2 Double Counting

Exercise 2.1 Let a and b be integers. Develop a formula for

$$\sum_{j=1}^{n} a + jb.$$

Solution: Using the double counting pattern, the sum can also be represented as

$$S_n = \sum_{j=1}^{n} a + (N+1)b - jb$$

Then we have

$$2S_n = \sum_{j=1}^{n} 2a + (N+1)b$$

$$S = Na + \tfrac{N(N+1)}{2}b$$

Exercise 2.2 How many vertices does an icosahedron have?

Solution: The number of vertex-faces is 60 since there are 20 faces and 3 vertices per face. Then

$$5V = 60$$

since each vertex is connected to 5 faces. The number of vertices is 12.

Exercise 2.3 If p is a prime and not 2, and a is an integer, show that 2 divides into $a^p - a$. (Try to construct a bijection on the set of size p orbits which pairs them.)

Solution: We can regard $(a^p - a)/p$ as the number of size p orbits on strings of length p chosen from a different characters. Since $p \mid a^p - a$ and 2 doesn't divide p, 2 divides $a^p - a$ if and only if 2 divides $(a^p - a)/p$.

Let A be a size a set of characters. Define a bijection $k : A \to A$ such that at most one element of A is mapped to itself and $k(k(x)) = x \ \forall x$. If S is a string of length p, denote S_j as the j^{th} character of the string. Let K be a function on the set of strings of length p such that if $T = K(S)$ then $T_j = k(S_j)$ for $j \in [1, p]$. K is a bijection because k is a bijection. Additionally K has the property that for any S in an orbit of length greater than 1, $K(S)$ is in a different orbit, and if strings S and T are in the same orbit, then $K(S)$ and $K(T)$ are in the same orbit. Therefore K also defines a bijection on the set of orbits. K is a pairing on the set of orbits of length p, so the set of orbits of size p, of which there are $(a^p - a)/p$, is divisible by two. Then $2 \mid a^p - a$.

Exercise 2.4 Suppose we have three sorts of jewels and apply the arguments for Vandermonde's identity, what formula do we find?

Solution: There are two ways to count the ways to choose r jewels from l jewels of type 1, m jewels of type 2, and n jewels of type 3. The first is to take $\binom{l+m+n}{r}$. Or, by counting sequentially

$$\sum_{j=0}^{r} \sum_{k=0}^{r-j} \binom{l}{j} \binom{m}{k} \binom{n}{r-j-k}.$$

Equating the two experessions gives the formula

$$\binom{l+m+n}{r} = \sum_{j=1}^{r} \sum_{k=1}^{r-j} \binom{l}{j}\binom{m}{k}\binom{n}{r-j-k}.$$

3 The Pigeonhole Principle

Exercise 3.1 Show that if a lossless compression algorithm exists for any file, then it is possible to represent any file by a single bit of information.

Solution: A *compression algorithm* maps binary sequences of length m to binary sequences of length n, where $m > n$. A *lossless compression algorithm* is a compression algorithm which is an injective map. Any file can be represented by a single bit of information if there is an injective function from the set of binary sequences to the set of binary sequences of length 1.

Let $x \in \{0,1\}^k$ be an arbitrary binary sequence and suppose $T : \{0,1\}^\omega \to \{0,1\}^\omega$ is a lossless compression algorithm. The length of $T(x)$ must be less than the length of x, and T is injective, so T^k is a composition of injective functions and $T^k(x)$ is at most 1 bit in length.

Exercise 3.2 Show that there are two residents of London with the same number of hairs on their head.

Solution: Let L be the set of residents of London. The cardinality of L is 9 million. Let $h : L \to \mathbb{Z}$ be the map from each resident to the number of hairs on his or her head. Assuming that each resident has less than 9 million hair follicles (the average is 150,000) then $|Im(h)| < |L|$. Therefore the map $h' : L \to Im(h)$ is a map from a set of higher cardinality to lower cardinality. By the pigeonhole principle h is not injective, and there must be residents $a, b \in L, a \neq b$ such that a and b have the same number of hair follicles, $h(a) = h(b)$.

Exercise 3.3 Suppose N people are at a party. Some have met before and some have not. Show that there are two people who have met the same number of other people before.

Solution: Let S be the set of people at the party. The map $m : S \to \mathbb{Z}$ represents the number of other people at the party that each person has met before. Clearly $m(x) \in [0, N-1]$ since each person has met at least no other person, and there are $N-1$ other people at the party. Suppose $m(x) = 0$ for some $x \in S$. Then for all $y \in S, m(y) \neq N-1$. The cardinality of the image of m is less than N, and by the pigeonhole principle the map m is not injective.

If $m(x) \neq 0$ for all $x \in S$, then $Im(m)$ is a subset of integers from 1 to $N-1$ and has cardinality less than N. The pigeonhole principle says that the map is not injective.

Exercise 3.4 Suppose we take a set of 101 different integers between 1 and 200. Show that there is a pair such that one divides the other.

Solution: Let a *dividing pair* be a pair of integers such that one divides the other. Every integer is either a power of 2 or an odd number times a power of 2, so every integer between 1 to 200 is of the form $(2n+1)2^m$, where n and m are natural numbers including zero. If two different integers have the same odd coefficient, they take on the form $k2^m$ and $k2^{m'}$. If $m' < m$ then $k2^m = k2^{m'}(2^{m-m'})$ in which case $k2^{m'}$ divides $k2^m$. Otherwise if $m' > m$ then $k2^{m'} = k2^m(k2^{m'-m})$ and $k2^m$ divides $k2^{m'}$. Therefore two integers are a dividing pair if they have the same odd coefficient when represented in the form $(2n+1)2^m$. There are 100 odd numbers between 1 and 200, and by the pigeonhole principle a choice of 101 integers between 1 and 200 must include two with the same odd coefficient.

Exercise 3.5 Represent the following decimals as ratios of integers.

- $0.1212121212\ldots,$
- $0.123123123123123\ldots,$
- $0.456456456\ldots.$

Solution:

1. $10^2x - x = 12 \implies x = \frac{12}{99} = 0.1212121212\ldots,$
2. $10^3x - x = 123 \implies x = \frac{123}{999} = 0.123123123123123\ldots,$
3. $10^3x - x = 456 \implies x = \frac{456}{999} = 0.456456456\ldots.$

4 Divisions

Exercise 4.1 Find the highest common factors of the following number pairs.

- 1236 and 369.
- 144 and 900.
- 99 and 36.

Solution:

1. $1236 = 3 \underset{1107}{\times} 369 + 129$

   $369 = 2 \underset{258}{\times} 129 + 111$

   $129 = 1 \underset{111}{\times} 111 + 18$

   $111 = 6 \underset{108}{\times} 18 + 3$

   $18 = 6 \underset{18}{\times} 3 + 0.$

   The highest common factor is 3.

2. $900 = 6 \underset{864}{\times} 144 + 36$

   $144 = 4 \underset{144}{\times} 36 + 0$

   The highest common factor is 36.

3. $99 = 2 \underset{72}{\times} 36 + 27$

   $36 = 1 \underset{27}{\times} 27 + 9$

   $27 = 3 \underset{27}{\times} 9 + 0$

   The highest common factor is 9.

Exercise 4.2 Show that the highest common factor of n and $n+1$ is 1.

Solution: From Euclid's division algorithm, rewriting $m = qn + r$ in terms of n and $n+1$ gives

$$n + 1 = 1 \times n + 1$$

where $m = n + 1, q = 1$ and $r = 1$. Then the highest common factor between m and n must be the same as that between n and r. The only factor of 1 is 1, so $(n + 1, n) = (n, 1) = 1$.

Exercise 4.3 Show that the highest common factor of n and $n^2 + 1$ is 1.

Solution: The same argument as in Exercise 4.2 applies, but this time $m = n^2 + 1, q = n$ and $r = 1$. The highest common factor $(n^2 + 1, n) = (n, 1) = 1$.

Exercise 4.4 What are the possible highest common factors of n and $n^2 + k$ if $k < n$?

From the equality $(n^2 + k, n) = (n, k)$ the possible highest common factors of n and $n^2 + k$ are the factors of n.

5 Contrapositive and Contradiction

Exercise 5.1 Use proof by contradiction to show that $\sqrt{12}$ is irrational.

Solution: Assume $\sqrt{12}$ is a non-integer rational, then $\sqrt{12} = \frac{m}{n}$ for coprime integers m, n. Squaring both sides, $12 = \frac{m^2}{n^2}$. Let the prime factors of m be $\{\alpha\}$ such that $m = \prod \alpha^{i_\alpha} = \alpha_1^{i_1} \alpha_2^{i_2} \alpha_3^{i_3} \cdots$ and the factors of n likewise be $\{\beta\}$. The factorization of m^2 is $\alpha_1^{2i_1} \alpha_2^{2i_2} \alpha_3^{2i_3} \cdots$ so n and n^2 have the same prime factors with different exponents. By the same logic m and m^2 have the same prime factors. Therefore m^2 and n^2 are coprime, and the quotient $\frac{m^2}{n^2}$ is not an integer. This contradicts $12 = \frac{m^2}{n^2}$ which follows from the assumption that $\sqrt{12}$ is a rational number. Therefore $\sqrt{12}$ is irrational.

Exercise 5.2 What are the contrapositives of the following statements?

- Every differentiable function is continuous.
- Every infinite set can be placed in a bijection with the rationals.
- Every prime number is odd.
- Every odd number is prime.

What are their converses? Which are true?

Solution:

1. Every discontinuous function is not differentiable.
2. Every set that cannot be placed in a bijection with the rationals is finite.
3. Every even number is not prime.
4. Every number that is not prime is even.

The converse of the first statement is: every continuous function is differentiable. The converse of the second statement is: every set that can be placed in a bijection with the rationals is infinite. The converse of the third statement is the fourth statement and vice-versa. The first and third statements are true, while the second and fourth statements are false.

6 Intersection-Enclosure and Generation

Exercise 6.1 Check that the intersection-enclosure pattern does indeed apply to each of the examples of Sect. 6.2

Solution:

1. Additive subgroups by generation:
   Let $S_0 = \{1\}$, and let S_n be the set $S_{n-1} \cup \{x - y \mid x, y \in S_{n-1}\}$ generated by taking the difference of the elements of S_{n-1} for $n \geq 1$. For example, $S_1 = \{0, 1\}$ and $S_2 = \{-1, 0, 1\}$. Let $S_\infty = \bigcup_n S_n$. If $x, y \in S_\infty$ then for some i, j we have $x \in S_i$ and $y \in S_j$. If $k = \max\{i, j\}$ then S_k contains both x and y, and $S_{k+1} \subset S_\infty$ contains $x - y$. Therefore S_∞ is an additive subgroup. Let B be the smallest subgroup containing 1. S_∞ is an additive subgroup containing 1, so S_∞ contains B. Any additive subgroup containing 1 must contain S_k for all k because the elements of S_k are generated only by the number 1 and the additive subgroup condition, so B contains S_∞. Therefore $S_\infty = B$ is the smallest additive subgroup generated by 1.

   Additive subgroups by intersection-enclosure:
   Let $\{S_\alpha\}_{\alpha \in I}$ be the set of all additive subgroups containing 1. The intersection $S = \bigcap_{\alpha \in I} S_\alpha$ contains 1 because every S_α contains 1. If $x, y \in S$ then $x, y \in S_\alpha$ for every $\alpha \in I$. By definition $x - y \in S_\alpha$ for every $\alpha \in I$. Therefore $x - y \in S$ and S is an additive subgroup. Any additive subgroup containing 1 contains S, so S is the smallest additive subgroup containing 1.

2. Vector sub-spaces
   A non-empty subset B of $\mathbb{R}^n$ is a *vector sub-space* if the sum of any two elements of B is also in B, and any scalar multiple of an element of B is also in B. Let $\{S_\alpha\}_{\alpha \in I}$ be the set of vector spaces containing a set of vectors A. Let $S = \bigcap_{\alpha \in I} S_\alpha$ be the intersection of those sets. If $x, y \in S$ then $x, y \in S_a lpha$ for every $alpha$ in I, so $x + y$ is in every S_α and $x + y$ is in S. If $x \in S_\alpha$ for every $\alpha \in I$ then $\lambda x \in S_\alpha, \alpha \in I$ and $\lambda x \in S$. Therefore S satisfies the vector sub-space conditions.

3. Closure
   A subset B of $\mathbb{R}$ is said to be *closed* if the limit of every convergent sequence in B is also in B. Let S be the intersection of every closed set containing a subset A of $\mathbb{R}$. If the sequence $x_n \in S \forall n$, then x_n is in every closed set C_A containing A. If $x_n \to x \in \mathbb{R}$ converges, then its limit $x \in C_A$ by the definition of closed. As the intersection of every C_A, S too contains x. Therefore S is closed.

4. Convexity
   A subset, B, of $\mathbb{R}^n$ is said to be *convex* if the straight line between any two of its points is also contained in it. Let S be the intersection of every convex set containing a subset A of $\mathbb{R}^n$ and suppose $x, y \in S$. Then x, y are in every convex set containing A, and by the definition of convexity the straight line between x and y is in every convex set containing A. Therefore the straight line $\lambda x + (1 - \lambda y), \lambda \in (0, 1)$ between x and y is in S, and S is the *convex hull*, the smallest convex set containing A.

Exercise 6.2 Given 3 points in the plane, what is their convex hull? Distinguish according to whether the 3 points are collinear.

Solution: If the three points $x, y, z \in \mathbb{R}^2$ are collinear, their convex hull is a line. Otherwise, the convex hull consists of the points bounded by a triangle.

Exercise 6.3 Given a subset of the plane, how many steps are required for the generation algorithm for convexity to terminate?

Solution: Let $A_0 \subset \mathbb{R}^n$ be a subset of the plane, and A_{n+1} be the set $A_n \cup \{\lambda x + (1-\lambda)y \mid x, y \in A_n, \lambda \in (0,1)\}$. The question is under what conditions does

$$A_k = A_{k+1}.$$

Denote the convex hull of A_0 as A'. Suppose A_0 is one point. Then $A' = A_0$ and the generation algorithm terminates after zero steps. If A_0 is two points, then A_1 is the set $A_0 \cup \{\lambda x + (1-\lambda)y \mid x, y \in A_0, \lambda \in (0,1)\} = A'$, the straight line between the two points in A_0, and the generation algorithm terminates after one step. If $|A_0| \geq 3$ and the elements of A_0 are not collinear, then the first step of the generation algorithm A_1 consists of the boundary of A_0 and lines connecting the points of A_0. The second step of the generation algorithm is then the convex hull A', the lines connecting all points on the boundary of A_0. In general, for a subset of the plane with three or more non collinear points, the generation algorithm for convexity terminates after two steps.

Exercise 6.4 A subset of $\mathbb{R}$ is said to be binarily division invariant if dividing any element by 2 results in an element of the subset. Check that intersection-enclosure applies. Also analyze generation for this property and show that it terminates with A_∞.

Solution:

1. Check that intersection-enclosure applies.
   Let A is a subset of $\mathbb{R}$ and let B be the intersection of every binarily division invariant subset containing A. If $x \in B$ then x is in every binarily division invariant subset A_β containing A, so $x/2$ is in every A_β and $x/2 \in B$. Therefore binary division invariance holds under the intersection-enclosure pattern.

2. Analyze generation for this property.
   Let A_0 be a subset of $\mathbb{R}$ and let $A_{n+1} = A_n \cup \{x/2 \mid x \in A_n\}$. Then the set $A_\infty = \bigcup_n A_n$ is the result of the generation algorithm. If $x \in A$ then $x \in A_k$ for some k and $x/2 \in A_{k+1} \subset A_\infty$. Therefore A_∞ is binarily division invariant and the generation algorithm terminates at A_∞.

Exercise 6.5 A subset of $\mathbb{R}$ is said to be binarily division invariant and zero-happy if dividing any element by 2 results in an element of the subset, it is closed, and, in addition, if it contains zero then it also contains 2. Check that intersection-enclosure applies. Also analyze generation for this property and show that it does not always terminate with A_∞.

Solution:

1. Check that intersection-enclosure applies.
   Checking intersection-enclosure for closure was done in Exercise 6.1, and in Exercise 6.4 for binary division invariance. All that remains is to check the third property—if the set contains zero then it also contains 2. Let S be the intersection of all binarily division invariant and zero happy sets containing a subset A of $\mathbb{R}$. If $2 \in S$ then 2 is in every such set containing A, and by definition all such

sets contain zero as well. As S is the intersection of those binarily division invariant and zero happy sets containing A, S also contains zero. Therefore intersection-enclosure applies to all of the conditions for binarily division invariant and zero happy sets.

2. Analyze generation for this property.
   Let A_0 be a subset of $\mathbb{R}$ and let A_{n+1} be the set

$$A_n \cup \{x/2 \mid x \in A_n\} \cup \{x_n \to x \mid x_n \in A_n \text{ and } x \in \mathbb{R}\}$$

   with the additional property that $A_{n+1} = A_{n+1} \cup \{2\}$ if $0 \in A_n$. Let $A_\infty = \bigcup_n A_n$. All that remains is to check that the properties hold for A_∞. Let x be an element of A_∞. Then $x \in A_n$ for some A_n and $x/2 \in A_{n+1} \subset A_\infty$, so the first property holds. If $0 \in A_\infty$ then $0 \in A_n$ for some n so that $2 \in A_{n+1} \subset A_\infty$ and the third property holds. The closure condition requires that the limit of every convergent sequence in A_∞ is contained in A_∞, but the limit of a sequence such that $x_n \in A_n$ will not be in A_∞. For example, the output A_∞ of the generation algorithm for the set $A_0 = \{1\}$ will contain arbitrarily small binary divisions of 1, but the limit of the convergent sequence $\frac{1}{2^n} \to 0$ is not contained in any A_n. In such cases the generation algorithm does not terminate at infinity and must be repeated.

Exercise 6.6 A subset of $\mathbb{R}$ is said to be *open* if its complement is closed. Will there be a smallest open subset containing $[0, 1]$?

Solution: Let $\{S_\alpha\}_{\alpha\in I}$ be the set of open subsets containing $[0, 1]$ and let S be the intersection $\bigcap_{\alpha\in I} S_\alpha$. Let x_n be a sequence such that $x_n \notin S$. The openness property is satisfied if the limit of the sequence $x_n \to x \notin S$ for every such sequence. Consider the sequence $a_n = -\frac{1}{n}$. For every a_n there is an open subset $S_n = [-\frac{1}{n+1}, 1]$ such that S_n is an open subset containing $[0, 1]$ and $a_n \notin S_n$. Since S is the intersection of all such open subsets, $a_n \notin S$. But S contains the limit of a_n which is 0. Therefore the smallest open subset containing $[0, 1]$ does not exist.

Exercise 6.7 Show that if p is a polynomial with real coefficients and $p(z) = 0$ then $p(\bar{z})$ is also zero. Use this fact in conjunction with the fundamental theorem of algebra to show that every real polynomial of odd order has a real zero.

Solution: Let

$$p(x) = a_n x^n + a_{n-1} x^n - 1 + \cdots + a_1 x + a_0$$

with all the a_i's real and suppose $p(z) = 0$. Then if $z = a + bi$, we have

$$p(z) = a_n(a + bi)^n + a_{n-1}(a + bi)^{n-1} + \cdots + a_1(a + bi) + a_0 = 0.$$

This expression simplifies to a real expression in terms of powers of a, even powers of b and the a_i's, plus a complex term with a coefficient in terms of powers of a, odd powers of b and the a_i's. Since $z = a + bi$ is a zero, both the real expression and the coefficient of i evaluate to zero. The conjugate $\bar{z} = a - bi$ can be interpreted as either taking the negative of b or the negative of i. In the first interpretation, the real expression is unchanged because it only involves even powers of b. In the second interpretation, if the coefficient of i evaluates to zero then replacing i with $-i$ still results in a zero complex part. Therefore if z is a zero of a real polynomial p then its conjugate $\bar{z}$ is a zero as well.

7 Differences of Invariants

Exercise 7.1 What sizes of board $m \times n$, can be covered by 2×2 squares?

Solution: Boards of size $m \times n$ can be covered by 2×2 squares if and only if $2 \mid m$ and $2 \mid n$. Trivially a 2×2 square can cover a board such that $m = 2$ and $n = 2$. If $m = 2$ and $n = 2k$, then the board can be covered by a single row consisting of k copies of a 2×2 square. If $m = 2l$ and $n = 2k$ then the board can be covered by l copies of a $2 \times 2k$ board. Therefore if a board of size $m \times n$ such that $2 \mid m$ and $2 \mid n$ then it can be covered by 2×2 squares.

To show the converse, if a board of size $m \times n$ that can be covered by 2×2 squares then so can a board of size $m - 2 \times n - 2$ provided m and n are greater than 2. A $1 \times n$ or $m \times 1$ board cannot be covered by 2×2 squares, so any $m \times n$ board such that $m - 2k = 1$ or $n - 2k = 1$ cannot be covered. These are precisely the boards with odd m or n. Therefore if a $m \times n$ board can be covered by 2×2 squares then m and n are both even.

Exercise 7.2 Define multiplication on $\mathbb{Z}_k$ by taking xy and then taking its remainder on division by k. For what values of x and k does there exist a number y such that $xy = 1$?

Solution: If the remainder of xy divided by k is 1, then $xy + qk = 1$. Given x and k there exist integers y and q satisfying this equation if and only if 1 is the greatest common factor of x and k. Therefore there exists a number y such that $xy = 1$ on $\mathbb{Z}_k$ as defined if x and k are coprime.

Exercise 7.3 Let n be a positive integer. Let $f(n)$ denote the sum of its digits. Show that the remainder on division by 9 is invariant under passing from n to $f(n)$. Use this to show that if we keep summing up the digits of n until we get a number less than 10 then n is divisible by 9 if and only if this last number is 9. Repeat and reformulate the result for 3. More generally, does this approach work for any other numbers? What if we change our number base?

Solution: The base 10 representation of a positive integer n can be expressed by

$$n = n_k 10^k + n_{k-1} 10^{k-1} + \cdots + n_1 10 + n_0$$

where the digits of n are $n_k n_{k-1} \ldots n_1 n_0$. In terms of this expression, we can write $f(n) = n_0 + n_1 + \cdots + n_k$. Using the fact that $10 \equiv 1 \pmod 9$, we have

$$n \equiv n_k 1^k + n_{k-1} 1^{k-1} + \cdots + n_0 \equiv n_k + n_{k-1} + \cdots + n_0 \equiv f(n) \pmod 9$$

which shows that divisibility by 9 is invariant between n and $f(n)$. By repeatedly passing from n to $f(n)$ one can reduce n to a single digit number. The only non-zero single digit number divisible by nine is 9, so n is divisible by 9 if and only if the termination of this process results in 9.

Because $10 \equiv 1 \pmod 3$ the same result holds for 3. If the base 10 representation of n is congruent to zero modulo 3, so is the sum of its digits $f(n)$. At the termination of the procedure, if the single digit output $f^k(n)$ is either 3, 6 or 9 then we can conclude that 3 divides n.

In general, given a positive integer in a base b, divisibility by d is invariant under taking the sum of the digits if $b \equiv 1 \pmod d$. Only under those conditions does this process work.

8 Linear Dependence, Fields and Transcendence

Exercise 8.1 Show that if $y \in \mathbb{C}$ is algebraic and $x^k = y$ then x is also algebraic.

Solution: If $y \in \mathbb{C}$ is algebraic then for some polynomial p, $p(y) = 0$. Let q be the polynomial $p(y^k)$. Then $q(x) = p(x^k) = p(y) = 0$ and x is also algebraic.

Exercise 8.2 Show that if y is transcendental then y^k is transcendental for all counting numbers k. Show also that $y^{1/k}$ is transcendental.

Solution: It is sufficient to prove the contrapositive, that if y^k is algebraic then y is algebraic. Exercise 8.1 showed that if y is algebraic and $y = x^k$ then x is algebraic. Therefore for transcendental y, y^k is transcendental for all counting numbers k.

Algebraic numbers are closed under multiplication. If x is algebraic, then so is $x \cdot x = x^2$. If x^n is algebraic, then $x \cdot x^n = x^{n+1}$ is algebraic. By induction, if x is algebraic then for all counting numbers k, x^k is algebraic. In other words, $y^{1/k}$ is algebraic implies $y^{(1/k)^k} = y^{(1/k)k} = y$ is algebraic. Therefore the contrapositive is true and if y is transcendental then $y^{1/k}$ is transcendental.

Alternate proof: If $y^{1/k}$ is transcendental then the set

$$S_1 = \{y^{j/k} \mid j = 0, \ldots, n\}$$

is linearly independent over $\mathbb{Q}$ for all counting numbers n. Any subset of a linearly independent set is also linearly independent, so the set

$$S_2 = \{y^{j/k} \mid j = 0, k, 2k, \ldots, nk\} \subset S_1$$

is linearly independent, and there are no linear combinations of powers of y with rational coefficients that are equal to zero. Equivalently, y is transcendental.

Exercise 8.3 Show that the smallest sub-field of $\mathbb{R}$ which is closed under the taking of the powers $1/2$, $1/3$, $1/5$ is contained in the set of algebraic numbers.

Solution: Let $\{S_\alpha\}_{\alpha \in I}$ be the set of subfields of $\mathbb{R}$ closed under the taking of the powers $1/2$, $1/3$ and $1/5$ indexed by an arbitrary index set I. The smallest set satisfying the above properties is the intersection $\bigcap_{\alpha \in I} S_\alpha$ of all such sets. Any set satisfying the above properties therefore contains the smallest set satisfying those properties. A real number y is algebraic if there is a nonzero rational polynomial p such that $p(y) = 0$. If such a polynomial p exists then a polynomial q exists for $x = y^{1/2}$ with the definition $q(z) = p(z^2)$ such that $q(x) = p(x^2) = p(y) = 0$. Likewise if $x = y^{1/3}$ or $x = y^{1/5}$ then $q(z) = p(z^3)$ and $q(z) = p(z^5)$ respectively. Therefore algebraic numbers are closed under the taking of the power $1/2$, $1/3$, $1/5$ and the smallest such set is contained in the set of algebraic numbers.

9 Formal Equivalence

Exercise 9.1 Using a ruler and compass construction, duplicate the square.

Solution: To duplicate a square means given a square of side length x, construct a square of area $2x^2$. Let the points $ABCD$ form a square. Centered at point A, draw a circle of radius AB and extend the line AB through the circle. Let E be the intersection of this line with the circle. Also extend the line AD through the circle and call the intersection F. Draw lines BD, DE, EF and FB. The points $BDEF$ form a square. Each side is the length BD, the diagonal of the square $ABCD$. If each side of the original square is length x, then the length of BD is $x\sqrt{2}$ by the Pythagorean theorem. Therefore the area of $BDEF$ is $2x^2$ as desired.

Exercise 9.2 Show directly that if a, b, r are rationals with $r > 0$, then $a + b\sqrt{r}$ satisfies a polynomial with rational coefficients.

Solution: Let $p(x) = x^2 - 2ax - (b^2r - a^2)$. Then

$$\begin{aligned} p(a + b\sqrt{r}) &= (a + b\sqrt{r})^2 - 2a(a + b\sqrt{r}) - (b^2r - a^2) \\ &= a^2 + 2ab\sqrt{r} + b^2r - 2a^2 - 2ab\sqrt{r} - b^2r + a^2 \\ &= 0 \end{aligned}$$

10 Equivalence Extension

Exercise 10.1 Prove that if k is a positive integer and x is a positive real number then x has a kth root in the real numbers.

Solution: Let E be the set $E = \{y \mid y^k \leq x\}$. This set contains 0 so it is non-empty, and it is bounded above by the maximum of 1 and x. The real numbers observe the least upper bound property, so the set E has a least upper bound l. It is necessary to show that $l = \sqrt[k]{x}$.

Suppose $l^k < x$ and let the sequence

$$l_n = l + \frac{1}{n}$$

so that $l_n > l$ for all n. The sequence $\{l_n\}$ converges to l, so there is an N such that l_n is arbitrarily close to l when $n \geq N$. Choose N such that $l_n^k - l^k < x - l^k$ when $n \geq N$. This implies $l_n^k < x$ and $l_n \in E$. Thus l_n is an element of E greater than l which contradicts the assumption that l is an upper bound. Then we can conclude $l^k \geq x$.

Now suppose $l^k > x$. The sequence

$$l_n = l - \frac{1}{n}$$

which also converges to l poses the same problem. For all n we have $l_n < l$ but there exists n such that $l_n^k > x$. Then l_n is an upper bound of the set E greater than l, which contradicts the assumption that l is the least upper bound. We conclude that $l^k = x$ and l is the kth root of x.

Exercise 10.2 Prove that every subset of the reals that is bounded below has a greatest lower bound.

Solution: Let E be a subset of the real numbers that is bounded below and let L be the set of lower bounds of E, the set

$$L := \{l : l \leq x, \forall x \in E\}.$$

The set L is nonempty since E is bounded below, and L is bounded above by any x in E. Let u be the least upper bound of L in the real numbers guaranteed by the least upper bound property. We show that u is a lower bound of E and conclude that it is the greatest lower bound.

Suppose $u \notin L$, u is not a lower bound of E. Then there exists x in E such that $x < u$. For every lower bound l of E we have $l \leq x$. Then x is an upper bound of L. The existence of an upper bound of L less than u contradicts the assumption that u is the least upper bound of L. So we conclude $u \in L$.

We have by definition of the least upper bound that $u \geq l$ for all $l \in L$. Thus u is a lower bound of E and it is the greatest lower bound of E.

Exercise 10.3 Construct the positive rationals directly from the natural numbers, and then construct the negative rationals with them. Establish a bijection between the rationals constructed this way and the rationals constructed the original way. Ensure that the bijection is the identity on the natural numbers and that it commutes with multiplication and division.

Solution: Let S_1 be the set $\mathbb{N} \times (\mathbb{N} - \{0\})$ and define an equivalence relation $\sim$ on S such that $(p_1, q_1) \sim (p_2, q_2)$ if $p_1 q_2 = p_2 q_1$. The fact that the equivalence relation is reflexive and symmetric follows from those properties of the equality relation. We will show that $\sim$ is transitive as well. If

$$(p_1, q_1) \sim (p_2, q_2) \text{ and } (p_2, q_2) \sim (p_3, q_3)$$

then

$$p_1 q_2 = p_2 q_1 \text{ and } p_2 q_3 = p_3 q_2.$$

So

$$q_2 = p_2 q_1 / p_1 \text{ and } q_2 = p_2 q_3 / p_3$$

and this division is well defined since it has an answer in $\mathbb{N}$. This implies

$$p_1 p_2 q_3 = p_3 p_2 q_1.$$

If $p_2 = 0$ then $p_1 q_2 = p_2 q_1 = 0$. Since $q_1 \in \mathbb{N} - \{0\}$ it cannot be zero. Therefore $p_1 = 0$. Similarly $p_3 q_2 = p_2 q_3 = 0$ and $p_3 = 0$. We have that $p_1 q_3 = p_3 q_1 = 0$ and conclude $(p_1, q_1) \sim (p_3, q_3)$.

If p_2 is not zero, then the division

$$p_1 p_2 q_3 / p_2 = p_3 p_2 q_1 / p_2$$

is defined in $\mathbb{N}$ and we have

$$p_1 q_3 = p_3 q_1$$

so that $(p_1, q_1) \sim (p_3, q_3)$ as desired, and $\sim$ is an equivalence relation.

We call the set of equivalence classes of $\sim$ on $\mathbb{N} \times (\mathbb{N} - \{0\})$ the nonnegative rational numbers, Q^+. Let (p, q) be the representative element of each equivalence class $[(p, q)]$ such that p and q are co-prime. If some p and q have a common factor u then $[(p/u, q/u)] = [(p, q)]$ so the choice of representative is well defined.

Next we define addition and multiplication on Q^+. We define

$$[(p_1, q_1)] \cdot [(p_2, q_2)] = [(p_1 p_2, q_1 q_2)]$$

and show that $\cdot$ is well defined. That is, the product of two equivalence classes must be consistent across any choice of representative. We show that

$$(p_1, q_1) \sim (\tilde{p}_1, \tilde{q}_1), (p_2, q_2) \sim (\tilde{p}_2, \tilde{q}_2) \implies (p_1 p_2, q_1 q_2) \sim (\tilde{p}_1 \tilde{p}_2, \tilde{q}_1 \tilde{q}_2).$$

The definition of the equivalence relation gives

$$p_1 p_2 \tilde{q}_1 \tilde{q}_2 = (p_1 \tilde{q}_1)(p_2 \tilde{q}_2) = (\tilde{p}_1 q_1)(\tilde{p}_2 q_2) = \tilde{p}_1 \tilde{p}_2 q_1 q_2$$

which proves the result.

We define addition on Q^+ by

$$[(p_1, q)] + [(p_2, q)] = [(p_1 + p_2, q)]$$

and show that addition is well defined for any choice of q. First, there always exists such a q for if $x = [(p_1, q_1)]$ and $y = [(p_2, q_2)]$ then we can choose the representation

$$x = [(p_1 q_2, q_1 q_2)], y = [(p_2 q_1, q_1 q_2)]$$

such that $q = q_1 q_2$ and addition is defined for this q. Consider q', the least common multiple of q_1 and q_2. For any other choice of q we have $q = r q'$. If we have

$$x = [(s_1, q')], y = [(s_2, q')]$$

with the least common multiple in the second coordinate, then then sum is

$$x + y = [(s_1 + s_2, q')].$$

For any other choice of q we have

$$[(rs_1, rq')] + [(rs_2, rq')] = [(r(s_1 + s_2), rq')]$$

and the fact that

$$(r(s_1 + s_2), rq') \sim (s_1 + s_2, q')$$

shows that the representations agree. Order is defined on Q^+ by $[(p_1, q)] > [(p_2, q)]$ if $p_1 > p_2$. Every pair of equivalence classes has a representation such that the second coordinates agree, so the relation is well defined. Note that, as in $\mathbb{N}$, subtraction sum $x - y$ is only defined on Q^+ if $x \geq y$.

Let S_2 be the set $Q^+ \times Q^+$ and define the equivalence relation $\sim$ on S_2 such that $(x_1, y_1) \sim (x_2, y_2)$ if $x_1 + y_2 = x_2 + y_1$. As before, reflexivity and symmetry follow from those properties of the $=$ relation. For transitivity, if $(x_1, y_1) \sim (x_2, y_2)$ and $(x_2, y_2) \sim (x_3, y_3)$ then

$$y_2 = x_2 + y_1 - x_1 \text{ and } y_2 = x_2 + y_3 - x_3$$

and the subtraction is well defined because it has an answer in $\mathbb{Q}$, which is y_2. This gives

$$x_2 + y_1 - x_1 = x_2 + y_3 - x_3$$

and rearranging terms yields

$$x_3 + y_1 = x_1 + y_3$$

which is the equivalence condition $(x_1, y_1) \sim (x_3, y_3)$ and thus $\sim$ is transitive.

The equivalence classes are the pairs of rational numbers (x, y) with the same sum $x - y$. Let Q be the set of equivalence classes of S_2 under the equivalence relation $\sim$. Each equivalence class can be distinguished by a unique member in which at least one entry is zero. For if $x_1 > y_1$ then $x_1 - y_1$ is defined on Q^+ and we can write

$$(x_1 - y_1) + y_1 = x_1 + 0$$

which is the equivalence condition for

$$(x_1 - y_1, 0) \sim (x_1, y_1).$$

Likewise if $y_1 > x_1$ then $y_1 - x_1$ is defined on Q^+ and

$$0 + y_1 = x_1 + (y_1 - x_1)$$

which shows that

$$(0, y_1 - x_1) \sim (x_1, y_1).$$

In the case that $x_1 = y_1$, we have

$$0 + y_1 = x_1 + 0$$

so that

$$(0, 0) \sim (x_1, y_1).$$

We define addition on Q by addition of the representative elements:

$$[(x,0)]+[(y,0)]=[(x+y,0)]$$
$$[(x,0)]+[(0,y)]=[(x,y)]$$
$$[(0,x)]+[(0,y)]=[(0,x+y)].$$

We also define multiplication in terms of the representative elements:

$$[(x,0)]\cdot[(y,0)]=[(xy,0)]$$
$$[(x,0)]\cdot[(0,y)]=[(0,xy)]$$
$$[(0,x)]\cdot[(0,y)]=[(xy,0)].$$

Let the set $\mathbb{Q}$ be the rationals constructed the original way. We will show that there is a bijection between Q and $\mathbb{Q}$ that is compatible with multiplication and division. Let $j\colon Q^+\to\mathbb{Q}$ be such that

$$j\colon [(p,q)]\mapsto[(i(p),i(q))]$$

where i maps the natural number p to the integer $(p,0)$ defined in the original way. Let $k\colon Q\to\mathbb{Q}$ be the map

$$k\colon \begin{aligned}(x,0)&\mapsto j(x)\\(0,x)&\mapsto -j(x)\\(0,0)&\mapsto j(0)\end{aligned}$$

such that k is defined on the representative elements of Q. If n is a natural number in Q then it is in the equivalence class represented by $([(n,1)],0)$ and we have

$$k((n,1),0)=j(n,1)=i(n)-i(0)=n$$

as desired.

In the case of multiplication and division it is necessary to check that k commutes with those operations for every pair of representative elements:

$$k((p_1,q_1),0)\cdot k((p_2,q_2),0)=j(p_1,q_1)\cdot j(p_2,q_2)=(i(p_1),i(q_1))\cdot(i(p_2),i(q_2))$$

which is the product of two elements of $\mathbb{Q}+$ and is defined by

$$[(p_1,q_1)]\cdot[(p_2,q_2)]=[(p_1p_2,q_1q_2)].$$

Then we have

$$(i(p_1),i(q_1))\cdot(i(p_2),i(q_2))=(i(p_1p_2),i(q_1q_2))=j(p_1p_2,q_1q_2)=k((p_1p_2,q_1q_2),0)$$

which is

$$k(((p_1,q_1),0)\cdot((p_2,q_2),0))$$

as desired.

Multiplicative inverses on Q are defined in terms of the representative elements of its equivalence classes, such that

$$((p,q),0)^{-1}=((q,p),0)\text{ and}$$
$$(0,(p,q))^{-1}=(0,(q,p)).$$

Then under composition with k,

$$k(((p,q),0)^{-1}) = k((q,p),0) = j(q,p) = (i(q),i(p)) = (i(p),i(q))^{-1} = j(p,q)^{-1} = k((p,q),0)^{-1}$$

and division commutes with the bijection.

Exercise 10.4 Give examples of subsets of the reals that do not have greatest lower bounds.

Solution: Any subset of the reals that is bounded below has a greatest lower bound. Therefore the sets that do not have greatest lower bounds are not bounded below. Examples of such sets are:

$$\{x \mid x < 0\}$$
$$\{x \mid x \leq 0\}$$
$$\{x \mid x = (-1)^n n^2 \text{ for } n \in \mathbb{N}\}$$

Exercise 10.5 Show that a monotone decreasing sequence of real numbers is bounded below if and only if it is convergent.

Solution: Let $\{a_n\}$ be a monotone decreasing sequence in $\mathbb{R}$. Suppose it is bounded below, and let L be the set of lower bounds. The least upper bound property dictates that the supremum l of L is in $\mathbb{R}$. Given $\epsilon > 0$, we have that $l+\epsilon$ is not a lower bound of $\{a_n\}$. Then there is a N such that $a_N < l+\epsilon$. Since $\{a_n\}$ is monotone decreasing, if $n > N$ then $a_n - l < \epsilon$. Therefore the sequence $\{a_n\}$ is convergent and it converges to l.

Let $\{a_n\}$ be a monotone decreasing sequence in $\mathbb{R}$, and suppose it is not bounded below. Given l and ϵ, there is an a_N such that $a_N < l - \epsilon$. Because a_n is monotone decreasing, for $n > N$ we have that $a_n < a_N < l - \epsilon$ and by rearranging we have $l - a_n > \epsilon$. Therefore a_n cannot converge to l.

11 Proof by Classification

Exercise 11.1 Find all Pythagorean triples with one side equal to 12.

Solution: These are positive integer triples (k, u, v) such that one of

$$k(v^2 - u^2) = 12 \tag{1}$$
$$2kuv = 12 \tag{2}$$
$$k(u^2 + v^2) = 12 \tag{3}$$

is true, with u and v co-prime, $u + v$ odd and $u < v$. Looking at the first equation, we can see that

$$k(v^2 - u^2) = 12$$

if and only if $v^2 - u^2$ divides 12. We have $u + v \leq 12$ if $v^2 - u^2 = (v - u)(v + u) \leq 12$. There are 14 co-prime pairs u, v such that their sum is odd and less than or equal to 12: (1, 2) (1,4) (1,6) (1,8) (1,10) (2,3) (2,5) (2,7) (2,9) (3,4) (3,8) (4,5) (4,7) and (5,6). Of these, there are 3 such that their sum divides 12: (1,2) (1,5) (1,11). However, only (1,2) is such that $v^2 - u^2$ divides 12. Then the first equation gives (1,2) as a generating pair. That triple is:

$$(4, 1, 2) \mapsto (12, 16, 20).$$

The second equation $2kuv = 12$ requires co-prime pairs $(u, v), u < v, u + v$ odd such that uv divides 6, that is $uv = 1, 2, 3$ or 6. Valid pairs are (1,2) (2,3) and (1,6). These generate the following Pythagorean triples:

$$(3, 1, 2) \mapsto (9, 12, 15)$$
$$(1, 2, 3) \mapsto (5, 12, 13)$$
$$(1, 1, 6) \mapsto (35, 12, 37).$$

The third equation $k(u^2 + v^2) = 12$ specifies the length of the hypotenuse. Candidate pairs will be those such that u and v are positive integers with squares less than 12. Integers of this form are 1, 2 and 3 which generate (1,2) and (2,3) as candidate pairs. However, we have $1^2 + 2^2 = 5$ and $2^2 + 3^2 = 13$ neither of which divide 12. Therefore there are no Pythagorean triples with a hypotenuse of length 12.

In summary there are four Pythagorean triples with a side of length 12, and they are (12, 16, 20) (9, 12, 15) (5, 12, 13) and (35, 12, 37).

Another way to look at this is to look at generating triples with $k = 1$ such one of the three sides is less than 12. These are:

$$(1, 1, 2) \mapsto (3, 4, 5)$$
$$(1, 1, 4) \mapsto (15, 8, 17)$$
$$(1, 1, 6) \mapsto (35, 12, 37)$$
$$(1, 2, 3) \mapsto (5, 12, 13)$$
$$(1, 3, 4) \mapsto (7, 24, 25).$$

By varying k on those triples with side length that divides 12, we have

$$
\begin{aligned}
(3,1,2) &\mapsto (9,12,15)\\
(4,1,2) &\mapsto (12,16,20)\\
(1,1,6) &\mapsto (35,12,37)\\
(1,2,3) &\mapsto (5,12,13)
\end{aligned}
$$

which are the same triples as before.

Exercise 11.2 Suppose a, b, c, k are positive integers and

$$c^k = ab,$$

with a, b co-prime. Does it follow that a and b are the kth powers of positive integers?

Solution: In terms of prime factors, we can write

$$c = c_1^{\phi_1} \cdot c_2^{\phi_2} \cdot c_3^{\phi_3} \cdots c_n^{\phi_n}$$

and to the k power

$$c^k = c_1^{k\phi_1} \cdot c_2^{k\phi_2} \cdot c_3^{k\phi_3} \cdots c_n^{k\phi_n}.$$

Since $ab = c^k$, a and b are each the product of prime factors c_i of c such that, if $c_i^{\beta_i}$ is the greatest power of c_i that divides b, then $c_i^{k\phi_i - \beta_i}$ divides a and is the greatest power of c_i that divides a. This is to ensure that $c_i^{k\phi_i-\beta_i} c_i^{\beta_i} = c_i^{k\phi_i}$ divides $c^k = ab$. So we can write

$$
\begin{aligned}
a &= c_1^{k\phi_1 - \beta_1} \cdots c_n^{k\phi_n - \beta_n}\\
b &= c_1^{\beta_1} \cdots c_n^{\beta_n}.
\end{aligned}
$$

for $0 \le \beta_i \le k\phi_i$. Since a and b are co-prime, there is no factor c_i such that c_i divides both a and b. If $0 < \beta_j < k\phi_j$ for any j then β_j divides both a and b. Therefore either $\beta_i = 0$ or $\beta_i = k\phi_i$ for all i. Since every prime factor of a and b has a power that is a multiple of k, this implies that a and b are the kth power of positive integers.

Exercise 11.3 Suppose we want to find all rational Pythagorean triples, that is right-angled triangles with all sides rational. Can we classify these?

Solution: For any rational Pythagorean triple (x, y, z) we can multiply by the greatest common denominator to get an integer Pythagorean triple, (kx, ky, kz). Inverting this operation, we have that every rational Pythagorean triple is of the form $(a/k, b/k, c/k)$ where k is rational and (a, b, c) is an integer Pythagorean triple. Then the rational Pythagorean triples are also classified by positive integer triples (k, u, v) where $u < v$, $u + v$ odd, and u, v co-prime; however, rational triples are generated by rational k.

Exercise 11.4 Consider the map from Pythagorean triples to the length of the longest side as a natural number. Is this map surjective? Is it injective? What about for the other two sides?

Solution: The length of the longest side of a right triangle described by a Pythagorean triple (a, b, c) is the length of the hypotenuse c. We have that a Pythagorean triple (a, b, c) can take on values c for its longest side such that

$$c = k(u^2 + v^2)$$

for positive integer k and co-prime u, v such that $u + v$ is odd and $u < v$. The map from a Pythagorean triple to the length of its longest side, $h\colon (a, b, c) \mapsto c$ is therefore the composition of the map

$$g\colon (k, u, v) \mapsto k(u^2 + v^2)$$

and the bijection

$$f\colon (a, b, c) \mapsto \left(\gcd(a, b, c), \frac{\sqrt{(c-a)/2)}}{\gcd(a, b, c)}, \frac{\sqrt{(c+a)/2}}{\gcd(a, b, c)}\right)$$

i.e. the map $h = g \circ f$. So the possible values of c are in the range of $k(u^2 + v^2)$ with k, u, v having the above. properties. The map g is not surjective in the natural numbers, since the smallest values for (k, u, v) are $(1, 1, 2) \mapsto 5$. If $u = 1$, then $c = k(1 + v^2)$ where v is even. For any k of the form $(1 + 4p^2)$, we have that

$$g(k, u, v) = g(1 + 4p^2, 1, v) = (1 + (2p)^2)(1 + v^2) = g(1 + v^2, 1, 2p)$$

and g is not injective. For example, $g(17, 1, 2) = 17(1 + 2^2) = 5(1 + 4^2) = g(5, 1, 4)$.

For the other two sides we have the maps

$$g_a\colon (k, u, v) \mapsto k(v^2 - u^2)$$
$$g_b\colon (k, u, v) \mapsto 2kuv$$

where $h_a = g_a \circ f$ and $h_b = g_b \circ f$. Neither map is surjective, for 1 is not in the image of g_a or g_b over the range of allowable values (k, u, v). By a similar argument as above g_a is not injective, since if $k = 4p^2 - 1$ then $g_a(k, 1, v) = g_a(v^2 - 1, 1, 2p)$. Nor is g_b injective, for if $2u_1v_1$ divides u_2v_2 then there is a k such that $g_b(k, u_1, v_1) = g_b(1, u_2, v_2)$. An example is $g_b(10, 1, 2) = g_b(1, 4, 5)$.

Exercise 11.5 Suppose we allow triangles to have negative side lengths. What differences does it make to the classification of Pythagorean triples?

Solution: If (a, b, c) is a Pythagorean triple with integer, possibly negative, values then so too is $(|a|, |b|, |c|)$ a Pythagorean triple, where each coordinate is the absolute value. Therefore it (k, u, v) is such that u, v are co-prime, $u + v$ odd and $u < v$, and the map $f\colon (k, u, v) \mapsto (a, b, c)$ classifies positive integer Pythagorean triples, then

$$g : (a, b, c) \mapsto \{((-1)^i a, (-1)^j b, (-1)^k c) \mid i, j, k \in \{0, 1\}\}$$

is the map such that $g \circ f$ classifies Pythagorean triples with possibly negative side lengths.

Exercise 11.6 How many Pythagorean triples have two sides differing by 1?

Solution: We have that

$$c - b = k(u^2 + v^2) - 2kuv = k(u^2 + v^2 - 2kuv) = k(v - u)^2$$

which is 1 whenever $k = 1$ and $v = u + 1$. In such cases u and v are co-prime, $u + v = 2u + 1$ and $u < v$ so $(1, u, u + 1)$ generates a pythagorean triple with two sides differing by 1 for any integer $u > 0$. There are infinitely many such triples.

12 Specific-generality

Exercise 12.1 Suppose we can prove that every map in which all countries have at least k sides can be coloured with k colours. Show that every map can then be coloured with k colours.

Solution: The proof is by induction. A map with less than k countries can be colored with k colors by picking a different color for each country. Suppose a map with n countries can be colored with k colors. Any map with $n+1$ countries such that each country has k or more sides can be colored by assumption. Otherwise, if a map has $n+1$ countries and there is a country with p sides, $p < k$, merging the country with one of its neighbors produces a map with $n+1$ countries which can be colored by the inductive hypothesis. When the merged country is reinserted into the colored map, the p adjacent countries can have at most p different colors, and therefore there are at least $k-p$ choices of colorings for the reinserted country that produce a valid k-colored map. We have shown that, under the general assumption that maps in which countries have at least k sides can be colored, if any map with n countries can be k-colored, then any map with $n+1$ countries can be k-colored. The result follows by induction.

Exercise 12.2 Show that if there are no solutions to the Fermat equation with x, y, z pairwise co-prime then there are no solutions in general.

Solution: Suppose

$$x^n + y^n = z^n$$

is a solution to the Fermat equation, and there is some p that divides two of x, y, z. Assume p divides x snd p divides y. Then $x = pk$, $y = kl$ and

$$p^n(k^n + l^n) = z^n$$

so that p divides z, and $z = pm$. Dividing by p^n,

$$k^n + l^n = m^n$$

which is a solution to the Fermat equation where p is no longer a common divisor. This proves the contrapositive of the desired statement—if there is a solution to the Fermat equation in general, there is a solution with co-prime x, y, z.

13 Diagonal Tricks and Cardinality

Exercise 13.1 Suppose X and Y are the natural numbers. We define f to be multiplication by two from X to Y. We define g to be multiplication by three. What is the bijection h constructed by the Schröder–Bernstein theorem? What if both maps are multiplication by 2?

Solution: When $f(x) = 2x$ and $g(x) = 3x$ the Schröder–Bernstein theorem gives the bijection

$$h(x) = \begin{cases} 2x & \text{if } 3 \nmid x \text{ or } 2 \mid x \\ \frac{x}{3} & \text{if } 3 \mid x \text{ and } 2 \nmid x \end{cases}$$

and with $f(x) = 2x$ and $g(x) = 2x$ the bijection is

$$h(x) = \begin{cases} 2x & \text{if } 2 \nmid x \\ \frac{x}{2} & \text{if } 2 \mid x \end{cases}.$$

Exercise 13.2 Show that every real number can be represented by a decimal that goes on forever.

Solution: Let x be a real number and m be the least integer such that

$$x < 10^m.$$

Let the sequence a_n be defined such that

$$x - \sum_{i=1}^{n} a_i 10^{m-i} < 10^{m-n}.$$

Then for any ϵ, there is a n such that $\epsilon > 10^{m-n}$ and $x - \sum_{i=1}^{n} a_i 10^{m-i} < 10^{m-n} < \epsilon$. Therefore x is equal to the infinite sum specified by the sequence $\{a_n\}$, which is the decimal representation of x.

Exercise 13.3 A submarine starts at integer point n. It travels with integer speed k. So after turn m, its location is $n + mk$. A ship which does not know n and k drops a depth charge once a turn on an integer. Show that it is possible to design an algorithm so that the submarine is always hit eventually. What if the speeds are rational? What if they are real?

Solution: The set of possible integer starting point and speed ordered pairs (n, k) is the set $\mathbb{Z} \times \mathbb{Z}$. Since $\mathbb{Z} \times \mathbb{Z}$ is the product of countable sets, it is in bijection with the natural numbers. Let $h \colon \mathbb{N} \to \mathbb{Z} \times \mathbb{Z}$ be a bijection. At turn m, the function $f_{n,k}(m) = n + mk$ gives the location of the submarine for starting condition (n, k). Since h is surjective there exists $r \in \mathbb{N}$ such that $h(r) = (n, k)$. By dropping a depth charge on turn m at the location $f_{h(m)}(m)$ the submarine will be hit on turn r.

If the speeds are rational, then the set of starting conditions is $\mathbb{Q} \times \mathbb{Q}$ is still countable, and there is a bijection $g \colon \mathbb{Q} \times \mathbb{Q} \to \mathbb{N}$. So it is possible to design an algorithm using g that will hit always the submarine with starting conditions (n, k) after a finite number of turns given by $g(n, k)$.

However if the starting conditions are real, the set of starting conditions has larger cardinality than the natural numbers, any any map $k\colon \mathbb{N} \to \mathbb{R} \times \mathbb{R}$ is not surjective. Therefore there is some starting condition not in the image of k and it is not possible to design an algorithm that is guaranteed to hit the submarine.

Exercise 13.4 Consider the set of real-valued functions on the reals. What is the cardinality of this set?

Solution: A real valued function on the reals $f\colon \mathbb{R} \to \mathbb{R}$ can be described by an $\mathbb{R}$-dimensional vector with values in $\mathbb{R}$, a point in the space

$$\prod_{\alpha \in \mathbb{R}} \mathbb{R} = \mathbb{R} \times \mathbb{R} \times \cdots .$$

Let $\mathbb{R}^{\mathbb{R}}$ denote this set, and $F(\mathbb{R})$ the set of real-valued functions on the reals. Then we have $\varphi\colon F(\mathbb{R}) \to \mathbb{R}^{\mathbb{R}}$ such that

$$\varphi(f) = \prod_{\alpha \in \mathbb{R}} f(\alpha).$$

If $x \neq y \in \mathbb{R}^{\mathbb{R}}$ then for some $\alpha \in \mathbb{R}$ we have $x_\alpha \neq y_\alpha$, and the functions f and g for which $\varphi(f) = x$ and $\phi(g) = y$ are such that

$$x = f(\alpha) \neq g(\alpha) = y$$

and φ is injective. Any ordered tuple $K = (k_\alpha) \in \mathbb{R}^{\mathbb{R}}$ indexed by $\mathbb{R}$ corresponds to a function h, $h(\alpha) = k_\alpha$ such that $\varphi(h) = K$, so ϕ is surjective. Therefore ϕ defines a bijection and the cardinality of the set of real-valued functions on the reals is $\mathbb{R}^{\mathbb{R}}$.

Exercise 13.5 Let X be an infinite set show that $X \cup \mathbb{N}$ has the same cardinality as X.

Solution: Suppose X is countably infinite. Then there is an bijection between X and $\mathbb{N}$, and the elements of X can be written x_j. Let the function f be such that $f(x_j) = x_{2j}$ for $x \in X$ and $f(j) = x_{2j+1}$ for $j \in \mathbb{N}$. Then f is an bijection from $X \cup \mathbb{N}$ to X and the two sets have the same cardinality.

Suppose X is uncountably infinite. Then there is no injection from X to $\mathbb{N}$, but there is an injection $f\colon \mathbb{N} \to X$. Let $E \subset X$ be the image of f and let F be its complement in X such that $X = E \cup F$. Since E is countable we have just shown there exists a bijection $g\colon E \to E \cup \mathbb{N}$. Then the extension of g, $\bar{g}\colon E \cup F \to E \cup \mathbb{N} \cup F$ such that $\bar{g}$ is the identity on F, is a bijection between X and $X \cup \mathbb{N}$ and the two sets have the same cardinality.

14 Connectedness and the Jordan Curve Theorem

Exercise 14.1 Show that a subset of $\mathbb{R}^n$ has k components if and only if there exists an integer-valued continuous function that takes k different values.

Solution: An integer-valued continuous function is constant on any straight line in $E \subset \mathbb{R}^n$. If p and q are in the same connected component of E, then there exists a path consisting of a union of straight line segments from p to q. So an integer-valued continuous function $f\colon E \to \mathbb{Z}$ must have the property $f(p) = f(q)$. If E has k components then there are k connected regions E_k for which $f(p) = f(q)$ when $p, q \in E_k$. Let f take the value k on each E_k. Then f is integer-valued continuous and takes k different values.

If there exists an integer-valued continuous function g that takes k different values on E, then there are $p_i, p_j \in E, 1 \leq i, j \leq k$ such that $g(p_i) \neq g(p_j)$. By definition there must be no straight line between any p_i and p_j. If the image of f through E is the set of integers $\{n_k\}$, then the sets $f^{-1}(n_k) \subset E$ for each n_k are the k connected components of E. For if p and q are connected then $f(p) = f(q)$ and $p, q \in f^{-1}(n_k)$ for some k.

Exercise 14.2 Show that if a point, p, can be joined to a connected set, U, by a polygonal path then p and all points of U are in the same component

Solution: Let $r \in U$, then q and r are connected. If there is a path $P_{p,q}$ and a path $P_{q,r}$ then the union of the paths $P_{p,q} \cup Pq, r$ produces a path from p to r. Therefore p is connected to every point in U and they are in the same connected component.

Exercise 14.3 Will a finite intersection of connected sets be connected?

Solution: Let E_k be a finite collection of connected sets. If $\bigcap_k E_k$ is empty or consists of one point then it is trivially connected, so suppose $p, q \in \bigcap_k E_k$. Then p and q belong to E_k for every k, and p and q are connected.

Exercise 14.4 Suppose we have a collection of connected sets U_j such that for all $j > 1$, there is some $k < j$, such that

$$U_j \cap U_k \neq \emptyset,$$

does this guarantee that $\bigcup_j U_j$ is connected?

Solution: Suppose there are 2 sets U_j. Then for U_2, $U_2 \cap U_1 \neq \emptyset$. Therefore U_1 and U_2 are connected. If there are $n+1$ sets U_j and this property holds for up to n sets, then $\bigcup_{j<n+1} U_j$ is connected, and $U_n \cap U_k \neq \emptyset$ for some $k < n$, so U_n and U_k are connected. Therefore $\bigcup_j U_j$ is connected.

Exercise 14.5 Suppose we take the real line and subtract k distinct points, how many components will there be? Prove it.

Solution: Let R be the set $\mathbb{R} - \{p_k\}$. R is divided into $k-1$ convex, connected intervals (p_i, p_{i+1}) for $1 \leq i \leq k$. For if $x < y$ are in an interval (p_i, p_{i+1}), then the interval (x, y) is in (p_i, p_{i+1}) and x and y are connected. Now $R - \{(p_i, p_{i+1})\}$ consists of the points less than p_1 and greater than p_k. If $x, y < p_1$ then there are no points removed in the interval (x, y) and so x and y are connected. Likewise for $x, y > p_k$. Therefore $(-\infty, p_1)$ and (p_k, ∞) are connected components and R consists of $k+1$ components.

15 The Euler Characteristic and the Classification of Regular Polyhedra

Exercise 15.1 Suppose we make two W shapes out of cubes and join the three top bits together, what happens?

Solution: The number of faces is decreased by $3 \times 2 = 6$, the number of edges is decreased by $3 \times 4 = 12$, and the number of vertices is decreased by $3 \times 4 = 12$. The Euler characteristic becomes

$$\chi = \chi_1 + \chi_2 + ((-12) - (-12) - 6) = 2 + 2 - 6 = -2.$$

Exercise 15.2 If we attempt to apply our proof that the Euler characteristic is 2 to a ring shape, where does it fail?

Solution: A ring shape is not convex, so given a point p inside the ring, the projection $\theta_p(q)$ from a point q on the surface of the ring to a point on the surface of a sphere in which the ring is inscribed is not injective. The shape cannot be represented as a network of triangles in the plane.

Exercise 15.3 Suppose we take a square in the plane and cut it into small triangles, what is the Euler characteristic? Suppose we stick opposite sides together, what is the new Euler characteristic? What if we twist one pair of sides before gluing?

. **Solution:** A square has Euler characteristic

$$\chi = V - E + F = 4 - 4 + 1 = 1.$$

There are three ways to cut a square into smaller pieces with straight lines.

1. Draw an edge between two non-adjacent vertices
2. Draw an edge between a vertex and an edge
3. Draw an edge between two edges

In the first case, the number of edges increases by 1 and the number of faces increases by 1. In the second case, there are two added edges—one that was drawn and one from the edge that was bisected—one added vertex, and an additional face. In the third case, three edges, two vertices and one face are added. After applying these cuts to the square, there may still be non-adjacent vertices that share a face. If there are, they may be connected by drawing an edge until the shape consists only of triangles, adding one face and one edge each time. After any of these operations the Euler characteristic remains 1.

For a square, sticking opposite sides together removes the 4 vertices and 4 edges on the boundary, and the Euler characteristic is

$$\chi = 0 + 0 + 1 = 1$$

which is unchanged. Then if the square is cut into triangles, the Euler characteristic depends on the triangles inscribed within the square, that is triangles that do not share an edge with the boundary. Any network of connected triangles has Euler characteristic 1, and gluing the boundary of the square together adds a face that is the complement of the are inside the network. Therefore the new shape has Euler characteristic 2.

Deforming one pair of sides does not change the number of vertices, edges or faces, so the Euler characteristic after gluing is still 2.

Exercise 15.4 Suppose we take a regular polyhedron and form a new one by making each vertex the center of one of the old faces and joining vertices from neighboring faces with edges. What happens for each of the 5 regular polyhedra?

Solution: The five regular polyhedra are

1. Tetrahedron
2. Hexahedron (cube)
3. Octahedron
4. Dodecahedron
5. Icosahedron

In the case of the tetrahedron, joining the centers of neighboring faces with edges creates a smaller tetrahedron with inverse orientation.

Applying this process to the cube creates an octahedron.

Likewise, applying this process to the octahedron creates a cube

Joining the centers of a dodecahedron creates a icosahedron and joining the centers of an icosahedron creates a dodecahedron.

16 Discharging

Exercise 16.1 What happens if we attempt to apply Thurston's discharging proof to a torus? (i.e. the surface of a doughnut.)

Solution: A torus is not convex, so it cannot be projected onto the surface of a sphere.

Exercise 16.2 Is it possible to have a polyhedron whose faces are all pentagons or hexagons? How many pentagons must such a polyhedron have? Suppose we merge two adjacent faces what happens?

Solution: The formula is

$$\sum_k (6-k)N_k = 12$$

where k is the number of sides and N_k is the number of faces with k sides. So if the only k is 5, we have

$$(6-5)N_5 = 12$$

which shows that there must be 12 faces. This describes the character of a dodecahedron. If all of the faces have six sides, then

$$(6-6)N_6 = 12$$

has no solution. If there are faces which are pentagons as well, then

$$(6-5)N_5 + (6-6)N_6 = 12$$

which means there may be any number of hexagons added, but there must be 12 faces which are pentagons. If two adjacent sides are merged, and they are both hexagons, the Euler characteristic doesn't change since the coefficient of N_6 is zero. Merging a pentagon with any other side however will change the Euler characteristic, and the resulting shape will not be a polyhedron.

17 The Matching Problem

Exercise 17.1 The room-mate problem is to pair new students of the same sex into rooms. A pairing is stable if no two students prefer each other to their room-mate. Does the algorithm presented here apply?

Solution: The matching algorithm is not applicable to this problem because it requires the population to be split into two sets such that a valid matching contains one member of each set. When all of the room-mates are the same sex the preference ranking for each student is over the whole set, and there is no way to non-arbitrarily split the students into two groups such that the matching algorithm will generate a stable pairing.

Exercise 17.2 We can rate a matching of all students by summing all the ranks of each member of a pair for his/her partner. Show that there is a stable pairing which minimizes this rating amongst stable pairings. Will it be unique? Design an algorithm to find it.

Solution: Let X and Y be finite sets in a one two one correspondence and S be the set of stable pairings. If $f\colon S \to \mathbb{N}$ is the sum of the ranks of each member of a pair for his or her partner then the image $f(S)$ is a subset of the natural numbers and has a minimum element n. Then $f^{-1}(n)$ is the non-empty set of stable pairings that minimize f.

However, the stable pairing that minimizes f is not unique. In fact, every stable pairing has the same sum of rankings. for if S is a pairing and there is are two pairs that if switched would create a lower sum of rankings, then the switched pair prefer each other to their assigned partners and the pairing S is not stable.

18 Games

Exercise 18.1 Suppose Nim is played with 2 stacks and each player can only take from one of them each turn. Analyze this game.

Solution: If Nim is played with 2 stacks, and the winner is still the player who takes the last match, then the winning state is to have 1-3 matches on your turn. This is the case if your opponent has to choose from one stack 4 matches. Like regular Nim, if your opponent is left with one stack of $4k$ matches, you win. Assume stack 1 has $4k$ matches. If stack 2 has 1-3 matches on your turn, you win. Therefore if stack 2 has 4 matches on your opponent's turn, you win.

The winning strategy for Nim with 2 stacks depends on the starting condition. If both stacks start with $4k$ matches, the second player wins by playing optimally: removing matches from whichever stack has $4k + \{1, 2, 3\}$ matches on your turn. If only one stack has $4k$ matches, the first player wins by removing matches from the other stack so that both stacks have $4k$ matches on the second player's turn. If neither stack has $4k$ matches, the problem reduces to the case where each stack has less than 4 matches.

The first player wins if there is only one stack to take from, so the optimal strategy is to take from whichever stack is greater to leave the same number of matches on each stack. Therefore in the case when each stack 1 starts with $4k + n_1$ matches and stack 2 starts with $4k + n_2$, where $n_1, n_2 \in \{1, 2, 3\}$, the first player wins when $n_1 \neq n_2$, and the second player wins otherwise.

Exercise 18.2 Suppose we play a version of Nim on a clock face. We start with the hour hand at 12 and the first player to get it to 6 wins. Each turn a player can move the hour hand 1–3 integer hours. What happens? What if we add in the rule that the clock cannot show the same time twice? What if we delete any time that has been visited by moving the remaining times closer together? What if we allow a 24-h clock?

Solution: Assuming the clock can only move forward, the first player wins in all cases by moving the clock two hours, so assume the clock can be moved both clockwise and in the counter-clockwise directions. Under the first set of conditions, the game ends in a stalemate as each player can undo their opponent's move.

In the case that the clock cannot revisit the same time twice, notice that the second player wins when there are four previously visited times in the set $\{10, 11, 12, 1, 2\}$ and the clock hand is on the remaining time. Then the second player wins when there are two visited times that set and when there are zero. Since 10 and 2 are not reachable from either position, there are states in which the second player must make a specific move, but through optimal play the second player wins. When visited times are deleted, the restriction that 10 and 2 are mutually unreachable disappears and the second player still wins.

Allowing a 24 hour clock changes the set to $\{10, 11, 12, \ldots, 24, 1, 2\}$. If 16 of the 17 times in this set have been visited and the clock hand is on the remaining time, the second to play wins. Since the set still has an odd number of elements, and after each pair of turns the number of unvisited elements is invariant, the second player wins when 0 times times have been visited and the clock is set to 12, i.e. at the starting conditions.

Exercise 18.3 Suppose the ability to "pass" in chess is added. If both players "pass" then the game is drawn. Can we then prove that there is not an optimal strategy making the second player always win?

Solution: If there is an optimal strategy such that the second to play always wins, then the optimal strategy of the first player is to pass so that the second player is first to play. The optimal strategy for both players is to pass, so the game is drawn and by contradiction there is no optimal strategy making the second player always win.

Exercise 18.4 There are n lions in a cage. A piece of meat is thrown in. What happens under the following conditions?

- The lions are hungry,
- A lion that eats the meat falls asleep.
- An asleep lion is meat to the other lions.
- Each lion and the meat is closest to exactly one other lion.
- The lions are ultra-intelligent.
- The lions prefer to stay alive.

Solution: The winning state for any lion is to be the last lion in the cage with another lion who is asleep, at which point the winning lion both eats and stays alive. While there remains another lion that is not asleep, the optimal strategy is to pass, Since all the lions follow the optimal strategy, nothing happens.

19 Analytical Patterns

Exercise 19.1 Prove or disprove that the following sequences converge. If they converge, identify the limit.

- $x_k = k$.
- $x_k = 1 - \frac{1}{k}$.
- $x_k = \frac{1+k}{2+k^2}$.
- $x_k = \sqrt{k+1} - \sqrt{k}$.

Solution:

1. Suppose $x_k \to x$. Then x_k must get closer to x as k increases. Let n be the closest integer to $x+1$, then for $k > n$ $x_k > n - k$ and the sequence doesn't satisfy the definition of convergence.
2. For this sequence x_k converges to 1. We can write x_k as the sum of the convergent sequences $y_k = 1, y_k \to 1$ and $z_k = \frac{1}{k} \to 0$. So $x_k = y_k + z_k \to y + z = 1$.
3. Dividing by k, the sequence is $x_k = \frac{1/k+1}{2/k+k}$. The numerator converges to 1 and the denominator goes to infinity, so $x_K \to 0$.
4. We can show that x_k is convergent because it is a monotone bounded sequence. For the square root function, $\sqrt{x} > \sqrt{y}$ for all $x > y$, so $\sqrt{x+1} - \sqrt{x}$ is decreasing and bounded below by 0. Therefore it converges, and 0 is the greatest lower bound so it converges to 0.

Exercise 19.2 Suppose $f\colon \mathbb{R}^2 \to \mathbb{R}$, is continuous on the set

$$E = \{1 \leq |(x, y)| \leq 2\},$$

show that f has a maximum and minimum on E. Find an example where these occur on the boundary.

Solution: The region E represents the difference of a circle of radius $\sqrt{2}$ and of radius 1 in the plane. This can be split up into the regions where $\{1 \leq |(x, y) \leq 1.5\}$ and $1.5 \leq |(x, y)| \leq 2|$. Now if one region has an (x', y') such that $f(x', y') < f(x, y)$ for (x, y) in the other region, then let (x_0, y_0) be that (x', y') and split the region in two and repeat the process. This generates a sequence (x_i, y_i), and continuity of f gives that $f(x_i, y_i) \to f(x, y)$ which is the maximum. The same process gives the minimum by taking the infimum over f of each of the two regions at each step.

The maximum and minimum occur on the boundary for any radially monotonic function f. So for example $f(x, y) = x^2 + y^2$, as the radius increases, $\sqrt{x^2 + y^2}$ increases and so f increases. Then the minima of f on E occur for any (x, y) such that $|(x, y)| = 1$ and the maxima at $|(x, y)| = 2$.

Exercise 19.3 A function is said to have the intermediate value property on $\mathbb{R}$ if $f\colon \mathbb{R} \to \mathbb{R}$, and if $a < b, f(a) < x < f(b)$ implies that there exists $c \in (a, b)$ with $f(c) = x$. Let a function g have value 0 at 0 and value

$$g(x) = \sin \frac{1}{x}$$

for $x \neq 0$. Does g have the intermediate value property? Is g continuous?

Solution: The function g does have the intermediate value property but it is not continuous at zero, since the sequence $\frac{1}{n} \to 0$ but $g(1/n)$ doesn't converge.

Exercise 19.4 Does each of the following series converge? Prove or disprove.

- $x_k = \frac{1}{1+k}$.
- $x_k = \frac{1}{1+k^2}$.
- $x_k = \frac{1}{1+k^3}$.

Solution:

1. The harmonic series $a_n = 1/n$ doesn't converge, and $x_k = a_{n+1}$ so the series have the same terms. Therefore the x_k doesn't converge.
2. Each term of this series is less than the terms of the convergent series $\frac{1}{n^2}$, so it converges.
3. Each term of this series is less than the terms of the convergent series $\frac{1}{n^3}$, so it converges.

20 Counterexamples

Exercise 20.1 Find matrices A and B which are distinct and $AB = BA \neq 0$.

Solution: $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$. $B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$. $AB = BA = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$.

Exercise 20.2 Find a matrix will all entries non-zero which is not invertible

Solution: $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

Exercise 20.3 Do there exist square matrices, A, with $A^2 = 0$ and all entries nonzero?

Solution: Such a matrix gives the equations $a^2 = -bc$, $ab = -bd$, $ac = -cd$, and $d^2 = -bc$. If all entries are non-zero. then $a^2 = d^2$ and $a = -d$. Such a matrix: $A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$ and $A^2 = 0$.

Exercise 20.4 If A is a square matrix and there exists B such that $AB = I$, must there exist C such that

$$CA = I,$$

and if C does exist, must we have $C = B$?

Solution: Suppose $C = BA$, then $CB = BAB = BI = B$ so that $C = I$ and so $BA = I$. If there exists another C such that $CA = I$ then $CAB = B$ and therefore $C = B$.

Exercise 20.5 Which of the following functions are Lipschitz continuous?

- $f(x) = x^{2/3}$.
- $f(x) = x$.

Solution:

1. Let $t = 0$ in $|s^{2/3} - t^{2/3}|$, then if $|s|$ is arbitrarily small $|s^{2/3}|$ is arbitrarily bigger than $|s|$. So there can't exist a constant C where $|f(s) - f(t)| \leq C|s - t|$ and f is not Lipschitz continuous.

2. Since $|s - t| < C|s - t|$ for any $C > 1$, f is Lipschitz continuous.

Exercise 20.6 Suppose $f : (0, 1) \to \mathbb{R}$ is continuous. Must there be a global maximum in (0,1)? i.e. does there always exist x such that $f(y) \leq f(x)$ for all $y \in (0, 1)$?

Solution: No, consider the function $f(x) = \frac{1}{x}$.