infinite continued fraction α is irrational. Furthermore, any two different irrational numbers have different infinite continued fraction expansions.
   - Proof: For the first statement, suppose $\alpha = p/q$ were rational. By the proposition above, we know that $0 < \left|\frac{p}{q} - \frac{p_n}{q_n}\right| < \frac{1}{q_n^2}$, meaning that $0 < |pq_n - p_n q| < \frac{q}{q_n}$.
   - However, $\frac{q}{q_n}$ goes to zero as $n \to \infty$, since q is fixed but q_n is a strictly increasing sequence. This is impossible, since if $q_n > q$ the expression $|pq_n - p_n q|$ would be an integer between 0 and 1.
   - For the second statement, first observe that $C_0 < \alpha < C_1$, meaning that $a_0 < \alpha < a_0 + \frac{1}{a_1}$, so we see that $\lfloor \alpha \rfloor = a_0$.
   - Next, observe that $\alpha = \lim_{n\to\infty} [a_0, a_1, \cdots, a_n] = \lim_{n\to\infty} \left(a_0 + \frac{1}{[a_1, \cdots, a_n]}\right) = a_0 + \frac{1}{[a_1, a_2, \cdots]}$.
   - Now suppose $\beta = [b_0, b_1, \cdots]$ and $\beta = \alpha$. By taking floors, we see that $b_0 = a_0$.
   - Then $[b_1, b_2, \cdots] = \frac{1}{\beta - b_0} = \frac{1}{\alpha - a_0} = [a_1, a_2, \cdots]$. Taking floors again shows $b_1 = a_1$.
   - Repeating the argument yields $b_i = a_i$ for every i, so α and β are identical.

- So far, we have discussed the ideas behind infinite continued fractions, but we have not actually computed any!
  - It is not hard, from the above, to work out the procedure for converting an irrational number α into an infinite continued fraction $[a_0, a_1, a_2, \dots]$.
  - First, we must have $a_0 = \lfloor \alpha \rfloor$, as we observed above.
  - Then, as we also observed, $[a_1, a_2, \dots] = \frac{1}{\alpha - a_0}$, so if we define $\alpha_1 = \frac{1}{\alpha - a_0}$ (which is greater than 1 because $0 < \alpha - a_0 < 1$ by irrationality of α and the definition of the floor function), we must have $a_1 = \lfloor \alpha_1 \rfloor$.
  - Now we repeat: we set $\alpha_2 = \frac{1}{\alpha_1 - a_1}$ and take $a_2 = \lfloor \alpha_2 \rfloor$.
  - In general, we obtain the terms recursively, via the relations $a_0 = \lfloor \alpha \rfloor$, $\alpha_i = \frac{1}{\alpha_{i-1} - a_{i-1}}$, and $a_i = \lfloor \alpha_i \rfloor$. As noted above, each of the a_i will be a positive integer, because α_i will always be greater than 1.
  - We should verify that the resulting continued fraction $[a_0, a_1, a_2, \dots]$ actually converges to α. To do this, we observe (essentially by the definition) that $\alpha = [a_0, a_1, \dots, a_n, \alpha_{n+1}]$, and then compute

$$\begin{aligned} |\alpha - [a_0, a_1, \dots, a_n]| &= \left|\frac{p_n \alpha_{n+1} + p_{n-1}}{q_n \alpha_{n+1} + q_{n-1}} - \frac{p_n}{q_n}\right| \\ &= \frac{1}{q_n(q_n \alpha_{n+1} + q_{n-1})} \\ &< \frac{1}{q_n q_{n-1}} \end{aligned}$$

  because α_{n+1} is positive. Since this tends to zero as $n \to \infty$, we see that α is indeed equal to $\lim_{n\to\infty} [a_0, a_1, \dots, a_n]$.

- Example: Find the continued fraction expansion of $\sqrt{2}$.
  - With $\alpha = \sqrt{2}$, we find, successively,

n	0	1	2	$\cdots$
α_n	$\sqrt{2}$	$\sqrt{2}+1$	$\sqrt{2}+1$	$\cdots$
a_n	1	2	2	$\cdots$
$\alpha_n - a_n$	$\sqrt{2}-1$	$\sqrt{2}-1$	$\sqrt{2}-1$	$\cdots$

  and since each term after this will repeat, we see that $\sqrt{2} = \boxed{[1, 2, 2, 2, 2, \dots]}$.

- Example: Find the continued fraction expansion of $\sqrt{7}$.
  - With $\alpha = \sqrt{7}$, we find, successively,

$$\begin{array}{cccccc} n & 0 & 1 & 2 & 3 & 4 & \cdots \\ \alpha_n & \sqrt{7} & \dfrac{\sqrt{7}+2}{3} & \dfrac{\sqrt{7}+1}{2} & \dfrac{\sqrt{7}+1}{3} & \sqrt{7}+2 & \cdots \\ a_n & 2 & 1 & 1 & 1 & 4 & \cdots \\ \alpha_n - a_n & \sqrt{7}-2 & \dfrac{\sqrt{7}-1}{3} & \dfrac{\sqrt{7}-1}{2} & \dfrac{\sqrt{7}-2}{3} & \sqrt{7}-2 & \cdots \end{array}$$

    and since each term after this will repeat, we see that $\sqrt{7} = \boxed{[2, 1, 1, 1, 4, 1, 1, 1, 4, \dots]}$.

- Example: Find the first ten terms of the continued fraction expansion of π.
  - With $\alpha = \pi$, we find, numerically, that the first ten terms are $[3, 7, 15, 1, 292, 1, 1, 1, 2, 1, \dots]$. This is easy to do even with a hand calculator: simply subtract off the integer part, reciprocate, and repeat.
  - There is no apparent pattern, and the sequence does not seem to repeat, in the nice way that the two previous examples did.

- Two of the continued fractions in the examples above eventually begin repeating. We will give this situation a special name:

- Definition: An infinite continued fraction $[a_0, a_1, a_2, \dots]$ is (eventually) periodic if there is some integer n such that $a_r = a_{n+r}$ for all sufficiently large r. We employ the notation $[a_0, a_1, a_2, \dots, a_k, \overline{a_{k+1}, a_{k+2}, \dots, a_{k+n}}]$ to indicate that the block of integers under the bar repeats indefinitely.
  - This is the same notation used for repeating decimals. (This is reasonable, since it is essentially the same situation, too.)

- Example: Find the real number $\alpha = [\overline{1}]$ and find its first ten convergents.
  - By the periodicity of the expansion, we know that $\alpha = 1 + \dfrac{1}{\alpha} = \dfrac{\alpha + 1}{\alpha}$.
  - This yields a quadratic equation for α, namely $\alpha^2 = \alpha + 1$, whose solutions are $\alpha = \dfrac{1 \pm \sqrt{5}}{2}$.
  - Since $\alpha > 1$, we need the plus sign, so $\alpha = \boxed{\dfrac{1+\sqrt{5}}{2}}$. (This is the famous golden ratio.)
  - We can compute the convergents explicitly: the first ten are 1, 2, $\dfrac{3}{2}$, $\dfrac{5}{3}$, $\dfrac{8}{5}$, $\dfrac{13}{8}$, $\dfrac{21}{13}$, $\dfrac{34}{21}$, $\dfrac{55}{34}$, and $\dfrac{89}{55}$.
  - Notice that these are simply ratios of consecutive Fibonacci numbers, which (once noticed) follows easily from the definition of α, since we have $\dfrac{p_{n+1}}{q_{n+1}} = 1 + \dfrac{1}{p_n/q_n} = \dfrac{p_n + q_n}{p_n}$, and so we see $q_{n+1} = p_n$ and $p_{n+1} = p_n + q_n = p_n + p_{n-1}$, which along with $p_1 = q_1 = 1$ is precisely the definition of the Fibonacci numbers.
  - Remark: In fact, our results about the convergence of the convergents provide a proof that $\lim\limits_{n\to\infty} \dfrac{F_{n+1}}{F_n} = \dfrac{1+\sqrt{5}}{2}$.

- Example: Find the real number $\alpha = [\overline{2, 5}]$ and find its first ten convergents.
  - By the periodicity of the expansion, we know that $\alpha = 2 + \dfrac{1}{5 + \dfrac{1}{\alpha}} = 2 + \dfrac{\alpha}{5\alpha + 1} = \dfrac{11\alpha + 2}{5\alpha + 1}$.
  - This yields a quadratic equation for α, namely $\alpha(5\alpha + 1) = 11\alpha + 2$, whose solutions are $\alpha = \dfrac{5 \pm \sqrt{35}}{5}$.
  - Since $\alpha > 1$, we need the plus sign, so $\alpha = \boxed{\dfrac{5+\sqrt{35}}{5}}$.

- The first ten convergents are 2, $\frac{11}{5}$, $\frac{24}{11}$, $\frac{131}{60}$, $\frac{286}{131}$, $\frac{1561}{715}$, $\frac{3408}{1561}$, $\frac{18601}{8520}$, $\frac{40610}{18601}$, $\frac{221651}{101525}$.

- So far, all of the periodic continued fractions we have seen have been solutions of a quadratic polynomial in $\mathbb{Q}[x]$. This is not an accident:

- Theorem (Periodic Continued Fractions): If α has a periodic continued fraction, then α is an irrational root of a quadratic polynomial with integer coefficients. (We call such a number a quadratic irrational.)

  - Proof: Let $\alpha = [a_0, a_1, a_2, \ldots, a_k, \overline{a_{k+1}, a_{k+2}, \ldots, a_{k+n}}]$, and $\gamma = [\overline{a_{k+1}, a_{k+1}, \ldots, a_{k+n}}]$.
  - Then by the periodicity of the expansion, we have $\gamma = [a_{k+1}, \ldots, a_{k+n}, \gamma]$.
  - Expanding this out yields $\gamma = \frac{p_{n-1}\gamma + p_{n-2}}{q_{n-1}\gamma + q_{n-2}}$, which is a quadratic equation for γ.
  - Since γ is irrational (being an infinite continued fraction), we conclude that $\gamma = \frac{b + \sqrt{c}}{d}$ for some integers b, c, and d.
  - Then $\alpha = [a_0, a_1, a_2, \ldots, a_k, \gamma]$ is also a rational function in γ (and irrational), so clearing the denominator shows that $\alpha = \frac{e + \sqrt{f}}{g}$ for some integers e, f, and g, which is also a root of a quadratic polynomial.

- The converse of this theorem is true also, but requires quite a bit more work. There are various approaches, but we will use an approach motivated by the arithmetic of $\mathbb{Q}(\sqrt{D})$.

- Definition: Let α be a quadratic irrational. The minimal polynomial $m(x)$ of α is the unique quadratic polynomial of which α is a root having the form $ax^2 + bx + c$ for relatively prime integers a, b, c where $a > 0$. We also define the discriminant of α to be the value $b^2 - 4ac \in \mathbb{Z}$.

  - In other settings, the minimal polynomial is assumed to be monic and have rational coefficients. We take integer coefficients in our definition here because we want to work with properties that rely on having integer coefficients rather than rational coefficients.
  - We will observe that because α is real and irrational, the discriminant of α is always positive, since it is the term under the square root in the quadratic formula for the roots of $m(x)$.
  - Example: The minimal polynomial of $\sqrt{2}$ is $x^2 - 2$, of discriminant 8.
  - Example: The minimal polynomial of the golden ratio $(1 + \sqrt{5})/2$ is $x^2 - x - 1$, of discriminant 5.
  - Example: The minimal polynomial of $(3 + \sqrt{13})/7$ is $49x^2 - 42x - 4$, of discriminant 2548.

- We have a few additional definitions:

- Definition: If $\alpha = \frac{p + \sqrt{D}}{q}$ is a quadratic irrational, then the other root of its minimal polynomial is its conjugate $\overline{\alpha} = \frac{p - \sqrt{D}}{q}$. We say that α is reduced if $\alpha > 1$ and also $-1/\overline{\alpha} > 1$.

  - In general, the minimal polynomial of α will be $q^2(x-\alpha)(x-\overline{\alpha}) = q^2x^2 - 2pqx + (p^2 - D)$ up to scaling by a divisor of q (the coefficients need not be relatively prime, since $p^2 - D$ could have a factor in common with q).
  - Example: The conjugate of $\alpha = \sqrt{2}$ is $\overline{\alpha} = -\sqrt{2}$, so α is not reduced since $-1/\overline{\alpha}$ is not greater than 1.
  - Example: The conjugate of $\alpha = (1+\sqrt{5})/2$ is $\overline{\alpha} = (1-\sqrt{5})/2$, so α is reduced since both α and $-1/\overline{\alpha} = \alpha$ are greater than 1.

- Now we can prove that every quadratic irrational has a periodic continued fraction expansion:

- Theorem (Quadratic Irrationals and Continued Fractions): Let α be a quadratic irrational with discriminant D, and let α_n be the nth remainder term obtained in computing the continued fraction expansion of α, so that $\alpha_0 = \alpha$ and $\alpha_n = \frac{1}{\alpha_{n-1} - \lfloor \alpha_{n-1} \rfloor}$ for all $n \geq 1$. Then the following hold:

  1. The remainder term α_n has discriminant D for all $n \geq 1$.

- Proof: We first show α_1 has discriminant D, so suppose α has minimal polynomial $m(x) = ax^2+bx+c$ and write $\lfloor \alpha \rfloor = a_0$.
- Since $\alpha = a_0 + 1/\alpha_1$ this means $a(a_0+1/\alpha_1)^2 + b(a_0+1/\alpha_1) + c = 0$, whence $a(a_0\alpha_1+1)^2 + b(a_0\alpha_1 + 1)\alpha_1 + c(\alpha_1)^2 = 0$; equivalently, $(aa_0^2 + ba_0 + c)\alpha_1^2 + (2aa_0 + b)\alpha_1 + a = 0$.
- Since a, b, c are relatively prime, so are $aa_0^2 + ba_0 + c$, $2aa_0 + b$, and a.
- Thus up to sign, the minimal polynomial of α_1 is $(aa_0^2 + ba_0 + c)x + (2aa_0 + b)x + a$, and so its discriminant is $(2aa_0 + b)^2 - 4a(aa_0^2 + ba_0 + c) = b^2 - 4ac = D$, as claimed.
- The desired result then follows by a trivial induction on n.

2. If α is a reduced quadratic irrational, then α_n is also reduced.
   - Proof: As in (1) we show that if α is reduced then α_1 is reduced, and then apply a trivial induction.
   - If α is reduced, then $\alpha_1 = \dfrac{1}{\alpha - \lfloor \alpha \rfloor} > 1$ since $0 < \alpha - \lfloor \alpha \rfloor < 1$.
   - Also, $\overline{\alpha}_1 = \dfrac{1}{\overline{\alpha} - \lfloor \alpha \rfloor}$ is negative because $\overline{\alpha}$ is negative, and its absolute value is between 0 and 1 because $\lfloor \alpha \rfloor \geq 1$. Thus, $-1/\overline{\alpha}_1 > 1$ as required, and so α_1 is reduced.
3. There are only finitely many reduced quadratic irrationals of discriminant D.
   - Proof: Suppose α is a reduced quadratic irrational of discriminant D and minimal polynomial $m(x) = ax^2 + bx + c$, where $b^2 - 4ac = D$ and $a > 0$.
   - Since $\alpha = \dfrac{-b + \sqrt{D}}{2a}$ is reduced, we have $-1/\overline{\alpha} > 1$ and so $-1 < \overline{\alpha} < 0$. Thus $\alpha + \overline{\alpha} = -b/a$ is positive, so since $a > 0$ that means $b < 0$.
   - Furthermore, $\overline{\alpha} = \dfrac{-b - \sqrt{D}}{2a}$ and $a > 0$, this requires $-b - \sqrt{D} < 0$ and so $b > -\sqrt{D}$. Thus $-\sqrt{D} < b < 0$ and so there are finitely many possible b.
   - But then since $\alpha = \dfrac{-b + \sqrt{D}}{2a}$ must have $\alpha > 1$, we see that $a < -b + \sqrt{D} < 2\sqrt{D}$. Since a is positive, there are finitely many possible a.
   - Then, finally, since $c = (b^2 - D)/(4a)$, there are finitely many possible triples (a, b, c) and thus finitely many possible α.
4. The remainder term α_n is reduced for sufficiently large n.
   - Proof: By definition, for any $n \geq 1$, we have $\alpha_n = \dfrac{1}{\alpha_{n-1} - \lfloor \alpha_{n-1} \rfloor} > 1$. It remains to obtain a bound on $-1/\overline{\alpha_n}$.
   - First, by definition we have $\alpha = [a_0, a_1, \ldots, a_n, \alpha_n]$, so if we set $[a_0, a_1, \ldots, a_n] = p_n/q_n$, then so that $\alpha = \dfrac{p_n\alpha_n + p_{n-1}}{q_n\alpha_n + q_{n-1}}$.
   - Conjugating yields $\overline{\alpha} = \dfrac{p_n\overline{\alpha}_n + p_{n-1}}{q_n\overline{\alpha}_n + q_{n-1}}$ since the p_i and q_i are integers hence unchanged by conjugating.
   - Rearranging this last expression gives $-\dfrac{1}{\overline{\alpha}_n} = -\dfrac{q_n\overline{\alpha} - p_n}{q_{n-1}\overline{\alpha} - p_{n-1}} = \dfrac{q_n}{q_{n-1}} \cdot \dfrac{\overline{\alpha} - p_n/q_n}{\overline{\alpha} - p_{n-1}/q_{n-1}}$.
   - For large n, as we have shown, $p_n/q_n \to \alpha$, and thus the second term approaches $\dfrac{\overline{\alpha} - \alpha}{\overline{\alpha} - \alpha} = 1$ (note that the denominator is nonzero because α is irrational). The first term q_n/q_{n-1} is always greater than 1, and its limit cannot equal 1 because $q_n \geq q_{n-1} + q_{n-2}$, so dividing by q_{n-1} and taking the limit would give $1 \geq 1 + 1$, impossible.
   - Therefore, for sufficiently large n, we see $-1/\overline{\alpha}_n > 1$, and so α_n is reduced.
5. The continued fraction expansion of a real number α is periodic if and only if α is a quadratic irrational.
   - Proof: We proved earlier that if α has a periodic continued fraction expansion, then α is a quadratic irrational.
   - For the converse, suppose α is a quadratic irrational of discriminant D. Then by (1), every remainder term in the continued fraction expansion of α has discriminant D.

- By (4), the nth remainder term is reduced for sufficiently large n. But by (3), there are only finitely many such remainder terms, so by the pigeonhole principle there must be at least one repetition somewhere.
- But once a remainder term repeats, the rest of the expansion will be the same, and so the expansion is periodic, as claimed.

6. The continued fraction expansion of a real number α is purely periodic (i.e., is of the form $\alpha = [\overline{a_0, a_1, \ldots, a_n}]$) if and only if α is a reduced quadratic irrational.
   - Proof: First suppose α has a purely periodic expansion. Then $\alpha = [a_0, a_1, \ldots, a_{kn}, \alpha]$ for every positive integer k. Since by (4) the remainders are eventually all reduced, this means α must be reduced.
   - Conversely, suppose α is reduced. By (5) we know that the continued fraction expansion is eventually periodic, say with $\alpha_{k+n} = \alpha_k$ for some k and n.
   - We first show that $\alpha_{k+n-1} = \alpha_{k-1}$, so suppose α is reduced and $\alpha_{k+n} = \alpha_k$. Then both α_{k+n} and α_k are reduced by (2). By definition we have $\alpha_{k+n} = \dfrac{1}{\alpha_{k+n-1} - a_{k+n-1}}$ and $\alpha_n = \dfrac{1}{\alpha_{n-1} - a_{n-1}}$, so conjugating and inverting yields $-\dfrac{1}{\overline{\alpha}_{n+k}} = a_{k+n-1} - \overline{\alpha}_{k+n-1}$ and $-\dfrac{1}{\overline{\alpha}_n} = a_{n-1} - \overline{\alpha}_{n-1}$.
   - Since both $\overline{\alpha}_{k+n-1}$ and $\overline{\alpha}_{n-1}$ are between -1 and 0, we see $a_{k+n-1} = \lfloor -\dfrac{1}{\overline{\alpha}_{n+k}} \rfloor = \lfloor -\dfrac{1}{\overline{\alpha}_n} \rfloor = a_{n-1}$, as claimed.
   - By iterating this fact (equivalently, by a trivial induction), this implies $\alpha_{j+n} = \alpha_j$ for all $j \geq 0$.
   - Then we see immediately that α has a periodic continued fraction expansion, as $a_{j+n} = \lfloor \alpha_{j+n} \rfloor = \lfloor \alpha_j \rfloor = a_j$ for all $j \geq 0$.

- Example: Find the continued fraction expansion of $(3 + \sqrt{13})/4$.
  - Notice here that $\alpha = (3 + \sqrt{13})/4 > 1$ has $-1/\overline{\alpha} = -4/(3 - \sqrt{13}) = 3 + \sqrt{13} > 1$, so α is reduced. Per (6) above, its continued fraction expansion will be purely periodic.
  - With $\alpha = (3 + \sqrt{13})/4$, we find, successively,

n	0	1	2	3	4	5
α_n	$(3+\sqrt{13})/4$	$(1+\sqrt{13})/3$	$(2+\sqrt{13})/3$	$(1+\sqrt{13})/4$	$3+\sqrt{13}$	$(3+\sqrt{13})/4$
a_n	1	1	1	1	6	
$\alpha_n - a_n$	$(-1+\sqrt{13})/4$	$(-2+\sqrt{13})/3$	$(-1+\sqrt{13})/3$	$(-3+\sqrt{13})/4$	$-3+\sqrt{13}$	

  and we can see at this point each term will repeat. Therefore, the continued fraction expansion is $\boxed{[\overline{1, 1, 1, 1, 6}]}$, which is indeed periodic.

6.2.4 Rational Approximation Via Continued Fractions

- One of our main goals in discussing continued fractions was to use them to give rational approximations. Here are some results in this direction:

- Proposition (Rational Approximation and Continued Fractions): Suppose α is any irrational real number and p/q is any rational number. Then the following hold:

1. If p_n/q_n is the nth continued fraction convergent to α, and $\left|\alpha - \dfrac{p}{q}\right| < \left|\alpha - \dfrac{p_n}{q_n}\right|$, then $q > q_n$. In fact, if $|q\alpha - p| < |q_n\alpha - p_n|$, then $q \geq q_{n+1}$.
   - Observe that the first statement says that the best rational approximation to α, among all terms in the Farey sequence of level q_n, is the convergent p_n/q_n.
   - Proof: Consider the Farey sequence of level q_n: since $|p_{n-1}q_n - p_nq_{n-1}| = 1$, we see that $\dfrac{p_{n-1}}{q_{n-1}}$ and $\dfrac{p_n}{q_n}$ are consecutive in this sequence.

- Hence, there is no rational number with denominator less than q_n that lies between them.
- For the second statement, suppose that $q < q_{n+1}$. By basic linear algebra, there exist integers x and y such that $p = xp_n + yp_{n+1}$ and $q = xq_n + yq_{n+1}$. (They are integers because the determinant $p_nq_{n+1} - p_{n+1}q_n$ of the associated coefficient matrix is ± 1 by our results on the convergents of the continued fraction.)
- Notice that since $q < q_{n+1}$, the second equation requires that one of x, y be positive and the other is negative. Since $\alpha - \dfrac{p_n}{q_n}$ and $\alpha - \dfrac{p_{n+1}}{q_{n+1}}$ also have opposite signs, we conclude that $x\left(\alpha - \dfrac{p_n}{q_n}\right)$ and $y\left(\alpha - \dfrac{p_{n+1}}{q_{n+1}}\right)$ have the same sign.
- Then we can write

$$\begin{aligned} |q\alpha - p| &= |(xq_n + yq_{n+1})\alpha - (xp_n + yp_{n+1})| \\ &= |x(q_n\alpha - p_n) + y(q_{n+1}\alpha - p_{n+1})| \\ &= |x| \cdot |q_n\alpha - p_n| + |y| \cdot |q_{n+1}\alpha - p_{n+1}| \\ &\geq |q_n\alpha - p_n| \end{aligned}$$

  which establishes the contrapositive of the desired result.

2. There are infinitely many distinct rational numbers $\dfrac{p}{q}$ such that $\left|\alpha - \dfrac{p}{q}\right| < \dfrac{1}{2q^2}$.
   - Remark: The constant 2 is not sharp. In fact, it is a theorem of Hurwitz that the 2 can be replaced with $\sqrt{5}$, but with no larger constant.
   - Proof: As usual let $\dfrac{p_n}{q_n}$ be the nth continued fraction convergent to α.
   - We claim that at least one of $\dfrac{p_n}{q_n}$ and $\dfrac{p_{n+1}}{q_{n+1}}$ satisfies the desired inequality, so suppose that neither does.
   - Then since α lies between $\dfrac{p_n}{q_n}$ and $\dfrac{p_{n+1}}{q_{n+1}}$, we have

$$\begin{aligned} \left|\frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}}\right|^2 &= \left(\left|\frac{p_n}{q_n} - \alpha\right| + \left|\frac{p_{n+1}}{q_{n+1}} - \alpha\right|\right)^2 \\ &> 4\left|\frac{p_n}{q_n} - \alpha\right| \cdot \left|\frac{p_{n+1}}{q_{n+1}} - \alpha\right| \\ &\geq 4 \cdot \frac{1}{2q_n^2} \cdot \frac{1}{2q_{n+1}^2} = \frac{1}{q_n^2 q_{n+1}^2}. \end{aligned}$$

     where in the middle step we used the inequality $(x+y)^2 \geq 4xy$ (which is equivalent to $(x-y)^2 \geq 0$, and equality cannot hold in our case because α is irrational).
   - Taking the square root gives $\left|\dfrac{p_n}{q_n} - \dfrac{p_{n+1}}{q_{n+1}}\right| > \dfrac{1}{q_nq_{n+1}}$, but this is false since these quantities are equal.

3. If $\dfrac{p}{q}$ is a rational number such that $\left|\alpha - \dfrac{p}{q}\right| < \dfrac{1}{2q^2}$, then in fact $\dfrac{p}{q}$ is a continued fraction convergent to α.
   - Proof: Suppose by way of contradiction that p/q is not a convergent and that $\left|\alpha - \dfrac{p}{q}\right| < \dfrac{1}{2q^2}$.
   - Let n be such that $q_n \leq q < q_{n+1}$.
   - By (1), it must be the case that $|q_n\alpha - p_n| < |q\alpha - p| = q\left|\alpha - \dfrac{p}{q}\right| < \dfrac{1}{2q}$.
   - Thus, we conclude that $\left|\alpha - \dfrac{p_n}{q_n}\right| < \dfrac{1}{2qq_n}$.
   - Now we get $\dfrac{1}{qq_n} \leq \left|\dfrac{p_nq - pq_n}{qq_n}\right| = \left|\dfrac{p}{q} - \dfrac{p_n}{q_n}\right| \leq \left|\dfrac{p}{q} - \alpha\right| + \left|\alpha - \dfrac{p_n}{q_n}\right| < \dfrac{1}{2qq_n} + \dfrac{1}{2q_n^2}$.
   - But this implies $q < q_n$, which is a contradiction.

- We will remark at this point that the continued fraction expansion provides a good way to detect approximations of rational numbers using a decimal expansion: simply compute the continued fraction of the decimal, and then round off appropriately.
- Example: Find a rational number of small denominator with decimal expansion $0.4614379084967\ldots$.
  - We compute the continued fraction expansion of $\alpha = 0.4614379084967$, which is easy to do with a calculator or computer.
  - We obtain the exact expression $\alpha = [0, 2, 5, 1, 57, 1, 53354674, 4, 1, 1, 6, 4]$.
  - We truncate just before the huge term in the middle to get a guess of $\alpha = [0, 2, 5, 1, 57, 1] = \dfrac{353}{765}$. Indeed, we can calculate that $\dfrac{353}{765} \approx 0.461437908496732$.
  - From our results above, we can see that any rational number that is a closer approximation will have denominator roughly on order of the next convergent $[0, 2, 5, 1, 57, 1, 53354674] = \dfrac{18834200269}{40816326362}$, so the rational number we found is clearly the simplest.
  - It is interesting to note that the period of the decimal expansion of $\dfrac{353}{765}$ is 16, so in fact we have identified the rational number before the expansion started repeating!
- We can also use these results, along with some of our facts about the Farey sequences to find the best rational approximation to a given real number α having a denominator below a given fixed bound N.
  - Our starting point is to calculate the last two convergents p_{n-1}/q_{n-1} and p_n/q_n whose denominators are less than N.
  - Since the convergents alternate being above and below α, this means α lies between these two convergents. Furthermore, from the relation $|p_n q_{n-1} - p_{n-1} q_n| = 1$ and our results on the Farey sequences, we see that p_{n-1}/q_{n-1} and p_n/q_n are consecutive terms in the Farey sequence of level q_n.
  - We can then generate all of the terms between these of level $\leq N$ by taking mediants, and from this short list we can identify the best approximation of α.
- Example: Find the rational number with denominator less than 100 that is closest to $\sqrt{7}$.
  - Earlier, we computed the continued fraction expansion $\sqrt{7} = [2, \overline{1, 1, 1, 4}]$.
  - The first few convergents are then 2, 3, 5/2, 8/3, 37/14, 45/17, 82/31, 127/48, 590/223.
  - The last two convergents with denominator less than 100 are 82/31 and 127/48. The only term between them in the Farey sequence of level 99 is their mediant, 209/79.
  - We can then compute that $\sqrt{7} - 82/31 \approx 5.9 \cdot 10^{-4}$, $\sqrt{7} - 127/48 \approx -8.2 \cdot 10^{-5}$, and $\sqrt{7} - 209/79 \approx 1.8 \cdot 10^{-4}$. Thus, the best approximation is $\boxed{127/48}$.
- In many situations the best approximation of α will be one of its continued fraction convergents, but this need not always be the case:
- Example: Find the rational number with denominator less than 10 that is closest to $\sqrt{7}$.
  - From above, the last two convergents with denominator less than 10 are 5/2 and 8/3. The terms between them in the Farey sequence of level 9 are 18/7, 13/5, and 21/8.
  - We can then compute that $\sqrt{7} - 5/2 \approx 0.1458$, $\sqrt{7} - 18/7 \approx 0.0743$, $\sqrt{7} - 13/5 \approx 0.0458$, $\sqrt{7} - 21/8 \approx 0.0208$, and $\sqrt{7} - 8/3 \approx -0.0209$.
  - Thus, the best approximation (by a quite small margin!) is $\boxed{21/8}$, which, we remark, is not a continued fraction convergent to $\sqrt{7}$.

6.2.5 Irrationality and Transcendence

- We can also use some of these properties of rational approximation we have developed to prove the irrationality of various quantities, and by suitably extending these results, we can even prove transcendence in some cases.
- One easy observation is that the continued fraction expansion of a real number α terminates in a finite number of steps if and only if α is rational.
    - Thus, taking the contrapositive shows that the continued fraction of α is infinite if and only if α is irrational. We could therefore establish irrationality by computing the continued fraction expansion of a given real number.
    - However, as a practical matter, this is not so easy to do. The easiest infinite continued fractions to compute are the periodic expansions, and as we proved, these are the expansions of quadratic irrationals. However, these are quite easy to prove irrational, since their irrationality is ultimately equivalent to the irrationality of $\sqrt{D}$ where D is a squarefree integer not equal to 1.
    - If we flip our approach around, we can say, for example, that the real number with continued fraction expansion $[1, 2, 3, 4, 5, 6, \dots]$ is irrational. However, we have no simple way of giving a closed-form formula for this real number. (As it happens, this number can be written in terms of values of a modified Bessel function, though this is not so easy to prove.)
- Our second method is to use another of our earlier results: as we showed, if α is irrational, then there are infinitely many distinct rational numbers p/q such that $|\alpha - p/q| < 1/q^2$.
    - Our main idea is that the converse of this statement holds as well:
- Proposition (Irrationality and Approximation): A real number α is irrational if and only if there exist infinitely many distinct rational numbers p/q such that $|\alpha - p/q| < 1/q^2$.
    - Proof: We established the forward direction earlier, so now suppose $\alpha = a/b$ is a fixed rational number.
    - Then $|\alpha - p/q| = |aq - bp| / (bq)$. If $q \leq b$ then there are only finitely many possible p/q with $|\alpha - p/q| < 1/q^2$ since there are only finitely many possible denominators q and finitely many p that work for any given q.
    - If $q > b$ then we would have $|aq - bp| / (bq) < 1/q^2$ so that $|aq - bp| < b/q < 1$. But since $|aq - bp|$ is an integer, it would then have to be zero, in which case p/q would equal a/b.
    - Putting these two cases together shows that if α is rational, then there are only finitely many distinct rational numbers p/q such that $|\alpha - p/q| < 1/q^2$, as claimed.
- In principle, we could try to use this result to establish the irrationality of an arbitrary irrational number. However, this can be quite cumbersome in practice.
    - The numbers for which it will be effective are those that we can write as an infinite sum of rational numbers whose terms decrease rapidly in size: we can then obtain the desired rational approximations by taking partial sums of the series.
    - As long as the tail of the series is very small (i.e., less than $1/q^2$) relative to the denominator q of each partial sum, we will be able to conclude that the sum of the series is irrational.
- Example: Show that $\alpha = \sum_{k=0}^{\infty} 10^{-3^k}$ is irrational.
    - Let $p_n/q_n = \sum_{k=0}^{n} 10^{-3^k}$ be the nth partial sum of the series. We observe that $q_n = 10^{3^k}$ since each of the other terms has a denominator dividing 10^{3^k}.
    - Furthermore, it is easy to see (e.g., from the decimal expansion of α) that the size of the tail $\sum_{k=n+1}^{\infty} 10^{-3^k}$ is at most $2 \cdot 10^{-3^{k+1}}$.
    - Then we have an easy bound $|\alpha - p_n/q_n| < 2 \cdot 10^{-3^{k+1}} < (10^{-3^k})^2 = 1/q_n^2$. Since all of the partial sums of this series are distinct, we obtain infinitely many such p_n/q_n, and therefore by our result above, α is irrational.

- As first observed by Liouville, we can extend this criterion to exclude algebraic numbers that are roots of higher-degree polynomials by increasing the exponent of q in the bound on the right-hand side.
  - We say that a number $\alpha \in \mathbb{C}$ is algebraic if α is the root of some nonzero polynomial $p(x)$ with rational coefficients.
  - If we consider all of the possible polynomials in $\mathbb{Q}[x]$ of which α is a root, by the well-ordering principle we can see that there is some polynomial of minimal degree d of which α is a root.
  - We refer to this degree d as the algebraic degree of α over $\mathbb{Q}$. There is a unique monic polynomial of this degree d of which α is a root; this polynomial is called the minimal polynomial of α over $\mathbb{Q}$.
  - Example: Quadratic irrationals have algebraic degree 2 over $\mathbb{Q}$, since they are roots of quadratic polynomials but not any polynomial of lower degree.
  - Example: The number $\sqrt[4]{2}$ has minimal polynomial $x^4 - 2$ over $\mathbb{Q}$ (although this is not completely trivial to prove) and thus has algebraic degree 4.
  - We will remark that the minimal polynomial is always irreducible (if it had a factorization, whichever factor had α as a root would have smaller degree) and cannot have any repeated roots (if it did, then m and its derivative m' would have a factor $x - \alpha$ in common).
  - Suppose α is algebraic. We may clear denominators in its minimal polynomial to see that α is the root of some polynomial $c_d x^d + c_{d-1} x^{d-1} + \cdots + c_0$ where the c_i are integers: this means $c_d \alpha^d + c_{d-1} \alpha^{d-1} + \cdots + c_0 = 0$.
  - If we then set $\beta = c_d \alpha$, by rescaling we can see that β is a root of the polynomial $x^d + c_{d-1} c_d x^{d-1} + \cdots + c_0 c_d^{d-1}$, which is monic and has integer coefficients.
  - Thus, up to an integer factor, any algebraic number is the root of a monic polynomial with integer coefficients.
- With these preliminaries finished, we can now give Liouville's result:
- Theorem (Liouville's Approximation Theorem): Suppose α is algebraic of degree $n > 1$ over $\mathbb{Q}$. Then there exists a positive real number A such that $|\alpha - p/q| \geq A/q^n$ for any rational number p/q.
  - The idea of the proof is to rescale α so its minimal polynomial has integer coefficients: then $m(p/q)$ is $1/q^n$ times an integer while $m(\alpha) = 0$ and so the mean value theorem will allow us to bound $|\alpha - p/q|$ from below.
  - Proof: First, by rescaling α by an appropriate positive integer, we may assume that the minimal polynomial of α has integer coefficients. This does not affect the result that needs to be shown, since if $|(N\alpha) - p/q| \geq A/q^n$ then $|\alpha - p/(Nq)| \geq (A/N)/q^n$ so taking $A' = A/N$ yields the necessary bound for α.
  - So now suppose α is algebraic of degree $n > 1$ over $\mathbb{Q}$ and that its minimal polynomial $m(x)$ has integer coefficients and factors as $m(x) = (x - \alpha)(x - \beta_1)(x - \beta_2) \cdots (x - \beta_{n-1})$ over $\mathbb{C}$. Note that the β_i are distinct from α because $m(x)$ cannot have repeated roots.
  - Now define M be the maximum value of $|m'(x)|$ on the interval $[\alpha - 1, \alpha + 1]$, and set $A = \min(1, 1/M, |\alpha - \beta_i|)$ over all of the roots β_i. We claim this value of A satisfies the given inequality.
  - To show this, suppose otherwise, so that p/q is rational and has $|\alpha - p/q| < A/q^n$. Then because $A \leq 1$, we have $p/q \in (\alpha - 1, \alpha + 1)$.
  - Also, because $A \leq |\alpha - \beta_i|$, we see that p/q cannot equal any of the β_i, and that there is no root of $m(x)$ between α and p/q.
  - If we write $m(x) = x^d + c_{d-1} x^{d-1} + \cdots + c_0$, then $m(p/q) = (p/q)^d + c_{d-1}(p/q)^{d-1} + \cdots + c_0 = (1/q^d) \cdot [p^d + c_{d-1} p^{d-1} q + \cdots + c_0 q^d]$.
  - Thus we have $|m(p/q)| \geq 1/q^d \cdot \left|p^d + c_{d-1} p^{d-1} q + \cdots + c_0 q^d\right| \geq 1/q^d$ because the term inside the absolute values is an integer and it cannot be zero since $m(p/q) \neq 0$.
  - Now, by the mean value theorem, there exists x_0 in the interval with endpoints p/q and α such that $m(\alpha) - m(p/q) = m'(x_0) \cdot (\alpha - p/q)$; rearranging yields $|m(\alpha) - m(p/q)| \,/\, |m'(x_0)| = |\alpha - p/q|$.
  - By assumption we have $1/|m'(x_0)| \geq M \geq A$, and also $m(\alpha) = 0$ and $|m(p/q)| \geq 1/q^d$. Plugging all of these in immediately yields the desired inequality $|\alpha - p/q| = |m(p/q)| \,/\, |m'(x_0)| \geq A/q^d$, as claimed.

- Using Liouville's theorem, we can give explicit constructions of transcendental numbers similar to the one given earlier.
  - The idea is that if α is an irrational real number such that there exists a constant $c > 0$ and a sequence of rational numbers p_n/q_n such that $|\alpha - p_n/q_n| < c/q_n^n$, then α is transcendental.
  - The point is that this sequence of rational numbers contradicts the assertion that α is algebraic of degree n for every n, by the theorem above, and so α must be transcendental.
  - We can construct such an α and corresponding rational approximations p_n/q_n by taking α to be an infinite series whose terms drop in size very quickly: instead of the example above yielding irrationality, now we want the tail after the nth partial sum p_n/q_n to be on the order of $1/q_n^n$ rather than $1/q_n^2$.
- Example: Show that $\alpha = \sum_{k=0}^{\infty} 10^{-k!}$ is transcendental.
  - Let $p_n/q_n = \sum_{k=0}^{n} 10^{-k!}$ be the nth partial sum of the series. We observe that $q_n = 10^{k!}$ since each of the other terms has a denominator dividing $10^{-k!}$.
  - Furthermore, it is easy to see (e.g., from the decimal expansion of α) that the size of the tail $\sum_{k=n+1}^{\infty} 10^{-k!}$ is at most $2 \cdot 10^{-(n+1)!}$.
  - Then we have an easy bound $|\alpha - p_n/q_n| < 2 \cdot 10^{-(n+1)!} = 2(10^{-n!})^{n+1} = 2/q_n^{n+1} < 1/q_n^n$. Since all of the partial sums of this series are distinct, we obtain infinitely many such p_n/q_n, and therefore by our result above, α is transcendental.

6.3 Pell's Equation

- Equations of the form $x^2 - Dy^2 = r$, for D a positive squarefree integer and r an arbitrary integer, are often referred to under the general heading of Pell's equation, named after the English mathematician Pell.
  - However, this name is a misattribution by Euler, and it is quite possible that Pell never actually studied these equations. Equations of this type have been studied throughout history, with notable early contributions made by the Indian scholars Brahmaguptra, Bhaskara II, and Narayana. Certain instances of Pell's equation (most notably $D = 2$) were also studied by the ancient Greeks, including Diophantus.
  - What we would like to be able to do is find a recipe for generating solutions to Pell's equation in the situations that they do exist, and to understand more about the structures of these solutions. The general approach we will follow is similar to the treatment developed by Lagrange in the mid-1700s.

6.3.1 Motivation and Small Examples

- Let us start by exploring the case $D = 2$ for various small r: thus, we are seeking integer solutions to the Diophantine equation $x^2 - 2y^2 = r$ for small values of r.
  - We can do a search by plugging in small nonnegative values of x and y from 0 to 20 and looking for pairs where $x^2 - 2y^2$ is close to zero. Collecting them via the value of r yields the following solutions:

r	1	2	3	4	5	6	7
(x, y)	$(1,0)$,$(3,2)$,$(17,12)$	$(2,1)$,$(10,7)$	none	$(2,0)$,$(6,4)$	none	none	$(3,1)$,$(5,3)$,$(13,9)$
r	-1	-2	-3	-4	-5	-6	-7
(x, y)	$(1,1)$,$(7,5)$	$(0,1)$,$(4,3)$	none	$(2,2)$,$(14,10)$	none	none	$(1,2)$,$(5,4)$,$(11,8)$

  - We can see that for some values of r (namely, $r = \pm 3$) there seem to be no solutions, while for other small values of r there are solutions.
  - By working mod 8, we can show that there are no solutions to $x^2 - 2y^2 = r$ when r is congruent to 3 or 5 modulo 8 (which includes the cases $r = \pm 3$ and ± 5 above). Explicitly, this follows because $x^2 \in \{0, 1, 4\}$ mod 8 and $-2y^2 \in \{0, 6\}$ mod 8, and so $x^2 - 2y^2 \in \{0, 1, 2, 4, 6, 7\}$ mod 8, meaning that it cannot be congruent to 3 or 5 mod 8.
  - By working mod 3, we can also show that there are no solutions to $x^2 - 2y^2 = r$ when r is congruent to 3 or 6 modulo 9 (which includes the cases $r = \pm 3$ and ± 6 above). Explicitly, this follows because $x^2 - 2y^2 \equiv x^2 + y^2$ mod 3, and this can equal zero mod 3 only when $(x, y) \equiv (0, 0)$ mod 3, but in such cases $x^2 - 2y^2 = r$ is divisible by 9, and thus cannot be 3 or 6 modulo 9.

- Another pattern we can observe from the examples above is that $x^2 - 2y^2 = r$ seems to have a solution if and only if $x^2 - 2y^2 = -r$ does, and that some of the solutions seem to be related.
- If we look for a relationship between the pairs $(1, 1)$, $(7, 5)$ from $r = -1$ and $(3, 2)$, $(17, 12)$ from $r = +1$, along with corresponding pairs for the other r, it is not hard to spot that if (a, b) is a solution with $a^2 - 2b^2 = -r$, then $(a + 2b, a + b)$ seems to be a solution with $a^2 - 2b^2 = r$.
- Indeed, this is true: if $a^2 - 2b^2 = -r$, then $(a + 2b)^2 - 2(a + b)^2 = -(a^2 - 2b^2) = r$.
- We can see quite easily that if we start with any solution to $x^2 - 2y^2 = r$ other than the "trivial" solutions $(\pm 1, 0)$ to $x^2 - 2y^2 = 1$, we can generate new solutions to $x^2 - 2y^2 = \pm r$ by applying this rule mapping $(a, b) \mapsto (a + 2b, a + b)$.
- For example, starting with $(1, 0)$ we obtain $(1, 0) \mapsto (1, 1) \mapsto (3, 2) \mapsto (7, 5) \mapsto (17, 12) \mapsto (41, 29) \mapsto (99, 70) \mapsto (239, 169) \mapsto \cdots$. The odd terms in the sequence are solutions to $x^2 - 2y^2 = 1$ while the even terms are solutions to $x^2 - 2y^2 = -1$.
- If we iterate the rule twice, mapping $(a, b) \mapsto (a + 2b, a + b) \mapsto (3a + 4b, 2a + 3b)$, we obtain a recipe for generating new solutions to $x^2 - 2y^2 = r$ from old solutions.

- Let's also examine the case $D = 3$ for small r: now we are seeking integer solutions to $x^2 - 3y^2 = r$.
  - We can do a search by plugging in small nonnegative values of x and y from 0 to 40 and looking for pairs where $x^2 - 2y^2$ is close to zero. Collecting them via the value of r yields the following solutions:

r	1	2	3	4	5	6	7
(x, y)	$(1, 0)$,$(2, 1)$,$(7, 4)$,$(26, 15)$	none	none	$(2, 0)$,$(4, 2)$,$(14, 8)$	none	$(3, 1)$,$(9, 5)$,$(33, 19)$	none
r	-1	-2	-3	-4	-5	-6	-7
(x, y)	none	$(1, 1)$,$(5, 3)$,$(19, 11)$	$(0, 1)$,$(3, 2)$,$(12, 7)$	none	none	none	none

  - By working mod 3, we can see that there are no solutions to $x^2 - 3y^2 = r$ when r is congruent to 2 modulo 3 (which includes the cases $r = -7, -4, -1, 2, 5$ above). This follows because $r \equiv x^2 \pmod 3$, which has no solution when $r \equiv 2 \pmod 3$ since 2 is not a quadratic residue.
  - We can use these results to see that there are also no solutions in some other cases: for example, if $r \equiv 3$ mod 9 then if we had a solution to $x^2 - 3y^2 = r$ then x would be divisible by 3. If $x = 3a$ then cancelling the factor of 3 yields $3a^2 - y^2 = (r/3)$ so that $y^2 - 3a^2 = -(r/3) \equiv 2 \bmod 3$, but as we just showed above, there is no solution to this Diophantine equation.
  - Also here, quite unlike in the case with $D = 2$, it seems that if there is a solution to $x^2 - 3y^2 = r$ then there is no solution to $x^2 - 3y^2 = -r$.
  - If we search for a recipe (like in the case with $D = 2$) to generate new solutions to $x^2 - 3y^2 = 1$ from old ones, we can eventually stumble upon the map $(a, b) \mapsto (2a + 3b, a + 2b)$, which maps $(1, 0) \mapsto (2, 1) \mapsto (7, 4) \mapsto (26, 15) \mapsto (97, 56) \mapsto (362, 209) \mapsto \cdots$.

- We can explain many of the patterns witnessed above by using properties of the ring $\mathbb{Z}[\sqrt{D}] = \{a + b\sqrt{D} : a, b \in \mathbb{Z}\}$ and the associated norm map $N(a + b\sqrt{D}) = a^2 - Db^2$.
  - The key observation is to notice that solving the equation $x^2 - Dy^2 = r$ is equivalent to solving $N(x + y\sqrt{D}) = r$.
  - Because the norm map is multiplicative, if $\alpha = x + y\sqrt{D}$ and $\beta = z + w\sqrt{D}$ where we have $N(\alpha) = r$ and $N(\beta) = s$, then the element $\alpha\beta = (x + y\sqrt{D})(z + w\sqrt{D}) = (xz + Dyw) + (xw + yz)\sqrt{D} \in \mathbb{Z}[\sqrt{D}]$ will have norm rs.
  - This is the idea underlying the "recipes" identified in the examples above for generating new solution to Pell equations from other solutions: in the particular situation where $N(\beta) = \pm 1$, we can see that $N(\alpha\beta^k) = (-1)^k r$, and so multiplying the element α by β, β^2, β^3, ... will yield more solutions to $x^2 - Dy^2 = \pm r$.
  - Indeed, we can generate such a sequence whenever we can identify the elements in $\mathbb{Z}[\sqrt{D}]$ of norm ± 1, which are precisely the units of $\mathbb{Z}[\sqrt{D}]$. (Recall the reason for this: we have $N(\alpha) = \alpha\overline{\alpha}$ where $\overline{\alpha} = x - y\sqrt{D}$ is the conjugate of α, so if $N(\alpha) = \pm 1$ then $\overline{\alpha}/N(\alpha)$ is a multiplicative inverse of α in $\mathbb{Z}[\sqrt{D}]$, and conversely if $\alpha\beta = 1$ then taking norms gives $N(\alpha)N(\beta) = N(\alpha\beta) = N(1) = 1$ so that $N(\alpha) = \pm 1$.)

- For example, in $\mathbb{Z}[\sqrt{2}]$ we have $N(1+\sqrt{2}) = (1+\sqrt{2})(1-\sqrt{2}) = -1$. Therefore, if $a+b\sqrt{2}$ has norm r, then $(a+b\sqrt{2})(1+\sqrt{2}) = (a+2b)+(a+b)\sqrt{2}$ will have norm $-r$. This is precisely the map $(a,b) \mapsto (a+2b, a+b)$ we identified above.
- Likewise, in $\mathbb{Z}[\sqrt{3}]$ we have $N(2+\sqrt{3}) = (2+\sqrt{3})(2-\sqrt{3}) = 1$. Therefore, if $a+b\sqrt{3}$ has norm r, then $(a+b\sqrt{3})(2+\sqrt{3}) = (2a+3b)+(a+2b)\sqrt{2}$ will also have norm r. This is precisely the map $(a,b) \mapsto (2a+3b, a+2b)$ we identified above.

- All of this discussion suggests that should start by looking for the solutions of $x^2 - Dy^2 = \pm 1$, which is equivalent to determining the units in $\mathbb{Z}[\sqrt{D}]$.
  - Based on our (admittedly small) searches above for solutions of $x^2 - Dy^2 = \pm 1$, it would appear that the units all have the form $\pm\alpha^n$ where α is the "smallest" solution to $x^2 - Dy^2 = \pm 1$ in the sense that $\alpha = x + y\sqrt{D}$ with $x, y > 0$ and where x is minimal.
  - This discussion suggests the following useful definition:

- Definition: For a fixed positive squarefree D, a fundamental solution (x_1, y_1) to Pell's equation is a pair (x_1, y_1) of positive integers such that $x_1^2 - Dy_1^2 = \pm 1$ and $x_1 + y_1\sqrt{D}$ is minimal among all solutions to $x^2 - Dy^2 = 1$. The fundamental unit of $\mathbb{Z}[\sqrt{D}]$ is $u = x_1 + y_1\sqrt{D}$.
  - Note that this fundamental solution and the fundamental unit are well defined: there will be a unique minimal positive value for $x_1 + y_1\sqrt{D}$ over all pairs (x_1, y_1) satisfying $x_1^2 - Dy_1^2 = \pm 1$.
  - Examples: By searching for solutions to $x^2 - Dy^2 = \pm 1$ we can generate fundamental units for various small squarefree D:

D	2	3	5	6	7	10	11	13	14
Fund. Unit	$1+\sqrt{2}$	$2+\sqrt{3}$	$2+\sqrt{5}$	$5+2\sqrt{6}$	$8+3\sqrt{7}$	$3+\sqrt{10}$	$10+3\sqrt{11}$	$18+5\sqrt{13}$	$15+4\sqrt{14}$
Norm	-1	1	-1	1	1	-1	1	-1	1

- One of the other key ideas for solving Pell's equation is the observation that if $x^2 - Dy^2$ is small and x, y are positive, then x/y is a good approximation to $\sqrt{D}$.
  - To illustrate, suppose we have a solution of $x^2 - Dy^2 = 1$.
  - Dividing by y^2 yields $(x/y)^2 - D = 1/y^2$, and now solving for x/y gives $x/y = \sqrt{D + 1/y^2} = \sqrt{D} \cdot \sqrt{1 + 1/(Dy^2)} \approx \sqrt{D} \cdot (1 + 1/(2Dy^2)) = \sqrt{D} + 1/(2y^2\sqrt{D})$ using the linearization $\sqrt{1+z} \approx 1 + z/2$.
  - In fact, the linearization is an overestimate since $(1+z/2)^2 = 1 + z + z^2/4 > 1 + z$. Thus, we obtain the inequality $\left|\dfrac{x}{y} - \sqrt{D}\right| < \dfrac{1}{2y^2\sqrt{D}}$.
  - The point here is that x/y is a good approximation to $\sqrt{D}$. In fact, it is extremely good: from our results on continued fractions and rational approximation, we know that if α is irrational and $\dfrac{p}{q}$ has the property that $\left|\alpha - \dfrac{p}{q}\right| < \dfrac{1}{2q^2}$, then in fact $\dfrac{p}{q}$ is a continued fraction convergent to α.
  - Here, since $\sqrt{D} > 1$, we see immediately that any solution to $x^2 - Dy^2 = 1$ must arise as a continued fraction convergent to $\sqrt{D}$.
  - Indeed, we can observe this numerically in the case $D = 2$: we have $\sqrt{2} = [1, \overline{2}] = [1, 2, 2, 2, \dots]$ with convergents 1/1, 3/2, 7/5, 17/12, 41/29, 99/70, ... , which (as ordered pairs) have $x^2 - 2y^2$ respectively equal to -1, 1, -1, 1, -1, 1, ...: they are precisely the solutions to $x^2 - 2y^2 = \pm 1$ we identified earlier. The period of the continued fraction expansion here is equal to 1 and the fundamental unit corresponds to the convergent [1].
  - For $D = 3$ we have $\sqrt{3} = [1, \overline{1, 2}] = [1, 1, 2, 1, 2, \dots]$ with convergents 1/1, 2/1, 5/3, 7/4, 19/11, 26/15, 71/41, ... , which (as ordered pairs) have $x^2 - 3y^2$ respectively equal to -2, 1, -2, 1, -2, 1, ... Here, we can see that we do not obtain any solutions to $x^2 - 3y^2 = -1$ (since in fact there are none as we proved earlier) but we do obtain solutions to $x^2 - 3y^2 = -2$ and $x^2 - 3y^2 = 1$. The period of the continued fraction expansion here is equal to 2, while the fundamental unit corresponds to the convergent $[1, 2]$.

- For $D = 7$ we have $\sqrt{7} = [2, \overline{1, 1, 1, 4}] = [2, 1, 1, 1, 4, 1, 1, 1, 4, \dots]$ with convergents 2/1, 3/1, 5/2, 8/3, 37/14, 45/17, 82/31, 127/48, 590/223, ... , which (as ordered pairs) have $x^2 - 7y^2$ respectively equal to -3, 2, -3, 1, -3, 2, -3, 1, -3, ... Here again we obtain no solutions to $x^2 - 3y^2 = -1$ but we do obtain solutions to $x^2 - 3y^2 = -3$, $x^2 - 3y^2 = 2$, and $x^2 - 3y^2 = 1$. The period of the continued fraction expansion here is equal to 4, while the fundamental unit corresponds to the convergent $[2, 1, 1, 1]$.
- For $D = 13$ we have $\sqrt{13} = [3, \overline{1, 1, 1, 6}] = [3, 1, 1, 1, 1, 6, 1, 1, 1, 1, 6, \dots]$ with convergents 3/1, 4/1, 7/2, 11/3, 18/5, 119/33, 137/38, 256/71, 393/109, 649/180, ... , which (as ordered pairs) have $x^2 - 13y^2$ respectively equal to -4, 3, -3, 4, -1, 4, -3, 3, -4, 1, ... Here we obtain solutions to $x^2 - 13y^2 = r$ for $r = -4, -3, -1, 1, 3, 4$. The period of the continued fraction expansion here is equal to 4, while the fundamental unit corresponds to the convergent $[3, 1, 1, 1, 1]$.
- It appears that the fundamental unit is obtained after one period of the continued fraction expansion, regardless of whether it has norm 1 or -1, and the continued fraction expansion of $\sqrt{D}$ also always seems to be of the form $[a_0, \overline{a_1, \dots, a_n, 2a_0}]$

6.3.2 Proofs of Main Results

- Let us now prove all of these various observations we have made in the examples:
- Theorem (Pell's Equation): Let D be a positive squarefree integer. Then the following hold:

1. Let r be an integer with $r^2 + |r| < D$. If x and y are positive integers with $x^2 - Dy^2 = r$, then $\frac{x}{y}$ is a continued fraction convergent to $\sqrt{D}$.
   - Proof: Suppose r is an integer with $r^2 + |r| < D$ and x and y are positive integers with $x^2 - Dy^2 = r$.
   - We want to show that $\left|\frac{x}{y} - \sqrt{D}\right| < \frac{1}{2y^2}$, which by our previous results will show that $\frac{x}{y}$ is a continued fraction convergent to $\sqrt{D}$.
   - The assumptions imply $r/y^2 + D$ is at most $|r|^2$ and also that $\sqrt{D} \leq |r|$.
   - Then $\left|\frac{x}{y} - \sqrt{D}\right| = \frac{\left|x^2 - Dy^2\right|}{|y|^2 \left|x/y + \sqrt{D}\right|} = \frac{|r|}{y^2 \left|\sqrt{r/y^2 + D} + \sqrt{D}\right|} \leq \frac{|r|}{y^2 \left|\sqrt{|r|^2} + |r|\right|} = \frac{1}{2y^2}$, as desired.
2. The equation $x^2 - Dy^2 = 1$ always has a nontrivial solution in integers (x, y).
   - Proof: If $\frac{p}{q}$ is a continued fraction convergent to $\sqrt{D}$, then $\frac{p}{q}$ is within $1/q^2 \leq 1$ of $\sqrt{D}$, so $\left|\frac{p}{q} - \sqrt{D}\right| < \frac{1}{q^2}$ and $\left|\frac{p}{q} + \sqrt{D}\right| < 1 + 2\sqrt{D}$.
   - Then $|p^2 - Dq^2| = q^2 \left|\frac{p}{q} - \sqrt{D}\right| \left|\frac{p}{q} + \sqrt{D}\right| < q^2 \cdot \frac{1}{q^2} \cdot (1 + 2\sqrt{D}) = 1 + 2\sqrt{D}$.
   - Since $\sqrt{D}$ is irrational, there are an infinite number of convergents but only a finite number of possible values for $p^2 - Dq^2$, so by the pigeonhole principle there is some r such that $p^2 - Dq^2 = r$ has infinitely many solutions.
   - Choose such an r: then there are only finitely many possible pairs for the reduction of (p, q) modulo r, so again by the pigeonhole principle there exist two distinct convergents x/y and s/t such that $x^2 - Dy^2 = s^2 - Dt^2 = r$, $x \equiv s \pmod{r}$, and $y \equiv t \pmod{r}$.
   - Now we compute $u = \frac{x + y\sqrt{D}}{s + t\sqrt{D}} = \frac{xs - Dyt}{r} + \frac{-xt + ys}{r}\sqrt{D}$, and observe that both $xs - Dyt \equiv x^2 - Dy^2 \equiv 0 \pmod{r}$ and $-xt + ys \equiv 0 \pmod{r}$, so the quotient u is of the form $a + b\sqrt{D}$ where $a, b \in \mathbb{Z}$.
   - But now $N(u) = \frac{N(x + y\sqrt{D})}{N(s + t\sqrt{D})} = 1$, so u is a unit in $\mathbb{Z}[\sqrt{D}]$, and so $\left(\frac{xs - Dyt}{r}, \frac{-xt + ys}{r}\right)$ is a nontrivial solution to Pell's equation.
3. The ring $\mathbb{Z}[\sqrt{D}]$ has a well-defined fundamental unit $u = x_1 + y_1\sqrt{D}$. Furthermore, if w is an arbitrary unit in $\mathbb{Z}[\sqrt{D}]$, then $w = \pm u^n$ for some integer n (possibly negative).

- Remark (for those who like group theory): This result says that the unit group structure of $\mathbb{Z}[\sqrt{D}]$ is isomorphic to $(\mathbb{Z}/2\mathbb{Z}) \times \mathbb{Z}$: the $\mathbb{Z}/2\mathbb{Z}$ factor represents the $\pm$ sign while the $\mathbb{Z}$ factor represents the power n of the fundamental unit u.
- Proof: The fundamental unit is well-defined by (2), since we are assured of the existence of at least one solution to $x^2 - Dy^2 = \pm 1$. Observe (trivially) that because $u = x_1 + y_1\sqrt{D}$ with x_1, y_1 positive, we have $u > 1$.
- If w is any arbitrary unit, then by scaling by -1 if necessary, we may assume w is positive. Then there exists a unique integer n such that $w \in [u^n, u^{n+1})$ since u is a real number greater than 1 and these intervals $[u^n, u^{n+1})$ partition the interval $(0, \infty)$.
- Now observe that $w \cdot u^{-n} \in [1, u)$, and $w \cdot u^{-n}$ is also a unit in $\mathbb{Z}[\sqrt{D}]$.
- If this unit $x + y\sqrt{D}$ were not equal to 1, then (possibly after flipping signs on one of its terms) it would yield a positive solution (x, y) to Pell's equation $x^2 - Dy^2 = \pm 1$ such that $x + y\sqrt{D} < u$.
- But this contradicts the minimality of u, so in fact we must have $w \cdot u^{-n} = 1$, whence $w = u^n$. Since we chose the sign of w to be positive, the units in $\mathbb{Z}[\sqrt{D}]$ are then of the form $\pm u^n$, as claimed.

4. If $u = x_1 + y_1\sqrt{D}$ is the fundamental unit in $\mathbb{Z}[\sqrt{D}]$, then if we define $x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n$ for nonnegative integers n, then (x_n, y_n) is a solution to $x^2 - Dy^2 = \pm 1$, and these are all of the solutions up to changing the signs of x_n or y_n.
   - Proof: This is merely a rewriting of (3) in terms of solutions to $x^2 - Dy^2 = \pm 1$ rather than units in $\mathbb{Z}[\sqrt{D}]$.
5. The continued fraction expansion of $\sqrt{D}$ is periodic and of the form $[a_0, \overline{a_1, a_2, \cdots, a_{k-1}, 2a_0}]$ with $a_0 = \lfloor \sqrt{D} \rfloor$.
   - Proof: Consider instead the continued fraction expansion of $\alpha = a_0 + \sqrt{D}$ where $a_0 = \lfloor \sqrt{D} \rfloor$: we claim that it is $[\overline{2a_0, a_1, a_2, \ldots a_{k-1}}]$ for some positive integer k. The integer part is $\lfloor \alpha \rfloor = \lfloor a_0 + \sqrt{D} \rfloor = a_0 + \lfloor \sqrt{D} \rfloor = 2a_0$, as claimed.
   - It remains to see that the expansion is purely periodic. By our results on purely periodic expansions, this is equivalent to saying that $\alpha = a_0 + \sqrt{D}$ is reduced. Clearly $\alpha > 1$, and we also have $-1/\overline{\alpha} = \dfrac{1}{\sqrt{D} - a_0} > 1$ because $0 < \sqrt{D} - a_0 < 1$ by the definition of a_0 and the fact that $\sqrt{D}$ is irrational.
   - Therefore, $\alpha = a_0 + \sqrt{D}$ is reduced, so its continued fraction is periodic with even starting term as claimed. The claims about the expansion of $\sqrt{D}$ are then immediate.
6. Let α_n be the nth remainder term and $a_n = \lfloor \alpha_n \rfloor$ be the nth term in the continued fraction expansion, so that $\sqrt{D} = [a_0, a_1, \ldots, a_n, \alpha_{n+1}]$, and take $p_n/q_n = [a_0, a_1, \ldots, a_n]$ to be the nth convergent. Define the sequences A_n and C_n by setting $A_0 = 0$ and $C_0 = 1$, and for $n \geq 1$ set $A_{n+1} = a_n C_n - A_n$ and $C_{n+1} = (D - A_{n+1}^2)/C_n$. Then A_n and C_n are integers, $\alpha_n = \dfrac{A_n + \sqrt{D}}{C_n}$, $p_n p_{n-1} - D q_n q_{n-1} = (-1)^n A_{n+1}$, and $p_n^2 - D q_n^2 = (-1)^{n+1} C_{n+1}$.
   - Proof: We show all of these simultaneously by induction on n. The base case $n = 0$ is trivial, since $\alpha_0 = \sqrt{D} = \dfrac{0 + \sqrt{D}}{1}$. We now do the inductive steps for each argument.
   - First, clearly A_{n+1} is an integer. For C_{n+1}, plugging in for A_{n+1} and expanding yields $C_{n+1} = \dfrac{D - (a_n C_n - A_n)^2}{C_n} = (2a_n A_n - a_n^2 C_n) + \dfrac{D - A_n^2}{C_n}$ and the fraction at the end is simply C_{n-1}. Thus A_{n+1} and C_{n+1} are integers. We also obtain the formula $C_{n+1} = 2a_n A_n - a_n^2 C_n + C_{n-1}$.
   - Second, suppose $\alpha_n = \dfrac{A_n + \sqrt{D}}{C_n}$. Then $\alpha_{n+1} = \dfrac{1}{\alpha_n - a_n} = \dfrac{C_n}{-A_{n+1} + \sqrt{D}} = \dfrac{A_{n+1} + \sqrt{D}}{(D - A_{n+1}^2)/C_n} = \dfrac{A_{n+1} + \sqrt{D}}{C_{n+1}}$ as claimed.
   - For the last two statements, suppose $p_n p_{n-1} - D q_n q_{n-1} = (-1)^n A_{n+1}$ and $p_n^2 - D q_n^2 = (-1)^{n+1} C_{n+1}$ and recall that $p_{n+1} = a_{n+1} p_n + p_{n-1}$ and $q_{n+1} = a_{n+1} q_n + q_{n-1}$.
   - We then have $p_{n+1} p_n - D q_{n+1} q_n = (a_{n+1} p_n + p_{n-1}) p_n - D(a_{n+1} q_n + q_{n-1})(q_n) = a_{n+1}(p_n^2 - D q_n^2) + (p_n p_{n-1} - D q_n q_{n-1})$. The first term equals $a_{n+1}(-1)^{n+1} C_{n+1}$ by the second inductive hypothesis $p_n^2 - D q_n^2 = (-1)^{n-1} C_{n+1}$ while the second term equals $(-1)^n A_{n+1}$ by the first inductive hypothesis.

- Thus, we see $p_{n+1}p_n - Dq_{n+1}q_n = (-1)^{n+1}(a_{n+1}C_{n+1} - A_{n+1}) = (-1)^{n+1}A_{n+2}$ by the definition of A_{n+2}.
- In a similar way, we also have $p_{n+1}^2 - Dq_{n+1}^2 = (a_{n+1}p_n + p_{n-1})^2 - D(a_{n+1}q_n + q_{n-1})^2 = a_{n+1}^2(p_n^2 - Dq_n^2) + 2a_{n+1}(p_np_{n-1} - Dq_nq_{n-1}) + (p_{n-1}^2 - Dq_{n-1}^2)$. The first term is $a_{n+1}^2(-1)^{n+1}C_{n+1}$ by the second inductive hypothesis, the second term is $2a_{n+1}(-1)^nA_{n+1}$ by the first inductive hypothesis, and the third term is $(-1)^nC_n$ by the second inductive hypothesis applied to $n-1$.
- Thus, $p_{n+1}^2 - Dq_{n+1}^2 = a_{n+1}^2(-1)^{n+1}C_{n+1} + 2a_{n+1}(-1)^nA_{n+1} + (-1)^nC_n = (-1)^{n+1}[a_{n+1}^2C_{n+1} - 2a_{n+1}A_{n+1} - C_n]$, and the term in brackets is equal to $-C_{n+2}$ by the calculation noted earlier. This means $p_n^2 - Dq_n^2 = (-1)^{n+1}C_{n+1}$ as claimed.

7. If $\sqrt{D} = [a_0, \overline{a_1, a_2, \ldots, a_{k-1}, 2a_0}]$ and $p_{k-1}/q_{k-1} = [a_0, a_1, \ldots, a_{k-1}]$, then the fundamental unit of $\mathbb{Z}[\sqrt{D}]$ is $p_{k-1} + q_{k-1}\sqrt{D}$. Its norm is -1 when k is odd and its norm is $+1$ when k is even.
   - Proof: Suppose that $\sqrt{D} = [a_0, \overline{a_1, a_2, \ldots, a_{k-1}, 2a_0}]$. Then since the expansion is periodic, we have $a_0 + \sqrt{D} = [2a_0, a_1, \ldots, a_{k-1}, a_0 + \sqrt{D}]$, so we see that $\alpha_{k+1} = \sqrt{D} - a_0$.
   - By the second part of (6), this means $\dfrac{A_k + \sqrt{D}}{C_k} = -a_0 + \sqrt{D}$, and so since $\sqrt{D}$ is irrational the only way this can occur is when $C_k = 1$. Then by the last part of (6), this means $p_{k-1}^2 - Dq_{k-1}^2 = (-1)^kC_k = (-1)^k$. Thus, $p_{k-1} + q_{k-1}\sqrt{D}$ is a unit in $\mathbb{Z}[\sqrt{D}]$.
   - Conversely, suppose that $p_n + q_n\sqrt{D}$ is a unit in $\mathbb{Z}[\sqrt{D}]$ so that $p_n^2 - Dq_n^2 = \pm 1$.
   - By (1), p_n/q_n is a convergent to $\sqrt{D}$. Then by (6), we have $p_n^2 - Dq_n^2 = (-1)^{n+1}C_{n+1}$ and so we must have $C_{n+1} = 1$ and $(-1)^{n+1}$ equal to the norm of $p_n + q_n\sqrt{D}$.
   - But if $C_{n+1} = 1$, since all remainders are between 0 and 1, we must have $\alpha_{n+1} = \sqrt{D} - \lfloor\sqrt{D}\rfloor = \alpha_0$. By periodicity, the only way this can occur is if $n+1$ is a multiple of k.
   - The fundamental unit corresponds to the smallest possible value of n, which (per the calculation above) is $n = k - 1$.
   - Thus, the fundamental unit of $\mathbb{Z}[\sqrt{D}]$ is indeed $p_{k-1} + q_{k-1}\sqrt{D}$ as claimed, and its norm is -1 when k is odd and its norm is $+1$ when k is even, also from the calculation above.

- Example: Find the fundamental unit of $\mathbb{Z}[\sqrt{2}]$ and identify all the units of $\mathbb{Z}[\sqrt{2}]$.
  - As we computed earlier, the fundamental unit of $\mathbb{Z}[\sqrt{2}]$ is $u = 1 + \sqrt{2}$. Thus, the units of $\mathbb{Z}[\sqrt{2}]$ are the elements $\pm(1+\sqrt{2})^n$ for $n \in \mathbb{Z}$.
  - For example, taking the fifth power yields the element $41 + 29\sqrt{2}$, and we can indeed compute that $41^2 - 2 \cdot 29^2 = -1$.

- Example: Find the fundamental unit in $\mathbb{Z}[\sqrt{7}]$.
  - Earlier, we computed the expansion as $\sqrt{7} = [2, \overline{1, 1, 1, 4}]$. This has the desired form with $k = 4$, so we conclude there is no solution to $x^2 - 7y^2 = -1$.
  - Then the desired convergent is $C_4 = [2, 1, 1, 1] = \dfrac{8}{3}$, and we can indeed verify that $8^2 - 7 \cdot 3^2 = 1$. We conclude that the fundamental unit of $\mathbb{Z}[\sqrt{7}]$ is $\boxed{8 + 3\sqrt{7}}$.

- Example: Find the fundamental unit in $\mathbb{Z}[\sqrt{13}]$.
  - We compute the continued fraction expansion of $\sqrt{13} = [3, \overline{1, 1, 1, 1, 6}]$. This has the desired form with $k = 5$, so we conclude there is a solution to $x^2 - 13y^2 = -1$.
  - Then the desired convergent is $C_5 = [3, 1, 1, 1, 1] = \dfrac{18}{5}$, and we can indeed verify that $18^2 - 13 \cdot 5^2 = -1$. We conclude that the fundamental unit of $\mathbb{Z}[\sqrt{13}]$ is $\boxed{18 + 5\sqrt{13}}$.
  - If we want a solution to Pell's equation $x^2 - 13y^2 = 1$ instead, we simply square the fundamental unit: $(18 + 5\sqrt{13})^2 = 649 + 180\sqrt{3}$: then the minimal nontrivial solution is given by $(649, 180)$.

6.3.3 The Super Magic Box

- We can calculate the sequences $\{A_n\}$ and $\{C_n\}$ described in (6) above in our theorem, and thereby find the fundamental unit as described in (7), using a computational procedure that is sometimes referred to as the "super magic box". It works as follows:
  - The rows in the table are the sequences A_n, C_n, a_n, p_n, q_n, and $p_n^2 - Dq_n^2$.
  - We compute the sequences a_n, A_n, C_n via the recurrences[6] $A_{n+1} = a_nC_n - A_n$, $C_{n+1} = (D - A_{n+1}^2)/C_n$, and $a_{n+1} = \lfloor (A_{n+1} + a_0)/C_{n+1} \rfloor$ with initial conditions $A_0 = 0$, $C_0 = 1$, and $a_0 = \lfloor \sqrt{D} \rfloor$. Once we reach a term with $C_k = 1$ we stop, since we will have finished computing the full continued fraction expansion in the previous step.
  - We can then evaluate the convergents p_n/q_n using the recurrence relations $p_n = a_np_{n-1} + p_{n-2}$ and $q_n = a_nq_{n-1} + q_{n-2}$ with initial conditions $p_{-1} = 1$, $p_0 = a_0$, $q_{-1} = 0$, $q_0 = 1$.

- Example: Find the fundamental unit in $\mathbb{Z}[\sqrt{14}]$ using the super magic box.
  - Here is the result of doing the super magic box calculation:

n	-1	0	1	2	3	4
$A_n = a_{n-1}C_{n-1} - A_{n-1}$		0	3	2	2	3
$C_n = (D - A_n^2)/C_{n-1}$		1	5	2	5	$\boxed{1}$
$a_n = \lfloor (A_n + a_0)/C_n \rfloor$		3	1	2	1	6
$p_n = a_np_{n-1} + p_{n-2}$	1	3	4	11	15	101
$q_n = a_nq_{n-1} + q_{n-2}$	0	1	1	3	4	27
$p_n^2 - 14q_n^2$		-5	2	-5	1	-5

  - From this calculation we can see in fact that $\sqrt{14} = [3, \overline{1, 2, 1, 6}]$ and that the fundamental unit in $\mathbb{Z}[\sqrt{14}]$ is $15 + 4\sqrt{14}$ with norm 1.
  - In the bottom row of the table, we have also calculated the value of $p_n^2 - 14q_n^2$ explicitly in the bottom row: we can in particular see that $p_n^2 - 14q_n^2 = (-1)^{n+1}C_{n+1}$, as we proved was the case in (7).

- Example: Find the smallest nontrivial solution to the Pell equation $x^2 - 29y^2 = 1$.
  - Here is the result of doing the super magic box calculation for $D = 29$:

n	-1	0	1	2	3	4	5
$A_n = a_{n-1}C_{n-1} - A_{n-1}$		0	5	3	2	3	5
$C_n = (D - A_n^2)/C_{n-1}$		1	4	5	5	4	$\boxed{1}$
$a_n = \lfloor (A_n + a_0)/C_n \rfloor$		5	2	1	1	2	10
$p_n = a_np_{n-1} + p_{n-2}$	1	5	11	16	27	70	
$q_n = a_nq_{n-1} + q_{n-2}$	0	1	2	3	5	13	
$p_n^2 - 29q_n^2$		-4	5	-5	-4	-1	

  - From this calculation we can see that the fundamental unit of $\mathbb{Z}[\sqrt{29}]$ is $70 + 13\sqrt{29}$ having norm -1.
  - Thus, the smallest nontrivial solution to the Pell equation $x^2 - 29y^2 = 1$ corresponds to the square of the fundamental unit, which is $(70 + 13\sqrt{29})^2 = 9801 + 1820\sqrt{29}$, yielding the solution $(x, y) = \boxed{(9801, 1820)}$.
  - We will remark here that the super magic box calculation is quite short and easy to do by hand, quite unlike a brute-force search for solutions to $x^2 - 29y^2 = 1$!

- As it turns out, we can use the ideas from the super magic box algorithm to give an integer factorization algorithm, as first proposed by Lehmer and Powers in 1931. So suppose that D is some large composite integer.
  - The idea, as with other factorization algorithms such as the quadratic sieve, is to find a solution to the congruence $x^2 \equiv y^2 \pmod{D}$ where $x \not\equiv \pm y \pmod{D}$.

[6] Note that by definition we actually have $a_n = \lfloor \alpha_n \rfloor = \lfloor (A_n + \sqrt{D})/C_n \rfloor$; however, because A_n and C_n are integers, we may replace $\sqrt{D}$ with its greatest integer $a_0 = \lfloor \sqrt{D} \rfloor$ without affecting the floor calculation used to find a_n. The advantage to the recurrence we gave is that we do not need to estimate $\sqrt{D}$ at any point beyond finding its greatest integer.

- In such a case, $(x+y)(x-y)$ is divisible by D. But the gcd of $x+y$ and D cannot be 1 (since then necessarily D would divide $x-y$), and it also cannot be D (since then necessarily D would divide $x+y$): this means $1 < \gcd(x+y, D) < D$, and so $\gcd(x+y, D)$ is a nontrivial common divisor of n. Note that we can rapidly calculate this gcd via the Euclidean algorithm.
- If we use the super magic box algorithm to compute the sequences A_n, C_n, a_n, as above, for the continued fraction expansion of $\sqrt{D}$, then we know that $p_n^2 - Dq_n^2 = (-1)^{n+1}C_{n+1}$, and so modulo D we see $p_n^2 \equiv (-1)^{n+1}C_{n+1} \pmod{D}$.
- In particular, if we are able to find a convergent such that n is odd and C_{n+1} is a perfect square, we will obtain a congruence of the form $p_n^2 \equiv k^2 \pmod{D}$, which will allow us to find a factorization as long as it turns out that $p_n \not\equiv \pm k \pmod{D}$.

- Example: Use the super magic box to find a factorization of $D = 1271$.
  - Here is the result of doing the super magic box calculation for $D = 1271$:

n	-1	0	1	2	3	4	5	$\cdots$
$A_n = a_{n-1}C_{n-1} - A_{n-1}$		0	35	11	14	29	31	$\cdots$
$C_n = (D - A_n^2)/C_{n-1}$		1	46	$\boxed{25}$	43	10	31	$\cdots$
$a_n = \lfloor (A_n + a_0)/C_n \rfloor$		35	1	1	1	6	2	$\cdots$
$p_n = a_n p_{n-1} + p_{n-2}$	1	35	36	71	107	713	1533	$\cdots$
$q_n = a_n q_{n-1} + q_{n-2}$	0	1	1	2	3	20	43	$\cdots$
$p_n^2 - 1271q_n^2$		-46	25	-43	10	-31	31	$\cdots$

  - Here, we see that $C_2 = 25$ is a perfect square. Therefore, $p_1^2 = 36^2$ will be congruent to C_n modulo D, so we see $36^2 \equiv 5^2 \pmod{1271}$.
  - We can easily see $\gcd(36+5, 1271) = 41$, and so we obtain the factorization $1271 = 41 \cdot 31$.

- Of course, this procedure requires some amount of luck to find a factorization quickly, since there is no guarantee that we will find a term with $(-1)^{n+1}C_{n+1}$ equal to a perfect square early on in the calculation.
  - We can improve the method by combining it with the ideas from Dixon's method and the quadratic sieve algorithm.
  - All we actually require are two terms whose squares are congruent modulo D. Since $|C_{n+1}| < 2\sqrt{D}$, this means if we compute $4\sqrt{D}$ terms of the continued fraction expansion, we will be guaranteed to find two values of $(-1)^{n+1}C_{n+1}$ that are congruent modulo D, and thus we will obtain two convergents whose numerators satisfy $p_m^2 \equiv p_n^2 \pmod{D}$.
  - Of course, it could happen that the numerators of these terms have $p_m \equiv \pm p_n \pmod{D}$, in which case we will need to search for other tuples until we find a pair such that $p_m \not\equiv \pm p_n \pmod{D}$.
  - However, if we combine these ideas with those of the quadratic sieve, we can improve the speed of the sieving method: the improvement comes from the fact that the continued fraction convergents all have $|p_n|^2$ quite small modulo D (specifically, it is always less than $\sqrt{D}$), and so they are comparatively much easier to factor than arbitrarily-chosen squares modulo D.
  - One may show that, suitably optimized, the resulting sieving algorithm will find a factorization of D in approximately $e^{\sqrt{2 \ln n \ln \ln n}}$ time.

6.4 An Assortment of Other Diophantine Equations

- In this section we discuss a number of other unrelated Diophantine equations.
  - Our goal here is to illustrate some of the very wide variety of other elementary techniques that can be used to solve Diophantine equations.

6.4.1 Assorted Diophantine Equations

- Proposition: The Diophantine equation $\frac{1}{x}+\frac{1}{y}=\frac{1}{2021}$ has exactly 5 solutions (x,y) with $0<x<y$: $(x,y)=(2022,4086462)$, $(2064,97008)$, $(2068,88924)$, $(3870,4230)$, and $(4042,4042)$.
  - The idea of this proof is to rearrange the equation and factor.
  - Proof: Note that $x,y\geq 2022$. Clearing denominators yields $2021y+2021x=xy$, which we can rearrange and factor as $(x-2021)(y-2021)=2021^2$.
  - We can see that $2021^2=43^2\cdot 47^2$ has 9 possible factorizations as the product of two positive integers.
  - These factorizations yield five possible pairs $(x-2021,y-2021)=(1,2021^2)$, $(43,94987)$, $(47,86903)$, $(43^2,47^2)$, $(2021,2021)$, and these give the five solutions listed above.
- Proposition: There are no solutions to the Diophantine equation $x^2+y^2+z^2=4^a(8b+7)$.
  - The idea of this proof is to use modular arithmetic.
  - Proof: We prove the result by induction on a.
  - For the base case $a=0$, consider the equation modulo 8.
  - Each square is either 0, 1, or 4 mod 8, so it is not possible to obtain a sum of 7 mod 8 by adding three of them.
  - Now suppose there are no solutions for $a\leq k$, and take $a=k+1$.
  - Consider the equation $x^2+y^2+z^2=4^{k+1}(8b+7)$ modulo 4. Each of the squares is 0 or 1, while the term $4^{k+1}(8b+7)$ is 0 mod 4, so all of the squares must be 0 mod 4.
  - Then $\left(\frac{x}{2}\right)^2+\left(\frac{y}{2}\right)^2+\left(\frac{z}{2}\right)^2=4^k(8b+7)$, but this is a contradiction since by induction, this equation has no solutions.
  - Remark: In fact, these are the only integers that cannot be written as a sum of three squares, as first proven by Legendre. (Gauss gave a formula for the number of such representations, similar to Fermat's formula for the number of ways of writing an integer as a sum of two squares.)
- Proposition: The Diophantine equation $y^2=x^4+4x^3+x^2+2x+1$ has the solutions $(x,y)=(-4,\pm 3)$, $(0,\pm 1)$, $(1,\pm 3)$, $(6,\pm 47)$, and no others.
  - The idea of this result is to attempt to complete the square of the x-terms, and then use some simple inequalities to bound how big x and y can be.
  - Proof: We complete the square of the x-terms and obtain $x^4+4x^3+x^2+2x+1=\left(x^2+2x-\frac{3}{2}\right)^2+8x-\frac{5}{4}$.
  - Thus, we should try comparing y^2 to $(x^2+2x-2)^2=x^4+4x^3-8x+4$ and $(x^2+x-1)^2=x^4+2x^3+2x^2-4x+1$.
  - We see that $y^2-(x^2+x-2)^2=x^2+10x-3$ is positive outside $[-10.3,0.3]$, while $(x^2+x-1)^2-y^2=x^2-6x$ is positive outside $[0,6]$.
  - Hence, if $x\notin[-10,6]$, then we have the strict inequalities $(x^2+x-2)<y^2<(x^2+x-1)^2$, which is impossible if x and y are both integers.
  - Thus it must be the case that $x\in[-10,6]$. It is then a straightforward calculation (trivial to implement by computer) to check the 17 cases to find the solutions as listed above.
  - Remark: More generally, one can adapt this proof method to show that there are only finitely many solutions to any equation of the form $y^2=x^4+ax^3+bx^2+cx+d$ for fixed integers a,b,c,d, as long as the right-hand side is not a perfect square.
- Proposition: The solutions to the Diophantine equation $x^2+y^2=z^3$ with $\gcd(x,y)=1$ are of the form $(x,y,z)=(a^3-3ab^2,\ 3a^2b-b^3,\ a^2+b^2)$ for relatively prime integers a,b of opposite parity.
  - The idea of this proof is to exploit the arithmetic of the Gaussian integers $\mathbb{Z}[i]$.

- Proof: First observe that if x, y were both odd, then $z^3 \equiv 2 \pmod 4$, but 2 is not a cube modulo 4.
- Since x, y are not both even since $\gcd(x, y) = 1$, we conclude that one is even and the other is odd.
- Now, over $\mathbb{Z}[i]$, factor the equation as $(x + iy)(x - iy) = z^3$.
- We claim that $x + iy$ and $x - iy$ are relatively prime: any common divisor would divide both $2x$ and $2y$, hence divide 2. But $1 + i$ (the only Gaussian prime dividing 2) does not divide $x + iy$, since x, y are of opposite parity.
- Now, by the uniqueness of prime factorization in $\mathbb{Z}[i]$, we conclude that $x + iy$ must be a unit times a cube.
- But since each unit in $\mathbb{Z}[i]$ is actually a cube, we conclude that $x + iy = (a + bi)^3$ for some $a + bi \in \mathbb{Z}[i]$.
- Equating real and imaginary parts yields $x = a^3 - 3ab^2$, $y = 3a^2b - b^3$, and then $z = (a + bi)(a - bi) = a^2 + b^2$, as claimed.
- Remark: One can use a similar argument to write down the solutions to $x^2 + y^2 = z^d$ for any positive integer d.

- Corollary: The only solution to the Diophantine equation $y^2 = x^3 - 1$ is $(x, y) = (1, 0)$.
  - Proof: Clearly, $\gcd(x, y) = 1$. Rearranging the equation into the form $1 + y^2 = x^3$ and applying the previous result shows that $1 = a^3 - 3ab^2$ for $a, b \in \mathbb{Z}$.
  - Factoring this gives $1 = a(a^2 - 3b^2)$. Clearly, $a = \pm 1$, and then the only solution is easily seen to be $(a, b) = (1, 0)$, yielding $(x, y) = (1, 0)$.

- We do not need to restrict to considering Diophantine equations where the terms involved are polynomials in the variables: we can also include variables in the exponents.

- Proposition: The Diophantine equation $7^a - 4^b = 3$ has the unique solution $(a, b) = (1, 1)$.
  - The idea of this result is to use congruence conditions.
  - Proof: Clearly a and b must be nonnegative, else the denominators of the rational numbers involved could not be equal. Clearly $b = 0$ fails, and $b = 1$ gives $a = 1$.
  - Now suppose $b \geq 2$ and consider the equation modulo 8: we obtain $7^a \equiv 3 \pmod 8$. However, there are no solutions to this equation, because 7^a can only be 7 or 1 modulo 8.
  - Therefore, the only solution is $(a, b) = (1, 1)$.

- Proposition: The Diophantine equation $3^a - 2^b = 1$ has the two solutions $(a, b) = (1, 1)$, $(2, 3)$, and no others.
  - The idea of this result is to use congruence conditions.
  - Proof: Clearly a and b must be nonnegative, else the denominators of the rational numbers involved could not be equal. Clearly $b = 0$ fails, and $b = 1$ gives $a = 1$.
  - Now suppose $b \geq 2$ and consider the equation modulo 4: we obtain $3^a \equiv 1 \pmod 4$, meaning that a is even, say, $a = 2k$.
  - Then we have $2^b = 3^{2k} - 1 = (3^k + 1)(3^k - 1)$, so $3^k + 1$ and $3^k - 1$ must both be powers of 2.
  - But their difference is 2, and so they must be 4 and 2 respectively. Thus, the only other solution is $(a, b) = (2, 3)$.
  - Remark: This result is a special case of a result called Catalan's conjecture (proven in 2002 by Mihailescu) that 8 and 9 are the only perfect powers that are consecutive: in other words, the only solutions to $x^a - y^b = 1$ in integers greater than 1 is $(a, b, x, y) = (2, 3, 3, 2)$.

- Proposition: The Diophantine equation $y^2 = x^3 + 7$ has no solutions.
  - The idea of this result is to rewrite the equation slightly, exploit congruence conditions, and then quadratic reciprocity to obtain a contradiction.
  - Proof: If x is even, then this equation yields $y^2 \equiv 3 \pmod 4$, which is not possible.
  - Then $x^3 + 7 \equiv 0 \pmod 4$ meaning $x \equiv 1 \pmod 4$.

- Now we write the equation as $y^2 + 1 = x^3 + 8 = (x + 2)(x^2 - 2x + 4)$.
- By our results about quadratic residues, any prime divisor of $y^2 + 1$ must be congruent to 1 modulo 4, since $y^2 + 1 \equiv 0 \pmod{p}$ is equivalent to -1 being a quadratic residue modulo p.
- Thus, any prime divisor, and therefore *any* divisor prime or otherwise, of $y^2 + 1 = x^3 + 8$ must be congruent to 1 modulo 4.
- But $x + 2$ is a divisor of $x^3 + 8$ congruent to 3 modulo 4, so we have a contradiction.
- Remark: This result is notable because there always is a solution to the equation $y^2 = x^3 + 7$ modulo p for every prime p. (This is not trivial to prove.)

- Proposition: There are infinitely many perfect squares that are the sum of two other consecutive perfect squares.
  - The idea of this result is to rearrange the equation and use properties of Pell's equation.
  - Proof: Suppose that $a^2 = b^2 + (b + 1)^2$ so that $a^2 = 2b^2 + 2b + 1$. Multiplying both sides by 2 and completing the square on the right-hand side yields $2a^2 = (2b + 1)^2 + 1$, so that $(2b + 1)^2 - 2a^2 = -1$.
  - This is a Pell equation of the form $x^2 - 2y^2 = -1$, where $x = 2b + 1$.
  - Since the fundamental unit of $\mathbb{Z}[\sqrt{2}]$ is $u = 1 + \sqrt{2}$ which has norm -1, we know that $x^2 - 2y^2 = -1$ will have infinitely many solutions given by odd powers of u: $x + y\sqrt{2} = (1 + \sqrt{2})^{2k+1}$ for $k \geq 0$.
  - Therefore, since x is always odd in such solutions, each of these infinitely many solutions yields a different pair (a, b) with $a^2 = b^2 + (b + 1)^2$.
  - For example, the first few pairs (a, b) are $(a, b) = (1, 0), (5, 3), (29, 20), (169, 119), (985, 696), (5741, 4059)$, and so forth.
  - Remark: It is also possible to approach this problem using our characterization of the Pythagorean triples, however, it is quite easy to get lost in the resulting morass of variables.

6.4.2 The Fermat Equation $x^n + y^n = z^n$

- One of the most famous Diophantine equations is Fermat's equation $x^n + y^n = z^n$, for a fixed integer $n \geq 3$. Clearly, there are solutions if one of the variables is equal to 0: the question is whether this equation possesses any other solutions.
  - It is enough to prove that there are no solutions in the cases $n = 4$ and $n = p$ where p is an odd prime, since any $n > 2$ is divisible by 4 or an odd prime.

- This result was famously conjectured by Fermat in 1637, who wrote (in the margin of his book, in Latin) "It is impossible to separate a cube into two cubes, or a fourth power into two fourth powers, or in general, any power higher than the second, into two like powers. I have discovered a truly marvellous proof of this, which this margin is too narrow to contain."
  - It is now believed that Fermat probably did not have a correct proof of this result.
  - As we will discuss more in later chapters, a substantial amount of number theory and abstract algebra was developed in the mid-19th and early-20th centuries in an attempt to establish the nonexistence of nontrivial integer solutions to $x^n + y^n = z^n$.

- One of the easier cases is the case with $n = 4$, which is in fact the subject of one of Fermat's very few theorems for which he gave an actual proof:

- Theorem (Fermat): The Diophantine equation $x^4 + y^4 = z^2$ has no solutions with $xyz \neq 0$. In particular, there are no solutions to $x^4 + y^4 = z^4$ with $xyz \neq 0$.
  - This result is originally due to Fermat. We show the result using a technique equivalent to induction often called Fermat's method of infinite descent (indeed, it first appeared in the proof of this very result).
  - The idea is to consider the smallest nontrivial solution of the equation in positive integers and use it to construct a smaller solution: the well-ordering principle of the integers then yields a contradiction, since we cannot have an infinite decreasing sequence of positive integers.

- Proof: Suppose the equation has nontrivial solutions and let u be the smallest positive integer such that $x^4 + y^4 = u^2$ has a solution.
- First note $\gcd(x, y) = 1$; otherwise we could replace x, y, u with x/d, y/d, u/d^2 to get a smaller solution.
- By reducing both sides modulo 4, we see that one of x, y is even and the other is odd: without loss of generality, assume x is even.
- Then (x^2, y^2, u) is a primitive Pythagorean triple, so from our parametrization we see that $x^2 = 2st$, $y^2 = s^2 - t^2$, and $u = s^2 + t^2$ for some integers $s > t > 0$ of opposite parity.
- Since $y^2 = s^2 - t^2$, it must be the case that s is odd and t is even: otherwise, $y^2 = s^2 - t^2$ would be congruent to -1 modulo 4.
- If we set $t = 2k$, then we see that $\left(\frac{x}{2}\right)^2 = sk$ where $\gcd(s, k) = 1$, so each of s and k is a perfect square by the uniqueness of prime factorizations.
- Setting $s = a^2$ and $k = b^2$ yields the system $y^2 = s^2 - t^2 = a^4 - 4b^4$, so that $y^2 + (2b^2)^2 = a^4$.
- Then $(y, 2b^2, t)$ is also a primitive Pythagorean triple, so there exist relatively prime integers m and n such that $2b^2 = 2mn$, $y = m^2 - n^2$, and $a^2 = m^2 + n^2$.
- The first equation gives $b^2 = mn$, so m and n are both squares: say, $m = v^2$ and $n = w^2$.
- Then, at last, we see that $a^2 = v^4 + w^4$, meaning that we have a new solution (v, w, a) to the original equation. Clearly $a \leq a^2 = s < s^2 + t^2 = u$, so this solution is smaller. We have obtained a contradiction, so in fact there cannot exist any solutions to the original equation.

- We will later establish the $n = 3$ case of Fermat's equation using properties of the ring $\mathbb{Z}[\omega]$ where $\omega = \frac{-1+\sqrt{-3}}{2}$.
  - Specifically, the idea is to factor the equation $x^3 + y^3 = z^3$ as $(x + y)(x + \omega y)(x + \omega^2 y) = z^3$, and then to show that, up to small factors, the terms $x + y$, $x + \omega y$, $x + \omega^2 y$ are relatively prime in $\mathbb{Z}[\omega]$. Up to these small factors, each of these terms must therefore be a perfect cube, which we can eventually use to derive a contradiction.

- The argument in the $n = 3$ case lends itself to a natural generalization, namely, factoring $x^n + y^n = z^n$ over the ring $\mathbb{Z}[\zeta_n]$ where $\zeta_n = e^{2\pi i/n}$ is an nth root of unity.
  - However, quite unfortunately, for most n, the ring $\mathbb{Z}[\zeta_n]$ does not have unique factorization!
  - So (alas!) this technique does not work in general. However, determining when this approach can succeed was one of the original motivations for studying unique factorization in general rings.

- The cases $n = 5$ and $n = 7$ were shown in the 1800s by various mathematicians using various techniques. A number of other cases were shown individually, and then results of Germain and others established infinite classes of prime n for which there are no nontrivial solutions to the equation.

- However, the lack of a solution to Fermat's equation for every $n > 2$ was not established until 1995, with Andrew Wiles's celebrated proof of the Taniyama-Shimura-Weil conjecture. (Wiles announced his result in 1993, but a gap was discovered later that year. Wiles, working with Richard Taylor, closed the gap by 1994.)
  - One of the initial steps in Wiles's proof stemmed from an observation made by Frey in 1984, which connects the solutions to $a^p + b^p = c^p$ to a certain elliptic curve.
  - Such a curve would have a number of unusual properties, and (in particular) is what is called a semistable elliptic curve, and it would also fail to be modular.
  - Wiles's results proved that every semistable elliptic curve is modular, which, when combined with Frey's observations, shows that the Fermat equation cannot have a solution in nonzero integers.
  - Over the next few chapters, we will develop some more of the background necessary to understand the structure of this result. But we will close by noting that, as with most major mathematical advances, the fundamental ideas put forward in Wiles's work are just as important as the end result of his proof.

Well, you're at the end of my handout. Hope it was helpful.
Copyright notice: This material is copyright Evan Dummit, 2014-2026. You may not reproduce or distribute this material without my express permission. No aspect of this material was created or edited by a large-language model or similar technology ("generative AI"), and this material may not be used for training of such technology.