Informally, a *set* is any collection of objects. The ‘objects’ may be mathematical objects such as numbers, functions and even sets, or letters or symbols of any sort, or objects of thought, or objects that we observe in the real world. For the purposes of this course we are primarily interested in sets whose objects are mathematical in nature, but we will also consider other kinds of sets for illustrative purposes.

The objects of the set are called the *elements* of the set.

**Notation 1.1.** If $S$ is a set, and $a$ is an object, then we write $a \in S$ to mean that $a$ is an element of the set $S$ (we also say that $a$ belongs to $S$) and $a \notin S$ means that $a$ is not an element of the set $S$.

**Example 1.2.** Let $S$ be the set of all people and $T$ the set of all living people. Then

$$Julius\,\,Caesar \in S \quad \text{but} \quad Julius\,\,Caesar \notin T.$$ 

**Notation 1.3.** We use braces ({} and {}) to describe sets in an explicit way. The set whose elements are the letters $a$, $b$ and $c$ and the numbers $3$, $5$ and $7$ (and which has no other elements) is denoted

$$\{a, b, c, 3, 5, 7\}.$$ 

(As we will see, in describing a set the order in which elements are listed is irrelevant).

Here are some standard sets in mathematics:

1. The set of natural numbers $\mathbb{N} := \{1, 2, 3, \ldots\}$. Thus $1 \in \mathbb{N}$ but $-1 \notin \mathbb{N}$.
2. The set of integers $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$: $-57 \in \mathbb{Z}$, $\frac{1}{2} \notin \mathbb{Z}$.
3. The set of rational numbers $\mathbb{Q}$: $\frac{1}{2} \in \mathbb{Q}$, $58 \in \mathbb{Q}$, $\sqrt{2} \notin \mathbb{Q}$.
4. The set of all real numbers $\mathbb{R}$: $\sqrt{2} \in \mathbb{R}$, $\sqrt{-2} \notin \mathbb{R}$

**Example 1.4.** The elements of sets may be sets themselves. The following set has three elements, one of which is a set:

$$A := \{1, 2, \{3, 4\}\}.$$ 

Thus

$$1 \in A, \quad 2 \in A, \quad \{3, 4\} \in A, \quad \text{but} \quad 3 \notin A, \quad 4 \notin A.$$
1.1. Subsets.

Notation 1.5. If $A$ and $B$ are sets we write $A \subset B$ (and say ‘$A$ is contained in $B$’ or ‘$A$ is a subset of $B$’) if every element of $A$ is also an element of $B$: i.e.

$$A \subset B \iff \forall x \in A \Rightarrow x \in B.$$ 

Example 1.6. We have $N \subset Z$, $Z \subset Q$, $Q \subset R$.

Example 1.7. Let $A$ be the set $\{1, 2, \{3, 4\}\}$, as above. Then $\{1, 2\} \subset A$. However, $\{3, 4\} \not\subset A$ (since $3 \not\in A$, $4 \not\in A$).

Of course, $\{3, 4\} \in A$.

Example 1.8. Let $B$ be the set $\{1, 2, \{1, 2\}\}$. Then $\{1, 2\} \subset B$ and $\{1, 2\} \in B$. (It is, needless to say, a very rare situation when a set is both an element and a subset of another set.)

Here we prove an obvious statement:

Lemma 1.9. If $A$, $B$ and $C$ are sets and if $A \subset B$, $B \subset C$ then $A \subset C$.

Proof. Let $a$ be any element of $A$. Since $A \subset B$, $a \in B$. Since $B \subset C$ every element of $B$ is an element of $C$. In particular, $a \in C$. We’ve shown that every element of $A$ is also an element of $C$. Thus $A \subset C$. \qed

Remark 1.10. Note that, by definition, any set $A$ is a subset of itself (since it is always true that $x \in A \Rightarrow x \in A$). Thus, it is always true that $A \subset A$.

Definition 1.11. If $A$ is a set, then by a proper subset of $A$ we mean a subset other than $A$ itself. If $B$ is a proper subset of $A$ we sometimes denote this by

$$B \not\subset A.$$ 

1.2. The empty set. There is a unique set which has no elements. It is called the empty set or the nullset. It is denoted $\emptyset$. Thus the following statement follows:

$$\forall x, \ x \not\in \emptyset.$$

Note furthermore that the empty set is a subset of every other set (since the statement $x \in \emptyset \Rightarrow x \in A$ is vacuously true for any object $x$: the hypothesis $x \in \emptyset$ is always false). Thus for any set $A$ we have $\emptyset \subset A$.

Remark 1.12. Note that the empty set is itself a mathematical object; it is a particular example of a set. It is something rather than nothing. It may help to visualize it as an empty bag.

Example 1.13. Let $X = \{\emptyset\}$.

Then $X \not= \emptyset$, since $X$ has an element, namely $\emptyset$.

Observe that $\emptyset \subset X$ (true for every set) and $\emptyset \in X$ (rare).
1.3. **Unions of sets.** If $A$ and $B$ are sets then we form a new set $A \cup B$ (‘$A$ union $B$’) by pooling together the elements of $A$ and the elements of $B$ into a single set. Thus

$$x \in A \cup B \iff x \in A \text{ or } x \in B.$$ 

Note that any given element can occur in a set only once.

**Example 1.14.** If $A = \{1, 2, 3\}$ and $B = \{3, 5, 7\}$ the $A \cup B = \{1, 2, 3, 5, 7\}$. 

Similarly 

$$\{1, 2, a, b, c\} \cup \{1, 3, b, d, f\} = \{1, 2, 3, a, b, c, d, f\}.$$

**Example 1.15.** If $A$ is the set of all Irish people and $B$ is the set of all astronomers, then $A \cup B$ is the set of all people who are either Irish or astronomers.

**Remark 1.16.** Note that if $A$ and $B$ are any two sets, then $A \subset A \cup B$, $B \subset A \cup B$.

More generally, if $A_1$, $A_2$, $A_3$... is any list of sets, we define

$$A_1 \cup A_2 \cup A_3 = (A_1 \cup A_2) \cup A_3$$

to be the set consisting of those elements which belong to (at least) one of the three sets and

$$A_1 \cup A_2 \cup A_3 \cup A_4 = ((A_1 \cup A_2) \cup A_3) \cup A_4$$

e tc.

1.4. **Intersection of Sets.** If $A$ and $B$ are sets then we form a new set $A \cap B$ (‘$A$ intersect $B$’ or the intersection of $A$ and $B$) by including only the elements which belong both to $A$ and to $B$. Thus

$$x \in A \cap B \iff x \in A \text{ and } x \in B.$$ 

**Example 1.17.** If $A = \{1, 2, 3\}$ and $B = \{3, 5, 7\}$ the $A \cap B = \{3\}$.

$$\{1, 2, a, b, c\} \cap \{1, 3, b, d, f\} = \{1, b\}.$$ 

**Example 1.18.** If $A$ is the set of all Irish people and $B$ is the set of all astronomers, then $A \cap B$ is the set of all people who are both Irish and astronomers; i.e. the set of all Irish astronomers.

**Remark 1.19.** Note that if $A$ and $B$ are any two sets, then $A \cap B \subset A$, and $A \cap B \subset B$.

As in the case of unions, we can take intersections of three or more sets:

$$A_1 \cap A_2 \cap A_3 = (A_1 \cap A_2) \cap A_3$$

is the set consisting of those elements which belong to all three sets. Etc.
1.5. Disjoint sets.

Definition 1.20. We say that two sets $A$ and $B$ are disjoint if $A \cap B = \emptyset$; i.e. if they have no elements in common.

Example 1.21. The set of all cats is disjoint from the set of all dogs.

Example 1.22. The set $\mathbb{N}$ is disjoint from the set $\mathbb{R}_{<0} := \{x \in \mathbb{R} \mid x < 0\}$.

Example 1.23. The set of all integers divisible by 3 is not disjoint from the set of all integers divisible by 4, since the integer 12 belongs to both sets.

Definition 1.24. We say that a set $A$ is the disjoint union of two subsets $B$ and $C$ if

$$A = B \cup C \text{ and } B \cap C = \emptyset.$$ 

We often denote this by $A = B \sqcup C$.

Example 1.25. $\mathbb{Z}$ is the disjoint union of the set of even integers and the set of odd integers.

More generally we say that $A$ is the disjoint union of the subsets $A_1, \ldots, A_n$ if

1. $A = A_1 \cup \cdots \cup A_n$ and
2. $A_i \cap A_j = \emptyset$ when $i \neq j$ (we say that the collection $A_1, \ldots, A_n$ is pairwise disjoint)

and we denote this by $A = A_1 \sqcup A_2 \sqcup \cdots \sqcup A_n$.

Example 1.26. Let $A_0 = \{m \in \mathbb{Z} \mid m \equiv 0 \pmod{3}\}$, $A_1 = \{m \in \mathbb{Z} \mid m \equiv 1 \pmod{3}\}$ and $A_2 = \{m \in \mathbb{Z} \mid m \equiv 2 \pmod{3}\}$.

Then

$$\mathbb{Z} = A_0 \sqcup A_1 \sqcup A_2.$$ 

1.6. Equality of Sets. A set is completely determined by its elements: i.e. two sets are equal if they have precisely the same elements. Thus if $A$ and $B$ are sets then

$$A = B \iff A \subset B \quad \text{and} \quad B \subset A.$$ 

Thus if we wish to prove that $A = B$ we must prove that $A \subset B$ and that $B \subset A$, or equivalently that for any $x$, $x \in A \implies x \in B$ and $x \in B \implies x \in A$.

Here is a simple example of proving a (fairly obvious) statement using these definitions:

Lemma 1.27. Let $A$ and $B$ be any two sets. Prove that

$$A \subset B \iff B = A \cup B.$$
Proof. $\implies$: Suppose that $A \subset B$. We must show that $B = A \cup B$.

Now it is always true that $B \subset A \cup B$. It remains to prove that $A \cup B \subset B$:

Let $x \in A \cup B$. Then, by definition, $x \in B$ or $x \in A$. If $x \in B$ then $x \in B$. Otherwise $x \in A$ and hence $x \in B$ since $A \subset B$. So either way, $x \in B$ and we are done.

$\iff$: Conversely, suppose that $B = A \cup B$. Then $A \subset A \cup B$ (always true) $\implies A \subset B$. $\square$

1.7. Subsets defined by properties. Any given set has many subsets in general. In practice most are either uninteresting or impossible to describe or imagine. (It will never be possible to describe or fully imagine all the infinitely many subsets of the set $\mathbb{R}$ of real numbers.) The interesting or useful subsets are those which correspond to some property or list of properties.

Let $P(x)$ denote a statement about a variable $x$; eg, ‘$x$ is an astronomer’, ‘$x > 0$’, ‘$x$ is prime’.

If $a$ is the name of an object then we can substitute it in for $x$ to get a statement $P(a)$ which may be either true or false: ‘Kepler is an astronomer’, ‘$5 > 0$’, ‘$12$ is prime’.

If $A$ is any set, then we write \( \{a \in A \mid P(a)\} \) to denote the subset of $A$ consisting of those elements $a$ for which the statement $P(a)$ is true; it is the subset of $A$ described by the property $P$.

Example 1.28. $\mathbb{N} = \{m \in \mathbb{Z} \mid m > 0\}$.

Example 1.29. The interval $[1, 2]$ consisting of all real numbers between (or equal to) 1 and 2 is the set

\[ \{x \in \mathbb{R} \mid x \geq 1 \text{ and } x \leq 2\} = \{x \in \mathbb{R} \mid 1 \leq x \leq 2\}. \]

Example 1.30. If $A$ and $B$ are any sets then

\[ A \cap B = \{x \in A \mid x \in B\} = \{x \in B \mid x \in A\}. \]

Example 1.31. $\{m \in \mathbb{Z} \mid 2|m\}$ is the set of all even integers.

$\{m \in \mathbb{Z} \mid 3|m \text{ and } m > 11\}$ is the set of all integers which are greater than 11 and divisible by 3.

1.8. Differences and complements. If $A$ and $B$ are any sets then the (set-theoretic) difference $A \backslash B$ is defined by

\[ A \backslash B := \{x \in A \mid x \notin B\}. \]

Example 1.32. Let $A = \{1, 2, 3, a, b, c\}$ and $B = \{2, 4, 6, a, c, e\}$. Then $A \backslash B = \{1, 3, b\}$ and $B \backslash A = \{4, 6, e\}$.

Remark 1.33. Note that $A \cap B$ is always disjoint from $A \backslash B$ and that $A = (A \cap B) \sqcup (A \backslash B)$. 

Example 1.34. Let \( A = \{ m \in \mathbb{Z} \mid 2|m \} \) and \( A = \{ m \in \mathbb{Z} \mid 3|m \} \). Then
\[
A \setminus B = \{ m \in \mathbb{Z} \mid 2|m \text{ and } 3 \not|m \} = \{ \ldots, -2, 2, 4, 8, 10, 14, \ldots \}
\]
while
\[
A \cap B = \{ m \in \mathbb{Z} \mid 2|m \text{ and } 3|m \} = \{ m \in \mathbb{Z} \mid 6|m \} = \{ \ldots, -12, -6, 0, 6, 12, \ldots \}.
\]

Remark 1.35. If \( A \) and \( B \) are any sets and if \( x \in A \cup B \) then one and only one of the following three statements about \( x \) is true:

(1) \( x \) belongs to \( A \) but not \( B \): \( x \in A \setminus B \).

(2) \( x \) belongs to \( B \) but not \( A \): \( x \in B \setminus A \).

(3) \( x \) belongs to both \( A \) and \( B \): \( x \in A \cap B \).

Thus
\[
A \cup B = (A \setminus B) \sqcup (A \cap B) \sqcup (B \setminus A).
\]

Exercise 1.36. Draw a picture illustrating this last remark.

1.9. The Power Set of a set.

Definition 1.37. If \( A \) is any set then the power set of \( A \), which we will denote \( \mathcal{P}(A) \), is the set whose elements are precisely the subsets of \( A \).

Example 1.38. Let \( A = \{ a \} \). Then \( A \) has precisely two subsets; namely, \( \emptyset \) and \( A \). Thus
\[
\mathcal{P}(A) = \{ \emptyset, A \} = \{ \emptyset, \{ a \} \}.
\]

Remark 1.39. If \( A \) is any nonempty set, then \( A \) always has the two subsets \( \emptyset \) and \( A \). So \( \mathcal{P}(A) \) always has at least two elements.

Example 1.40. Let \( A = \{ 1, 2 \} \). Then \( A \) has four subsets: \( \emptyset \), \( \{ 1 \} \), \( \{ 2 \} \) and \( A = \{ 1, 2 \} \). Thus
\[
\mathcal{P}(A) = \{ \emptyset, \{ 1 \}, \{ 2 \}, \{ 1, 2 \} \}.
\]

Remark 1.41. The set \( \{ a \} \) whose only element is the object \( a \) should not be confused the object \( a \) itself.

Example 1.42. Note that \( \mathcal{P}(\emptyset) = \{ \emptyset \} \). This is not the empty set.

Exercise 1.43. Let \( A = \{ 1, 2, 3 \} \). Then \( A \) has eight subsets in total; \( \mathcal{P}(A) \) has eight elements. List them.

1.10. Cartesian products.

Definition 1.44. Let \( A \) and \( B \) be any nonempty sets (we allow \( A = B \)). The cartesian product of \( A \) and \( B \), denoted \( A \times B \) is the set whose elements are all ordered pairs of the form \( (a, b) \) where \( a \in A \) and \( b \in B \):
\[
A \times B := \{ (a, b) \mid a \in A, b \in B \}.
\]
Remark 1.45. The key property of ordered pairs is that order of the elements matters: If $a \neq b$ then $(a, b) \neq (b, a)$.

More generally the rule governing ordered pairs is the following:

The ordered pair $(a, b)$ is equal to the ordered pair $(c, d)$ if and only if $a = c$ and $b = d$.

Example 1.46. If $A = \{1, 2\}$ and $B = \{a, b, c\}$ then $A \times B = \{(1, a), (1, b), (1, c), (2, a), (2, b), (2, c)\}$. So $A \times B$ has $2 \cdot 3 = 6$ elements.

Example 1.47. $\mathbb{R} \times \mathbb{R}$, also denoted $\mathbb{R}^2$, is the cartesian plane consisting of all pairs $(x, y)$ of real numbers $x$ and $y$. It is visualized as an infinite plane containing two perpendicular lines (or ‘axes’), one horizontal and the other vertical, in which the element $(x, y)$ labels the point whose (perpendicular) distance to the vertical axis is $x$ and whose distance to the horizontal axis is $y$. The numbers $x$ and $y$ are the coordinates of the point. (See your calculus course for details.)

Example 1.48. $\mathbb{Z} \times \mathbb{Z} = \mathbb{Z}^2 \subset \mathbb{R}^2$ is the integer lattice in $\mathbb{R}^2$; it is the set of all points both of whose coordinates are integers.

Definition 1.49. If we have three sets $A$, $B$, $C$ we can form the product $A \times B \times C$ whose elements are all ordered triples $(a, b, c)$ with $a \in A$, $b \in B$ and $c \in C$.

More generally, if $A_1, \ldots, A_n$ is a list of sets then $A_1 \times \cdots \times A_n$ is the set whose elements are all ordered $n$-tuples $(a_1, \ldots, a_n)$ with $a_i \in A_i$ for $i = 1, 2, \ldots, n$.

Remark 1.50. Clearly we can identify the set $(A \times B) \times C$ with the set $A \times B \times C$. The elements of the first have the form $((a, b), c)$ and those of the second look like $(a, b, c)$. Either way, these are just ordered triples, and it is both convenient and reasonable to regard them as being the same.

2. Cardinality and Counting

Definition 2.1. A set with finitely many elements is called a finite set. Otherwise it is an infinite set.

Example 2.2. The set $\{1, 2, 3\}$ is finite.

The sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ are all infinite.

Example 2.3. The interval $[1, 2] = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$ is an infinite set, since it has infinitely many elements. (However it is a finite interval since it has finite length.)
**Definition 2.4.** If $A$ is a finite set then the number of elements in $A$ is called the cardinality of $A$ and is denoted $|A|$ (and in some texts as $\#A$).

**Example 2.5.** The cardinality of $\{1, 2, 3\}$ is 3:

$$|\{1, 2, 3\}| = 3.$$  

Similarly, $|\{a, b, c\}| = 3$.

**Example 2.6.** $|\{a, b, c, d\}| = 4$ but $|\{a, b, \{c, d\}\}| = 3$.

**Example 2.7.** $|\emptyset| = 0$.

$|\mathcal{P}(\emptyset)| = 1$ (see Example 1.42).

We will use repeatedly below the following important (and ‘obvious’) **basic counting principle**:

If $A$ is a finite set and if

$$A = A_1 \cup A_2 \cup \cdots \cup A_n$$

then

$$|A| = |A_1| + |A_2| + \cdots + |A_n|.$$  

We’ll begin with some simple applications:

**Lemma 2.8.** Let $A_1$ and $A_2$ be any two finite sets. Then

$$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|.$$  

**Proof.** Recall that $A_1 \cup A_2 = (A_1 \setminus A_2) \cup (A_1 \cap A_2) \cup A_2 \setminus A_1$. Thus

(1) $$|A_1 \cup A_2| = |A_1 \setminus A_2| + |A_1 \cap A_2| + |A_2 \setminus A_1|.$$  

But

$$A_1 = (A_1 \setminus A_2) \cup (A_1 \cap A_2)$$

$$\Rightarrow |A_1| = |A_1 \setminus A_2| + |A_1 \cap A_2|$$

$$\Rightarrow |A_1 \setminus A_2| = |A_1| - |A_1 \cap A_2|.$$  

Similarly $|A_2 \setminus A_1| = |A_2| - |A_1 \cap A_2|$. Substituting these back into equation (1) gives

$$|A_1 \cup A_2| = (|A_1| - |A_1 \cap A_2|) + |A_1 \cap A_2| + (|A_2| - |A_1 \cap A_2|).$$  

We can extend to this to a union of three sets:

**Lemma 2.9.** If $A_1$, $A_2$ and $A_3$ are any three finite sets then

$$|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|.$$  

Proof. Applying Lemma 2.8 to two sets $A_1 \cup A_2$ and $A_3$ we obtain
(2) \[ |A_1 \cup A_2 \cup A_3| = |A_1 \cup A_2| - (A_1 \cup A_2) \cap A_3| . \]
Now \( (A_1 \cup A_2) \cap A_3 = (A_1 \cap A_3) \cup (A_2 \cap A_3) \) (Prove this!)
Thus applying Lemma 2.8 to $A_1 \cap A_3$ and $A_2 \cap A_3$ gives
\[ |(A_1 \cup A_2) \cap A_3| = |A_1 \cap A_3| + |A_2 \cap A_3| - |A_1 \cap A_2 \cap A_3| . \]

Also, by Lemma 2.8,
\[ |A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2| . \]
Substituting these formulae back into equation 2 gives the result. \( \square \)

Remark 2.10. These last two lemmas are examples of the Inclusion-Exclusion principle for counting.

Exercise 2.11. Figure out the general formula for counting $|A_1 \cup \cdots \cup A_n|$. Prove it by induction on $n$. (This is difficult. Congratulations if you can do it by yourself.)

Here are some examples of the Inclusion-Exclusion principle in action:

Example 2.12. A large group of people contains 100 Irish people and 200 students. Of these, 43 are Irish students. How many people in the group are either Irish or students.

Solution: Let $I$ be the set of Irish people in the group, and $S$ the set of students. Then
\[ |I \cup S| = |I| + |S| - |I \cap S| \]
\[ = 100 + 200 - 43 = 257. \]

Example 2.13. How many positive integers less than or equal to 2000 are divisible by either 3 or 4 or 5?

Solution: We begin by recalling that the number of positive integers less than or equal to $N$ which are divisible by $d$ is $[N/d]$.

Let $A_1 = \{m \leq 2000 \mid 3|m\}$, $A_2 = \{m \leq 2000 \mid 4|m\}$ and $A_3 = \{m \leq 2000 \mid 5|m\}$. Thus we are asked to calculate $|A_1 \cup A_2 \cup A_3|$. Now $|A_1| = [2000/3] = 666$. $|A_2| = [2000/4] = 500$ and $|A_3| = [2000/5] = 400$.

$A_1 \cap A_2$ is the set of all positive integers less than or equal to 2000 which are divisible by both 3 and 4. Since $(3, 4) = 1$, an integer is divisible by 3 and 4 if and only if it is divisible by $3 \cdot 4 = 12$. Thus $A_1 \cap A_2 = \{m \leq 2000 \mid 12|m\}$ and $|A_1 \cap A_2| = [2000/12] = 166.$
Similarly, \( A_1 \cap A_3 = \{ m \leq 2000 \mid 15|m \} \) and hence \( |A_1 \cap A_3| = \lfloor 2000/15 \rfloor = 133 \).

\( A_2 \cap A_3 = \{ m \leq 2000 \mid 20|m \} \) and hence \( |A_1 \cap A_3| = \lfloor 2000/20 \rfloor = 100 \).

Finally \( A_1 \cap A_2 \cap A_3 = \{ m \leq 2000 \mid 60|m \} \) and thus \( |A_1 \cap A_2 \cap A_3| = \lfloor 2000/60 \rfloor = 33 \).

Therefore

\[
|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3| \\
= 666 + 500 + 400 - 166 - 133 - 100 + 33 \\
= 1200.
\]

Thus exactly 1200 of the numbers from 1 to 2000 are divisible either by 3, 4 or 5 (and therefore the remaining 800 are divisible by none of these three numbers).

**Lemma 2.14.** Let \( A \) and \( B \) be finite sets. Then \( |A \times B| = |A| \cdot |B| \).

**Proof.** Let \( m = |B| \). We label the elements of \( B \): \( b_1, \ldots, b_m \). Then

\[
A \times B = A \times \{b_1\} \sqcup A \times \{b_2\} \sqcup \cdots \sqcup A \times \{b_m\}.
\]

Therefore

\[
|A \times B| = |A \times \{b_1\}| + \cdots + |A \times \{b_m\}| \\
= |A| + \cdots + |A| \\
= |A| \cdot m = |A| \cdot |B|.
\]

\( \square \)

**Corollary 2.15.** Let \( A_1, \ldots, A_n \) be finite sets. Then

\[
|A_1 \times \cdots \times A_n| = |A_1| \cdot |A_2| \cdots |A_n|.
\]

**Proof.** We will prove this by induction on \( n \geq 2 \).

The case \( n = 2 \) is the Lemma just proved.

Suppose the result is known for \( n \) and that \( A_1, \ldots, A_n, A_{n+1} \) are sets. Then

\[
|A_1 \times \cdots \times A_n \times A_{n+1}| = |(A_1 \times \cdots \times A_n) \times A_{n+1}| \\
= (|A_1 \times \cdots \times A_n|) \cdot |A_{n+1}| \text{ by the case } n = 2 \\
= (|A_1| \cdots |A_n|) \cdot |A_{n+1}| \text{ by our ind. hyp.} \\
= |A_1| \cdots |A_n| \cdot |A_{n+1}| \text{ as required.}
\]

\( \square \)
If $A$ is any set we let
\[ A^n := A \times \cdots \times A \]

Thus, taking $A_1 = A_2 = \cdots = A_n = A$ in the last corollary we obtain:

**Corollary 2.16.** Let $A$ be any finite set. Then $|A^n| = |A|^n$.

**Example 2.17.** The elements of the set $\{0, 1\}^n$ are called binary strings of length $n$.

For example, $\{0, 1\}^3 = \{0, 1\} \times \{0, 1\} \times \{0, 1\}$. Thus, the binary strings of length 3 are: $(0, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1), (1, 0, 0), (1, 0, 1), (1, 1, 0), (1, 1, 1)$. There are eight binary strings of length 3.

In general, by the last corollary, the number of binary strings of length $n$ is
\[ |\{0, 1\}^n| = |\{0, 1\}|^n = 2^n. \]

### 3. Functions

**Notation 3.1.** Let $A$ and $B$ be any two sets. Informally, a function (or ‘map’) $f$ from $A$ to $B$ associates to each element of $A$ one and only one element of $B$. The set $A$ is called the domain of the function $f$, and the set $B$ is called the codomain or target of $f$. (We will sometimes refer to elements of the domain as ‘inputs’ of the function.)

If $a \in A$, the element of the set $B$ which the function $f$ associates to $a$ is called the value of $f$ at $a$ and is usually denoted $f(a)$. (We will sometimes refer to this as the ‘output’ of $f$ at $a$, but this is not really standard mathematical terminology.)

We often encapsulate this information with the notation:
\[ f : A \rightarrow B, \ a \mapsto f(a). \]

**Example 3.2.** Let $A = \{1, 2, 3\}$ and let $B = \{a, b\}$. We can define or construct maps from $A$ to $B$ at will, simply by specifying the value of the function at each input, 1, 2 and 3:

For example, let $f : A \rightarrow B$ be the following function:
\[
\begin{align*}
f(1) &= b, \\
f(2) &= b, \\
f(3) &= a. 
\end{align*}
\]
Let $g : A \rightarrow B$ be the function
\[
\begin{align*}
g(1) &= b \\
g(2) &= a \\
g(3) &= b.
\end{align*}
\]

(Q: How many functions are there from $A$ to $B$?)

Similarly, we can construct functions from $B$ to $A$: Let $h : B \rightarrow A$ be the function
\[
\begin{align*}
h(a) &= 2 \\
h(b) &= 3.
\end{align*}
\]

(Q: How many functions are there from $B$ to $A$?)

If the domain of a function is infinite, then, of course, we cannot describe the function, input by input, in this way.

The functions of mathematical interest in such cases are usually constructed via a formula or procedure of some kind: i.e. we describe a procedure for detereming the value of the function at a typical element of the domain $x$.

**Example 3.3.** The **squaring function** on $\mathbb{R}$ is the function
\[
f : \mathbb{R} \rightarrow \mathbb{R}, \ x \mapsto x^2.
\]
Thus for any real number $x$, the value of this function is $x^2$; i.e. $\forall x \in \mathbb{R}, \ f(x) = x^2$.

Thus
\[
f(0) = 0^2 = 0, \ f(5) = 5^2 = 25, \ f(\pi) = \pi^2, \ f(-5) = (-5)^2 = 25, \ldots
\]

**Example 3.4.** The **identity function** on $\mathbb{R}$ is the function
\[
f : \mathbb{R} \rightarrow \mathbb{R}, \ x \mapsto x.
\]
Thus $f(x) = x$ for all real numbers $x$.

**Example 3.5.** In fact, we can do the same for any set $A$.

The identity function on a set $A$ is the function whose value at any input $a$ is again $a$.

This is one of the most basic and important examples of a function. We will use the notation $\text{Id}_A$ for this function.

Thus, the domain and codomain of $\text{Id}_A$ are both $A$:
\[
\text{Id}_A : A \rightarrow A, \ a \mapsto a.
\]
If we take the set \( A = \{1, 2, 3\} \), for example, then
\[
\text{Id}_A(1) = 1 \\
\text{Id}_A(2) = 2 \\
\text{Id}_A(3) = 3
\].

**Example 3.6.** We have already met the floor \( \lfloor x \rfloor \) of a real number \( x \): it is the largest integer which is less than or equal to \( x \). We can use it to construct a function:
\[ f : \mathbb{R} \to \mathbb{Z}, \ x \mapsto \lfloor x \rfloor. \]
So \( f(\pi) = \lfloor \pi \rfloor = 3 \), \( f(78.2) = \lfloor 78.2 \rfloor = 78 \), \( f(28/6) = \lfloor 28/6 \rfloor = 4 \) etc.

**Example 3.7.** Let \( A \) be any set and let \( B \) be a subset of \( A \).

The indicator function of \( B \) in \( A \) is the function, \( 1_B \), which sends \( a \) to 1 if \( a \in B \) and to 0 otherwise:
\[
1_B : A \to \{0, 1\}
\]
defined by
\[
\forall a \in A, \ 1_B(a) = \begin{cases} 
1, & a \in B \\
0, & a \not\in B.
\end{cases}
\]
For example, if \( A = \{a, b, c, d\} \) and \( B = \{a, c\} \) then
\[
1_B(a) = 1 \\
1_B(b) = 0 \\
1_B(c) = 1 \\
1_B(d) = 0.
\]

**Example 3.8.** Let \( B \subset A \). The the map \( f : B \to A \), \( f(b) = b \) is called the inclusion map of \( B \) in \( A \). (Note that if \( B \neq A \) it is not equal to the identity function \( \text{Id}_B \); it has a different codomain.)

3.1. **Composition of Functions.** Suppose that \( A \), \( B \) and \( C \) are any sets (not necessarily distinct from each other) and that \( f : A \to B \) and \( g : B \to C \) are functions. Thus if \( a \in A \), the value of \( f \) at \( a \), \( f(a) \), is an element of \( B \), and therefore is an admissible input for the function \( g \). The value of \( g \) at \( f(a) \), \( g(f(a)) \) is an element of \( C \). Thus, using \( f \) and \( g \), we have a method – by doing \( f \) first, and then \( g \) – to associate to any element \( a \in A \) an element \( g(f(a)) \) in \( C \); i.e. by doing \( g \) after \( f \), we have a new function from \( A \) to \( C \).

**Definition 3.9.** Given two functions \( f : A \to B \) and \( g : B \to A \), \( g \) composed with \( f \), denoted \( g \circ f \) is the function described as follows:
\[
g \circ f : A \to C, \ a \mapsto g(f(a)).
\]
Example 3.10. Let $A = \{1, 2, 3\}$, $B = \{a, b\}$ and $C = \{5, 6, 7\}$. Let $f : A \to B$ be the function

$$f(1) = a$$
$$f(2) = b$$
$$f(3) = b$$

and let $g : B \to C$ be the function

$$g(a) = 5$$
$$g(b) = 7.$$  

Then $g \circ f : A \to C$ is the function

$$(g \circ f)(1) = g(f(1)) = g(a) = 5$$
$$(g \circ f)(2) = g(f(2)) = g(b) = 7$$
$$(g \circ f)(3) = g(f(3)) = g(b) = 7.$$  

Example 3.11. Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by $f(x) = x^3$, and let $g : \mathbb{R} \to \mathbb{R}$ be the function $g(x) = x + 3$.

Then the function $g \circ f : \mathbb{R} \to \mathbb{R}$ is given by

$$(g \circ f)(x) = g(f(x)) = g(x^3) = x^3 + 3 \quad \forall x \in \mathbb{R}.$$  

On the other hand, the composite $f \circ g : \mathbb{R} \to \mathbb{R}$ is also defined, and it is given by

$$(f \circ g)(x) = f(g(x)) = f(x + 3) = (x + 3)^3 \quad \forall x \in \mathbb{R}.$$  

Note that the functions $g \circ f$ and $f \circ g$ are quite different.

Remark 3.12. Observe that if $f$ and $g$ are any two functions then the composite function $g \circ f$ is defined only if the codomain of $f$ is the same as the domain of $g$. When this happens we say that $g$ can be composed with $f$.

3.2. Equality of Functions. It will be important to us below to understand exactly what is meant by saying that the function $f$ equals the function $g$:

Definition 3.13. Let $f$ and $g$ be any two functions. The $f = g$ if and only if

1. The domain of $f$ is equal to the domain of $g$,
2. The codomain of $f$ is equal to the codomain of $g$ and
3. $f(a) = g(a)$ for all elements $a$ of the domain.

Example 3.14. Let $f : \mathbb{R} \to \mathbb{Z}$ be the function $f(x) = \lfloor x \rfloor$ and let $g : \mathbb{R} \to \mathbb{R}$ be the function $g(x) = \lfloor x \rfloor$. Then $f \neq g$ since the codomain of $f$ is not the same as the codomain of $g$. 
Example 3.15. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be the function $f(x) = x^2$ and let $g : [0, \infty) \rightarrow \mathbb{R}$ be the function $g(x) = x^2$. (Here $[0, \infty)$ is a notation for the interval $\{x \in \mathbb{R} \mid x \geq 0\}$ of all nonnegative real numbers.) Then $f \neq g$ since they have different domains.

Example 3.16. Let $A = \{1, 2, 3\}$, $B = \{a, b, c\}$, $C = A = \{1, 2, 3\}$. Consider the following four functions:

\[ f_1 : A \rightarrow B, \quad f_1(1) = a, \quad f_1(2) = b, \quad f_1(3) = c \]
\[ f_2 : A \rightarrow B, \quad f_2(1) = b = f_2(2), \quad f_2(3) = c \]
\[ g_1 : B \rightarrow C, \quad g_1(a) = 1 = g_1(b), \quad g_1(c) = 2 \]
and
\[ g_2 : B \rightarrow C, \quad g_2(a) = 3, \quad g_2(b) = 1, \quad g_2(c) = 2. \]

Note that $f_1 \neq f_2$ and $g_1 \neq g_2$.

However, I Claim that $g_1 \circ f_1 = g_2 \circ f_2$.

First note that both of these functions have the same domain ($A$) and the same codomain ($C$). We need finally to check that they take the same value at each input:
\[ (g_1 \circ f_1)(1) = g_1(f_1(1)) = g_1(a) = 1, \quad (g_2 \circ f_2)(1) = g_2(f_2(1)) = g_2(b) = 1 \]
\[ (g_1 \circ f_1)(2) = g_1(f_1(2)) = g_1(b) = 1, \quad (g_2 \circ f_2)(2) = g_2(f_2(2)) = g_2(b) = 1 \]
\[ (g_1 \circ f_1)(3) = g_1(f_1(3)) = g_1(c) = 2, \quad (g_2 \circ f_2)(3) = g_2(f_2(3)) = g_2(c) = 2 \]

Lemma 3.17. Let $f : A \rightarrow B$ be any function. Then
\[
(1) \quad f \circ \text{Id}_A = f.
(2) \quad \text{Id}_B \circ f = f.
\]

Proof. (1) First note that $f$ can be composed with $\text{Id}_A$ since the codomain of $\text{Id}_A$ is $A$, which is the domain of $f$.

The functions $f$ and $f \circ \text{Id}_A$ both have domain $A$ and codomain $B$.

Finally, for any $a \in A$, we have
\[
(f \circ \text{Id}_A)(a) = f(\text{Id}_A(a)) = f(a).
\]

So the two functions are equal.

(2) Similar. Do it as an exercise.

Remark 3.18. Note that if $f : A \rightarrow B$ is a function and if $A \neq B$, then the composites $\text{Id}_A \circ f$ and $f \circ \text{Id}_B$ are not defined.
3.3. Surjective maps.

**Definition 3.19.** Given a function \( f : A \to B \), the image (or range) of \( f \), denoted \( \text{Image}(f) \) or \( f(A) \), is the subset of \( B \) consisting of all values (or outputs) of the function \( f \); i.e. it is the set \( \{ b \in B \mid b = f(a) \text{ for some } a \in A \} \).

**Definition 3.20.** A map \( f : A \to B \) is surjective or onto if \( \text{Image}(f) = B \); i.e. \( f \) is surjective if for every \( b \in B \) there is at least one \( a \in A \) with \( f(a) = b \).

**Remark 3.21.** Another informal way to say this is that the function \( f \) is surjective if every element of the target gets hit.

**Example 3.22.** The doubling function \( f : \mathbb{R} \to \mathbb{R}, \ f(x) = 2x \) is surjective.

**Proof.** To prove this we must show that any element, \( y \) say, in the codomain \( \mathbb{R} \) lies in the image of this function; i.e. we must show that every real number \( y \) is twice some other real number.

We spell out the argument:

Let \( y \in \mathbb{R} \). Then

\[
y = 2 \cdot \frac{y}{2} = f \left( \frac{y}{2} \right).
\]

□

**Example 3.23.** The squaring function \( f : \mathbb{R} \to \mathbb{R}, \ f(x) = x^2 \) is not surjective. The number \(-1 \in \mathbb{R}\) does not lie in the image of this map.

**Example 3.24.** The function \( f : \mathbb{R} \to [0, \infty) \) is onto: Let \( y \in [0, \infty) \). Then \( y \geq 0 \). So \( y = (\sqrt{y})^2 = f(\sqrt{y}) \) lies in \( \text{Image}(f) \).

**Example 3.25.** The map \( f : \mathbb{R} \to \mathbb{R}, \ x \mapsto \sin(x) \) is not surjective (for example 2 is not in the image).

But the map \( f : \mathbb{R} \to [-1, 1], \ x \mapsto \sin(x) \) is surjective.

For real functions – functions \( f : \mathbb{R} \to \mathbb{R} \) – surjectivity can often be read off from the graph of the function. Recall that the graph of the function is the set \( \{(x, y) \in \mathbb{R}^2 \mid y = f(x)\} \subset \mathbb{R}^2 \). We have the following criterion:

**Lemma 3.26.** Let \( f : \mathbb{R} \to \mathbb{R} \) be a real function. Then \( f \) is surjective if and only if every horizontal line cuts the graph at least once.

**Proof.** Suppose \( f \) is surjective. Let \( L \) be a horizontal line. The equation of \( L \) is \( y = b \) for some real number \( b \). Since \( f \) is surjective, \( b = f(a) \) for some \( a \in \mathbb{R} \). Thus the point \((a, b)\) lies on the graph and the line \( L \).

Conversely suppose every horizontal line cuts the graph. Let \( b \in \mathbb{R} \). The horizontal line \( y = b \) cuts the graph at some point; i.e. the graph contains
the point \((a, b)\) for some \(a \in \mathbb{R}\). By definition of the graph, it follows that \(b = f(a)\). Thus \(b \in \text{Image}(f)\) and \(f\) is surjective. \(\square\)

The following is an important general criterion for the surjectivity of a map:

**Theorem 3.27.** Let \(f : A \to B\) be a function. The following statements are equivalent:

1. \(f\) is surjective.
2. There exists a function \(g : B \to A\) with the property that \(f(g(b)) = b\) for all \(b \in B\).

**Proof.** (1)\(\implies\) (2): Suppose \(f\) is surjective. We describe how to construct a function \(g : B \to A\) with the required property:

For each \(b \in B\) choose some \(a \in A\) with \(f(a) = b\) and define \(g(b) = a\). Then, by definition, \(f(g(b)) = f(a) = b\).

(2)\(\implies\) (1): Suppose that a function \(g\) exists with the stated property. Let \(b \in B\). Let \(a = g(b)\). Then \(f(a) = f(g(b)) = b\). So \(b \in \text{Image}(f)\) and \(f\) is surjective. \(\square\)

**Example 3.28.** Let \(f : \mathbb{R} \to [0, \infty)\) be the function \(f(x) = x^2\). We have seen already that \(f\) is surjective. Let \(g : [0, \infty) \to \mathbb{R}\) be the function \(g(x) = \sqrt{x}\). Then \(g\) has the property stated in the theorem: Let \(b \in [0, \infty)\). Then \(f(g(b)) = f(\sqrt{b}) = (\sqrt{b})^2 = b\).

However \(g\) is not the only function with this property. For example the function \(h : [0, \infty) \to \mathbb{R}, x \mapsto -\sqrt{x}\) also has this property: Let \(b \in [0, \infty)\). Then \(f(h(b)) = (-\sqrt{b})^2 = b\).

**Example 3.29.** Let \(f : \mathbb{R} \to \mathbb{Z}\) be the floor function \(x \mapsto \lfloor x \rfloor\). \(f\) is surjective. In fact it is easy to find a function \(g\) as in the theorem. For example, let \(g : \mathbb{Z} \to \mathbb{R}\) be the inclusion map defined by \(g(n) = n\). Then, for all \(n \in \mathbb{Z}\), \(f(g(n)) = f(n) = \lfloor n \rfloor = n\).

However there are an infinity of other functions \(\mathbb{Z} \to \mathbb{R}\) with the required property. For example, fix some number \(a\) lying between 0 and 1. We use this number to define a function from \(\mathbb{Z}\) to \(\mathbb{R}\) which we’ll name \(g_a:\)

\[g_a : \mathbb{Z} \to \mathbb{R}, \ n \mapsto n + a.\]

Then for every \(n \in \mathbb{Z}\)

\[f(g_a(n)) = f(n + a) = \lfloor n + a \rfloor = n \text{ since } n < n + a < n + 1.\]

Let us consider the situation described in the Theorem: We have two functions \(f : A \to B\) and \(g : B \to A\) satisfying \(f(g(b)) = b\) for all \(b \in B\). Note that this later statement is entirely equivalent to the equation of functions

\[f \circ g = \text{Id}_B.\]
Thus we can rephrase the Theorem as saying that \( f : A \to B \) is surjective if and only if there is a function \( g : B \to A \) satisfying \( f \circ g = \text{Id}_B \).

**Definition 3.30.** Suppose given two functions \( f : A \to B \) and \( g : B \to A \) which satisfy \( f \circ g = \text{Id}_B \). Then \( g \) is said to be a right inverse of \( f \) and \( f \) is said to be a left inverse of \( g \).

With this terminology, the theorem says that a map is surjective if and only if it has a right inverse. Note, however, that our examples show that a surjective function can have many – even infinitely many – distinct right inverses.

### 3.4. Injective maps.

**Definition 3.31.** A function \( f : A \to B \) is said to be injective or 1-to-1 if it has the property that unequal elements of the domain have unequal values: For all \( a_1, a_2 \in A \) if \( a_1 \neq a_2 \) then \( f(a_1) \neq f(a_2) \).

Note that an equivalent – and often more useful – way to state this is: \( f \) is injective if it has the property that, for all \( a_1, a_2 \in A \), \( f(a_1) = f(a_2) \implies a_1 = a_2 \).

**Example 3.32.** The doubling function \( f : \mathbb{R} \to \mathbb{R}, x \mapsto 2x \) is injective: For suppose \( f(a_1) = f(a_2) \). Then \( 2a_1 = 2a_2 \). Dividing by 2, we get \( a_1 = a_2 \) as required.

**Example 3.33.** Let \( f : \mathbb{R} \to \mathbb{R} \) be the map \( f(x) = x^2 \). Then \( f \) is not injective: \(-1 \neq 1 \) in \( \mathbb{R} \), but \( f(-1) = 1 = f(1) \). So two different inputs yield the same output.

**Example 3.34.** However, the function \( f : [0, \infty) \to \mathbb{R}, x \mapsto x^2 \) is injective. For suppose that \( f(a_1) = f(a_2) \). Then \( a_1^2 = a_2^2 \). It follows that \( \sqrt{a_1^2} = \sqrt{a_2^2} \). But \( a_1 = \sqrt{a_1^2} \) and \( a_2 = \sqrt{a_2^2} \) since \( a_1, a_2 \geq 0 \). Thus \( a_1 = a_2 \) as required.

**Example 3.35.** The map \( f : \{1, 2, 3\} \to \{1, 2, 3\} \) defined by \( f(1) = 1 \), \( f(2) = 3 \) and \( f(3) = 1 \) is not injective since \( 1 \neq 3 \) but \( f(1) = f(3) \).

**Lemma 3.36.** Let \( f : \mathbb{R} \to \mathbb{R} \) be a function. Then \( f \) is injective if and only if every horizontal line in \( \mathbb{R}^2 \) intersects the graph at most once.

**Proof.** Suppose that \( f \) is injective and that \( L \) is a horizontal line. The equation of \( L \) is \( y = b \) for some \( b \in \mathbb{R} \). Suppose, for the sake of contradiction that \( L \) intersects the graph at two points. Thus there are distinct points \((a_1, b)\) and \((a_2, b)\) lying on \( L \) and on the graph of \( f \). By definition of the graph, we have \( f(a_1) = f(a_2) \), which contradicts injectivity of \( f \).

Conversely, suppose that every horizontal line cuts the graph at most once. Suppose that \( f(a_1) = f(a_2) \) for some \( a_1, a_2 \in \mathbb{R} \). We must show that \( a_1 = a_2 \). Let \( b = f(a_1) = f(a_2) \). The points \((a_1, b)\) and \((a_2, b)\) lie on the
We also have the following general criterion of injectivity:

**Theorem 3.37.** A function \( f : A \to B \) is injective if and only if it has a left inverse \( g : B \to A \); i.e. if and only if there is a function \( g : B \to A \) satisfying \( g \circ f = \text{Id}_A \).

**Proof.** Observe that the condition \( g \circ f = \text{Id}_A \) is equivalent to \( g(f(a)) = a \) for all \( a \in A \).

Suppose now that \( f \) is injective. We construct a right inverse \( g : B \to A \) as follows: Choose once and for all some fixed element, call it \( p \), in \( A \). Let \( b \in B \). We must describe the element \( g(b) \in A \). There are two possibilities:

1. If \( b \in \text{Image}(f) \), then \( b = f(a) \) for some unique \( a \in A \). We set \( g(b) = a \).
2. If \( b \not\in \text{Image}(f) \), we simply let \( g(b) = p \).

Now let \( a \) be any element of \( A \). Let \( b = f(a) \). Then \( b \in \text{Image}(f) \) and \( a \) is the unique element of \( A \) which is mapped to \( b \) by \( f \). So \( g(b) = a \) by definition. Thus \( g(f(a)) = g(b) = a \) as required.

Conversely, suppose that \( f \) has a left inverse \( g : B \to A \). We will show that \( f \) is injective: Suppose that \( a_1, a_2 \in A \) satisfy \( f(a_1) = f(a_2) \). Then \( a_1 = g(f(a_1)) = g(f(a_2)) = a_2 \).

**Example 3.38.** Let \( f : \mathbb{Z} \to \mathbb{R} \) be the inclusion map \( n \mapsto n \). This is clearly injective. Note that the floor function \( g : \mathbb{R} \to \mathbb{Z} \), \( x \mapsto \lfloor x \rfloor \) is a left inverse: For all \( n \in \mathbb{Z} \), \( g(f(n)) = g(n) = \lfloor n \rfloor = n \).

There are, however, (infinitely) many other left inverses of \( f \). For example, let \( h : \mathbb{R} \to \mathbb{Z} \) be the function described as follows:

\[
h(x) = \begin{cases} 
  x, & x \in \mathbb{Z} \\
  0, & x \not\in \mathbb{Z}
\end{cases}
\]

Then \( h(f(n)) = h(n) = n \) for all \( n \in \mathbb{Z} \); i.e. \( h \circ f = \text{Id}_\mathbb{Z} \).

3.5. **Bijective maps.**

**Definition 3.39.** A map \( f : A \to B \) is said to be bijective if it is both injective and surjective. We often call such a map a bijection or a bijective correspondence.

Thus \( f : A \to B \) is bijective if every element \( b \) of \( B \) is the value of \( f \) at one and only one element \( a \) and \( A \); i.e. the function \( f \) pairs the elements of \( A \) with the elements of \( B \) in such a way that each element of \( A \) gets paired with one and only one element \( B \) and vice versa.
Example 3.40. The map \( f : \{1, 2, 3\} \rightarrow \{a, b, c\} \) given by \( f(1) = a \), \( f(2) = b \) and \( f(3) = c \) is one example of a bijection between these two sets.

The map \( g : \{1, 2, 3\} \rightarrow \{a, b, c\} \) given by \( g(1) = b \), \( g(2) = c \) and \( g(3) = a \) is another bijection between these two sets.

Example 3.41. For any set \( A \) the identity map \( \text{Id}_A \) is a simple example of a bijection from \( A \) to itself.

Example 3.42. The map \( f : \{1, 2, 3\} \rightarrow \{1, 2, 3\} \) given by \( f(1) = 2 \), \( f(2) = 1 \) and \( f(3) = 3 \) is bijection from the set \( \{1, 2, 3\} \) to itself.

Definition 3.43. A permutation of a finite set \( A \) is a bijection from \( A \) to itself.

Putting together Lemmas 3.26 and 3.36 we get:

Lemma 3.44. A real function \( f : \mathbb{R} \rightarrow \mathbb{R} \) is bijective if and only if every horizontal line cuts the graph exactly once.

On the other hand, by Theorems 3.27 and 3.37 an arbitrary function \( f : A \rightarrow B \) is bijective if and only if \( f \) has both a right inverse \( g : B \rightarrow A \) and a left inverse \( h : B \rightarrow A \).

Theorem 3.45. Suppose that \( f : A \rightarrow B \) has both a right inverse and a left inverse. Then they are equal to each other and \( f \) has only one right or left inverse.

Proof. First Proof Let \( g : B \rightarrow A \) be a right inverse of \( f \), and let \( h : B \rightarrow A \) be a left inverse.

\( h \) and \( g \) have the same domain and codomain, so it remains to show that \( g(b) = h(b) \) for all \( b \in B \).

Let \( b \in B \). Since \( f \) is bijective, there is a unique \( a \in A \) satisfying \( f(a) = b \). Since \( g \) is a right inverse, we have \( f(g(b)) = b \) also. By uniqueness of \( a \), this implies \( g(b) = a \). On the other hand, since \( h \) is a left inverse, we have \( h(f(a)) = a \). But \( f(a) = b \). So \( h(b) = a = g(b) \). Thus \( g = h \) as required.

If \( g' \) is another right inverse, then we also have \( g' = h \) and hence \( g' = g \). So there is only one right inverse. Similarly there is only one left inverse (which is also a right inverse).

Second Proof Let \( g : B \rightarrow A \) be a right inverse. Then \( f \circ g = \text{Id}_B \).

Let \( h : A \rightarrow B \) be a left inverse. Then \( h \circ f = \text{Id}_A \).

We show that \( g = h \) as follows:

\[
h = h \circ \text{Id}_B = h \circ (f \circ g) = (h \circ f) \circ g = \text{Id}_A \circ g = g.
\]
Definition 3.46. Given a function \( f : A \to B \), a function \( g : B \to A \) is called a (2-sided) inverse if it is both a left and a right inverse.

Note that if \( g \) is a 2-sided inverse of \( f \) then \( f \) is a 2-sided inverse of \( g \) by symmetry.

If a 2-sided inverse exists then there is only one (by Theorem 3.45). We denote it \( f^{-1} \).

We conclude:

Theorem 3.47. A function \( f \) is bijective if and only it has a 2-sided inverse \( f^{-1} \).

If \( f : A \to B \) is a bijection, then the inverse map is the map which sends every element of \( B \) back to where it came from.

Example 3.48. Let \( f : \{1, 2, 3\} \to \{a, b, c\} \) be the bijection \( f(1) = a \), \( f(2) = b \) and \( f(3) = c \).

The inverse map \( f^{-1} : \{a, b, c\} \to \{1, 2, 3\} \) is then the map \( f^{-1}(a) = 1 \), \( f^{-1}(b) = 2 \) and \( f^{-1}(c) = 3 \).

We often prove that a map is bijective by constructing a (2-sided) inverse.

Example 3.49. I claim that the map \( f : [0, \infty) \to [0, \infty), x \mapsto x^4 \) is bijective.

Proof. Let \( g : [0, \infty) \to [0, \infty) \) be the map \( g(x) = \sqrt[4]{x} \).

For all \( x \in [0, \infty) \) we have \( f(g(x)) = f(\sqrt[4]{x}) = (\sqrt[4]{x})^4 = x \). So \( g \) is a right inverse of \( f \).

For all \( x \in [0, \infty) \) we have \( g(f(x)) = g(x^4) = \sqrt[4]{x^4} = x \) (since \( x \geq 0 \)). So \( g \) is a left inverse of \( f \).

Thus \( f \) is bijective and \( g = f^{-1} \).

Example 3.50. Let \( f : \mathbb{R} \to \mathbb{R} \) be the map \( f(x) = 5x \) and let \( g : \mathbb{R} \to \mathbb{R} \) be the map \( g(x) = x/5 \). Then \( g = f^{-1} : f(g(x)) = f(x/5) = 5 \cdot (x/5) = x \) and \( g(f(x)) = g(5x) = (5x)/5 = x \) for all \( x \in \mathbb{R} \).

Lemma 3.51. If \( f : A \to B \) is a bijection and \( g : B \to C \) is a bijection, then \( g \circ f : A \to C \) is a bijection.

Proof. We have inverses \( g^{-1} : C \to B \) and \( f^{-1} : B \to A \). We show that \( f^{-1} \circ g^{-1} \) is an inverse of \( g \circ f \):
\[(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ Id_B \circ f = f^{-1} \circ f = Id_A.\]

Similarly, \((g \circ f) \circ (f^{-1} \circ g^{-1}) = Id_C.\]

\[\square\]

4. SOME more counting

Remark 4.1. Note that a finite set \(A\) has cardinality \(n\) if and only if there is a bijection from the set \(A\) to the set \(\{1, 2, \ldots, n\}\).

This leads to our second important counting principle: Two finite sets have the same cardinality if and only if there is a bijection between them.

We will see many applications of this principle below. Here we use it to count the number of subsets (i.e. the cardinality of the power set) of a finite set \(A\).

Lemma 4.2. Let \(A\) be a finite set. Then \(|\mathcal{P}(A)| = 2^{\left|A\right|}\).

Proof. We will prove this by induction on \(n := |A|\).

If \(n = 1\), \(A\) has a single element and hence \(\mathcal{P}(A) = \{\emptyset, A\}\). So \(|\mathcal{P}(A)| = 2 = 2^1\). Thus the result is true for \(n = 1\).

Suppose that the result is known for sets of cardinality \(n\) and that \(|A| = n + 1\).

Fix some element \(p \in A\). Let \(B = A \setminus \{p\}\). So \(|B| = n\).

Every subset of \(A\) either contains the element \(p\) or it doesn’t. If a subset doesn’t contain \(p\) then it is actually contained in \(B\). Thus the power set of \(A\) is the disjoint union \(\mathcal{P}(B) \sqcup C\) where \(C = \{X \in \mathcal{P}(A) \mid p \in X\}\) (i.e. \(C\) is the collection of subsets of \(A\) which contain the element \(p\)).

It follows that \(|\mathcal{P}(A)| = |\mathcal{P}(B)| + |C|\).

Now we show that there is a bijection from \(C\) to \(\mathcal{P}(B)\):

Let \(f : C \to \mathcal{P}(B)\) be the map sending the set \(X\) (which contains \(p\)) to \(X \setminus \{p\}\).

Let \(g : \mathcal{P}(B) \to C\) be the map sending \(X \subset B\) to \(X \cup \{p\} \in C\).

Then for all \(X \in C\), \(g(f(X)) = g(X \setminus \{p\}) = (X \setminus \{p\}) \cup \{p\} = X\).

Furthermore, for all \(Y \in \mathcal{P}(B)\), \(f(g(Y)) = f(Y \cup \{p\}) = (Y \cup \{p\}) \setminus \{p\} = Y\).

Thus \(g\) is a 2-sided inverse of \(f\) and hence \(f\) is a bijection.
It follows that $|C| = |\mathcal{P}(B)|$, and hence that
$$|\mathcal{P}(A)| = |\mathcal{P}(B)| + |\mathcal{P}(B)| = 2 \cdot |\mathcal{P}(B)| = 2 \cdot 2^n = 2^{n+1}$$
(since $|\mathcal{P}(B)| = 2^n$ by our inductive hypothesis).

\[\square\]

Notation 4.3. If $A$ and $B$ are any two sets, $A^B$ denotes the set whose elements are precisely the functions from $B$ to $A$.

Example 4.4. Let $A = \{1, 2\}$ and $B = \{a, b\}$. Then $A^B$ has four elements; i.e. there are four functions from $B$ to $A$. We list them:

\[
\begin{align*}
  f_1 : B &\rightarrow A, & f_1(a) &= 1, f_1(b) &= 1 \\
  f_2 : B &\rightarrow A, & f_2(a) &= 1, f_2(b) &= 2 \\
  f_3 : B &\rightarrow A, & f_3(a) &= 2, f_3(b) &= 1 \\
  f_4 : B &\rightarrow A, & f_4(a) &= 2, f_4(b) &= 2
\end{align*}
\]

Lemma 4.5. Let $A$ be any set. Then there is a natural bijection $A^n \leftrightarrow A^{\{1,\ldots,n\}}$.

Proof. First we wish to define a natural map, let’s call it $F$, from $A^n \rightarrow A^{\{1,\ldots,n\}}$; i.e. given an ordered $n$-tuple $a = (a_1, \ldots, a_n)$ we want to associate to it a function, let’s call it $f_a$ from $\{1, \ldots, n\}$ to $A$. There’s really only one natural way to do this: Let $f_a(i) = a_i$ for $i = 1, 2, \ldots, n$. Thus we define $f_a : \{1, \ldots, n\} \rightarrow A$.

So $F : A^n \rightarrow A^{\{1,\ldots,n\}}$ is defined by $F(a) = f_a$.

Now we will define a map $G : A^{\{1,\ldots,n\}} \rightarrow A^n$ and show that it is both a left and a right inverse of $F$: Given $f \in A^{\{1,\ldots,n\}}$, let $G(f) := (f(1), \ldots, f(n)) \in A^n$.

Let $a = (a_1, \ldots, a_n) \in A^n$. We must show that $G(F(a)) = a$. Now by definition of $F$ and $G$

\[G(F(a)) = G(f_a) = (f_a(1), \ldots, f_a(n)) = (a_1, \ldots, a_n) = a.\]

Conversely, we must show that $F(G(f)) = f$ for every $f \in A^{\{1,\ldots,n\}}$. Now given $f \in A^{\{1,\ldots,n\}}$, $G(f) = a = (a_1, \ldots, a_n)$ where $a_i = f(i)$. Thus

\[F(G(f)) = F(a) = f_a\]

and for each $i$, by definition $f_a(i) = a_i = f(i)$ and hence $f_a = f$ and we’re done. \[\square\]

Corollary 4.6. Let $A$ and $B$ be any two finite sets. Then

$$|A^B| = |A|^{|B|}.$$
Proof. Suppose that $|B| = n$. Then there is a bijection from $B$ to $\{1, \ldots, n\}$. Thus

$$|A^B| = |A^{\{1, \ldots, n\}}| = |A^n| \text{ by Lemma 4.5} = |A|^n \text{ by Corollary 2.16} = |A|^{|B|}.$$ 

$\square$

Example 4.7. The number of functions from the set $\{1, 2, 3, 4\}$ to the set $\{1, 2, 3, 4, 5\}$ is $5^4 = 625$.

The following Lemma gives a useful dictionary between subsets of a set $A$ and maps from $A$ to $\{0, 1\}$.

Lemma 4.8. Let $A$ be any set. Then there is a natural bijection

$$\mathcal{P}(A) \leftrightarrow \{0, 1\}^A.$$ 

Proof. Given any subset $X$ of $A$ we can associate to it the corresponding indicator function $1_X$: Let $F: \mathcal{P}(A) \rightarrow \{0, 1\}^A$ be the map $F(X) = 1_X$.

Conversely, we define $G: \{0, 1\}^A \rightarrow \mathcal{P}(A)$ by $G(f) = \{x \in A \mid f(x) = 1\}$.

We leave it as an exercise to verify that $G$ is both a right and a left inverse of $F$. $\square$

Remark 4.9. Note that this gives another proof that $|\mathcal{P}(A)| = 2^{|A|}$ when $A$ is finite.

Corollary 4.10. Let $n \geq 1$ and let $S = \{1, \ldots, n\}$. There is a natural bijection from the power set of $S$ to the set $B_n = \{0, 1\}^n$ of binary sequences of length $n$.

Proof. We have a bijection $\mathcal{P}(S) \leftrightarrow \{0, 1\}^S$ by Lemma 4.8 and a bijection $\{0, 1\}^S \leftrightarrow \{0, 1\}^n$ by Lemma 4.5. $\square$

Exercise 4.11. Describe this bijection explicitly.

We have seen how to count the number of functions from $B$ to $A$ when $A$ and $B$ are finite sets. A more difficult task is to count the number of injective functions from $B$ to $A$.

Lemma 4.12. If $A$ and $B$ are finite sets with $|A| = n$ and $|B| = m$ and $1 \leq m \leq n$, then the number of injective maps from $B$ to $A$ is $n \cdot (n - 1) \cdots (n - m + 1)$.
Proof. If $X$ and $Y$ are sets, we will let $\text{Inj}(X,Y)$ denote the set of injective maps from $X$ to $Y$. We wish to prove that

$$|\text{Inj}(B,A)| = n \cdot (n-1) \cdots (n-m+1).$$

We will prove this by induction on $m$. If $m = 1$, $B$ has only one element and there are exactly $n$ maps from $B$ to $A$, all injective. So the result holds for $m = 1$.

Suppose the result is known for $m$, and $B$ has $m+1$ elements. Without loss of generality (renaming elements if necessary), we can take $B = \{1, \ldots, m, m+1\}$ and $A = \{1, \ldots, n\}$ (where $n \geq m+1$ of course). Let $C = \text{Inj}(B,A)$. We wish to prove that $|C| = n \cdot (n-1) \cdots (n-m)$.

For each $i \leq n$, let $C_i = \{f \in C \mid f(m+1) = i\}$. Then

$$C = C_1 \sqcup \cdots \sqcup C_n.$$ 

Now I Claim that there is a bijection $C_i \leftrightarrow \text{Inj}(\{1, \ldots, m\}, A \setminus \{i\})$. Namely: If $f \in C_i$, define $\tilde{f} : \{1, \ldots, n\} \to A \setminus \{i\}$ by $\tilde{f}(j) = f(j)$ for $j \leq n$. (Note that this makes sense because $f$ is injective: since $f(m+1) = i$ it follows that $f(j) \neq i$ when $j \leq m$.) Conversely, given $h \in \text{Inj}(\{1, \ldots, m\}, A \setminus \{i\})$, let $\tilde{h} \in C$ be the map defined by

$$\tilde{h}(j) = \begin{cases} h(j), & j \leq m \\ i, & j = m+1. \end{cases}$$

These correspondences are inverse to each other (check this) and hence define a bijection as claimed.

It follows from the Claim that for each $i$

$$|C_i| = |\text{Inj}(\{1, \ldots, m\}, A \setminus \{i\})| = (n-1) \cdots (n-m)$$

by our inductive hypothesis.

Thus

$$|C| = |C_1| + \cdots + |C_n| = n \cdot (n-1) \cdots (n-m)$$

proving the result. \qed

Remark 4.13. It is easy to see that under the bijective correspondence $A^{\{1, \ldots, m\}} \leftrightarrow A^m$, injective maps correspond to those $m$-tuples $(a_1, \ldots, a_m)$ which have the property that $a_i \neq a_j$ if $i \neq j$.

Thus the lemma we have just proved can be restated as follows:

For any finite set $A$ with $n \geq m$ elements, the number of ordered $m$-tuples of distinct elements of $A$ is $n \cdot (n-1) \cdots (n-m+1)$.

Example 4.14. The number of injective maps from $\{1, 2, 3, 4\}$ to $\{a, b, c, d, e, f, g, h\}$ is $8 \cdot 7 \cdot 6 \cdot 5 = 1680$.

Equivalently, the number of ordered 4-tuples with distinct entries taken from the set $\{a, b, c, d, e, f, g, h\}$ is 1680.
Let $A$ be a finite set with $n$ elements. An injective map $f : A \to A$ is necessarily surjective, since the image must contain $n$ distinct elements. Thus an injective map from $A$ to $A$ is automatically a bijection; i.e. a permutation.

Taking $B = A$ (and $m = n$) in Lemma 4.12, we therefore deduce:

**Corollary 4.15.** Let $A$ be a finite set of cardinality $n$. The number of permutations of $A$ is $n \cdot (n-1) \cdots 2 \cdot 1 = n!$.

## 5. Binomial Coefficients

Recall that for $n \geq 1$, $n! = 1 \cdot 2 \cdots (n-1) \cdot n$. We also define $0! = 1$.

**Definition 5.1.** Let $0 \leq m \leq n$. Then we define the binomial coefficient, $\binom{n}{m}$ by

$$\binom{n}{m} := \frac{n!}{m! (n-m)!}.$$

**Remark 5.2.** On the face it the binomial coefficients are rational numbers. We will see below that they are always integers.

**Remark 5.3.** Note, from the definition, that for all $m, n$ we have

$$\binom{n}{m} = \binom{n}{n-m}.$$

**Example 5.4.**

$$\binom{4}{2} = \frac{4!}{2! \cdot 2!} = \frac{4 \cdot 3}{1 \cdot 2} = 6$$

**Example 5.5.**

$$\binom{6}{4} = \frac{6!}{2! \cdot 4!} = \frac{6 \cdot 5}{1 \cdot 2} = 15$$

**Example 5.6.** For all $n \geq 1$

$$\binom{n}{1} = \binom{n}{n-1} = \frac{n!}{1! (n-1)!} = \frac{n!}{(n-1)!} = n.$$

**Example 5.7.** For all $n$, 

$$\binom{n}{0} = \binom{n}{n} = \frac{n!}{0! \cdot n!} = 1.$$

**Theorem 5.8.** For all $1 \leq m \leq n$ we have

$$\binom{n}{m-1} + \binom{n}{m} = \binom{n+1}{m}.$$
Proof.
\[
\binom{n}{m-1} + \binom{n}{m} = \frac{n!}{(m-1)! (n-m+1)!} + \frac{n!}{m! (n-m)!}
\]
\[
= \frac{n!}{(m-1)!} \left[ \frac{1}{(n-m+1)!} + \frac{1}{m \cdot (n-m)!} \right]
\]
\[
= \frac{n!}{(m-1)!} \cdot \frac{m + (n-m+1)}{m \cdot (n-m+1)!}
\]
\[
= \frac{n!}{m! (n-m+1)!} \cdot \frac{(n+1)}{(n+1-m)!}
\]
\[
= \frac{(n+1)!}{m! (n+1-m)!} = \binom{n+1}{m}.
\]

The statement of this theorem can be visualized in Pascal’s Triangle: We organize the binomial coefficients in a triangular array in which the \(n\)-th row consists of the \(n+1\) binomial coefficients \(\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n}\), and in such a way that the coefficient \(\binom{n+1}{m-1}\) in row \(n+1\) lies half-way between the coefficients \(\binom{n}{m-1}\) and \(\binom{n}{m}\) in the previous row:

\[
\begin{array}{cccccc}
1
& & & & & \\
& 1 & & & & \\
& & 1 & 2 & & \\
& & & 1 & 3 & 3 & 1 \\
& & & 1 & 4 & 6 & 4 & 1 \\
& 1 & 5 & 10 & 10 & 5 & 1 \\
\end{array}
\]

The theorem then says that in this array any term is calculated by adding together the two terms immediately above it (as indicated by the arrows).

**Corollary 5.9.** For all \(0 \leq m \leq n\), \(\binom{n}{m} \in \mathbb{Z}\).

Proof. If \(m = 0\), then \(\binom{n}{m} = 1 \in \mathbb{Z}\). So it remains to prove the result when \(1 \leq m \leq n\).

We’ll prove this by induction on \(n \geq 1\).
When \( n = 1 \), necessarily \( m = 1 \) and \( \binom{n}{m} = \binom{1}{1} = 1 \in \mathbb{Z} \).

Suppose now that the result is known for \( n \) – i.e. that \( \binom{n}{m} \in \mathbb{Z} \) for all \( m \) – and that \( 1 \leq m \leq n + 1 \). If \( m \leq n \), then by Theorem 5.8

\[
\binom{n+1}{m} = \binom{n}{m-1} + \binom{n}{m}
\]

which is an integer by our inductive hypothesis. Otherwise, \( m = n + 1 \) and \( \binom{n+1}{m} = \binom{n+1}{n+1} = 1 \in \mathbb{Z} \).

\[\square\]

The binomial coefficients count the number of \( m \)-element subsets of a set with \( n \) elements:

**Theorem 5.10.** Let \( S \) be a finite set with \( |S| = n \). For \( 0 \leq m \leq n \) let

\[
\mathcal{P}_m(S) := \{ X \in \mathcal{P}(S) \mid |X| = m \}
\]

be the set of all subsets of \( S \) of cardinality \( m \). Then \( |\mathcal{P}_m(S)| = \binom{n}{m} \).

**Proof.** We’ll proceed by induction on \( n \geq 0 \).

If \( n = 0 \) (and \( m = 0 \)) or if \( n = 1 \) (and \( m = 0 \) or \( 1 \)) the result is immediate.

Suppose now that the result is known for sets of size \( n \geq 1 \) and that \( |S| = n + 1 \).

Let \( p \in S \) be some fixed element of \( S \). Let \( T = S \setminus \{p\} \). So \( |T| = n \).

Let \( R = \{ X \in \mathcal{P}_m(S) \mid p \in X \} \). Then \( \mathcal{P}_m(S) = \mathcal{P}_m(T) \sqcup R \) and hence

\[
|\mathcal{P}_m(S)| = |\mathcal{P}_m(T)| + |R|.
\]

However there is a natural bijection

\[
F : R \to \mathcal{P}_{m-1}(T), \ F(X) = X \setminus \{p\}
\]

with inverse

\[
G : \mathcal{P}_{m-1}(T) \to R, \ G(Y) = Y \cup \{p\}.
\]

Thus \( |R| = |\mathcal{P}_{m-1}(T)| \) and hence

\[
|\mathcal{P}_m(S)| = |\mathcal{P}_m(T)| + |\mathcal{P}_{m-1}(T)|
\]

\[
= \binom{n}{m} + \binom{n}{m-1} \quad \text{by our ind. hyp.}
\]

\[
= \binom{n+1}{m} \quad \text{by Theorem 5.8}.
\]

\[\square\]
Example 5.11. In a lottery, six numbers are chosen from the numbers 1, 2, \ldots, 42. In how many ways can this be done?

Solution: The number of 6-element subsets of \( S = \{1, \ldots, 42\} \) is

\[
\binom{42}{6} = \frac{42 \cdot 41 \cdot 40 \cdot 39 \cdot 38 \cdot 37}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} = 7 \cdot 41 \cdot 13 \cdot 38 \cdot 37 = 5245786.
\]

(Thus the chances of winning the lottery with a single choice of 6 numbers is 1/5245786.)

Remark 5.12. Let \( B_n = \{0, 1\}^n \) be the set of binary strings of length \( n \). So \( |B_n| = 2^n \).

Let \( S \) be a set of cardinality \( n \). For simplicity, we take \( S = \{1, \ldots, n\} \). Then we have a bijection \( \mathcal{P}(S) \leftrightarrow B_n \) by Corollary 4.10.

Under this correspondence, it is easily seen that \( \mathcal{P}_m(S) \) corresponds to the set, call it \( B_{n,m} \), of all binary strings of length \( n \) with exactly \( m \) 1s (and \( n - m \) 0s).

Thus the theorem also tells us that the number of binary strings of length \( n \) with exactly \( m \) 1s is \( \binom{n}{m} \): i.e.

\[
|B_{n,m}| = \binom{n}{m}.
\]

This interpretation of the binomial coefficients has many uses, as we will see.

Corollary 5.13. For any \( n \geq 1 \)

\[
\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{m} + \cdots + \binom{n}{n} = 2^n.
\]

Proof. Let \( S \) be a set with \( n \) elements. Then

\[
\mathcal{P}(S) = \mathcal{P}_0(S) \sqcup \mathcal{P}_1(S) \sqcup \cdots \sqcup \mathcal{P}_m(S) \sqcup \cdots \sqcup \mathcal{P}_n(S)
\]

and hence

\[
2^n = |\mathcal{P}(S)| = |\mathcal{P}_0(S)| + |\mathcal{P}_1(S)| + \cdots + |\mathcal{P}_m(S)| + \cdots + |\mathcal{P}_n(S)|
\]

\[
= \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{m} + \cdots + \binom{n}{n}.
\]

\( \square \)
5.1. **The Binomial Theorem.** Let $x$ and $y$ be numbers or variables. Let’s calculate $(x + y)^3$:

\[
(x + y)^3 = (x + y) \cdot (x + y) \cdot (x + y) \\
= (x^2 + 2xy + y^2) \cdot (x + y) \\
= (x^2 + 2xy + y^2) \cdot x + (x^2 + 2xy + y^2) \cdot y \\
= (x^3 + 2x^2y + xy^2 + x^2y + 2xy^2 + y^3) \\
= x^3 + 3x^2y + 3xy^2 + y^3.
\]

Similarly

\[
(x + y)^4 = (x + y)^3 \cdot (x + y) \\
= (x^3 + 3x^2y + 3xy^2 + y^3) \cdot x + (x^3 + 3x^2y + 3xy^2 + y^3) \cdot y \\
= x^4 + 3x^3y + 3x^2y^2 + xy^3 + x^3y + 3x^2y^2 + 3xy^3 + y^4 \\
= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4.
\]

In both cases, the coefficients which arise are exactly the binomial coefficients. This is what always happens:

**Theorem 5.14 (The Binomial Theorem).** If $x, y$ are numbers, functions or variables, and if $n \geq 1$, we have

\[
(x + y)^n = \binom{n}{0} x^n + \binom{n}{1} x^{n-1}y + \cdots + \binom{n}{m} x^{n-m}y^m + \cdots + \binom{n}{n} y^n.
\]

**Proof.** We’ll prove this by induction on $n \geq 1$.

When $n = 1$ the statement is

\[
x + y = (x + y)^1 = \binom{1}{0} x + \binom{1}{1} y
\]

which is true.
Suppose the result has been proved for \( n \). Then

\[
(x + y)^{n+1} = (x + y)^n \cdot (x + y)
\]

\[
= \left[ \binom{n}{0} x^n + \cdots + \binom{n}{m} x^{n-m} y^m + \cdots + \binom{n}{n} y^n \right] \cdot (x + y)
\]

by our ind. hyp

\[
= \left[ \binom{n}{0} x^n + \cdots + \binom{n}{m} x^{n-m} y^m + \cdots + \binom{n}{n} y^n \right] \cdot x
\]

\[
+ \left[ \binom{n}{0} x^n + \cdots + \binom{n}{m} x^{n-m} y^m + \cdots + \binom{n}{n} y^n \right] \cdot y
\]

\[
= \binom{n}{0} x^{n+1} + \binom{n}{1} x^n y + \cdots + \binom{n}{m} x^{n+1-m} y^m + \cdots + \binom{n}{n} x y^n
\]

\[
+ \binom{n}{0} x^n y + \cdots + \binom{n}{m-1} x^{n-m+1} y^m + \binom{n}{m} x^{n-m} y^{m+1} + \cdots + \binom{n}{n} y^{n+1}
\]

\[
= \binom{n}{0} x^{n+1} + \left( \binom{n}{1} + \binom{n}{0} \right) x^n y + \cdots + \left( \binom{n}{m} + \binom{n}{m-1} \right) x^{n+1-m} y^m
\]

\[
+ \cdots + \binom{n}{n} y^{n+1}
\]

\[
= \binom{n+1}{0} x^{n+1} + \cdots + \binom{n+1}{m} x^{n+1-m} y^m + \cdots + \binom{n+1}{n+1} y^{n+1}
\]

as required. \( \Box \)

**Notation 5.15** (\( \Sigma \)-notation). This is a useful notation for expressing long sums – such as those occurring in the binomial theorem – in a more succinct way.

Suppose that \( F_1, \ldots, F_k \) is a sequence of numbers or functions or expressions (anything that can sensibly be added), then \( \sum_{n=1}^{k} F_n \) denotes the sum \( F_1 + F_2 + \cdots + F_k \).

Thus

\[
\sum_{n=1}^{5} n = 1 + 2 + 3 + 4 + 5 = 15.
\]

\[
\sum_{n=1}^{3} (n^2 + n) = (1^2 + 1) + (2^2 + 2) + (3^2 + 3) = 2 + 6 + 12 = 20.
\]

\[
\sum_{n=1}^{4} 2nx^n = 2x + 4x^2 + 6x^3 + 8x^4.
\]
More generally, we can start our sum at any number:

\[
\sum_{n=0}^{3} 2^n = 2^0 + 2^1 + 2^2 + 2^3 = 1 + 2 + 4 + 8 = 15.
\]

\[
\sum_{n=3}^{5} n^2 x^n = 3^2 x^3 + 4^2 x^4 + 5^2 x^5 = 9x^3 + 16x^4 + 25x^5.
\]

With this notation, the binomial theorem says: For any \( n \geq 1 \),

\[
(x + y)^n = \sum_{m=0}^{n} \binom{n}{m} x^{n-m} y^m.
\]

**Example 5.16.** Find the coefficient of \( y^4 \) in \((1 + y)^6\).

**Solution:** Let \( x = 1 \) and \( n = 6 \) in the Binomial Theorem. So

\[
(1 + y)^6 = \sum_{m=0}^{6} \binom{6}{m} y^m.
\]

When \( m = 4 \) the coefficient is \( \binom{6}{4} = 15 \).

**Example 5.17.** Find the coefficient of \( x^4 \) in the expansion of \((2 + 3x)^9\).

**Solution:**

\[
(2 + 3x)^9 = \sum_{m=0}^{9} \binom{9}{m} 2^{9-m} (3x)^m
\]

\[
= \sum_{m=0}^{9} \binom{9}{m} 2^{9-m} \cdot 3^m x^m.
\]

Taking \( m = 4 \), the coefficient of \( x^4 \) is

\[
\binom{9}{4} \cdot 2^5 \cdot 3^4 = 326592.
\]

**Example 5.18.** What is the coefficient of \( x^{23} \) in the expansion of \( \left( 4x^3 + \frac{3}{x^2} \right)^{11} \)?

**Solution:**

\[
\left( 4x^3 + \frac{3}{x^2} \right)^{11} = \sum_{m=0}^{11} \binom{11}{m} (4x^3)^{11-m} \cdot \frac{3^m}{x^{2m}} = \sum_{m=0}^{11} \binom{11}{m} (4)^{11-m} \cdot 3^m \cdot x^{33-5m}.
\]

Now \( 33 - 5m = 23 \) when \( m = 2 \). So the coefficient of \( x^{23} \) is

\[
\binom{11}{2} 4^9 \cdot 3^2 = 3^2 \cdot 4^9 \cdot 5 \cdot 11.
\]
We can use the binomial theorem to give an alternative proof of Corollary 5.13 above:

**Lemma 5.19.** For any $n \geq 1$

\[
\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{m} + \cdots + \binom{n}{n} = 2^n.
\]

*Proof.* Letting $x = y = 1$ in the proof of the binomial theorem gives

\[2^n = (1 + 1)^2 = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{m} + \cdots + \binom{n}{n}.\]

\[\square\]

Here is another consequence:

**Lemma 5.20.** For any $n \geq 1$

\[
\binom{n}{0} - \binom{n}{1} + \cdots + (-1)^m \binom{n}{m} + \cdots + (-1)^n \binom{n}{n} = 0.
\]

*Proof.* Let $x = 1, y = -1$ in the binomial theorem. \[\square\]

6. **Appendix: The mathematical definition of Function**

In Section 3 above, we gave an informal account of what a function is, good enough for our purposes. But we did not say precisely what a function is. The word ‘associates’ has no precise mathematical meaning and we didn’t attempt to explain it. Our informal definition does not address questions such as:

(1) Is a function itself a kind of set or some new kind of mathematical object, and if so, what kind of object?

(2) How can we determine whether something is or is not a function?

In modern mathematics, we define the concept of function as follows:

**Definition 6.1.** Let $A$ and $B$ be sets (possibly equal). A function $f$ from $A$ to $B$ is a subset of $A \times B$ with the following property:

For every $a \in A$, there is one and only one $b \in B$ such that $(a, b) \in f$.

The set $A$ is called the *domain* of the function $f$, and $B$ is called the *codomain* or *target*. We denote this by $f : A \to B$.

For any $a \in A$ the unique $b \in B$ with the property that $(a, b) \in f$ is called the *value* of $f$ at $a$ and is denoted $f(a)$.
Example 6.2. For example, consider the subset
\[ f := \{(1, a), (2, a), (3, b), (4, c)\} \]
of \{1, 2, 3, 4\} \times \{a, b, c, d\}. Then \( f \) is a function from \{1, 2, 3, 4\} to \{a, b, c, d\} since each of 1, 2, 3 and 4 occur exactly once as the first entry (or coordinate) of an element of \( f \).

We can equally describe the function \( f \) as follows:
\[ f : \{1, 2, 3, 4\} \rightarrow \{a, b, c, d\}, \ f(1) = a, \ f(2) = a, \ f(3) = b, \ f(4) = c. \]

Example 6.3. The subset
\[ g := \{(1, a), (1, b), (2, c), (3, d), (4, d)\} \]
of \{1, 2, 3, 4\} \times \{a, b, c, d\} is not a function from \{1, 2, 3, 4\} to \{a, b, c, d\} since both (1, a) and (1, b) belong to it.