Math 127: Finite Cardinality

Mary Radcliffe

1 Basics

Now that we have an understanding of sets and functions, we can leverage those definitions to an understanding of size. In general, the question we will be considering is this: given a set S, how big is it?

Definition 1. Given a nonempty set X, we say that X is *finite* if there exists some $n \in \mathbb{N}$ for which there exists a bijection $f : \{1, 2, \dots, n\} \to X$. The set $\{1, 2, \dots, n\}$ is denoted by $[n]$. If there exists a bijection $f : [n] \to X$, we say that X has *cardinality* or *size* n, and we write $|X| = n$. If X is not finite, then we say that X is *infinite*. By convention, the empty set is presumed to be finite, and $|\emptyset| = 0$.

All this is to say: our definition of finiteness is based on an understanding of finiteness in the natural numbers. Effectively, we declare that the following sets are finite:

$$\emptyset, \{1\}, \{1, 2\}, \{1, 2, 3\}, \{1, 2, 3, 4\}, \{1, 2, 3, 4, 5\}, \dots$$

Then, if we have any different set, we say that it is finite if it corresponds, bijectively, to one of these. Our notion of size is determined by the size of these sets, and the size of these sets is determined by their largest elements.

Throughout these notes, we will give various theorems and shortcuts to determining the size of a finite set. Fundamentally, though, it all comes back to bijections. This is particularly useful when, in the next notes, we consider the size of infinite sets. Since it is difficult (nay, impossible!) to literally count the number of things in an infinite set, we will use the idea of bijection to determine when things are the "same" size.

Before we start developing theorems, let's get some examples working with the definition of finite sets.

Example 1. Fix $m \in \mathbb{N}$. Let $X_m = \{q \in \mathbb{Q} \mid 0 \leq q \leq 1, \text{ and } mq \in \mathbb{Z}\}$. Prove that X is finite, and determine its cardinality.

Solution. To prove that X_m is finite, by definition we need a natural number n chosen so that we can construct a bijection from $[n]$ to X_m. To figure out what n might be, let's take a look at a few examples of X_m.

$$\begin{array}{ll} \text{When } m = 1: & X_m = \{0, 1\} \\ \text{When } m = 2: & X_m = \{0, \frac{1}{2}, 1\} \\ \text{When } m = 3: & X_m = \{0, \frac{1}{3}, \frac{2}{3}, 1\} \\ \text{When } m = 4: & X_m = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\} \end{array}$$

So it appears that $|X_m|$ should be $m + 1$. Indeed, this makes some mathematical sense also; if $mq \in \mathbb{Z}$, then we should be able to think of q with denominator m, so that the denominators cancel out. The number of different rationals between 0 and 1 that have denominator m is certainly $m+1$.

Now, to formally prove that $|X_m| = m+1$, we need to construct a bijection from $[m+1]$ to X_m. Define $f : [m+1] \to X_m$ by $f(k) = \frac{k-1}{m}$. Note that $mf(k) = k-1 \in \mathbb{Z}$, and for all $k \in [m+1]$, $0 \leq f(k) \leq 1$, so f is a well-defined function. We wish to establish that f is bijective.
Injectivity: Let $k, j \in [m+1]$ with $f(k) = f(j)$. Then $\frac{k-1}{m} = \frac{j-1}{m}$, so clearly $k = j$. Therefore f is injective.
Surjectivity: Let $q \in X_m$. Then $mq \in \mathbb{Z}$; let $k = mq$, so $q = \frac{k}{m}$. By definition of X_m, we must have that $0 \leq k \leq m$. Hence, $1 \leq k+1 \leq m+1$, and hence $k+1 \in [m+1]$. Moreover, $f(k+1) = \frac{k+1-1}{m} = \frac{k}{m} = q$. Therefore, f is surjective.
Since f is both injective and surjective, it is bijective, and thus X_m is finite, with $|X_m| = m+1$. □

We note that here, we constructed a bijection explicitly to the set $[m+1]$, but that is not strictly necessary. Indeed, we can get away with just constructing a bijection to another set, whose size we know.

Theorem 1. *Let X and Y be finite sets. Then $|X| = |Y|$ if and only if there exists a bijection $h : X \to Y$.*

Indeed, this theorem can be taken as the definition of sets having equal cardinality, rather than the definition being taken as having a bijection to $[n]$. This is helpful, as it allows us to compare the sizes of various sets without having to directly construct bijections into $[n]$, but just between each other.

Proof. [Proof of Theorem 1] Suppose that X and Y are finite sets with $|X| = |Y| = n$. Then there exist bijections $f : [n] \to X$ and $g : [n] \to Y$. Taking $h = g \circ f^{-1}$, we get a function from X to Y. Moreover, as f^{-1} and g are bijections, their composition is a bijection (see homework) and hence we have a bijection from X to Y as desired.

For the other direction, suppose that X and Y are finite sets, and that there exists a bijection $h : X \to Y$. Suppose that $|Y| = n$, so that there is a bijection $g : [n] \to Y$. Then the function $f = h^{-1} \circ g$ is a bijection from $[n]$ to X, and hence $|X| = n$. □

The next theorem may seem a bit silly, but it is in fact quite important. In order for finiteness to be a meaningful idea, it must also be true that infiniteness is a real thing; that is, we'd like to confirm that infinite sets exist, and that proving something is finite actually matters. So we have:

Theorem 2. *The set $\mathbb{N}$ is infinite.*

Proof. Let us suppose, to the contrary, that $\mathbb{N}$ is finite. Then there exists $n \in \mathbb{N}$ having a bijection $g : [n] \to \mathbb{N}$. For simplicity of notation, write $g_i = g(i)$ for $1 \leq i \leq n$. We claim the following:

> **Claim 1.** The set $S = \{g_1, g_2, \ldots, g_n\}$ has a maximum.
>
> **Proof.** [Proof of Claim] We work by induction on n. If $n = 1$, then $S = \{g_1\}$, and clearly g_1 is a maximum for S.
>
> Now, suppose that any set containing n natural numbers has a maximum. Consider the set $S = \{g_1, g_2, \ldots, g_n, g_{n+1}\}$. By the inductive hypothesis, we know that the subset $T = \{g_1, g_2, \ldots, g_n\}$ has a maximum; let $j \in [n]$ be such that g_j is a maximum for T, so $g_j \geq g_i$ for all $1 \leq i \leq n$. We now consider two cases. If $g_{n+1} \leq g_j$, then $g_j \geq g_i$ for all $1 \leq i \leq n+1$, and hence g_j is a maximum for S. On the other hand, if $g_{n+1} > g_j$, then we have that for all $1 \leq i \leq n$, $g_{n+1} > g_j \geq g_i$, and hence g_{n+1} is a maximum for S. In any case, S has a maximum element.
>
> Therefore, by induction, the set $S = \{g_1, g_2, \ldots, g_n\}$ has a maximum. □

Now, notice that $g([n]) = \{g_1, g_2, \ldots, g_n\}$. Let g_j be the maximum element of $g([n])$. Let $m = g_j + 1 \in \mathbb{N}$. Then as g_j is the maximum element of $g([n])$, we must have that $m \notin g([n])$. Then g is not surjective. But, by assumption, g is bijective, and hence g is surjective.

This is clearly a contradiction, and hence we must have that $\mathbb{N}$ is infinite. $\square$

We note that the claim made in this proof is actually an item of independent interest. We already know, from the Well-Ordering Principle, that any nonempty subset of $\mathbb{N}$ has a minimum. The claim made in the proof of Theorem 2 shows that any finite nonempty subset of $\mathbb{N}$ also has a maximum. This is important enough that, just to emphasize it, we'll break it out into its own theorem.

Theorem 3. *Any finite nonempty subset of* $\mathbb{N}$ *contains a maximum.*

1.1 Basics of finite sets

Now, let's take a look at how size of sets relates to our favorite set operations. In this section, we'll focus almost exclusively on finite sets; a discussion of infinite sets will occur in the next set of notes. Here, we will develop some theorems to help us count finite sets, and in the next few sections of these notes, we will use these theorems and others to develop, understand, and count basic combinatorial objects.

In order to prove some of these theorems, we shall require that the bijections we construct between sets we wish to count and $[n]$ take certain specified structures. So we start with the following Lemma, which will appear throughout these theorems as a tool.

Lemma 1. *Let X be a nonempty finite set with $|X| = n$, and let $S \subseteq X$. Then there exists a bijection $f : [n] \to X$, such that $f^{-1}(S) = [k]$ for some $1 \le k \le n$.*

Before we prove Lemma 1, let's decode a little what we are doing here. The Lemma tells us really two things. First, that the subset S is finite; indeed, that it has size $k \le n$ (this is explicitly stated as Corollary 1). Second, that we can construct our bijection in such a way as to ensure that the elements of S appear first. This can be useful as we consider putting sets together; we can specify not only that S is part of the image of the bijection, but which part of the image it really is.

Proof. [Proof of Lemma 1] We work by induction on n. First, consider the base case that $n = 1$. Let $f : [1] \to X$ be a bijection, so that $X = \{f(1)\}$. There are two cases for S: either $S = \emptyset$ or $S = \{f(1)\}$. In either case, the result of the theorem is trivially true.

Now, let us assume for induction that the result holds when $|X| = n$, for any subset of X.

Suppose that $|X| = n + 1$, and let $S \subseteq X$. If S is empty, the result is trivial, so let us presume that S is nonempty. By definition of cardinality, there exists a bijective function $g : [n + 1] \to X$. We consider two cases, according as whether $g(n + 1) \in S$.

For the first case, suppose that $g(n + 1) \notin S$. Define $X' = X \backslash \{g(n + 1)\}$, and notice that $S \subseteq X'$. Moreover, the function $h : [n] \to X'$ defined by $h(k) = g(k)$ for all $1 \le k \le n$ is a bijection, so $|X'| = n$. By the inductive hypothesis, then, there exists a bijection $h' : [n] \to X'$ having $(h')^{-1}(S) = [k]$ for some k.

Construct $f : [n + 1] \to X$ by

$$f(m) = \begin{cases} h'(m) & \text{if } 1 \le m \le n \\ g(n+1) & \text{if } m = n + 1. \end{cases} \tag{1}$$

We make the following claim:

Claim 1. The function f defined in equation (1) is bijective.

A formal proof of this claim is a homework exercise.

Note that f is bijective, and that $f^{-1}(S) = h^{-1}(S) = [k]$ by construction. Therefore, the result is satisfied.

For the second case, suppose $g(n+1) \in S$. Define $X' = X \backslash \{g(n+1)\}$, and define $S' = S \backslash \{g(n+1)\}$. Notice that $S' \subseteq X'$. Moreover, the function $h : [n] \to X'$ defined by $h(k) = g(k)$ for all $1 \le k \le n$ is a bijection, so $|X'| = n$. By the inductive hypothesis, then, there exists a bijection $h' : [n] \to X'$ having $(h')^{-1}(S') = [k]$ for some k.

Construct $f : [n+1] \to X$ by

$$f(m) = \begin{cases} g(n+1) & \text{if } m = 1 \\ h'(m-1) & \text{if } 2 \le m \le n+1. \end{cases} \tag{2}$$

We make the following claim:

Claim 2. The function f defined in equation (2) is bijective.

A formal proof of this claim is a homework exercise.

Note that f is bijective, and that $f^{-1}(S) = \{1\} \cup \{m | m-1 \in h^{-1}(S')\} = [k+1]$ by construction. Therefore, the result is satisfied.

Therefore, in either case, the inductive step is true.

Hence, by induction, the result holds for any finite set X. □

We note that this lemma has many useful corollaries, some of which are to be proven in homework. For example:

Corollary 1. *If X is finite and $S \subseteq X$, then S is finite and $|S| \le |X|$.*

Corollary 2. *If $g : X \to Y$ is injective, and Y is finite, then X is also finite, and $|X| \le |Y|$.*

This corollary can be proven by considering $g(X)$ as a subset of Y, and applying Lemma 1. The fact that g is injective ensures that when we restrict the codomain of g to just $g(X)$, we get a bijection, so that $|X| = |g(X)|$ by an application of Theorem 1.

Corollary 3. *If $g : X \to Y$ is surjective, and X is finite, then Y is also finite, and $|Y| \le |X|$.*

Proof. Suppose that $g : X \to Y$ is surjective, and that X is finite. For each $y \in Y$, select an element $a_y \in g^{-1}(\{y\})$; note that such an element exists due to surjectivity of g. Define a set $S = \{a_y \in X \mid y \in Y\}$, the set of all chosen preimage points. Note that $S \subseteq X$, so by Corollary 1, we have that S is finite and $|S| \le |X|$.

Moreover, let us consider the restriction of g to S, $g|_S$. Note that for all $y \in Y$, $a_y \in S$ and $g|_S(a_y) = y$, so $g|_S$ is surjective. Moreover, if $a_y, a_z \in S$, then $g|_S(a_y) = y$ and $g|_S(a_z) = z$. By definition, if $y = z$, we must have that $a_y = a_z$, and thus $g|_S$ is also injective. Therefore, $g|_S$ is a bijection between S and Y, and therefore by Theorem 1, we must have that $|Y| = |S| \le |X|$. □

Corollary 4. *If X is finite or Y is finite, then $X \cap Y$ is finite.*

The proof of this corollary is a homework exercise.

Corollary 5. *If X is a finite set, then for all sets A, $X \backslash A$ is finite and $|X \backslash A| \le |X|$.*

This is immediate, since by definition for any set A, $X \backslash A$ is a subset of X, so an application of Corollary 1 is all that is needed.

Now, let us turn our attention to proving something a bit more complex using the result of Lemma 1. In particular, we will prove the following theorem about the cardinality of the union of two sets.

Theorem 4. *Let X and Y be finite sets. Then $|X \cup Y| = |X| + |Y| - |X \cap Y|$.*

This theorem is a small case of what's known as the Inclusion-Exclusion Principle. The full Inclusion-Exclusion Principle is as follows.

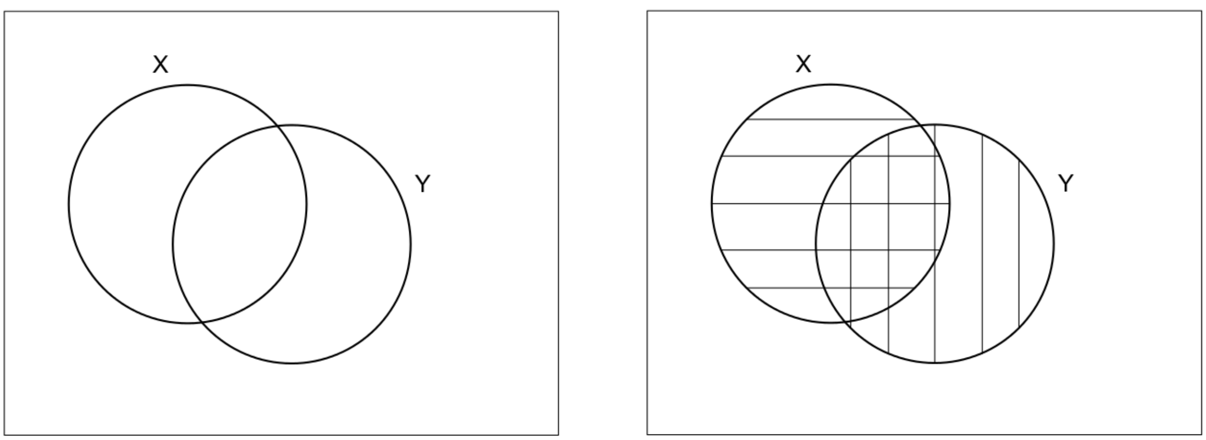
Theorem 5. *Let $X_1, X_2, \ldots X_n$ be finite sets. Then*

$$\begin{aligned}
|X_1 \cup X_2 \cup \cdots \cup X_n| \quad = \quad & |X_1| + |X_2| + \cdots + |X_n| \\
& -|X_1 \cap X_2| - |X_1 \cap X_3| - \cdots - |X_{n-1} \cap X_n| \quad \textit{(all intersections of two sets)} \\
& +|X_1 \cap X_2 \cap X_3| + \cdots + |X_{n-2} \cap X_{n-1} \cap X_n| \quad \textit{(all intersections of three sets)} \\
& \ \vdots \\
& +(-1)^{n-1}|X_1 \cap X_2 \cap \cdots \cap X_n|
\end{aligned}$$

This is a little confusing, so before we hit up the proof of Inclusion Exclusion, let's decode this a bit.

First, let's consider the case that you have two sets, X and Y, and wish to determine $|X \cup Y|$, as depicted in Figure 1a. We want to count how many elements are inside the circles.

The first idea, then, is to count the elements in X, and then also count the elements in Y, and add them up. We can see that in Figure 1b, where the horizontal hatches represent elements in X, and the vertical hatches represent elements in Y. But notice that any points in the intersection $X \cap Y$ get counted two times by such a strategy, which is obviously not going to give us an accurate count of the elements in $X \cup Y$. So, we can fix up our count by subtracting away this overcounting: adding on a $-|X \cap Y|$. This gives us exactly the result we see in Theorem 4: $|X \cup Y| = |X| + |Y| - |X \cap Y|$.



(a) (b)

Figure 1

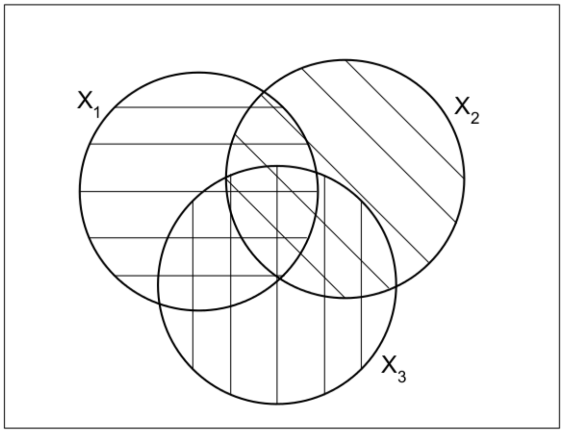
Now, what about Theorem 5? That seems a lot more complicated. To understand what's happening let's just add one more set to our picture, and look at $|X_1 \cup X_2 \cup X_3|$. We'll try the same strategy as before: first, we add up all the sizes of the three sets, represented by horizontal, vertical, and diagonal hatching, as shown in Figure 2a.

But as before, we are overcounting on the intersections. So we'll try to take away one count from each intersection, by adding on the term $-|X_1 \cap X_2| - |X_1 \cap X_3| - |X_2 \cap X_3|$. To represent this, I'll remove one set of hatching from each intersection of two sets, as shown in Figure 2b. Specifically, remove the horizontal hatching from $X_1 \cap X_2$, remove the vertical hatching from $X_1 \cap X_3$, and remove the diagonal hatching from $X_2 \cap X_3$.

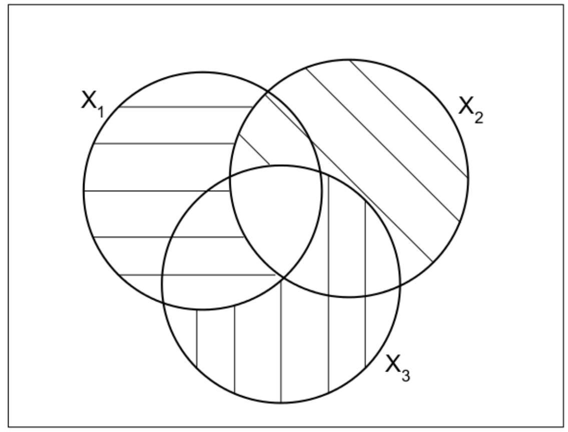
Well, we can see that we fixed up some of the overcount, but we also created a problem. We removed ALL the count from the very center, where all three sets intersect each other. So we need to add that back in, with a term of the form $+|X_1 \cap X_2 \cap X_3|$. If we put all this together, we come to the result of

Theorem 5:

$$|X_1 \cup X_2 \cup X_3| = |X_1| + |X_2| + |X_3| - |X_1 \cap X_2| - |X_1 \cap X_3| - |X_2 \cap X_3| + |X_1 \cap X_2 \cap X_3|.$$



(a)



(b)

Figure 2

To get the whole result of Theorem 5, this pattern would continue repeating itself. Every time we add a set, we add a layer to our counting. First, add up the sizes of the sets. Then, subtract off the sizes of the intersections of pairs. Then, add up the sizes of the intersections of triples. Then, subtract off the sizes of the intersections of quadruples. Yada, yada, yada.

Example 2. Call a number "prime-looking" if it is not divisible by 2, 3, or 5, but is composite. How many prime-looking numbers are there below 200?

Solution. We shall use in this proof the fact that there are 46 prime numbers between 1 and 200: 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113, 127, 131, 137, 139, 149, 151, 157, 163, 167, 173, 179, 181, 191, 193, 197, and 199.
We shall first determine how many numbers up to 100 are not divisible by 2, 3, or 5. Let

$$\begin{aligned} X_2 &= \{n \in \mathbb{N} \mid n < 200 \text{ and } n \text{ is divisible by } 2\} \\ X_3 &= \{n \in \mathbb{N} \mid n < 200 \text{ and } n \text{ is divisible by } 3\} \\ X_5 &= \{n \in \mathbb{N} \mid n < 200 \text{ and } n \text{ is divisible by } 5\}. \end{aligned}$$

Notice that $X_2 \cup X_3 \cup X_5$ is exactly those numbers up to 200 that are divisible by at least one of $2, 3, 5$, so the number of integers up to 200 that are NOT divisible by any of $2, 3, 5$ is $200 - |X_2 \cup X_3 \cup X_5|$. Notice, further, that if we consider X_d to be the set of integers up to 100 that are divisible by d, we have $|X_d| = \lfloor \frac{200}{d} \rfloor$.
Now, using the Principle of Inclusion-Exclusion, as in Theorem 5, we obtain

$$\begin{aligned} |X_2 \cup X_3 \cup X_5| &= |X_2| + |X_3| + |X_5| - |X_2 \cap X_3| - |X_2 \cap X_5| - |X_3 \cap X_5| + |X_2 \cap X_3 \cap X_5| \\ &= |X_2| + |X_3| + |X_5| - |X_6| - |X_{10}| - |X_{15}| + |X_{30}| \quad (*) \\ &= \left\lfloor \frac{200}{2} \right\rfloor + \left\lfloor \frac{200}{3} \right\rfloor + \left\lfloor \frac{200}{5} \right\rfloor - \left\lfloor \frac{200}{6} \right\rfloor - \left\lfloor \frac{200}{10} \right\rfloor - \left\lfloor \frac{200}{15} \right\rfloor + \left\lfloor \frac{200}{60} \right\rfloor \\ &= 100 + 66 + 40 - 33 - 20 - 13 + 3 \\ &= 143. \end{aligned}$$

where line (*) follows because a number divisible by both 2 and 3 is divisible by 6, a number divisible by both 2 and 5 is divisible by 10, etc.
Hence, the number of integers between 1 and 200 that are NOT divisible by any of 2, 3, 5 is $200 - 143 = 57$.
Finally, we eliminate those numbers between 1 and 200 that are not divisible by any of 2, 3, 5, but are also not composite; this would be any primes other than 2, 3, 5, and also the number 1. Hence, we have that the number of prime-looking numbers up to 200 is $57 - 43 - 1 = 13$. □

Now that we have a handle on what's going on here, let's prove the Inclusion-Exclusion Theorems. To do so, we will make use of the following Lemma:

Lemma 2. *If X and Y are finite sets, and $X \cap Y = \emptyset$, then $|X \cup Y| = |X| + |Y|$.*

Proof. Suppose that X and Y are finite sets, and $X \cap Y = \emptyset$. By definition, there exist $n, m \in \mathbb{N}$ and bijections $f : [n] \to X$ and $g : [m] \to Y$.

Define a function $h : [n + m] \to X \cup Y$ by

$$h(k) = \begin{cases} f(k) & \text{if } 1 \leq k \leq n \\ g(k-n) & \text{if } n+1 \leq k \leq m \end{cases}$$

We wish to prove that h is a bijection. First, suppose that $j, k \in [n + m]$ with $h(j) = h(k)$. Since $X \cap Y = \emptyset$, we cannot have that one of $h(j), h(k)$ is in X and the other from Y; hence both $h(j)$ and $h(k)$ are in the same set, either X or Y. Because f and g are themselves bijections, we therefore must have that $j = k$. Hence, h is injective.

Now, for surjectivity. Let $y \in X \cup Y$. If $y \in X$ then as f is a bijection, $\exists k \in [n]$ such that $f(k) = y$. Moreover, $h(k) = f(k) = y$. If $y \in Y$ than as g is a bijection, $\exists j \in [m]$ such that $g(j) = y$. Moreover, $h(j + n) = g(j) = y$. In either case, we have that y is in the image of h. Hence, h is surjective.

Therefore, h is bijective, and $|X \cup Y| = n + m = |X| + |Y|$. □

Now, let us prove Theorems 4 and 5.

Proof. [Proof of Theorem 4 To recall the statement of the Theorem, we wish to prove that for any finite sets X, Y we have $|X \cup Y| = |X| + |Y| - |X \cap Y|$.

First, note that $X \cup Y = X \cup (Y \backslash X)$, and that X and $Y \backslash X$ are disjoint. Therefore, $|X \cup Y| = |X| + |Y \backslash X|$ by Lemma 2. Solving for $|Y \backslash X|$, we obtain $|Y \backslash X| = |X \cup Y| - |X|$.

In addition, $Y = (Y \backslash X) \cup (X \cap Y)$, which are also disjoint sets. Therefore, $|Y| = |Y \backslash X| + |X \cap Y|$ by Lemma 2. Again solving for $|Y \backslash X|$, we obtain $|Y \backslash X| = |Y| - |X \cap Y|$.

Taking these two equations together, we thus obtain $|X \cup Y| - |X| = |Y| - |X \cap Y|$, and thus $|X \cup Y| = |X| + |Y| - |X \cap Y|$. □

Proof. [Proof of Theorem 5 To prove the full Inclusion-Exclusion Principle, we work by induction on n, the number of sets in the union.

The case that $n = 1$ is trivial. The case that $n = 2$ is proven in Theorem 4.

For the inductive hypothesis, suppose that for some $n \in \mathbb{N}$, we have the result:

$$\begin{aligned} |X_1 \cup X_2 \cup \cdots \cup X_n| &= |X_1| + |X_2| + \cdots + |X_n| \\ &\quad -|X_1 \cap X_2| - |X_1 \cap X_3| - \cdots - |X_{n-1} \cap X_n| \\ &\quad +|X_1 \cap X_2 \cap X_3| + \cdots + |X_{n-2} \cap X_{n-1} \cap X_n| \\ &\quad \vdots \\ &\quad +(-1)^{n-1}|X_1 \cap X_2 \cap \cdots \cap X_n| \end{aligned}$$

Consider the union $X_1 \cup X_2 \cup \cdots \cup X_n \cup X_{n+1}$.

Let $Y = X_1 \cup X_2 \cup \cdots \cup X_n$. Then by Theorem 4, we have that

$$|X_1 \cup X_2 \cup \cdots \cup X_n \cup X_{n+1}| = |Y \cup X_{n+1}| = |Y| + |X_{n+1}| - |Y \cap X_{n+1}|.$$

For $1 \leq k \leq n$, let $Y_k = X_k \cap X_{n+1}$. Notice that $Y \cap X_{n+1} = Y_1 \cup Y_2 \cup \cdots \cup Y_n$. We therefore can rewrite the above statement as

$$|X_1 \cup X_2 \cup \cdots \cup X_n \cup X_{n+1}| = |Y \cup X_{n+1}| = |X_1 \cup X_2 \cup \cdots \cup X_n| + |X_{n+1}| - |Y_1 \cup Y_2 \cup \cdots \cup Y_n|.$$

We can now apply the inductive hypothesis to the first and last terms. This yields the following:

$$\begin{aligned} |X_1 \cup X_2 \cup \cdots \cup X_n \cup X_{n+1}| &= |X_1| + |X_2| + \cdots + |X_n| \\ &\quad -|X_1 \cap X_2| - |X_1 \cap X_3| - \cdots - |X_{n-1} \cap X_n| \\ &\quad +|X_1 \cap X_2 \cap X_3| + \cdots + |X_{n-2} \cap X_{n-1} \cap X_n| \\ &\quad \vdots \\ &\quad +(-1)^{n-1}|X_1 \cap X_2 \cap \cdots \cap X_n| \\ &\quad +|X_{n+1}| \\ &\quad -|Y_1| - |Y_2| - \cdots - |Y_n| \\ &\quad +|Y_1 \cap Y_2| + |Y_1 \cap Y_3| + \cdots + |Y_{n-1} \cap Y_n| \\ &\quad -|Y_1 \cap Y_2 \cap Y_3| - \cdots - |Y_{n-2} \cap Y_{n-1} \cap Y_n| \\ &\quad \vdots \\ &\quad -(-1)^{n-1}|Y_1 \cap Y_2 \cap \cdots \cap Y_n|. \end{aligned}$$

Now, by definition, as $Y_k = X_k \cap X_{n+1}$, we have that for any $j_1, j_2, \ldots, j_\ell$, that

$$\begin{aligned} Y_{j_1} \cap Y_{j_2} \cap \cdots \cap Y_{j_\ell} &= (X_{j_1} \cap X_{n+1}) \cap (X_{j_2} \cap X_{n+1}) \cap \cdots \cap (X_{j_\ell} \cap X_{n+1}) \\ &= X_{j_1} \cap X_{j_2} \cap \cdots \cap X_{j_\ell} \cap X_{n+1}. \end{aligned}$$

Thus, any intersection of ℓ Y_j sets is in fact an intersection of $\ell + 1$ X_j sets. Moreover, any intersection of X_j sets that includes X_{n+1} can be expressed as an intersection involving the Y_j sets, and any intersection of X_j sets that does not include X_{n+1} can be expressed as intersection without the Y_j sets. By rearranging the large equation above, we thus obtain the desired result:

$$\begin{aligned}
|X_1 \cup X_2 \cup \cdots \cup X_n \cup X_{n+1}| &= |X_1| + |X_2| + \cdots + |X_n| + |X_{n+1}| \\
&\quad -|X_1 \cap X_2| - \cdots - |X_{n-1} \cap X_n| - |Y_1| - \cdots - |Y_n| \\
&\quad +|X_1 \cap X_2 \cap X_3| + \cdots + |X_{n-2} \cap X_{n-1} \cap X_n| + |Y_1 \cap Y_2| + \cdots + |Y_{n-1} \cap Y_n| \\
&\quad \vdots \\
&\quad +(-1)^n |Y_1 \cap Y_2 \cap \cdots \cap Y_n|. \\
&= |X_1| + |X_2| + \cdots + |X_n| + |X_{n+1}| \\
&\quad -|X_1 \cap X_2| - \cdots - |X_{n-1} \cap X_n| - |X_1 \cap X_{n+1}| - \cdots - |X_n \cap X_{n+1}| \\
&\quad +|X_1 \cap X_2 \cap X_3| + \cdots + |X_{n-2} \cap X_{n-1} \cap X_n| + |X_1 \cap X_2 \cap X_{n+1}| + \cdots + |X_{n-1} \cap X_n \cap \\
&\quad \vdots \\
&\quad +(-1)^n |X_1 \cap X_2 \cap \cdots \cap X_n \cap X_{n+1}|.
\end{aligned}$$

□

Finally, we close this basic finite counting with a theorem for Cartesian Products.

Theorem 6. *Let X and Y be finite sets. Then $|X \times Y| = |X||Y|$.*

Proof. First, if either X or Y is empty, then $X \times Y$ is also empty, and hence $|X \times Y| = 0 = |X||Y|$, and the result holds.

We proceed by induction on $|X|$. Let Y be any finite nonempty set, having $|Y| = n$, so that there exists a bijection $f : [n] \to Y$.

If $|X| = 1$, let $X = \{x\}$. Define $g : [n] \to X \times Y$ by $g(k) = (x, f(k))$. This is clearly bijective, and hence $|X \times Y| = n = 1 \cdot n = |X||Y|$.

Now, suppose that it is known that if $|X| = n$, then $|X \times Y| = |X||Y|$.

Let X be a set with $|X| = n + 1$. Then there exists $h : [n + 1] \to X$, a bijection. For simplicity of notation, write $x_k = h(k)$ for all $k \in [n + 1]$.

Notice that $X \times Y = (X \backslash \{x_{n+1}\} \times Y) \cup (\{x_{n+1}\} \times Y)$, and that these sets are disjoint. Applying the inductive hypothesis, the base case, and Lemma 2, we thus obtain

$$\begin{aligned}
|X \times Y| &= |(X \backslash \{x_{n+1}\} \times Y) \cup (\{x_{n+1}\} \times Y)| \\
&= |(X \backslash \{x_{n+1}\} \times Y)| + |(\{x_{n+1}\} \times Y)| \quad \text{(by Lemma 2)} \\
&= |X \backslash \{x_{n+1}\}||Y| + |Y| \\
&= n|Y| + |Y| = (n+1)|Y| = |X||Y|,
\end{aligned}$$

as desired.

Hence, by induction, the result holds for all finite sets X and Y. □