1. **Fundamentals**

A *set* is a collection of distinct objects, called *elements* of the set. The elements of a set can be anything whatsoever—thus, \{emu, frog, newt\} and \{?, !, ‘\} are valid sets, the first consisting of the English words for specific animals and the second of punctuation marks—as long as it is understood that no two elements of a set can be exactly the same. It is well-known to mathematicians that this informal definition of sets can lead to some bizarre logical contradictions (consider the set of all sets ?!@*#?), but in this course, we will never need to venture near such strange beasts.¹

Sets are the most fundamental objects in mathematics in the sense that all other mathematical ideas are either built from sets or use them.² Indeed, many entire branches of mathematics, including probability theory, involve studying the properties of a set adorned with some additional structure. For example, we can define two operations + and × on the set of real numbers³ that each take a pair of numbers to a new number via certain rules (addition and multiplication). The set of real numbers adorned with these operations forms what mathematicians call a *field*. In probability theory, we consider a set of outcomes for some random experiment and adorn this set with a measure of how likely the different outcomes are to occur.

Given the idea of a set, we would like to be able to do something with it. At minimum we must be able to define new sets, determine if two sets are equal or if one lies “inside” the other, decide whether or not an object is an element of a particular set, and describe the number of elements a set contains.

¹ Mathematicians have developed a complete axiomatic theory of sets that circumvents these problems.

² What about numbers, you might ask? It turns out that numbers themselves can be defined as sets. For example, the integer zero is defined as the empty set; the integer one is the set whose single member is the empty set; the integer \(k\) is the set whose \(k\) members are the sets representing \(0, 1, \ldots, k - 1\). This same construction can be extended to all integers, then to the rational and real numbers, and beyond.

³ For now, I will assume that the set of real numbers and the set of integers are passingly familiar; the former contains all “numbers” between \(-\infty\) and \(\infty\) and the latter contains those real numbers with no fractional part. See Section 2 below for a more complete discussion.
In mathematical text, we will denote sets by enclosing the collection of elements in braces \( \{ \} \)’s. Throughout this course, we will use braces only to represent sets, so whenever you see something so enclosed you will know what type of object it is. If the set has only a few elements, it is easy to specify it completely by listing the elements explicitly. For example, \( \{2, 4, 6, 8\} \) is the set of positive, even integers less than 10. There is no preferred order in the listing of a set since the set embodies the collection of elements not an ordered list of elements. Hence, \( \{8, 2, 4, 6\}, \{4, 8, 2, 6\}, \{6, 8, 2, 4\} \), and any other permuted listing all represent the same set. Most sets have too many elements to list explicitly, so we describe them by defining the conditions that an object must satisfy to be an element of the set. This description is enclosed in \( \{ \} \)’s as usual. For example, \( \{\text{real numbers } x \text{ such that } x < 2 \text{ and } x > -16\} \) is the set of all real numbers that are strictly smaller than 2 and strictly bigger than -16, and \( \{\text{integers } p \text{ such that } p \text{ is prime}\} \) is the set of prime numbers. The variables \( x \) and \( p \) in these descriptions are dummy variables; they have meaning only within the definition of the set as labels for an arbitrary candidate (in this case, a real number or an integer) that make it easier to specify the defining conditions. Thus, the set \( \{\text{real numbers } z \text{ such that } -16 < z < 2\} \) is exactly the same set as the former because it defines the same set of numbers. It is irrelevant that in the latter we use a different dummy variable to specify the defining conditions or that we state those conditions differently.

Once we start defining sets, it makes our lives easier if we can give them more succinct names by which to refer to them later. (Naming is an activity that both mathematicians and computer scientists enjoy.) We name sets by associating a symbol with the set using an = sign. For example, if I write “let \( \mathcal{A} = \{\text{real numbers } u \text{ such that } (u + 1)^2 < 1\} \)” then I can use \( \mathcal{A} \) or the complete specification of the set interchangeably. If I am going to refer to this set many times, it is much more convenient to use \( \mathcal{A} \) than to repeat the definition. Notice that, on the face of it, the symbol \( \mathcal{A} \) gives no indication that it refers to a set. This is a disadvantage of using names: we have to discern from context (or just remember) what a particular symbol means. The same phenomenon occurs in computer programs: the variables \( x \) and \( \mathcal{A} \) can stand for objects of any type, only context and memory help us deter-
mine what they are. Programmers make this task easier for someone reading their program by using mnemonic names and by sticking to a naming convention that makes it easier to identify types. We will do the same thing; see Appendix N for a complete description of our conventions. In particularly, we will use script Roman capital letters to denote sets.

When defining a set, one must be precise and unambiguous. Pictures of sets, although not always precise, are nonetheless often helpful for understanding what a definition really means. For example, consider two sets $B = \{\text{real numbers } x \text{ such that } x^2 - x < -\frac{3}{16}\}$ and $C = \{\text{real ordered pairs } (x, y) \text{ such that } 0 < x \leq y < 1\}$. With some practice, it will become easier to picture these sets in your mind, but for now, take a look at the corresponding figures. In the first case, $B$ contains all real numbers bigger than $1/4$ and smaller than $3/4$ as can be seen by referring to the superimposed quadratic curve $x^2 - x + \frac{3}{16} = (x - 1/4)(x - 3/4)$. In the second case, the set consists of ordered pairs, which we can interpret as coordinates in the x-y plane. The set $C$ then contains all those points inside the unit square ({$(x, y)$ such that $0 < x < 1$ and $0 < y < 1$}) above the main diagonal (i.e., x-coordinate smaller than y-coordinate). The pictured triangle results. Pictures of sets can also be used more abstractly to help understand the relationship between different sets. In such pictures, called Venn diagrams, the specific definition of the sets is not important; you look at Venn diagrams to gain an instant visual interpretation. Two sets which share no elements in common can be thought of as “non-overlapping”; a set whose elements are all elements of a second set can be thought of as “contained in” that second set. Both of these relationships are immediately clear from the Venn diagram and hold true quite generally.

**Thought Question** Is $\{2, 2, 4, 6, 8\}$ a set according to the definition? Why or why not?

**Thought Question** Is $\{\emptyset, \{0, 1\}, \{0, 1, 2\}\}$ a set according to the definition? Why or why not?
1.1. Comparing Sets

Two sets \( \mathcal{E} \) and \( \mathcal{F} \) are said to be equal,\(^7\) denoted \( \mathcal{E} = \mathcal{F} \), if and only if they have exactly the same elements. If \( \mathcal{E} \) and \( \mathcal{F} \) are not equal, we write \( \mathcal{E} \neq \mathcal{F} \). Figure 6 illustrates three different ways that \( \mathcal{E} \) and \( \mathcal{F} \) can be unequal: non-overlapping, overlapping but containing distinct elements, or one contained in the other. Given two sets \( \mathcal{E} \) and \( \mathcal{F} \), we say that \( \mathcal{E} \) is a subset of \( \mathcal{F} \), denoted \( \mathcal{E} \subseteq \mathcal{F} \), if every element of \( \mathcal{E} \) is also an element of \( \mathcal{F} \).\(^8\) (For symmetry, we also say in this case that \( \mathcal{F} \) is a superset of \( \mathcal{E} \), denoted \( \mathcal{F} \supseteq \mathcal{E} \).) As I have defined it above, the following two statements are true:

- Every set is a subset of itself.
- \( \mathcal{E} = \mathcal{F} \) if and only if both \( \mathcal{E} \subseteq \mathcal{F} \) and \( \mathcal{F} \subseteq \mathcal{E} \) are true.

You should convince yourself that both these facts follow from the definition.\(^9\)

1.2. Membership

To indicate that an object \( x \) is an element of a set \( \mathcal{E} \), we write \( x \in \mathcal{E} \). The operator \( \in \) here is read in one of several equivalent ways “element of”, “member of”, “in”, “belongs to”, and so forth. Sometimes it is useful to indicate the opposite, that \( x \) is not an element of \( \mathcal{E} \); we do this by writing \( x \notin \mathcal{E} \).\(^{10}\) Both operators \( \in \) and \( \notin \) take an object on the left and a set on the right. Thus, \( 2 \in \{2, 4, 6, 8\} \) and \( 0 \notin \{\text{real numbers } z \text{ such that } -16 < z < 2\} \). \( 0 \notin \{2, 4, 6, 8\} \) and \( 2 \notin \{\text{real numbers } z \text{ such that } -16 < z < 2\} \).

A very important technique to master is the argument that logically proves two sets equal using the second bulleted result above. As an example, consider the two sets

\[
\mathcal{E} = \{\text{ordered pairs } (x, y) \text{ such that } x^2 + y^2 \leq 1\}
\]

\[
\mathcal{F} = \{\text{ordered pairs } (r \cos \theta, r \sin \theta) \text{ such that } 0 \leq \theta < 2\pi, 0 \leq r \leq 1\}.
\]

To show that \( \mathcal{E} = \mathcal{F} \), I need to show that both \( \mathcal{E} \subseteq \mathcal{F} \) and \( \mathcal{F} \subseteq \mathcal{E} \) are true. To do this, I need to show that for any element \( p \in \mathcal{E} \) that I choose, it is also true that \( p \in \mathcal{F} \). Choose \( p \in \mathcal{E} \). Because this is an arbitrary point,\(^{11}\) all we know about it is that \( p \) satisfies the conditions defining \( \mathcal{E} \); namely, \( p = (x, y) \) with \( x^2 + y^2 \leq 1 \). Let \( r = \sqrt{x^2 + y^2} \) and let
\[ \theta = \arccos(x/r) \text{ (or 0 if } r = 0). \] (In other words, \( \theta \) is the angle such that \( \cos \theta = x/r \) which we can find because \( x/r \) is between 0 and 1.) Then, 
\[ x = r \cos \theta, \] and because \( x^2 + y^2 = r^2 \) and \((\cos \theta)^2 + (\sin \theta)^2 = 1\), we also have that \( y = r \sin \theta \). Hence, \( p = (x, y) = (r \cos \theta, r \sin \theta) \in \mathcal{F} \). Since the point \( p \) was a completely arbitrary element of \( \mathcal{E} \), this argument works for any point in \( \mathcal{E} \). We have thus established that \( \mathcal{E} \subseteq \mathcal{F} \). Now choose an arbitrary point \( p \in \mathcal{F} \). Again because this is an arbitrary point, all we know about it is that \( p \) satisfies the conditions defining \( \mathcal{F} \); namely, \( p = (r \cos \theta, r \sin \theta) \) with \( 0 \leq r \leq 1 \) and \( 0 \leq \theta < 2\pi \). But as we saw above, the values \( x = r \cos \theta \) and \( y = r \sin \theta \) satisfy \( x^2 + y^2 \leq r^2 \). Thus, \( \mathcal{F} \subseteq \mathcal{E} \), and we have shown that \( \mathcal{E} = \mathcal{F} \). (See the picture of this set at the right.) It is quite common for two completely different descriptions of the same set to arise, and when this happens, we need to apply this argument to show that they are indeed the same. To summarize: given two sets \( \mathcal{E} \) and \( \mathcal{F} \), I have to show that (i) for any element \( x \in \mathcal{E} \), it is also true that \( x \in \mathcal{F} \), and (ii) for any element \( x \in \mathcal{F} \), it is also true that \( x \in \mathcal{E} \).

**Thought Question** Is \{real numbers \( x \) such that \( 0 < x < 1 \}\} \( \subset \) \{real numbers \( x \) such that \( 0 \leq x \leq 1 \}\}? Are the two sets equal?

### 1.3. The Size of a Set

The *cardinality* of a set is the number of items it contains. We denote the cardinality of a set \( \mathcal{S} \) by \( \# \mathcal{S} \). This can be finite or infinite but must be non-negative. We will see examples of both below, including an answer to the question of whether infinity plus 1 is bigger than infinity.

The size of a set can be defined in many ways. A *measure* is a mathematical object that quantifies the size of sets. Each distinct measure embodies a different way to assess how big a set is. Measures are the subject of Appendix M.

### 1.4. Upper and Lower Bounds

A set \( \mathcal{A} \) of real numbers is said to be *bounded above* if there is a \( u < \infty \) such that every \( x \leq u \) for every \( x \in \mathcal{A} \). This means that if you move a certain finite distance to the right along the real line, then every element of \( \mathcal{A} \) is to the left of you. A set \( \mathcal{A} \) of real numbers is said to
be *bounded below* if there is a \( \ell > -\infty \) such that every \( \ell \leq x \) for every \( x \in A \). This means that if you move a certain finite distance to the left along the real line, then every element of \( A \) is to the right of you. A set is *bounded* if it bounded above and bounded below, otherwise it is *unbounded*.

The number \( u \) in the definition of “bounded above” just given is called an upper bound of \( A \). If \( A \) is bounded above, an *upper bound* of \( A \) is any real number that is \( \geq \) every element of \( A \). If \( A \) is not bounded above, then \( \infty \) is the only upper bound for \( A \). The smallest possible upper bound for a set \( A \) of real numbers is called the *least upper bound* (or *supremum*) and is denoted \( \sup A \). The least upper bound of \( A \) satisfies the following:

- \( \sup A \) is an upper bound for \( A \),
- If \( u \) is any upper bound for \( A \), then \( \sup A \leq u \).

If \( A \) has a largest value then it equals \( \sup A \). For instance, any finite set of numbers has a maximum, which satisfies the two defining properties of the least upper bound. However, some sets do not have a largest value, including the set of all real numbers and \( \{ \text{real numbers } x \text{ such that } 0 < x < 1 \} \).

Similarly, the number \( \ell \) in the definition of “bounded below” just given is called a lower bound of \( A \). If \( A \) is bounded below, a *lower bound* of \( A \) is any real number that is \( \leq \) every element of \( A \). If \( A \) is not bounded below, then \( -\infty \) is the only lower bound for \( A \). The biggest possible lower bound for a set \( A \) of real numbers is called the *greatest lower bound* (or *infimum*) and is denoted \( \inf A \). The greatest lower bound of \( A \) satisfies the following:

- \( \inf A \) is a lower bound for \( A \),
- If \( \ell \) is any lower bound for \( A \), then \( \ell \leq \sup A \).

If \( A \) has a smallest value then it equals \( \inf A \). For instance, any finite set of numbers has a minimum, which satisfies the two defining properties of the greatest lower bound. Analogously to the above, some sets do not have a smallest value, including the real numbers and \( \{ \text{real numbers } x \text{ such that } 0 < x < 1 \} \).

**Thought Question** If \( A \) is a set of real numbers and \( \#A < \infty \), can \( A \) be unbounded? Why or why not?
Thought Question What is the least upper bound of the set of real numbers $x$ such that $0 < x < 1$? What is the least upper bound of the set of real numbers $x$ such that $0 \leq x \leq 1$? What is the least upper bound of the set of real numbers? What is $\inf \{0\}$? What is $\inf \{1/n \text{ such that } n \in \mathbb{Z}_+\}$?

Thought Question If $\mathcal{A}$ is a set of real numbers and $\mathcal{B} = \{-x \text{ such that } x \in \mathcal{A}\}$, how do $\inf \mathcal{A}$ and $\sup \mathcal{A}$ relate to $\sup \mathcal{B}$ and $\inf \mathcal{B}$?

1.5. Collections of Sets

We will frequently deal with a collection of many related sets at once, so it is useful to have a convention for naming them. Consider the following collections of sets:

- Sets $\mathcal{A}_k = \{2, \ldots, 2k\}$ for each integer $k \geq 1$,
- Sets $\mathcal{B}_x = \{\text{real } z \text{ such that } 0 \leq z \leq x\}$ for each real $x > 0$,
- Sets $\mathcal{C}_{\Letter} = \{a, \ldots, \Letter\}$ for each letter $\Letter$ in $\{a, \ldots, z\}$.

There are two ingredients that define each of these collections. First, we define an index set; the value of the index determines which set in the collection we are considering. In these cases, the integers, real numbers, and lowercase letters respectively are the index sets. Second, we define the one set in the collection for each element of the index set. In practice, we refer to the sets in the collection using a common letter symbol (e.g., $\mathcal{A}$, $\mathcal{B}$, etc.) with the index as a subscript. Typically, we will use the integers or non-negative integers as the index set, and in that case the collection is also called a sequence of sets. When we want to refer to the collection as a whole we will informally speak of the $\mathcal{A}_k$'s or the $\mathcal{B}_x$'s. If we want to be more formal in discussing an entire collection $\mathcal{A}_\gamma$ with $\gamma \in \mathcal{I}$, we will write $(\mathcal{A}_\gamma)_{\gamma \in \mathcal{I}}$.

2. Some Important Examples

2.1. The Empty Set

The simplest possible set is the set that contains no elements, called the empty set. Because it is so important, the empty set has its own symbol $\emptyset$ which we will use because it is so common. Naturally, $\#\emptyset = 0$. The empty set is a subset of every set. ¹⁴ Why? Consider a set $\mathcal{A}$. Because $\emptyset$ contains no elements, it is true by default that every element of $\emptyset$ is also an element of $\mathcal{A}$. This argument works for any set $\mathcal{A}$.
2.2. Numbers

The next most important sets, and those perhaps most familiar to you, are sets of numbers. The first numbers you probably came across were the positive integers 1, 2, 3, ..., also called the natural numbers or counting numbers. Certain fundamental sets of numbers, like the positive integers, arise so frequently that we give them special symbols. We denote the set of positive integers by \( \mathbb{Z}^+ \), where the \( \mathbb{Z} \) represents for the German word for integer "Zahl" and the '+' indicates positive numbers. Once you learned about zero, your horizons had expanded to the set of non-negative integers, or whole numbers, \( \mathbb{Z} = \{0, 1, 2, 3, \ldots\} \), the \( \oplus \) here gives a mnemonic for "positive and zero = non-negative". Having discovered zero, the next step is to realize that negative numbers make sense. (If you borrow 4 apples from me and eat them, how many apples do you have?) Combine the negative integers \( -1, -2, -3, \ldots \) with the non-negative integers and you get the \textit{integers} \( \mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\} \). The sets \( \mathbb{Z}, \mathbb{Z}^+, \) and \( \mathbb{Z}^\oplus \) are all infinite sets, and none of these three sets contains a largest element, \(^{16}\) though \( \mathbb{Z}^+ \) and \( \mathbb{Z}^\oplus \) do have a smallest element.

The set of \textit{rational numbers} \( \mathbb{Q} \) is defined by

\[
\mathbb{Q} = \left\{ \frac{p}{q} \text{ such that } p \in \mathbb{Z} \text{ and } q \in \mathbb{Z}^+ \right\}.
\]

Remember that the elements of a set are unique, so the ratios \( p/q \) and \( 2p/2q \), which are equal, appear only once in \( \mathbb{Q} \). The set \( \mathbb{Q} \) contains the numbers whose decimal representation has a fixed (or repeating) number of digits. All numbers representable on a computer are rational, for instance. Because there are infinitely many integers, \( \mathbb{Q} \) is also an infinite set. What may be more surprising is the following property of \( \mathbb{Q} \): between any two rational numbers \( q_1 < q_2 \), there are infinitely many more rational numbers.\(^{17}\) Try to devise an argument to prove this.

However, most numbers, including such favorites as \( e \), \( \pi \), and \( \sqrt{2} \), cannot be represented as rational numbers; they require an infinite, non-repeating sequence of digits in their decimal representation. These are called \textit{irrational numbers}.\(^{18}\) The set of \textit{real numbers} \( \mathbb{R} \) consists of all the rational and all the irrational numbers. We often refer to \( \mathbb{R} \) colloquially as "the real line". As we will see below, the cardinality of \( \mathbb{R} \) is much greater than the cardinality of \( \mathbb{Z} \); that is, there are many
more real numbers than integers. (Oddly enough, the sets $\mathbb{Z}, \mathbb{Z}_+, \mathbb{Z}_\mathbb{N}$, and $\mathbb{Q}$ all have the same cardinality; we will see why later on.) It is also useful to define $\mathbb{R}_+$ to be the set of positive real numbers and $\mathbb{R}_\geq$ to be the set of non-negative real numbers. By definition, we have that $\mathbb{R}_+ \subset \mathbb{R}_\geq \subset \mathbb{R}.$

The set of complex numbers $\mathbb{C}$ consists of all combinations $a + ib$ where $a, b \in \mathbb{R}$ and $i = \sqrt{-1}$ (electrical engineers use $j$). By taking $b = 0$, we just get the real numbers back, so $\mathbb{R} \subset \mathbb{C}.$ Complex numbers arise as the solutions to certain polynomial equations. For example, the equation $x^2 + 1 = 0$ has two complex solutions, $\pm i$, but no real solutions.

These sets that we have introduced are contained one within the other: $\mathbb{Z}_+ \subset \mathbb{Z}_\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$ Most of the time, we will deal with real numbers, and if a number’s type is unspecified, you can assume that it is real by default. Hence, if we write \{x \text{ such that } x > 30\} we mean that $x \in \mathbb{R}.$ (Of course, it does not hurt to be precise, so feel free to write \{x \in \mathbb{R} \text{ such that } x > 30\}. And remember to indicate when you mean something specific, like $x \in \mathbb{Z}.$)

**Thought Question** Is $\sqrt{25}$ a rational number?

**Thought Question** Can you find a real number that is not a complex number? Can you find a complex number that is not a real number?

2.3. Intervals

**An interval** of real numbers is a subset of $\mathbb{R}$ containing all numbers in a certain range. There are four types of intervals distinguished only by which boundary points are part of the set. Closed intervals $[a, b]$ are defined for $a, b \in \mathbb{R}$ with $a < b$ by

$$[a, b] = \{x \in \mathbb{R} \text{ such that } a \leq x \leq b\}.$$

The orientation of the brackets is intended to help you remember that closed intervals contain both their end-points. A special case that arises often is the **unit interval** $[0, 1].$ Open intervals $(a, b]$ contain neither of their endpoints; they are defined for $a, b \in \mathbb{R}$ with $a < b$ by

$$(a, b] = \{x \in \mathbb{R} \text{ such that } a < x \leq b\}.$$
Again, note how the bracket orientation encodes the definition. (We can also define an open interval with \( a = -\infty \) and \( b = \infty \); this \([ -\infty, \infty[\) is just the set of real numbers.) The two types of half-open intervals are \([a, b[\) and \( ]a, b]\). The former are defined for \( a < b \) with \( a \in \mathbb{R} \) and \( b \) in \( \mathbb{R} \) or equal to \( \infty \); the latter are defined for \( a < b \) with \( a \in \mathbb{R} \) or equal to \( \infty \) and \( b \) in \( \mathbb{R} \). These are given by

\[
[a, b[ = \{ x \in \mathbb{R} \text{ such that } a \leq x < b \}
\]

\[
]a, b] = \{ x \in \mathbb{R} \text{ such that } a < x \leq b \}.
\]

Thus, \([0, 1[, [0, 2[, \text{ and } [0, \infty[\) are all valid intervals; the last of which contains all non-negative numbers in \( \mathbb{R} \). Occasionally, we wish to allow for the possibility that a number is \( \infty \) or \(-\infty \), so we will use infinities in interval specifications as a convenient shorthand. For example, \([0, \infty]\) is the set of non-negative real numbers union the set \{\( \infty \}\).

**Thought Question** If \( a \) is a real number, what is the set \([a, a]\)?

**Thought Question** Under what situations is an interval of the form \([a, b[\) bounded?

**Thought Question** What are the least upper bound and greatest lower bound of the interval \([a, b[\)?

### 2.4. Vectors

A **vector**\(^{19}\) is an ordered sequence of objects. We denote a vector as a comma separated list surrounded by *parentheses*. For example, \((2, 4), (1, 2, 5), \text{ and } (3, 4.2, 19, -20, 4.24, 10^{-10})\) are vectors of real numbers with length 2, 3, and 6. If a vector has length \( k \), we call it a \( k \)-vector. In the example above, \((1, 2, 5)\) is a 3-vector and \((2, 4)\) is a 2-vector. Another name for a 2-vector is **ordered pair**; an ordered pair of real numbers is just the familiar object from your mathematics classes used to denote two-dimensional coordinates. A 1-vector is just the object itself (the parentheses are superfluous in this case). It is also possible to have vectors of infinite length such as \((1, e, e^2, e^3, \ldots)\) or \((x_1, x_2, x_3, \ldots)\); an infinite vector is also called a **sequence**. Unlike sets, the order in which a vector’s elements are listed matters. While the sets \{1, 2, 5\}, \{5, 2, 1\}, and \{2, 1, 5\} are all the same, the vectors \((1, 2, 5), (5, 2, 1), \text{ and } (2, 1, 5)\)
are all different. We often need distinguish between vectors and single numbers in isolation. The latter are called **scalars**.

The elements in the list given by a vector are called its **components**. If \( x \) is a vector then by convention \( x_i \) is the \( i \)th component. Conversely, given numbers (i.e., scalars) \( x_1, x_2, \ldots, x_n \), we can construct the vector \( x \) by putting these components together: \( x = (x_1, x_2, \ldots, x_n) \). A key fact about vectors is that we can always move interchangeably between using the vector as an object itself (e.g., as the symbol \( x \)) and working with the components (e.g., the list \( x_1, \ldots, x_n \)).

It is often useful to consider sets of vectors. Consider three important examples. First, \( \mathbb{R}^1 \) is the set of 1-vectors of real numbers, but since a 1-vector \( (x) \) is just \( x, \mathbb{R}^1 = \mathbb{R} \), the real line. Second, \( \mathbb{R}^2 \) is the set of 2-vectors — ordered pairs — of real numbers; each element is of the form \( (x, y) \) for some real numbers \( x \) and \( y \). This yields the “x-y plane” of two-dimensional coordinates you have used in your mathematics classes. Each vector \( (x, y) \) corresponds to a point in the plane. Third, \( \mathbb{R}^3 \) is the set of 3-vectors — ordered triples — of real numbers; each element is of the form \( (x, y, z) \) for some real numbers \( x, y, z \). Each such vector gives the coordinates of a point in three-dimensional space; \( \mathbb{R}^3 \) is the set of all such points. Of course, there is no need to stop at three. For any \( k \in \mathbb{Z}_+ \), we define \( \mathbb{R}^k \) to be the set of \( k \)-vectors of real numbers. Each element of \( \mathbb{R}^k \) is of the form \( (x_1, \ldots, x_k) \), where the components are real numbers. We also write \( \mathbb{R}^\infty \) to denote the set of **sequences** of real numbers. For example, \( (1, 1/2, 1/3, 1/4, \ldots) \) is an \( \infty \)-vector whose \( n \)th component is the number \( 1/n \).

Vectors have the property that they can be added together componentwise and they can be scaled by a number (thus the name scalar). If we have two vectors of the same length \( x = (x_1, \ldots, x_n) \) and \( y = (y_1, \ldots, y_n) \), then the vector \( x + y \) has components \( x_i + y_i \). That is,

\[
x + y = (x_1 + y_1, \ldots, x_n + y_n).
\]

Similarly, the vector \( cx \) has components \( cx_i \). That is,

\[
cx = (cx_1, \ldots, cx_n).
\]

for example, \( (2, 1, 3) + (1, -1, 3) = (3, 0, 3) \) and \( 3(1, 2, 3) = (3, 6, 9) \). The two operations can also be combined. For instance, if \( x \) and \( y \) are
two vectors and $0 \leq \lambda \leq 1$ is a scalar, then $\lambda x + (1 - \lambda)y$ is the point $\lambda$ of the way along the line between $x$ and $y$. Similarly, $x - y$ is the vector $x + (-1)y$.

Vectors in $\mathbb{R}^k$ (for $k \in \mathbb{Z}_+$) are often represented by arrows from the origin of the coordinate system to the point represented by the vector. This can be an intuitively useful heuristic; it represents a vector as a direction and a magnitude. The vector $x + y$ is the arrow from the origin to the point obtained by following the $x$ arrow to $x$ and then following a copy of the $y$ arrow shifted to $x$. Similarly, $x - y$ is the arrow from $y$ to $x$. (You can remember this because $y + x - y = x$.)

We can also sometimes compare vectors in $\mathbb{R}^k$. If $a, b \in \mathbb{R}^k$, we say that $a \leq b$ if $a_i \leq b_i$ for all $i = 1, \ldots, k$. Analogous definitions work for $<, >$, and $\geq$. Notice that two vectors need not be comparable; that is, for $k > 1$, both $a \leq b$ and $a > b$ can be false.

If $x \in \mathbb{R}^n$, we can denote the Euclidean length of the vector by $|x|$, where

$$|x| = \sqrt{\sum_{i=1}^{n} x_i^2}. \quad \text{(S.1)}$$

In $\mathbb{R}$, this reduces to the ordinary absolute value. For a vector $x$, $|x|$ is the distance between the point (in $n$-dimensional space) represented by $x$ and the origin of the coordinate system. Indeed, we can use this to measure the distance between any two points. If $x, y \in \mathbb{R}^n$, then

$$|x - y| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}. \quad \text{(S.2)}$$

In $\mathbb{R}$, this reduces to the ordinary absolute difference between two numbers.

### 2.5. Balls

Balls are sets in $\mathbb{R}^k$ consisting of points at most a fixed Euclidean distance away from a specified point. The open ball around $x$ of radius $r$ is the set $\mathcal{B}_x(r)$ which contains within $r$ of $x$ except for those on the spherical surface.

$$\mathcal{B}_x(r) = \{x \text{ such that } |x - u| < r\}. \quad \text{(S.4)}$$

The analogous closed ball $\overline{\mathcal{B}}_x(r)$ contains in addition the spherical surface.

$$\overline{\mathcal{B}}_x(r) = \{x \text{ such that } |x - u| \leq r\}. \quad \text{(S.4)}$$
In \( \mathbb{R} \), the open ball is an interval of the form \([x, x + r] \) and the closed ball is an interval of the form \([x, x + r] \).

2.6. Product Sets

If \( \mathcal{A} \) is a set and \( k \) is a positive integer, we use \( \mathcal{A}^k \) to denote\(^{20} \) the set of \( k \)-vectors of elements of \( \mathcal{A} \). Each element of \( \mathcal{A}^k \) can be written in the form \((a_1, \ldots, a_k)\) where the components \( a_i \) are all in \( \mathcal{A} \).

**Thought Question** Describe what the set \([0, 1]^3 \) looks like.

We can generalize this notion by considering vectors whose elements are obtained from an arbitrary sequence of sets. For example, consider the “rectangle” in the x-y plane consisting of points with \( 0 \leq x \leq 10 \) and \( 0 \leq y \leq 5 \). This rectangle set consists of ordered pairs \((x, y)\) with \( x \in [0, 10] \) and \( y \in [0, 5] \). Similarly, the points \((x, y, z)\) in \( \mathbb{R}^3 \) with \( x \in [0, 1], y \in [0, 2], \) and \( z \in [0, 4] \) form an open rectangular solid. Such vector sets are called *product sets*. We denote the product by a \( \times \) operator between the sets. For example, the rectangle above is the set \([0, 10] \times [0, 5] \). In general, if \( \mathcal{A}_1, \ldots, \mathcal{A}_m \) are any sets, then the product set \( \mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_m \) is defined as follows:

\[
\mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_m = \{(a_1, \ldots, a_m) \text{ such that } a_i \in \mathcal{A}_i, i = 1, \ldots, m\}.
\]

The vector sets we saw previously are also product sets. For instance, \( \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}, \mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R} \). We get \( \mathcal{A}^k \) back from this operation by

\[
\mathcal{A}^k = \underbrace{\mathcal{A} \times \cdots \times \mathcal{A}}_{k \text{ times}}.
\]

**Hyper-Rectangles.** A special case of product sets are products of intervals. A product of \( k \) intervals forms what is called a \( k \)-dimensional hyper-rectangle. For example, \([0, 1] \times [0, 2] \times [0, 3] \times [0, 4] \) is a four-dimensional subset of \( \mathbb{R}^4 \). A 1-dimensional hyper-rectangle is just an interval in \( \mathbb{R} \); a 2-dimensional hyper-rectangle is just a rectangle in \( \mathbb{R}^2 \); a 3-dimensional hyper-rectangle is just a rectangular solid in \( \mathbb{R}^3 \). The higher dimensional cases serve as a useful generalization. In fact, we use interval notation to define hyper-rectangles in \( \mathbb{R}^k \). For example, if \( a, b \in \mathbb{R}^k \) with \( b \geq a \), \([a, b] \) is the set of all points \( x \) such that \( a \leq x \leq b \), meaning \( a_i \leq x_i \leq b_i \) for \( i = 1, \ldots, k \). The analogous definition works for \([a, b [ \), \(]a, b] \), and \( ooI(a, b) \) replacing \( \leq \) with \( < \) for open-ended cases.

\(^{20}\) Note that \( \mathcal{A}^1 = \mathcal{A} \). We will see later that this exponent notation actually makes sense in other ways as well.
### 3. How to Make New Sets from Old Ones

**Unions.** The word “union” implies a bringing together, and that is exactly what you should think when considering the union of sets. See the examples in the figure.\(^{21}\) If \(A\) and \(B\) are sets, then their union \(A \cup B\) is the set containing those elements that belong to either \(A\) or \(B\):

\[
x \in A \cup B \text{ if and only if } x \in A \text{ or } x \in B.
\]

The “or” here is a logical or; it is true if \(x\) is in either or both sets. As usual with the definition of sets, no element is repeated; as the logical formulation above shows, even if an object is an element of both \(A\) and \(B\) it occurs only once in the union. Hence, \([0, 2] \cup [1, 4] = [0, 4]\) and \(\{a, b, c\} \cup \{c, r, g\} = \{a, b, c, r, g\}\).

Suppose now that I have three sets \(A, B,\) and \(C\). One can form the union of these three sets in three different ways:

1. \((A \cup B) \cup C,\)
2. \(A \cup (B \cup C),\)
3. \((A \cup C) \cup B.\)

Fortunately, by the definition above, all of these lead to exactly the same set,\(^{22}\) which we can confidently refer to as the union of \(A, B,\) and \(C\):

\[
x \in A \cup B \cup C \text{ if and only if } x \in A \text{ or } x \in B \text{ or } x \in C.
\]

What works for two sets or three sets, works for any number of sets. As an example, consider the collection of sets \(A_1, \ldots, A_{10}\) defined by \(A_i = \{2^{i-2}, 2^{i-1}, 2^i\}\). The union of the sets in this collection is denoted \(\bigcup_{i=1}^{10} A_i\). Logically, a number is in the union if it is in at least one of the \(A_i\); hence, \(\bigcup_{i=1}^{10} A_i = \{1/2, 1, 2, 4, 8, \ldots, 1024\}\). Check this carefully yourself. The same logic also works if the collection of sets is infinite. Consider for instance the infinite collection of sets defined by \(A_i = [1/(i + 1), 1]\) for \(i \in \mathbb{Z}^+\). The union of all the \(A_i\)'s is \(\bigcup_{i=1}^{\infty} A_i\) which contains any point that is in at least one of the \(A_i\)'s.

**Thought Question** Is \(0 \in \bigcup_{i=1}^{\infty} A_i\)? Is \(1\)? What is this union?

In general if \(A_1, A_2, \ldots\) is a collection of sets, finite or infinite, we have

\[
x \in \bigcup_{i} A_i \text{ if and only if } x \in A_i \text{ for at least one } i.
\]
If the index set is understood, we will write leave the range of indices implicit, as in \( \bigcup_i \); when we need to be specific, we will include the indices explicitly as above.

**Intersections.** At the intersection of two roads is a patch of ground that is common to both of them; this is the image underlying the intersection of sets. If \( \mathcal{A} \) and \( \mathcal{B} \) are sets, then their intersection \( \mathcal{A} \cap \mathcal{B} \) is the set containing those elements that belong to both \( \mathcal{A} \) and \( \mathcal{B} \):

\[
x \in \mathcal{A} \cap \mathcal{B} \text{ if and only if } x \in \mathcal{A} \text{ and } x \in \mathcal{B}.
\]

See the examples in the figure. The “and” here is a logical and; it is true if \( x \) is in both sets.

What is the intersection of two sets with no elements in common? (Think about this before proceeding.) If two sets have no elements in common, then their intersection must be the set with no elements, namely the empty set.

**Thought Question** What is the intersection of a set with itself?

If \( \mathcal{A} \subseteq \mathcal{B} \), what is \( \mathcal{A} \cap \mathcal{B} \)?

Just as unions can be defined for arbitrary collections of sets, so can intersections. In the example above, where \( \mathcal{A}_i = [1/(i + 1), 1[ \) for \( i \in \mathbb{Z}_+ \), the intersection \( \bigcap_{i=1}^\infty \mathcal{A}_i \) is a set whose elements are those numbers that are in every \( \mathcal{A}_i \). This is just the set \( [1/2, 1[ \). Check this for yourself. In general, for a collection \( \mathcal{A}_1, \mathcal{A}_2, \ldots \), finite or infinite,

\[
x \in \bigcap_i \mathcal{A}_i \text{ if and only if } x \in \mathcal{A}_i \text{ for every } i.
\]

If the index set is understood, we will write leave the range of indices implicit, as in \( \cap_i \); when we need to be specific, we will include the indices explicitly as above.

**Differences and Complements.** Given two sets \( \mathcal{A} \) and \( \mathcal{B} \), the difference \( \mathcal{A} - \mathcal{B} \) is defined to be the set of elements of \( \mathcal{A} \) that are not in \( \mathcal{B} \):

\[
x \in \mathcal{A} - \mathcal{B} \text{ if and only if } x \in \mathcal{A} \text{ and } x \notin \mathcal{B}.
\]

See the figure. Notice that this definition treats \( \mathcal{A} \) and \( \mathcal{B} \) differently; just as in a numerical subtraction, the order matters.

Very often, we work exclusively with subsets of some larger universe set \( \mathcal{U} \), such as the set of real numbers or as we will see later the set of
outcomes of an experiment. When the universe set is clear from context, we speak of the complement of a set. If $\mathcal{A} \subset \mathcal{U}$, then the complement of $\mathcal{A}$, denoted by $\text{compl}(\mathcal{A})$, is just $\mathcal{U} - \mathcal{A}$. This contains all the points in the superset that are not in $\mathcal{A}$. The complement is convenient when there is an overarching superset; it is the set equivalent of a logical not. Notice that $\mathcal{A} \cap \text{compl}(\mathcal{A}) = \emptyset$; a set and its complement are always disjoint.\(^{27}\) Also, $\mathcal{A} \cup \text{compl}(\mathcal{A}) = \mathcal{U}$.

**Mixing the Operations.** We have already seen that we can take unions or intersections of collections of sets in any order and get the same result; in other words, no parentheses are needed. This is no longer true when we mix unions, intersections, and differences in the same expression. The following general relationships hold:

\[
\mathcal{A} \cap \left( \bigcup_{i} B_{i} \right) = \bigcup_{i} (\mathcal{A} \cap B_{i})
\]

\[
\mathcal{A} \cup \left( \bigcap_{i} B_{i} \right) = \bigcap_{i} (\mathcal{A} \cup B_{i})
\]

\[
\text{compl} \left( \bigcup_{i} B_{i} \right) = \bigcap_{i} \text{compl}(B_{i})
\]

\[
\text{compl} \left( \bigcap_{i} B_{i} \right) = \bigcup_{i} \text{compl}(B_{i})
\]

where the $B_{i}$ are some collection of sets, finite or infinite. You should try to reason out verbally why these statements are true. For instance, if a point $x$ is in $\mathcal{A}$ and at least one of the $B$’s, then there is some $i$ for which $x \in \mathcal{A} \cap B_{i}$; hence $x$ is in the union of these sets. The reverse argument works as well.

**Thought Question** Find a simpler expression for

\[
\text{compl} \left( \bigcap_{n=1}^{\infty} \left[ -1/n, 1/n \right] \right).
\]

**Disjoint Sets.** To say that two sets $\mathcal{A}$ and $\mathcal{B}$ are disjoint means that they have no elements in common.\(^{28}\) That is $\mathcal{A} \cap \mathcal{B} = \emptyset$. Similarly, if $\mathcal{A}_{i}$ is a collection of sets, it is said to be a pairwise disjoint collection if for every index $i$ and $j$ with $i \neq j$, $\mathcal{A}_{i} \cap \mathcal{A}_{j} = \emptyset$; in other words, every pair of sets in the collection is disjoint. The collection of sets $\mathcal{A}_{i} = [i, i + 1[$ for $i \in \mathbb{Z}$ is pairwise disjoint but the collection $\mathcal{B}_{j} = [0, j[\hspace{1em}$ for $j \in \mathbb{Z}_{+}$.
is not. If $\mathcal{A}_i$ is a disjoint collection, the $\bigcup_i \mathcal{A}_i$ is called a disjoint union. The property of disjointness arises frequently in probability theory, and disjoint unions are special only because they make some computations easier.

**Thought Question** If $(\mathcal{A}_i)$ is a collection of sets indexed by $\mathbb{Z}_+$, and $\bigcap_{i \in \mathbb{Z}_+} \mathcal{A}_i = \emptyset$, does it imply that the collection is pairwise disjoint?

**Thought Question** (Eventually and Infinitely Often)

Suppose we have a sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \ldots$. Describe the meaning of the following two sets:

$$\left\{ \mathcal{A}_i \text{ eventually as } i \to \infty \right\} = \bigcup_{i = 1}^{\infty} \bigcap_{j \geq i} \mathcal{A}_j$$

$$\left\{ \mathcal{A}_i \text{ infinitely often as } i \to \infty \right\} = \bigcap_{i = 1}^{\infty} \bigcup_{j \geq i} \mathcal{A}_j.$$

Hint: If $i = 10$, the set inside the union on the top line is $\bigcap_{j = 10}^{\infty} \mathcal{A}_j$, and the set inside the intersection on the bottom line is $\bigcup_{j = 10}^{\infty} \mathcal{A}_j$. What do these sets represent? Think about the expressions above from the inside out.

**The Power Set of a Set.** Given any set $\mathcal{S}$, we can form a new set that consists of subsets of $\mathcal{S}$. This is called the power set of $\mathcal{S}$ and is denoted by $2^{\mathcal{S}}$. For our purposes, this is mostly a curiosity, but it does provide a nice example of how mathematical notation internalizes meaning. Suppose that $\mathcal{S}$ has a finite cardinality, then it has $2^{\# \mathcal{S}}$ elements. (Why? Order the elements of $\mathcal{S}$ in some arbitrary way and for each subset of $\mathcal{S}$ construct a binary number with one bit per element which is 1 if that element is in the subset and 0 otherwise. There are exactly $2^{\# \mathcal{S}}$ such numbers.) If $\mathcal{S}$ is non-empty, then it can be shown that the cardinality of $2^{\mathcal{S}}$, called $2^{\# \mathcal{S}}$, is strictly larger than the cardinality of $\mathcal{S}$. This is not obvious for sets with infinite cardinality. Notice that this gives us a whole sequence of “infinities”, each larger than the next.
Suppose that \( \mathbb{Z} \) were finite. Then we could find some finite \( n \in \mathbb{Z}_+ \) such that \( \mathbb{Z} \) is in explicit correspondence to \( \{1, \ldots, n\} \). But take the maximum of the integers connected to 1 through \( n \) by the lines we’ve drawn and add one. This cannot be connected to any number in \( \{1, \ldots, n\} \) (because it is bigger than the maximum of such numbers) but is clearly an integer. Oops. Since we come to a contradiction if \( \mathbb{Z} \) is finite, \( \mathbb{Z} \) must be infinite.

If \( A \subset B \) and \( A \) is uncountable, then so is \( B \) because there is no way to put the subset \( A \) or therefore \( B \) into correspondence with \( \mathbb{Z} \).

4. Finite, Countable, and Uncountable Sets

A finite set is a set whose cardinality is a finite integer. We can express the same idea in a more complicated but equivalent way by constructing an explicit correspondence between a set and the first \( n \) integers for some integer \( n \). Consider a set \( S \). List out the elements of \( S \) on one side of a (potentially very big) page. If for some \( n \in \mathbb{Z}_+ \) you can list the numbers \( 1, \ldots, n \) on the other side of the page and connect each of these numbers to exactly one element of \( S \) with a line, then \( S \) is finite and has cardinality \( n \). Any set that is not finite is called an infinite set.

Why go through all this when you can simply count, you may ask. Two answers: (i) to a mathematician this process of listing and connecting is counting, and (ii) this formalized notion of cardinality generalizes better to infinite sets.

We have already seen that there is a whole sequence of infinities, but it is unclear what those numbers represent. To get a feel for this, let’s start with the integers, which are an infinite set. List the integers on the left side of the page (one per line), and list the even integers on the other side. Next, draw a line from integer \( k \) on the left to integer \( 2k \) on the right. As you look over your list, you will notice that every number on the left is connected to one and only one number on the right, and vice versa. We have constructed an explicit correspondence between the integers and even integers; by definition, the two sets have the same cardinality. This is true even though one is a subset of the other! (Infinite sets can be odd.) The same type of argument can be used to show that \( \mathbb{Z}, \mathbb{Z}_+, \mathbb{Z}_{\mathbb{Z}}, \) and \( \mathbb{Q} \) all have the same cardinality.

Any infinite set that can be put in explicit, one-to-one correspondence (every point in one set is connected to only one point in the other set and vice versa) with the integers is called countably infinite. A set that is either finite or countably infinite is called countable; even if it takes forever, we can count the elements of such a set. A set that is not countable is called uncountable or uncountably infinite. There are many more elements of such a set than there are integers.

Are there any uncountable sets? Yes, \( \mathbb{R} \) is uncountable. (\( \mathbb{C} \) and \( \mathbb{R}^k \) for any \( k \geq 1 \) are also uncountable, as a result.) To see this, we will show that the interval \( ]0, 1[ \) is uncountable, which implies that any interval including \( \mathbb{R} = ]-\infty, \infty[ \) is itself uncountable. The proof that this is
so is the famous Cantor’s diagonal argument, and it is a hoot.

Consider numbers in ]0, 1[. Every such number can be written as \( \sum_{k=1}^{\infty} a_k 2^{-k} \) for some sequence of \( a_k \)'s that are all 0 or 1.\(^{31}\) In other words, we can write each such number as a binary fraction of the form 0.a_1a_2a_3a_4 \ldots. If we throw away all terminating sequences of \( a_k \)'s, those with only finitely many 1's, then this binary fraction representation is unique. (There aren’t really that many of these in some sense, so this is no loss). For instance, 1/2 has a representation 0.011111111111111111\ldots since 0.100000000000000000\ldots, with only one 1, has been thrown away.

That’s all details; it can all be demonstrated but take my word for it for now. The real fun comes in the next step. Suppose that the interval ]0, 1[ is countable. Then, we can list out all these binary fractions and the integers and draw appropriate lines between the lists. Suppose we do this and suppose our list of the elements of ]0, 1[ looks as follows:

\[
\begin{align*}
0.a_{11}a_{12}a_{13}a_{14}a_{15} \ldots \\
0.a_{21}a_{22}a_{23}a_{24}a_{25} \ldots \\
0.a_{31}a_{32}a_{33}a_{34}a_{35} \ldots \\
0.a_{41}a_{42}a_{43}a_{44}a_{45} \ldots \\
0.a_{51}a_{52}a_{53}a_{54}a_{55} \ldots \\
\ldots \\
\ldots \\
\ldots
\end{align*}
\]

Take this list and construct a correspondence with the integers, which we can do by the assumption that ]0, 1[ is countable. But now make a new number 0.b_1b_2b_3\ldots where \( b_k = 1 - a_{kk} \). This new number is not in our original list because it differs from every number on that list in at least one digit. Ooops. What this means is that no matter how we attempt to construct a correspondence between ]0, 1[ and the integers we always miss some of the numbers in ]0, 1[. That is, ]0, 1[ is uncountable.

\(^{31}\) Note: \( \sum_{k=1}^{\infty} 2^{-k} = 1 \).