

1 Basic Combinatorics

1.1 Sets and sequences

Sets. A set is an unordered collection of distinct objects. The objects are called elements of the set. We use braces to denote a set, for example, the set with elements 1, 2 and 3 is denoted $\{1, 2, 3\}$. Since the elements are not ordered, we can rearrange the elements in the representation to get the same set, so $\{1, 2, 3\}$ and $\{3, 2, 1\}$ are the same set. The set with no elements is denoted $\{\}$ or \emptyset , and is called the empty set. A set A is a subset of a set B , denoted $A \subseteq B$, if every element of A is also an element of B . We write $a \in A$ to denote that a is an element of set A . If A is a set with finitely many elements, we write $|A|$ for the number of elements of the set A . For example, $\{1, 2, \dots, n\}$ is a set of size n , and we will denote it by $[n]$. Some standard infinite sets include \mathbb{Z} , the set of integers, $\mathbb{Z}_{\geq 0}$, the set of non-negative integers, and \mathbb{R} , the set of real numbers. Recall that if A and B are sets, then $A \cap B = \{a : a \in A \text{ and } a \in B\}$, and $A \cup B = \{a : a \in A \text{ or } a \in B\}$. These are the intersection and union of the sets A and B respectively. Two sets A and B are disjoint if $A \cap B = \emptyset$. Sets $A_i : i \in S$ are pairwise disjoint if $A_i \cap A_j = \emptyset$ for all $i, j \in S$ with $i \neq j$. We write $\bigcup_{i \in S} A_i$ to denote the union of all sets A_i such that $i \in S$, and similarly for intersections.

Sequences. A sequence is an ordered collection of (not necessarily distinct) objects. The objects are called entries of the sequence. We use brackets to denote a sequence, for example $(1, 1, 2)$ denotes the sequence with entries 1, 1 and 2. Since the entries are ordered, we can rearrange the elements in the representation to get a new sequence, so $(1, 1, 2)$ and $(1, 2, 1)$ are different sequences. When the entries are required to be distinct, the sequence is called a permutation of the set of its entries. For example, $(1, 2, 3)$ and $(2, 3, 1)$ are permutations of $\{1, 2, 3\}$. If a and b are sequences, then a is a subsequence of b if we can delete entries of b to get a . For example, $(1, 2, 3, 4)$ is a subsequence of $(1, 1, 2, 1, 3, 1, 4)$ obtained by deleting 1s. The length of a sequence with finitely many entries is the number of entries in the sequence. Here is some notation involving products and sums of elements of sets and sequences: when we want to sum up the values of a function $f(i)$ for $i \in S$, where S is a set or a sequence, we write $\sum_{i \in S} f(i)$. The symbol we use for products is \prod , so the product of $f(i)$ over $i \in S$ is denoted $\prod_{i \in S} f(i)$. We will be making extensive use of this notation.

Recurrences. In many instances, a sequence is given by a formula for the n th term of the sequence, or by a recurrence equation. For example, $(2^n : n \in \mathbb{Z}_{\geq 0})$ is the sequence of powers of two, and we can write it alternatively as the sequence $(a_n)_{n \geq 0}$ given by $a_0 = 1$ and $a_n = 2a_{n-1}$. The sequence $(a_n)_{n \geq 0}$ where $a_0 = a_1 = 1$ and $a_n = a_{n-1} + a_{n-2}$ is called the Fibonacci sequence. In principle, we can work out any value of a_n from the recurrence, for example $a_7 = 21$ in the Fibonacci sequence. However, it is not immediately clear what a_n is explicitly as a function of n , we will develop general methods later on for solving recurrences.

1.2 Counting sets and sequences

Basic combinatorial questions involve counting sequences of finite length and sets of finite size. The following theorem tells us the total number of subsets of an n -element set:

Theorem 1 *The number of subsets of an n -element set is 2^n .*

For example, if $n = 2$, the subsets of $[n]$ are $\{1, 2\}$, $\{1\}$, $\{2\}$ and \emptyset . The next natural question is how many sequences of length n can be formed from a k -element set? For example, from the set $\{a, b\}$, we can form the sequences (a, a) , (a, b) , (b, a) and (b, b) of length two. The answer is as follows:

Theorem 2 *The number of sequences of length n from a k -element set is k^n .*

It should already be plain why this theorem is true: there are k choices for each entry of the sequence, and so k^n choices to fill up the sequence. The same reasoning actually gives Theorem 1: observe that each subset of $[n]$ is a sequence of n decisions, yes or no, according to whether an element is in the set or not. For example, the set $\{1, 2, 4\}$ as a subset of $[5]$ corresponds to the sequence

$$\begin{array}{ccccc} Y & Y & N & Y & N \\ \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ 1 & 2 & 3 & 4 & 5 \end{array}$$

In other words, if we encode yes by 1 and no by 0, every set corresponds to a sequence of zeroes and ones (a binary sequence) of length n , and there are 2^n of these (the case $k = 2$ of Theorem 2). By this logic, counting permutations is just as easy:

Theorem 3 *The number of permutations of a set of size n is $n! := n(n-1)(n-2)\dots 1$.*

The notation $n!$ is read n factorial, and denotes the product of all integers from 1 to n . Again, there are n choices for the first entry of a permutation, but then only $n-1$ for the next, $n-2$ for the next, and so on until the last entry, since all the entries are distinct. By the same argument, the number of sequences of elements of $[n]$ of length k with all entries distinct (compare with Theorem 2) is $n(n-1)(n-2)\dots(n-k+1)$. The last thing to count is the number of subsets of size k in an n -element set. For example, the set $\{1, 2, 3\}$ has three subsets of size two: they are $\{1, 2\}$, $\{2, 3\}$, $\{1, 3\}$. The answer in general is given by the following theorem

Theorem 4 *The number of sets of size k in an n -element set is*

$$\binom{n}{k} := \frac{n(n-1)(n-2)\dots(n-k+1)}{k!}.$$

The numbers $\binom{n}{k}$ defined in this theorem are called binomial coefficients, for reasons which we shall see shortly. This theorem is only a bit trickier to prove than the ones before. We observed above that there are $n(n-1)(n-2)\dots(n-k+1)$ sequences of k distinct entries from $[n]$. If we take the set of the entries, we get a set of size k . However, the same set of size k is counted by many sequences of length k , since we can rearrange the entries of the sequence without affecting the set of entries. By Theorem 3, there are $k!$ ways to rearrange the entries of a sequence of length k , so we see that the number of sequences of length k with distinct entries from n , namely $n(n-1)(n-2)\dots(n-k+1)$, is exactly $k!$ times the number of sets of size k . This proves the above theorem.

1.3 Two Principles

All of the basic theorems in the last section have the same organizing principle, known as the multiplication principle. Informally, the multiplication principle says that if we want to know how many sequences (x_1, x_2, \dots, x_k) there are given that the number of choices for x_i is known, all we have to do is multiply together the number of choices (or decisions) for each x_i when x_1, x_2, \dots, x_{i-1} have already been chosen.

Principle 1 (The Multiplication Principle) *The number of sequences (x_1, x_2, \dots, x_k) such that there are a_i choices for x_i after having chosen x_1, x_2, \dots, x_{i-1} for each $i = 1, 2, \dots, n$ is exactly $a_1 a_2 \dots a_n$.*

Our argument for proving Theorem 1 uses the multiplication principle with two choices for each x_i , namely $x_i \in \{0, 1\}$ for all i , in which case the number of choices is 2^n . For Theorem 3, we use the multiplication principle noting that there are $n - i$ choices for the i th entry of a permutation after the first i entries have been chosen. An alternative way of stating the multiplication principle is using decision trees or Cartesian products of sets, but we will not require this formulation here. We will return to this topic later. It is instructive to do another example involving the multiplication principle.

Example. We determine the number of walks of k steps on an infinite square grid, starting at the origin. Such a walk is drawn in Figure 1 below. A walk is allowed to reverse along the same edge of the grid. We encode the walk as a sequence of choices: the first entry of the sequence is the direction taken in the first step of the walk. In general, at the i th step, we record the direction taken as the i th entry of the sequence. In this way, we produce a sequence (x_1, x_2, \dots, x_k) which represents the walk. We can easily count these sequences: there are four choices for x_1 , and then for each subsequent x_i , there are four choices. So by the multiplication principle, the total number of walks of k steps is exactly 4^k .

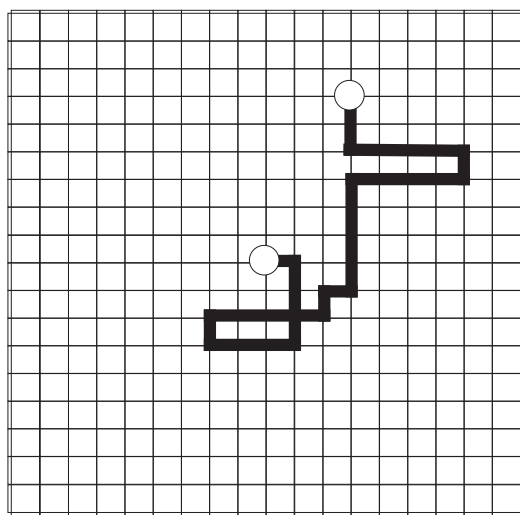


Figure 1 : A walk of length 29 on a 19×19 grid.

Two sets A and B are disjoint if $A \cap B = \emptyset$. A second principle we use often is to break down a counting problem into a number of disjoint parts which are easier to deal with. We will refer to this as the summation principle:

Principle 2 (The Summation Principle) *Let A_1, A_2, \dots, A_n be disjoint finite sets. Then*

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{i=1}^n |A_i|.$$

Here is a simple application of the summation principle. It illustrates a common use: if we want to count the number of sets with a certain property, it is often easier to add up over all integers k the number of sets of size k with that property.

Example. We determine the number of sequences (A, B) such that $A \subseteq B \subseteq \{1, 2, \dots, n\}$. For example, if $n = 2$, then the pairs (A, B) are

$$(\emptyset, \emptyset), (\emptyset, 1), (\emptyset, 2), (\emptyset, 12), (1, 1), (1, 12), (2, 2), (2, 12), (12, 12)$$

where for brevity we write 12 instead of the set $\{1, 2\}$. If we try to use the multiplication principle directly, we get in trouble: the number of choices of A depends on what B is. To use the summation principle, first consider all pairs (A, B) such that $|B| = k$, for each $k \leq n$. Those pairs can be counted with the multiplication principle: there are $\binom{n}{k}$ choices for B , and once B is chosen, there are 2^k subsets of B , so there are 2^k choices for A . So the multiplication principle tells us that there are $2^k \binom{n}{k}$ ways to choose (A, B) with $|B| = k$ and $A \subseteq B$. But k could be anything from zero to n , so the summation principle tells us that the number of (A, B) is

$$\sum_{k=0}^n 2^k \binom{n}{k}.$$

Later we will see that this sum is actually equal to 3^n (see section on bijections and the binomial theorem).

1.4 Inclusion-Exclusion

A basic course in mathematics confirms $|A \cup B| = |A| + |B| - |A \cap B|$. This is a special instance of the inclusion exclusion formula, or combinatorial sieve:

Principle 3 (Inclusion-Exclusion) *Let A_1, A_2, \dots, A_n be sets of finite size. Then*

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{S \subseteq [n]} (-1)^{|S|+1} \left| \bigcap_{i \in S} A_i \right|.$$

The proof of the principle is elementary. First note that $\sum_{|S|=1} |\bigcap_{i \in S} A_i|$ overestimates $|A_1 \cup A_2 \cup \dots \cup A_n|$ since all the elements which are in two or more sets are overcounted. So we have to subtract $\sum_{|S|=2} |\bigcap_{i \in S} A_i|$ – all intersections of two of the sets. But then we have subtracted too much: all elements in three or more of the sets are undercounted. So we have to add the size of all intersections of three sets, but then again we have overcounted, and so on. When the sets A_i are pairwise disjoint (they share no elements – meaning $A_i \cap A_j = \emptyset$ for all i, j), we get the summation principle.

Using inclusion-exclusion, we can prove the following nice result in number theory. In this theorem, two integers are coprime if their greatest common divisor (highest common factor) is 1. The number of positive integers coprime to n and less than or equal to n is denoted $\varphi(n)$. This function is called the Euler totient function, or the Euler φ function, and is fundamentally important in number theory. The following theorem tells us that we know the value of $\varphi(n)$ if we know the prime factors of n :

Theorem 5 *Suppose n is a positive integer and p_1, p_2, \dots, p_k are the distinct prime factors of n . Then*

$$\varphi(n) = n \prod_{i=1}^k \left(1 - \frac{1}{p_i}\right).$$

Proof \triangleright We prove this theorem using inclusion-exclusion. Let $S \subseteq [k]$ and let A_i denote the set of integers less than n which are divisible by prime p_i . Then $|A_i| = n/p_i$, and

$$\left| \bigcap_{i \in S} A_i \right| = \frac{n}{\prod_{i \in S} p_i}$$

since the left hand side is the number of integers less than n divisible by each p_i for $i \in S$, and hence divisible by $\prod_{i \in S} p_i$. By inclusion-exclusion, the number of integers less than n divisible by at least one of the primes p_i for $i \in [k]$ is

$$\left| \bigcup_{i=1}^k A_i \right| = \sum_{S \subseteq [k]} (-1)^{|S|+1} \frac{n}{\prod_{i \in S} p_i} = -n \sum_{S \subseteq [k]} \frac{1}{\prod_{i \in S} (-p_i)}.$$

Note that the sum is over all sets $S \subseteq [k]$ with S non-empty. Now we use the following formula valid for real numbers x_i :

$$\sum_{S \subseteq [k]} \prod_{i \in S} x_i = \prod_{i=1}^k (1 + x_i) - 1.$$

To see this formula, note that $\prod_{i \in S} x_i$ appears exactly once when we expand $\prod_{i=1}^k (1 + x_i)$, but we have to subtract 1 since S is not allowed to be empty. Putting $x_i = -1/p_i$, we see

$$\left| \bigcup_{i=1}^k A_i \right| = n - n \prod_{i=1}^k \left(1 - \frac{1}{p_i}\right).$$

This is the number of integers in $[n]$ which are not coprime to n . Subtracting from n , we get the number of integers which are coprime to n , and the required formula. \blacksquare

A fixed point of a permutation (x_1, x_2, \dots, x_n) of $[n]$ is an integer i for which $x_i = i$. A derangement of $[n]$ is a permutation of $[n]$ with no fixed points. For example the ten derangements of $[4]$ are listed below

$$\begin{array}{cccc} (2, 4, 1, 3) & (2, 3, 4, 1) & (2, 1, 4, 3) & (3, 4, 1, 2) \\ (3, 4, 2, 1) & (3, 1, 4, 2) & (4, 1, 2, 3) & (4, 2, 1, 3) \\ (4, 3, 2, 1) & (4, 3, 1, 2) & & \end{array}$$

The multiplication principle cannot be used to count derangements, since the number of choices for the i th entry depends on what the earlier entries were.

Theorem 6 *The number of derangements of $[n]$ is*

$$n! \sum_{k=0}^n \frac{(-1)^k}{k!}.$$

Proof \triangleright Let A_k be the set of permutations (x_1, x_2, \dots, x_n) with k as a fixed point. Then for any non-empty set $S \subseteq [n]$ of size k ,

$$\left| \bigcap_{i \in S} A_i \right| = (n - k)!$$

since once we have fixed $x_i = i$ for $i \in S$, there are $(n - k)!$ ways to complete the permutation. Now inclusion-exclusion gives

$$\begin{aligned} \left| \bigcup_{i=1}^n A_i \right| &= \sum_{S \subseteq [n]} (-1)^{|S|+1} (n - |S|)! \\ &= \sum_{k=1}^n \binom{n}{k} (-1)^{k+1} (n - k)! \\ &= - \sum_{k=1}^n (-1)^k \frac{n!}{k!}. \end{aligned}$$

Note the sum starts at $k = 1$ since we don't allow $S = \emptyset$. This is the number of permutations with at least one fixed point. So the number of derangements is

$$n! + \sum_{k=1}^n (-1)^k \frac{n!}{k!} = n! \sum_{k=0}^n \frac{(-1)^k}{k!}.$$

This proves the theorem. ■

For $n = 4$, we get

$$4! \left(1 - 1 + \frac{1}{2} - \frac{1}{6} + \frac{1}{24} \right) = 9$$

which agrees with the listing of nine derangements given above. There is a nice formula for derangements, using Taylor series. Recall that

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

is the Taylor series for e^x . Replacing x with -1 , we obtain that the number of derangements of $[n]$ is $\lfloor n!/e + 1/2 \rfloor$ for $n \geq 2$, where $\lfloor x \rfloor$ is the largest integer less than or equal to a given real number x . Note that $\lfloor n!/e + 1/2 \rfloor$ is the nearest integer to $n!/e$. There are many proofs of the above result, for example using generating functions; we have only chosen a very standard one.

1.5 Bijections and Combinatorial Proofs

Let A and B be sets. A function $f : A \rightarrow B$ is called an injection (or one-to-one) if whenever $x, y \in A$ are distinct, then $f(x) \neq f(y)$. The function f is a surjection (or onto B) if for every $b \in B$ there exists $x \in A$ such that $f(x) = b$. Finally, $f : A \rightarrow B$ is a bijection if f is an injection and a surjection. How would we check, given two sets A and B of finite size, that $|A| = |B|$? For each element $a \in A$, we would associate in some way an element $b \in B$, so that no other element of A is associated with b . In other words, we would element-by-element find a matching or pairing of the elements of A with the elements of B . This is exactly a bijection $f : A \rightarrow B$. For example, let A be the set of positive even integers and let B be the set of positive odd integers. Then there is a natural bijection $f : A \rightarrow B$ given by $f(a) = a - 1$ for each $a \in A$. In this section, we treat the question of finding bijections between two sets A and B . We have already seen an example of bijections: let A denote the set of all subsets of $[n]$ and let B be the set of binary strings (01 strings) of length n . For $b \in B$, define $f(b) = \{i : b_i = 1\}$. Then $f : B \rightarrow A$ is a bijection (check this) so $|A| = |B| = 2^n$, which proves Theorem 1.

Example. We claim that the number of sets in $X = \{1, 2, \dots, n\}$ of even size equals the number of sets of odd size. For example, if $n = 4$, then the number of sets of odd size is eight: they are

$$\{1\}, \{2\}, \{3\}, \{4\}, \{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 4\}, \{2, 3, 4\},$$

which is equal to the number of sets of even size. To prove this using bijections, we have to come up with a way of producing, for each set of even size, a unique set of odd size, and vice versa. This will define our bijection. This is easy if n is odd: for each set A , define $f(A) = A^c$, where

$$A^c = \{a \in X : a \notin A\}$$

is the complement of A . Clearly this has odd size whenever A has even size, and f is a bijection. But what if n is even? Then A^c has even size if A has even size, so the same function f fails. So we define $f(A)$ more carefully. If $n \in A$, then let $f(A) = A \setminus \{n\}$ (in other words, remove n from A). Definitely $f(A)$ has odd size in this case. If $n \notin A$, then let $f(A) = A \cup \{n\}$ (in other words, add n to A). Once more, $|f(A)|$ is odd. Now we must check that f is a bijection. If $f(A) = f(B)$, then either $A \setminus \{n\} = B \setminus \{n\}$ or $A \cup \{n\} = B \cup \{n\}$, and both of these mean $A = B$. So f is an injection. If B is a set of odd size, is there a set A of even size with $f(A) = B$? Yes: if $n \in B$, then $A = B \setminus \{n\}$ gives $f(A) = B$, and if $n \notin B$, then $A = B \cup \{n\}$ gives $f(A) = B$. So f is a surjection, and our claim is proved.

1.6 Combinatorial Identities

There is a mathematical way to express what was shown in the last example, namely that the number of sets in $[n]$ of even size equals the number of sets of odd size in $[n]$:

$$\sum_{k \text{ even}} \binom{n}{k} = \sum_{k \text{ odd}} \binom{n}{k}.$$

This is known as a combinatorial identity. Many combinatorial identities are proved by interpreting each side combinatorially, and then establishing a bijection between the two

objects. We have another example of a combinatorial identity: the number of subsets of an n -element set is 2^n , so

$$\sum_{k=0}^n \binom{n}{k} = 2^n.$$

If you were given the identity, the left hand side would be interpreted as the number of subsets of an n -element set, and the right hand side is the number of binary strings of length n , and we would establish the bijection we gave in a previous example.

Example. Prove the identity

$$\sum_{k=0}^n 2^k \binom{n}{k} = 3^n.$$

Fortunately we know what the left hand side represents by looking at a preceding example: it is the number of sequences (A, B) such that $A \subseteq B \subseteq [n]$. The right hand side may be interpreted as the number of sequences (x_1, x_2, \dots, x_n) such that $x_i \in \{0, 1, 2\}$ – there are 3^n such sequences, by the multiplication principle. We want a bijection

$$f : \{(A, B) : A \subseteq B \subseteq [n]\} \rightarrow \{(x_1, x_2, \dots, x_n) : x_i \in \{0, 1, 2\} \text{ for } i \in [n]\}.$$

This is now quite simple: define $f(A, B) = (x_1, x_2, \dots, x_n)$ where $x_i = 2$ if $i \in A \cap B$, $x_i = 1$ if $i \in B \setminus A$, and $x_i = 0$ otherwise. Then f is a bijection, since the sequence $f(A, B)$ tells us which elements are in A (the positions filled with a 2), which elements are in B but not A (the positions filled with a 1), and the elements in neither A nor B (the positions filled with a 0). This proves the identity. We will see later that the identity above is a special case of a theorem called the Binomial Theorem.

Example. Suppose we want to prove for $n \geq 1$ that

$$\sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n}.$$

The right hand side is just the number of subsets of a $2n$ -element set of size n , by definition. But what is the left side? Recall that

$$\binom{n}{k} = \binom{n}{n-k}$$

so the left hand side is a sum of $\binom{n}{k} \binom{n}{n-k}$. By the summation and multiplication principles, for two sets X and Y of size n , this is the number of pairs (A, B) such that $A \subseteq X$ and $B \subseteq Y$ and $|A| + |B| = n$ (see Figure 2). If X and Y are disjoint, then $|A \cup B| = |A| + |B| = n$ and $|X \cup Y| = 2n$. Let P be the set of subsets of $X \cup Y$ of size n . Now define the function

$$f : \{(A, B) : A \subseteq X, B \subseteq Y, |A| + |B| = n\} \rightarrow P$$

by $f(A, B) = A \cup B$. If we have two pairs (A, B) and (A', B') , then $f(A, B) \neq f(A', B')$ unless $A' = A$ and $B' = B$, since X and Y are disjoint. So f is an injection. Also f is a surjection: for each set $Z \in P$ of size n , we see that $f(Z \cap X, Z \cap Y) = Z$. Therefore f is a bijection, and we have proved the identity.

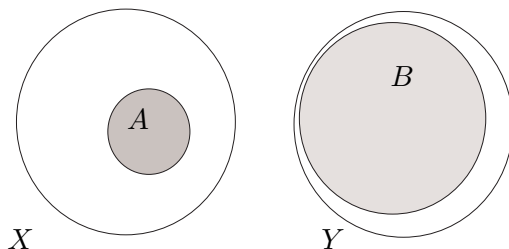


Figure 2 : Choosing $A \subseteq X$ and $B \subseteq Y$ with $|A| + |B| = n$.

Combinatorial identities are generally very hard to prove by finding bijections, since it often hard to see what each side of the identity actually represents.

1.7 Mathematical Induction

Let $P(n)$ denote a logical statement for each positive integer n . Thus for each integer n , we can determine whether $P(n)$ is true or $P(n)$ is false. For example, $P(n)$ might be the statement that there is a prime larger than n , and so on. In the most basic form, the principle of mathematical induction can be stated as follows:

Principle 4 (Mathematical Induction) *Let $P(n)$ be a statement for each positive integer n , and suppose that $P(1)$ is true, and $P(n) \rightarrow P(n+1)$ for each positive integer n . Then $P(n)$ is true for every positive integer.*

There are two steps in any induction. First one establishes the base case (in the terms above, one proves $P(0)$). Then, under the assumption that $P(n)$ is true, one attempts to prove that $P(n+1)$ must also be true (it is very important to get the order correct here – we are to show $P(n) \rightarrow P(n+1)$ and not $P(n+1) \rightarrow P(n)$).

Example. For every positive integer n , we prove that

$$\sum_{k=1}^n k = \binom{n+1}{2}.$$

First we verify the base case $n = 1$: clearly both sides equal 1 so we're done. Now suppose $\sum_{k=1}^n k = \binom{n+1}{2}$. We use this to show $\sum_{k=1}^{n+1} k = \binom{n+2}{2}$. By induction,

$$\sum_{k=1}^{n+1} k = \sum_{k=1}^n k + (n+1) = \binom{n+1}{2} + (n+1) = \binom{n+2}{2}$$

and this proves the statement.

In some instances, the following stronger form of induction is necessary:

Principle 5 (Strong Induction) *Let $P(n)$ be a statement for each positive integer n , and suppose that $P(1)$ is true, and $P(n) \wedge P(n-1) \wedge \dots \wedge P(1) \rightarrow P(n+1)$ for each positive integer n . Then $P(n)$ is true for every positive integer.*

To give a natural example of strong induction, we consider the notion of a tree. A tree is a collection of vertices and edges joining pairs of vertices such that every pair of vertices is connected by a path and no set of edges forms a cycle. A tree is drawn below (the edges are straight line segments, and vertices occur where the straight line segments intersect):

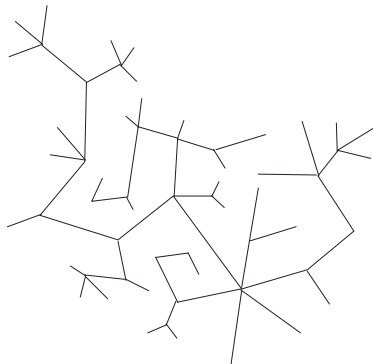


Figure 3 : A tree with 60 vertices.

In the figure, there are 59 edges. This is not just a coincidence: the number of edges in a tree is always one less than the number of vertices.

Theorem 7 *Let T be a tree with n vertices. Then T has $n - 1$ edges.*

Proof \triangleright The base case of our strong induction is $n = 1$. A tree with one vertex clearly has no edges, so the theorem is true in this case. Now suppose $n > 1$ and let T be a tree on n vertices. Suppose all trees with $m < n$ vertices have $m - 1$ edges. To prove that T has $n - 1$ edges, delete an edge from T . Evidently this splits T into two trees T_1 and T_2 , such that T_1 has n_1 vertices, T_2 has n_2 vertices, and $n_1 + n_2 = n$. By strong induction, T_1 has $n_1 - 1$ edges and T_2 has $n_2 - 1$ edges. Therefore T has $(n_1 - 1) + (n_2 - 1) + 1 = n - 1$ edges – the plus one is for the edge we took out of T . \blacksquare

1.8 Averaging*

One of the most fundamental principles in combinatorics is the pigeonhole principle. It is entirely straightforward:

Principle 6 (Pigeonhole Principle) *If a sequence of length $n+1$ has only n different entries, then two of the entries are the same.*

While the principle is straightforward, it has some remarkable applications. Here is one such application:

Theorem 8 *Amongst any set of $n+1$ integers in $[2n]$, there is one which divides the other.*

Proof \triangleright Every integer x can be written uniquely in the form $2^{f(x)}g(x)$, where $g(x)$ is the largest odd divisor of x . Then $g(x) \in \{1, 3, 5, \dots, 2n - 1\}$, so $g(x)$ can take only n different values. Since there are $n + 1$ integers in our set, there are two integers x and y with $g(x) = g(y)$, let's say $g(x) = g(y) = m$. But then $x = 2^{f(x)}m$ and $y = 2^{f(y)}m$, so if

$f(y) \geq f(x)$ then $x|y$ and if $f(x) \geq f(y)$ then $y|x$. We have found two integers one of which divides the other. ■

The next application of the pigeonhole principle is sometimes called the handshaking lemma.

Theorem 9 *In a group of n people, there are always two people who are acquainted with the same number of people in the group.*

Proof ▷ Notice that it is not possible that some person is acquainted with no people at the same time as some other person is acquainted with everyone. This means that the number of people a person is acquainted with is in the set $\{0, 1, 2, \dots, n-2\}$ or in the set $\{1, 2, \dots, n-1\}$. Since there are n people, two of these numbers must be equal – so two people are acquainted with the same number of people in the group. ■

The pigeonhole principle is sometimes referred to as Dirichlet's Box Principle, due to the following important result in number theory attributed to Dirichlet:

Theorem 10 (Dirichlet Box Principle) *Let N be a positive integer and let α be any real number. Then there exist positive integers $q \leq N$ and $p \leq \lceil \alpha \rceil N$ such that*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{qN}.$$

Proof ▷ Let $\{x\}$ denote the fractional part of the real number x . Then the $N+1$ numbers $\{i\alpha\}$ for $0 \leq i \leq N$ define points in the interval $[0, 1]$. By the pigeonhole principle, there must be two numbers $\{i\alpha\}$ and $\{j\alpha\}$ differing by less than $1/N$ such that $i, j \leq N$. Suppose $j > i$ and let $q = j - i$. Then, for some integer p , we have $|q\alpha - p| < 1/N$. In other words,

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{qN}$$

and that completes the proof. ■

This theorem is a basic key to many theorems in combinatorial and analytic number theory, such as the equidistribution of the real numbers $\{i\alpha\} : i \in \mathbb{Z}$ in the interval $[0, 1]$. This is beyond the scope of the course, so we do not press this further here. We will give the following application, but only sketch the proof:

Corollary 11 *The sum $\sum_{n=1}^{\infty} |\cos n|^n$ diverges.*

Proof ▷ We will show that $|\cos n|^n$ is at least $\frac{1}{2}$ infinitely often. After some computations, it can be seen that Dirichlet's Box Principle shows that $|q\pi - n| < 1/n$ for infinitely many n . It follows that $|\cos n| = |\cos(q\pi - n)| > \cos(1/n)$ for infinitely many n . Using the known fact $\cos x \geq 1 - \frac{1}{2}x^2$, $|\cos n| > 1 - 1/(2n^2)$. However, since¹

$$\lim_{n \rightarrow \infty} \left(1 - \frac{1}{2n^2}\right)^n = 1$$

we get $|\cos n|^n$ is at least $\frac{1}{2}$ for infinitely many n . ■

¹Check this limit using l'Hôpital's Rule.