2 Sets

Sets show up in virtually every topic in mathematics, and so understanding their basics is a necessity for understanding advanced mathematics. As far as we're concerned, the word 'set' means what you probably think it means: a collection of things.

At this point we will turn our attention to sets, what they are, how we communicate them, and what we can do with them. As we will see over the next few months, sets show up everywhere in mathematics. The good news is that as far as we're concerned, they are fairly easy to describe. The term **set** in mathematics pretty much means what you probably think it means: a collection of things. We will make this a bit more precise later, but this is a good picture to keep in mind for now. Usually we care about sets containing numbers, points, shapes, and functions, but we can also consider sets of less "mathematical" objects such as colors, places, or people.

2.1 Notation

Before we discuss sets, we make a brief remark about notation. One challenge of learning advanced mathematics is learning its notation – special symbols, terms, and even just conventions of the subject. Of course *all* subjects have their own special terms and conventions, but in many subjects, especially in the humanities, these terms are mostly English words, for which we all have at least a basic understanding of. In mathematics, though, there is much notation that is clearly *not* English. To the uninitiated eyes, the scribbles, squiggles, and symbols of mathematical discourse likely appear as a foreign language.

One goal of this semester is to make you more comfortable with mathematical notation. To this end, over the course of the semester, I will make an effort to explain any notation before I use it. Sometimes there is historical significance to a particular choice of notation, sometimes there is not. Sometimes, a particular choice of notation is universal – and virtually all scientists use the same symbols to refer to the same objects – and sometimes it is not. In all cases, though, as I've tried to emphasize before, the notation itself is **never** important. If we used the symbol 'ξ' instead of the symbol '5' to indicate the smallest whole number greater than 4, or if we used a double box around a number $\boxed{\boxed{16}} = 4$ to indicate its square root, instead of $\sqrt{16} = 4$, no idea in mathematics would change, even if the way in which we wrote it down would change. Despite its lack of substantive import, at the end of the day mathematical notation is necessary for communication. If you want to understand what people are talking about, and communicate with them, you need to first learn to read and write their language. Learning about mathematical notation will help us do that. And the sooner we get comfortable with it, the sooner we'll be able to ignore the unimportant, uninteresting symbols you see and instead focus on understanding the deep ideas that mathematics explores.

2.2 Set Basics

We begin with the basics of sets and their notation.

Definition 1. *A* **set** *is an unordered collection of unique things; each 'thing' is called an* **element** *or* **member** *of the set.*

We begin with a few examples.

$$\{1, 2, 3, 4, 5\}. \tag{15}$$

This is the set of whole numbers between 1 and 5. We always use curly braces to indicate that we are referring to a set; although this notation is arbitrary, it is universal, and use of round parentheses or square brackets will only cause confusion.

We can give the above set a name, which will be helpful when we want to reference it later:

$$S = \{1, 2, 3, 4, 5\}. \tag{16}$$

Now, instead of repeating all of these numbers each time we want to refer to this set, we can just reference S. Traditionally we use upper-case letters as names for sets, but this rule is not set in stone.

The set S has five elements: the numbers 1, 2, 3, 4, and 5. The number 7, 100, and 3.14 are not elements, or members, of this set. We often use the terms member and element interchangeably.

If some 'thing' is an element of a set, there are several ways we can communicate that. We can write the English sentence "5 is an element of the set S", but this feels tedious, and we might prefer a shorter way to communicate the same point. For at least a hundred years it has become accepted to use the symbol $\in$ to denote "is an element of the set", so that the idea from the above sentence can be communicated by writing $5 \in S$. Sometimes this is read as '5 is an element of S', or '5 is a member of S', or '5 belongs to S'; all of these mean the same thing. If some 'thing' does not belong to S, we usually use the same symbol with a line through it, as in $\notin$. So we can write $7 \notin S$ to indicate that the number 7 does not belong to S.

Let us consider another set

$$H = \{1, 2, 3, \ldots, 100\}. \tag{17}$$

This set H has one hundred elements, but we only write a few of them; the dots here are called an *ellipses* and stand for what you would expect them to, i.e., the unlisted integers between 3 and 100.

Sets do not have to contain numbers. We might consider the set of colors:

$$C = \{\text{red}, \text{yellow}, \text{green}, \text{blue}\}. \tag{18}$$

Here each of these colors is an element of the set. Or we might consider a list of our country's states:

$$U = \{\text{Alabama}, \text{Alaska}, \ldots, \text{Wyoming}\}. \tag{19}$$

As before, the ellipses indicates that there are more elements that we are not listing. We might consider all whole numbers greater than zero:

$$\mathbb{N} = \{1, 2, 3, \ldots\}. \tag{20}$$

These are known as the **natural numbers**, and the bold-faced $\mathbb{N}$ is typically used to denote this set. Notice that unlike in the previous cases, where the ellipses indicated a finite number of unlisted elements, here the ellipses indicates an infinite number of elements, namely all whole numbers greater than 3.

One important property of a set is that all of its elements are unique. If we see duplicate elements, we ignore them, so that:

$$S = \{1, 2, 2, 3, 4, 5\} \tag{21}$$

should be understood as

$$S = \{1, 2, 3, 4, 5\}. \tag{22}$$

A second important point is that elements of a set are not ordered. You can think of a set as a bag of objects – there is no "order" to objects in a bag, they're just all thrown in there. Of course when we list them we might need to introduce some kind of ordering, but that is something that we add for our convenience. We might arbitrarily order the letters in the alphabet, and then order states alphabetically, but there's no reason why the letters of the alphabet can't be ordered in some other manner.

$$S = \{5, 1, 4, 2, 2, 3, 2, 4, 5\}. \tag{23}$$

and

$$T = \{1, 2, 3, 4, 5\} \tag{24}$$

are identical sets.

Empty Set

There is a very simple set that gets its own name and symbol. It is the empty set – the set with no elements – and it is denoted by the crossed out zero $\emptyset$. It's important to remember that $\emptyset = \{\}$, the set with no elements. That is different from the set that contains just the number zero $\{0\}$, a set containing the empty set $\{\emptyset\}$, or the number 0 itself. The empty set contains zero elements, and is the only such set with that property. It might feel weird to talk about an empty set, but remember that for thousands of years, people also found it difficult to talk about the number 0. You might think of an empty set as a bag with nothing in it, and there is nothing unusual about that.

2.3 Unions and Intersections

The first slightly interesting thing we can do with sets is put them together to make new sets. How might you go about doing that? We will discuss two ways to make new sets given two other sets. Let's say we have these two sets:

$$S = \{1, 2, 3, 4, 5\}, \qquad T = \{0, 1\} \tag{25}$$

How can we use these to form a new set? We can consider the set which contains all elements that belong to either S or T. We call this set the **union** of the two sets S and T, and we use the symbol $\cup$ to denote it:

$$U = S \cup T = \{0, 1, 2, 3, 4, 5\}. \tag{26}$$

Notice that U contains six elements; the union of two sets always has at least as many elements as either of the two original sets.

How else can we use two sets to make a new set? We can consider the set of all elements that belong to both S and T. We call this the **intersection** of the two sets S and T, and we use the symbol $\cap$ to denote it:

$$I = S \cap T = \{1\}. \tag{27}$$

Note that I has only one element in it; the intersection of two sets always has at most as many elements as either of the two original sets.

2.4 Subsets

Perhaps as important as unions and intersections is the concept of subsets. Let's consider a new set, that of all integers, or whole numbers:

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}. \tag{28}$$

You may have noticed that for some sets considered above, for example $\mathbb{N}$, all of their elements are also contained in $\mathbb{Z}$. We want to have a concise way of saying this, so we don't have to write, for example:

All elements of $\mathbb{N}$ are also contained in $\mathbb{Z}$.

How can we say this in a shorter way? We write:

$$\mathbb{N} \subset \mathbb{Z}. \tag{29}$$

The symbol $\subset$ means that every element in $\mathbb{N}$ is also an element of $\mathbb{Z}$. More generally we have the following definition:

Definition 2. *If A and B are sets, and every element of A is also an element of B, then we say that A is a* **subset** *of B. We usually write this as $A \subset B$. Sometimes we say that B* **contains** *A.*

This definition allows us to express many mathematical facts. For example, if we use $\mathbb{P}$ to denote the set of prime numbers (e.g., 2, 3, 5, 7, 11, etc.), then we can write $\mathbb{P} \subset \mathbb{N}$ to express the fact that every prime number is also a natural number. If we let O denote the set of odd integers, and E the set of even integers, then we can write $E \subset \mathbb{Z}$ and $O \subset \mathbb{Z}$ to communicate that both the even numbers and odd numbers are subsets of the integers.

The empty set $\emptyset$ has a notable feature as far as subsets are concerned – it's a subset of every set! For any set A, all (zero) elements of the empty set are also elements of A. This is logically equivalent to saying that for any set A, there are no elements that belong to the empty set but do not belong to A. One consequence of this fact is that every set has at least one set, the empty set $\emptyset$.

Almost all sets are guaranteed to also have a second subset. In particular, for every set A, it is the case that $A \subset A$, since every element of A is, of course, also an element of A. The two subsets $\emptyset$ and A are distinct unless A is the empty set. These two subsets, A and $\emptyset$, are sometimes called *trivial* subsets, for understandable considerations.

NOTE 1: We sometimes want to indicate that a set A is a subset of B, but different from it. If that is the case, then we say that A is a **proper** subset of B. Some texts use two different symbols to indicate the property of being a subset, and the property of being a proper subset. More specifically, if A is a proper subset of B, then they will write $A \subset B$; if A might be equal to B, then they will write $A \subseteq B$. Other books use these two terms interchangeably. When these symbols are used to indicate different ideas, the distinction between $\subset$ and $\subseteq$ is analogous to that between $<$ and $\leq$; your understanding of subsets will benefit greatly from meditating on this analogy.

NOTE 2: Many students first learning about the symbols $\in$ and $\subset$ initially confuse them. Remember that $a \in A$ is used to indicate that a is an element of A. To be an element of a set, a itself need not be a set itself (though it can be). However, to be a subset of B, A must itself first be a set. Assuming a is not a set itself, it would make no sense to write $a \subset B$. Likewise, unless the set A happens to be an element of B, then it would be wrong to write $A \in B$ even if A is a subset of B. Work out several examples for yourself to make sure you understand the difference.

2.5 Set-builder notation

As described above, one way to describe a set is by listing all of its elements. On occasion this is difficult, and sometimes we use an ellipses . . . to indicate that there are additional elements that have not been explicitly listed. However, there are cases where even an ellipses cannot help us adequately describe a set. For example, consider the unit interval $I = [0, 1]$, the set of all numbers between 0 and 1. We cannot list all of its elements, even with the use of an ellipses. This inspires the use of **set-builder notation**, in which we provide rules for building a set, though without listing any of its elements explicitly. [Notes here are incomplete; additional information on this subject can be found on https://en.wikipedia.org/wiki/Set-builder_notation]