Contents

1 Sets, Relations, and Functions .................................................. 2
  1 Sets with one, two and three members ......................................... 2
  2 Finite sets ........................................................................... 2
  3 Ordered pairs ....................................................................... 5
  4 The Cartesian Product of Two Sets ............................................. 6
  5 Relations ............................................................................. 7
  6 Functions ............................................................................ 9

2 Vector spaces over \( \mathbb{R} \) .......................................................... 13
  1 Vectors and scalars ................................................................. 15
  2 The zero vector ..................................................................... 16
  3 The zero vector ..................................................................... 17
  4 Subtraction of vectors ............................................................. 21
  5 “Points” vs. “Vectors” .............................................................. 22
  6 Some geometry, at last! ............................................................. 23
  7 Lines ................................................................................... 23
  8 The line through two distinct points .......................................... 25
  9 Parallelograms ..................................................................... 27
 10 The three medians of a triangle .................................................... 28
 11 Linear independence of two vectors ........................................... 32
 12 Planes .................................................................................. 33

3 Homework assignment No. 3, due on Thu. Feb. 11 ...................... 39
1 Sets, Relations, and Functions

1 Sets with one, two and three members

Given any object \( a \), we can form a set that has exactly one member, namely, \( a \). This set is called the singleton of \( a \), and is denoted by the expression \( \{a\} \).

The formal definition is

\[
\{a\} = \{x : x = a\}.
\]

So the criterion for membership in \( \{a\} \) is very simple: given an object \( u \), \( u \) belongs to \( \{a\} \) if and only if \( u = a \).

Given two objects \( a, b \) (which may or may not be different) we can form a set whose members are \( a \) and \( b \). This set is denoted by the expression \( \{a, b\} \), and is called the unordered pair of \( a \) and \( b \).

The formal definition is

\[
\{a, b\} = \{x : x = a \lor x = b\}.
\]

So the criterion for membership in \( \{a, b\} \) is very simple: given an object \( u \), \( u \) belongs to \( \{a, b\} \) if and only if \( u = a \) or \( u = b \).

Notice that the set \( \{a, b\} \) is a two-member set if \( a \neq b \), but it is a one-member set if \( a = b \).

Given three objects \( a, b, c \) (which need not be different) we can form a set whose members are \( a, b, \) and \( c \). This set is denoted by the expression \( \{a, b, c\} \), and is called the unordered triple of \( a, b, \) and \( c \).

The formal definition is

\[
\{a, b, c\} = \{x : x = a \lor x = b \lor x = c\}.
\]

So the criterion for membership in \( \{a, b, c\} \) is very simple: given an object \( u \), \( u \) belongs to \( \{a, b, c\} \) if and only if \( u = a \) or \( u = b \) or \( u = c \).

Notice that the set \( \{a, b, c\} \) is a three-member set if \( a \neq b \), \( a \neq c \), and \( b \neq c \), but it is a two-member set if \( a = b \) and \( b \neq c \), or \( b = c \) and \( a \neq b \), or \( a = c \) and \( b \neq a \). And it is a one-member set if \( a = b = c \).

2 Finite sets

One could go on as in the preceding section, and define “unordered quadruple”, “unordered quintuple” and so on. The trouble with this is that, one we
get to, say, defining “set with “elements”, then the definition becomes logn
and cumbersome. It is very simple, but too long.

Furthermore, we would like to define once and for all, “set with n mem-
bers” for any natural number n. We can do this inductively:

**Definition 1.** A set with 1 member (or 1-member set) is set which the sin-
gleton \( \{a\} \) for some \( a \).

If \( n \) is a natural number or zero, then a set with \( n + 1 \) members is a set
\( S \) such that \( S = S_0 \cup \{a\} \) for some set \( S_0 \) with \( n \) members and some \( a \) such
that \( a \neq S_0 \).

In addition, the empty set is a zero-member set.

This is an **inductive definition:** it tells you what a “one-member set”
is, and then it tells you what an \( n + 1 \)-member set is if you know what an
\( n \)-member set is.

**Definition 2.** A **finite set** is a set which is an \( n \)-member set for some \( n \in \mathbb{N} \cup \{0\} \).

To see how it works, let us give an example.

**Example 1.** Let \( S \) be the set of all prime numbers \( p \) such that \( p \leq 24 \). Then
the members of \( S \) are 2, 3, 5, 7, 11, 13, 17, 19, and 23.

According to Definition 1, the set \( \{2\} \) is a 1-member set. Then, using
Definition 1 again, the set \( \{2, 3\} \) is a 2-member set. Then, using Definition
1 one more time, the set \( \{2, 3, 5\} \) is a 3-member set. Then the set \( \{2, 3, 5, 7\} \)
is a 4-member set. And

What do we mean by “the number of members of a set”? (The mathe-
matical word for this is “cardinality”: if \( S \) is a set with 25 members, then we
say that the cardinality of \( S \) is 25, or that \( S \) has cardinality 25.)

We would like to say that “the cardinality of a finite set is the \( n \) such that
\( S \) is an \( n \)-member set”. The problem with this is the following: suppose, for
example, that there was a set \( S \) such that

\(*
S \) is both a 53-member set and a 71-member set, according to Definition
1.
In that case, which number is the cardinality of \( S \)? Is it 53 or 71.

Naturally, you will say that “this can never happen. A set cannot both have 53 members and also 73 members”. But then, of course, being a mathematician, I will ask you: why not? And you will say something like “because everybody knows that a set cannot have at the same time one number of members and another number of members.” But that is just like saying: “it cannot happen because it cannot happen”, or “it cannot happen because everybody knows it cannot happen”.

In mathematics, a statement such as “(*) cannot happen” needs proof. It turns out that one can prove the following theorem:

**Theorem 1.** Let \( S \) be a finite set. Let \( m \in \mathbb{N} \cup \{0\} \) and \( n \in \mathbb{N} \cup \{0\} \) be such that \( S \) has \( m \) members and \( S \) has \( n \) members. Then \( m = n \).

**Proof.** The proof is an easy exercise is proving things by induction. I leave it as an exercise for you.

The basic idea of what you have to do is this: you do “induction on \( n \)”. That is, you consider the following proposition \( P(n) \):

\[
(\#) \text{ Whenever } m \in \mathbb{N} \cup \{0\} \text{ and } S \text{ is a finite set that has both } m \text{ members and } n \text{ members, it follows that } m = n.
\]

(Notice that the proposition

\[
(\%) \text{ If } S \text{ is a finite set that has both } m \text{ members and } n \text{ members, it follows that } m = n.
\]

is a statement about \( S, m \) and \( n \), so it has three open variables, and we could call it, say, \( Q(m,n,S) \). So we cannot even think of proving (\%) by induction, because induction enables us to prove statements of the form \((\forall n \in \mathbb{N})A(n)\), where \( A(n) \) is a statement about one variable \( n \).

However, (\#) (that is, \( P(n) \)) says that \((\forall S)(\forall m \in \mathbb{N} \cup \{0\})Q(m,n,S)\), and this is a statement about one variable, namely, \( n \), so this is exactly the kind of statement that we can prove by induction.

One last remark: You would use induction to prove \((\forall n \in \mathbb{N})P(n)\). But here you want to prove \((\forall n \in \mathbb{N} \cup \{0\})P(n)\). In order to deal with this problem, you should

1. Prove \( P(0) \) separately, and then
2. prove \((\forall n \in \mathbb{N})P(n)\) by induction.

And the rest is up to you. I strongly recommend that you do this proof.

Exercise 1. Prove Theorem 1

3 Ordered pairs

Given any two objects \(a, b\), we can form the **ordered pair** \((a, b)\). This is basically a “list of two items”: when I give you the ordered pair \((a, b)\), I am telling you that the first entry is \(a\) and the second entry is \(b\).

The ordered pair \((a, b)\) is completely different from \(\{a, b\}\), the set whose members are \(a\) and \(b\). This set is called the **unordered pair** with members \(a, b\). If I give you the set \(S = \{a, b\}\), then this tells you that the members of \(S\) are \(a\) and \(b\), but it doesn’t tell you which one is first. In other words, \(\{a, b\}\) and \(\{b, a\}\) are the same set. The formal mathematical definition is as follows.

**Definition 3.** Let \(a, b\) be arbitrary objects.\(^1\) The set \(\{a, b\}\) given by

\[
\{a, b\} = \{x : x = a \lor x = b\}
\]

is the ordered pair with members \(a, b\).

That is, \(\{a, b\}\) is the set consisting of all the things that are equal to \(a\) or to \(b\).

Clearly, \(\{b, a\}\) is exactly the same thing as \(\{b, a\}\), that is

\[
(\forall a)(\forall b)\{a, b\} = \{b, a\}.
\]

**The ordered pair** \((a, b)\) **is a different object altogether:** The ordered pair \((a, b)\) and the ordered pair \((b, a)\) are not the same set (unless \(a = b\)). So the ordered pair \((a, b)\) is not the same thing as the unordered pair \(\{a, b\}\).

---

\(^1\)So \(a\) and \(b\) could be anything: natural numbers, real numbers, members of \(\mathbb{R}^3\), functions from \(\mathbb{R}\) to \(\mathbb{R}\), matrices, cows, planets, stars, whatever. Or \(a\) could be a cow and \(b\) the number 324.
And now that I have told you what the ordered pair \((a, b)\) is not, I have to tell you what it is. The actual formal definition does not matter very much. All we need is to find some way to associate to two objects \(a, b\) a set \((a, b)\) is such a way that \((a, b)\) and \((b, a)\) are not the same (unless \(a = b\)). Here is one way to do it:

**Definition 4.** Let \(a, b\) be arbitrary objects.\(^2\) The set \((a, b)\) given by

\[
(a, b) = \{\{a\}, \{a, b\}\}.
\]

is the unordered pair with members \(a, b\). We call \(a\) the first entry of \((A, b)\), and we call \(b\) the second entry of \((a, b)\).

Clearly, the members of \((a, b)\) are not \(a\) and \(b\). They are the sets \(\{a\}\) and \(\{a, b\}\). Notice that the members of \((b, a)\) are not the same as those of \((a, b)\): the members of \((b, a)\) are the sets \(\{b\}\) and \(\{a, b\}\).

It follows from this definition that the following is true:

**Theorem 2.** If \(a, b, c, d\) are any objects, then \((a, b) = (c, d)\) if and only if \(a = c\) and \(b = d\).

**Proof.** Recommended exercise for you to do.

**Remark 1.** It does not matter very much how you define the ordered pair \((a, b)\), as long as in the definition \(a\) and \(b\) are treated differently, so you can distinguish between the first entry and the second entry. For example, we could have defined \((a, b)\) to be the set \(\{\{a\}, \{b\}\}\). Or we could have taken it to be \(\{\{a\}\}, \{\{b\}\}\}\).

Theorem 2 says that if you have two ordered pairs that are the same set, then their first entries are the same and their second entries are the same.

### 4 The Cartesian Product of Two Sets

**Definition 5.** If \(A, B\) are sets, the Cartesian product of \(A\) and \(B\) is the set of all ordered pairs whose first entry is a member of \(A\) and whose second entry is a member of \(B\). We use \(A \times B\) to denote the Cartesian product of \(A\) and \(B\). Then

\[
A \times B \overset{\text{def}}{=} \{ x : (\exists a)(\exists b)(a \in A \wedge b \in B \wedge x = (a, b)\}.
\]

\(^2\)As before, \(a\) and \(b\) could be anything: natural numbers, real numbers, members of \(\mathbb{R}^3\), functions from \(\mathbb{R}\) to \(\mathbb{R}\), matrices, cows, planets, stars, whatever.
Exercise 2. Prove that if $A$ and $B$ are finite sets, and $m$, $n$ are natural numbers (or zero) such that $A$ has $m$ members and $B$ has $m$ members, then $A \times B$ has $mn$ members.

HINT: First prove that if $X$, $Y$ are finite sets, $X \cap Y = \emptyset$, and $m$, $n$ are natural numbers (or zero) such that $X$ has $m$ members and $Y$ has $m$ members, then $X \cup Y$ has $m+n$ members. (You can do this by induction on $m$, or on $n$.) Then use this result to prove that $A \times B$ has $mn$ members by induction on $m$ (or on $n$).

Intuitively, a relation is something that admits “inputs” and for each input produces one or several outputs. We can characterize a relation by listing all the pairs $(x, y)$ such that the relation produces the output $y$ for the input $x$. This leads to the following formal definition:

5 Relations

Definition 6. A relation is a set of ordered pairs. That is:

$(1.7) \text{is a relation } \iff (\left( R \text{ is a set } \land (\forall u)(u \in R \implies (\exists a)(\exists b)\ u = (a, b)) \right)$

The above definition says exactly what we want: $R$ is a relation if and only if $R$ is a set and every member of $R$ is an ordered pair.

As explained before, if $R$ is a relation then we should think of the ordered pairs $(x, y)$ that belong to $R$ as “input-output pairs”: if we are told that a pair $(x, y)$ belongs to $R$, then this tells us that $R$ produces the output $y$ for the input $x$.

If $x$ is any object, then it may happen that $x$ occurs as the first entry of a pair $(x, y)$ that belongs to the relation $R$. In that case, we will say that $x$ is an admissible input for $R$, or that $R$ admits $x$ as an input, or that $R$ accepts $x$ as an input.

Definition 7. If $R$ is a relation, then the domain of $R$ is the set of all admissible inputs for $R$. We use Dom$(R)$ to denote the domain of $R$.

Formally,

$(1.8) \text{Dom}(R) \overset{\text{def}}{=} \{x : (\exists y)(x, y \in R)\}$

We can give a similar definition for the set of outputs produced by $R$:
Definition 8. If $R$ is a relation, then the range of $R$ is the set of all outputs produced by $R$. We use $\text{Rn}(R)$ to denote the range of $R$.

Formally,

\[
\text{Rn}(R) \overset{\text{def}}{=} \{ y : (\exists x)(x, y \in R) \}.
\]

Notational convention: If $R$ is a relation, we write “$xRy$” as an alternative for “$(x, y) \in R$”. So “$xRy$” is another way of saying that “$R$ produces the output $y$ for the input $x$.”

Example 2. Consider the relation “less than”. Let us use the symbol “$<$” to denote it. This relation produces, for a real number $x$ as input, all real numbers $y$ such that $x$ is less than $y$. Then the domain of $<$ is $\mathbb{R}$, the set of all real numbers, and the range of $<$ is also $\mathbb{R}$. The pair $(x, y)$ belongs to $<$ if and only if $x$ is less than $y$, so when we write “$x < y$” for “$(x, y) \in <$”, the expression “$x < y$” says precisely that $x$ is less than $y$, which is why you have been writing “$x < y$” all along to say that $x$ is less than $y$.

Example 3. Consider the relation “$x = y^2$”. Let us use the symbol $R$ to denote it. This relation produces, for a real number $x$ as input, all the real numbers $y$ such that $y^2 = x$. Clearly, the admissible inputs are the nonnegative\(^3\) And every real number is a possible output. So

\[
\text{Dom}(R) = \{ u \in \mathbb{R} : u \geq 0 \} \quad \text{and} \quad \text{Rn}(R) = \mathbb{R}.
\]

Example 4. Consider the relation “$y = x^2$”. Let us use the symbol $R$ to denote it. This relation produces, for a real number $x$ as input, all the real numbers $y$ such that $y = x^2$. Clearly, the admissible inputs are all the real numbers, because every real number can be squared. And the possible outputs are the nonnegative real numbers So

\[
\text{Dom}(R) = \mathbb{R} \quad \text{and} \quad \text{Rn}(R) = \{ u \in \mathbb{R} : u \geq 0 \}.
\]

Example 5. Consider the relation “divides\(^4\) Let us use the symbol “$|$” to denote this relation. This relation produces, for each integer $m$ as input, all integers $n$ such that $m$ divides $n$. Then the domain of $|$ is $\mathbb{Z}$, the set of all

---

\(^3\)In these notes, “positive” means “$> 0$” and “nonnegative” means “$\geq 0$”.

\(^4\)Recall that an integer $m$ divides an integer $n$ if there exists an integer $k$ such that $n = mk$. Equivalent ways to say that $m$ divides $n$ are: $m$ is a factor of $n$, $n$ is divisible by $m$, $n$ is a multiple of $m$. 
integers, and the range of $|\cdot|$ is also $\mathbb{Z}$. The pair $(m, n)$ belongs to $|$ if and only if $m$ divides $n$, so when we write “$m|n$” for “$(m, n) \in |$”, the expression “$m|n$” says precisely that $m$ divides $n$, which why one normally writes “$m|n$” to say that $m$ divides $n$.

6 Functions

We now want to define the concept of “function”. Basically, a function is a relation $R$ that has the extra property that for every admissible input $R$ produces only one output. Let us write this out in precise mathematical language:

**Definition 9.** We say that a relation $R$ has the unique output property if for every $x \in \text{Dom}(R)$ there exists a unique $t$ such that $xRy$.

Equivalently, $R$ has the unique output property if

$$(\forall x)(\forall y)(\forall z) \left( (xRy \land xRz) \implies y = z \right).$$

We are now ready to say what a function is:

**Definition 10.** A function is a relation that has the unique output property.

Let us rephrase this definition in more direct terms, without using the concept of relation or the unique output property

**Definition 11.** A function is a set $f$ of ordered pairs such that, whenever $x, y, z$ are such that $(x, y) \in f$ and $(x, z) \in f$, it follows that $y = z$.

**Definition 12.** Let $f$ be a function. Let $x$ be a member of the domain of $f$. Then the unique $y$ such that $(x, y) \in f$ is called the value of $f$ at $x$. We use $f(x)$ to denote the value of $f$ at $x$.

It follows from the above definition that, if $f$ is a function, then the three statements

$$(x, y) \in f$$

$xfy$$

y = f(x)$$

mean exactly the same thing.

**Example 6.**
1. The relation “<” of Example 2 is not a function. \textit{Proof:} take \( x = 2, y = 3, z = 4 \). Then \( x < y \) and \( x < z \), but \( y \neq z \).

2. The relation \( R \) of Example 3 is not a function. \textit{Proof:} take \( x = 1, y = 1, z = -1 \). Then \( xRy \) and \( xRz \), but \( y \neq z \).

3. The relation \( R \) of Example 4 is a function. \textit{Proof:} Let \( x, y, z \) be such that \( xRy \) and \( xRz \). Then \( y = x^2 \) and \( z = x^2 \). So \( y = z \).

4. The relation \( | \) of Example 5 is not a function. \textit{Proof:} take \( x = 1, y = 1, z = 2 \). Then \( xRy \) and \( xRz \), but \( y \neq z \).

\begin{boxedtext}
You should read the Wikipedia article on “function (mathematics)”. And you should also read the Wikipedia article entitled “History of the function concept”. Please regard these two articles as reading requirements for our course.
\end{boxedtext}

\textbf{Remark 2.} The concept of “function” emerged during the second half of the 17th Century. The word “function” was introduced by Leibniz in 1673. The notation “\( f(x) \)” for the value of a function was introduced approximately in 1734. At that time, what people called “functions” was what we now call “differentiable functions”. (So, for example, the “function” \( f : \mathbb{R} \rightarrow \mathbb{R} \) given by \( f(x) = |x| \) for \( x \in \mathbb{R} \) was not regarded as a function. Later, mathematicians thought that a function was something given by some kind of formula or rule.

The idea that “a function is a rule that for every value of a variable \( x \) produces a value \( f(x) \)” is very commonly held even today by people who make some use of mathematics but are unfamiliar with modern mathematics. \textit{This idea is wrong} for the following two reasons:

1. The concept of “rule” is not a precise mathematical notion. The language of modern mathematics is the language of \textit{sets}, in which one can talk about sets, members of sets, and nothing else\(^5\). If you could give

\(^5\)Sure, we can talk about other things. But those things are always abbreviations of statements about sets and members. For example, we can say that “\( X \) is a subset of \( Y \)”,
a precise definition of “rule” in terms of sets and set membership, then it might perhaps be possible to give a precise definition of “function” as a rule of some kind. But until this is done, the definition of “function” as a rule does not pass the main test for a mathematical definition: for a mathematical definition to be acceptable, it has to be precise.

2. If a function was a “rule”, then the functions \( f \), \( g \) defined by the following two rules would be different:

RULE 1: To compute \( y = f(x) \), multiply \( x \) by \( x \); the result of this multiplication is \( f(x) \).

RULE 2: To compute \( y = g(x) \), compute \( x - 1 \), compute \( x + 1 \), multiply the two numbers, and add 1. The result of these computations is \( g(x) \).

The function \( f \) is given by \( f(x) = x^2 \). The function \( g \) is given by \( g(x) = (x - 1)(x + 1) + 1 \). So, you see, \( f \) and \( g \) are given by two different “rules”. However, it is easy to see that \( (x - 1)(x + 1) + 1 = x^2 \) (because \( (x - 1)(x + 1) = x^2 - 1 \)). So \( g(x) \), although computed by a totally different “rule”, gives exactly the same value, for each \( x \), as \( f(x) \).

If a function was a “rule”, then \( f \) and \( g \) would be different functions, because they are computed by different rules. But \( f \) and \( g \) are the same function! So a function cannot be a “rule”.

The only alternative is to say something like “the function is not the rule; the function is what you compute using a rule; if two different rules give rise to the same values, then they are the same function.” That sound fine! But: what is that thing that you compute using the rule? It is the set of input-output pairs! So we would be much better

but this is just a shorter way of saying

\[
(\forall x)(x \in X \implies x \in Y),
\]

which is a statement involving just sets and set membership. And we can talk about integers and real numbers, but integers and real numbers are special kinds of sets. And we can talk about ordered pairs, but the ordered pair \((a, b)\) is just a set, whose members are the sets \(\{a\}\) and \(\{a, b\}\). And we can talk about functions, but functions are relations of a special kind, and relations are sets of a special kind.
off if we say “a function is a set of input-output pairs obtained by computing an output $f(x)$ for each $x$ in a certain domain according to some rule”. And then once you understand this, all you need to remark is that the reference to the “rule”, besides being too vague to make sense, is superfluous. We can just say: “a function as a set of ordered pairs, such that, for every input $x$, the output (i.e., the $y$ such that $(x, y) \in f$) is unique. And this is precisely the modern definition of “function”, which we presented in Definition 10, or its equivalent version Definition 11.

The modern definition of “function” (i.e., Definition 10) was first proposed\(^6\) by Norbert Wiener in 1914. (The set-theoretical definition of “ordered pair” that we now use (i.e., Equation (1.5) in Definition 4) was proposed by Kazimierz Kuratowski in 1921).

\[\square\]

**Definition 13.** Let $f$ be a function, and let $A, B$ be sets. We say that $f$ is a function from $A$ to $B$, and write

$$f : A \mapsto B,$$

if $\text{Dom}(f) = A$ and $\text{Rn}(f) \subseteq B$.

**Remark 3.** “$f : A \mapsto B$” is a sentence, that says “$f$ is a function from $A$ to $B$”. It follows that there are lots of things that you should not say.

For example, you should not say

\[(1) \quad f \text{ is a function from } A \mapsto B,
\]

because (1) just says “$f$ is a function from from $A$ to $B$”, which is obviously not what you want to say.

Nor should you say

\[(2) \quad f \text{ is a function } f : A \mapsto B,
\]

because (2) just says “$f$ is a function $f$ is a function from from $A$ to $B$”, which is obviously not what you want to say.\[\square\]

\(^6\)Actually, Bertrand Russell had considered that definition in 1903 but rejected it.
Remark 4. A function $f$ has a domain and a range. But *when we write* $f : A \mapsto B$, *this does not* means that $A$ is the domain of $f$ and $B$ is the range. It means that $A$ is the domain and the range is a subset of $B$.

You may think that this asymmetry is ugly, that it would be better to make “$f : A \mapsto B$” mean “$A$ is the domain and $B$ is the range”. This, however, is not a good idea, for the following reason. When you create a function, you have to tell me how you compute (or determine) $f(x)$ for each $x$. And, obviously, you will have to specify which $x$’s the function accepts as inputs. So you have to specify the domain $\text{Dom}(f)$, and the “rule” for determining $f(x)$ for each in the domain. Usually, you don’t know *a priori* in which set you are going to end. You may know, for example, that your computation is always going to produce a real number, but it may be hard to figure out exactly which real numbers will occur as outputs. Determining $\text{Rn}(f)$, the set of outputs, exactly may be a difficult thing to do, so it would be rather inconvenient if you had to determine the set $\text{Rn}(f)$ in order to even be able to write $f : A \mapsto B$.

Furthermore, in most cases we are interested in studying lots of different functions with values in a set, and we want to compare them. For example, if we want to study curves in the plane, we can say that “a plane curve” is a continuous function $\gamma : I \mapsto \mathbb{R}^2$, where $I$ is an open interval. This is a good way to say that $\gamma$ is a plane curve, i.e., that it takes values in $\mathbb{R}^2$. But of course it will never be the case that the range of $\gamma$ is the whole set $\mathbb{R}^2$! □

2 Vector spaces over $\mathbb{R}$

Definition 14. A **vector space over** $\mathbb{R}$ is a set $V$ equipped with

1. A binary operation called *addition*, that produces, for every pair $(v, w)$ of members of $V$, a member of $V$ called the *sum of $v$ and $w$*, and denoted by the expression “$v + w$”,

2. A binary operation called *scalar multiplication*, that produces, for every member $r$ of $\mathbb{R}$ and every member $v$ of $V$, a member of $V$ called the *product of $r$ and $v$*, and denoted by the expressions “$rv$” or “$r.v$”,

---

*I am about to explain to you why it is not a good idea. If you don’t find my explanation convincing, then just accept that “$f : A \mapsto B$” means what I say it means because I say so, and that’s it.*
in such a way that the following properties hold:

(1) (Associative law of addition) If \( u, v, w \) are arbitrary members of \( V \), then
\[
    u + (v + w) = (u + v) + w.
\]

(2) (Commutative law of addition) If \( u, v \) are arbitrary members of \( V \), then
\[
    u + v = v + u.
\]

(3) (Subtraction law\(^8\)) If \( u, v \) are arbitrary members of \( V \), then there exists a \( w \in V \) such that
\[
    u = v + w.
\]

(4) (Cancellation law) If \( u, v, w \) are arbitrary members of \( V \), then
\[
    u + w = v + w \implies u = v.
\]

(5) (Distributive laws) If \( u, v \) are arbitrary members of \( V \), and \( r, s \) are arbitrary real numbers, then
\[
    r(u + v) = ru + rv
\]
and
\[
    (r + s)u = ru + su.
\]

(6) (Mixed associative law) If \( u \) is an arbitrary member of \( V \), and \( r, s \) are arbitrary members of \( \mathbb{R} \), then
\[
    r(su) = (rs)u.
\]

(7) (The one-law) If \( u \) is an arbitrary member of \( V \), then
\[
    1 \cdot u = u.
\]

\(^8\)We call this the “subtraction law” because the \( w \) whose existence is asserted here clearly is \( u - v \).
1 Vectors and scalars

If we want to study and prove facts about a vector space $V$ over $\mathbb{R}$, we have to deal with two different kinds of objects: members of $V$, and real numbers. So it is convenient to have different names for those two kinds of objects:

1. The members of $V$ are called vectors.
2. The real numbers are called scalars.

Then it should be clear that many things can be done, and many other things cannot be done.

For example:

1. We can add two vectors (and the result is a vector).
2. We can add two scalars (and the result is a scalar).
3. We cannot add a vector and a scalar.
4. We can multiply two scalars (and the result is a scalar).
5. We cannot multiply two vectors.
6. We can multiply a scalar times a vector (and the result is a vector).

Furthermore:

a. There is no such thing as a “vector” unless we specify the vector space first. In particular, if a student writes “a vector space is a set of vectors”, then this is nonsense.

b. If we are working with two or more vector spaces at the same time, then when we say “vector” we have to specify which vector space our “vector” belongs to. For example, if we are working with two vector spaces $V, W$, we have to say “vector of $V$” or “vector of $W$”.
2 The zero vector

So far, I have not told you what the “zero” vector is. The zero vector should be a vector \( z \) such that \( v + z = v \) for every vector \( v \). We want to introduce such a vector into our discourse, and give it a name. (We will call it “zero”, of course.) In order to be able to do that, we have to prove that such a vector necessarily exists, and is unique. And, once we have proved that, we will have the right to talk about it, and give it a name.

**Remark 5.** If we prove that an object of a certain kind exists and is unique, then (and only then) we are allowed to introduce it into our universe of discourse and give it a name. And this name should be a new word or phrase, because otherwise this new object could be mistaken for some other object that we have already named.

**Example 7.** Once we have learned that Nicole Brown has been murdered, we are entitled to introduce the phrase “the murderer” to refer to the person that killed Nicole. But we are not entitled to calling that person “O. J.”, because “O. J.” is the name of somebody else, and if we called the murderer “O. J.” we would be presupposing that O. J. killed Nicole. If we find out that O.J. indeed killed Nicole, then we can start using the name “O.J.” for the murderer.

**Example 8.** Suppose we find out that there exists a real number \( x \) that satisfies \( x^2 = 2 \). (You must have seen the proof of the existence of such a number in Math 311.) Are we entitled to calling that number “\( \sqrt{2} \)”?

The answer is not until we prove that that number is unique. And in fact “the real number \( x \) that satisfies \( x^2 = 2 \)” is not unique. (Reason: If \( a \) is a real number that satisfies \( x^2 = 2 \), then \( -a \) is another number that also satisfies \( x^2 = 2 \). So, if there exists a real number whose square is 2, then there are at least two of them. So the number is not unique.)

Now consider a slightly different situation: suppose we ask for a real number \( x \) such that \( x^2 = 2 \) and \( x > 0 \). In this case, we can prove that the number \( x \) that satisfies those conditions is unique. (Proof: The existence is something you proved in Math 311. As for uniqueness, this can be proved in several different ways. For example, suppose there were two different positive real numbers \( a, b \) such that \( a^2 = 2 \) and \( b^2 = 2 \). Then \( a^2 = b^2 \), so \( a^2 - b^2 = 0 \), so \( (a - b)(a + b) = 0 \). Therefore \( a - b = 0 \) or \( a + b = 0 \). But “\( a + b = 0 \)” is impossible, because \( a > 0 \) and \( b > 0 \), so \( a + b > 0 \), so \( a + b \) is not equal to 0. So \( a - b = 0 \). But then \( a = b \). So \( a \) and \( b \) cannot be different after all.)
Now, finally, we are entitled to giving a name to the unique positive real number whose square is 2. And, as you well know, we call this number $\sqrt{2}$.

**Example 9.** In my Math 300 notes (notes for Lecture 5, Pages 11-12), you will find the “Peano paradox”. It’s a “proof” of the statement that 1 is the largest natural number. This statement is of course false. Then how come it can be proved? The answer is that the proof is incorrect. Where does the proof go wrong? It goes wrong in the first step, which says “Let $n$ be the largest natural number”. This is a forbidden step, because **we are not allowed to introduce an object and give it a name (such as “$n$”) unless we have proved that such an object exists.** And in this case we have not proved that a largest natural number exists, so we are not allowed to introduce one and give it a name.

The morale of this example is this: **if you allow yourself one incorrect step in a proof, then you can prove any kind of nonsense you want. Therefore, proofs with one incorrect step are totally useless, and worth nothing.**

**Exercise 3.** Prove that $2 + 2 = 3$, with a proof that starts with the following step:

**STEP 1.** Pick a member of the empty set and call it $a$.

*After that step, every step should be correct.* And your proof should end with “$2 + 2 = 3$”.

## 3 The zero vector

As explained before, we would like to introduce a vector $z$ (called “zero”) such that

\[ u + z = u \quad \text{for every vector } u. \tag{2.10} \]

**Definition 15.** Let $V$ be a vector space over $\mathbb{R}$. A zero vector of $V$ is a vector $z \in V$ such that (refzerovec) holds.

Let us prove that the zero vector exists and is unique.

**Theorem 3.** If $V$ is a vector space over $\mathbb{R}$, then there exists a unique vector $z$ such that

\[ u + z = u \quad \text{for every } u \in V. \]
Proof. We will first prove that for any fixed vector $u$ there exists a “zero vector” that works for that $u$. And then we will prove that this vector is the same for all $u$.

Fix a vector $u \in V$. Then by the subtraction law there exists a vector $z \in V$ such that
\begin{equation}
    u = u + z.
\end{equation}
We now show that this $z$ is unique. Suppose we have two $z$‘s that satisfy (2.11). Call them $z_1$ and $z_2$, so we have
\begin{align*}
    u &= u + z_1 \quad \text{and} \quad u = u + z_2.
\end{align*}
It follows from this that
\begin{align*}
    z_1 + u &= u + z_1 \\
    &= u \\
    &= u + z_2 \\
    &= z_2 + u,
\end{align*}
Then the cancellation law implies that $z_1 = z_2$. So, indeed, for our fixed vector $u$, the $z$ that satisfies (2.11) is unique.

There is, however, no obvious reason why the same $z$ should work for all $u$. So, for each $u \in V$, let us call the $z$ we have found $z_u$. We want to prove that $z_u$ is the same vector for all $u$.

Let $u_1, u_2$ be two arbitrary vectors. Then
\begin{align*}
    z_{u_1} + u_1 &= u_1 \quad \text{and} \quad z_{u_2} + u_2 = u_2.
\end{align*}
It follows thst
\begin{align*}
    (2.12) \quad z_1 + (u_1 + u_2) &= (z_1 + u_1) + u_2 \\
    (2.13) \quad &= u_1 + u_2 \\
    (2.14) \quad &= u_2 + u_1 \\
    (2.15) \quad &= (z_2 + u_2) + u_1 \\
    (2.16) \quad &= z_2 + (u_2 + u_1) \\
    (2.17) \quad &= z_2 + (u_1 + u_2),
\end{align*}
where the justifications of the six steps above are as follows:
1. Step (2.12) follows from the associative law of addition.

2. Step (2.13) follows from the fact that \( z_1 + u_1 = u_1 \).

3. Step (2.14) follows from the commutative law of addition.

4. Step (2.15) follows from the fact that \( z_2 + u_2 = u_2 \).

5. Step (2.16) follows from the associative law of addition.

6. Step (2.17) follows from the commutative law of addition.

So we have shown that
\[
z_1 + (u_1 + u_2) = z_2 + (u_1 + u_2),
\]
It then follows from the cancellation law that
\[
z_1 = z_2.
\]
So we have shown that the \( z \) that works for a vector \( u \in V \) is the same for all \( u \).

Q.E.D.

In view of Theorem 3, we can give a name to this special vector: we can call it “zero”, or “the origin of \( V \)” and use the symbol \( 0_V \) to denote it.

Then
\[
(2.18) \quad u + 0_V = u \quad \text{for every} \quad u \in V.
\]
In view of the commutative law of addition, it also follows that
\[
(2.19) \quad 0_V + u = u \quad \text{for every} \quad u \in V.
\]

Notice that

1. each vector space has its own zero vector;

2. the zero vector \( 0_V \) is a different object from the zero scalar, that we will go on calling 0.

Often, we will just write “0” rather than “\( 0_V \)”, and call this vector “the origin”, provided it is clear whether we are talking about the zero scalar or the zero vector, and which vector space is involved. So, for example, in Theorems 4, 5, and 6 below, it is clear which of the “zeros” mentioned in the statement must be the zero vector, and which one is the zero scalar, so we just write “0” rather than “\( 0_V \)”. 
I will now state three simple theorems about the zero vector, prove one, and ask you to prove the other two.

**Theorem 4.** Let $V$ be a vector space over $\mathbb{R}$. Then

$$r \cdot 0 = 0$$

for every scalar $r$.

**Proof.** We start by observing that

$$0 + 0 = 0,$$

because $0 + u = u$ for every vector $u$, so in particular $0 + 0 = 0$. If $r$ is a scalar, we have then

$$r \cdot (0 + 0) = r \cdot 0.$$

By one of the distributive laws:

$$r \cdot (0 + 0) = r \cdot 0 + r \cdot 0.$$

Hence

$$r \cdot 0 + r \cdot 0 = r \cdot 0.$$

But

$$0 + r \cdot 0 = r \cdot 0,$$

because $0 + u = u$ for every $u \in V$.

Therefore

$$r \cdot 0 + r \cdot 0 = 0 + r \cdot 0.$$

It then follows from the cancellation law that

$$r \cdot 0 = 0,$$

completing our proof. \[\text{Q.E.D.}\]

**Theorem 5.** Let $V$ be a vector space over $\mathbb{R}$. Then$^9$

$$0 \cdot u = 0$$

for every vector $u \in V$.

**Proof.** You do this proof.

$^9$In the formula “$0 \cdot u = 0$”, it is clear that the first zero is the zero scalar, because if it was the zero vector then “$0 \cdot u$” would not make sense, since there is no such thing as the product of a vector by another vector. And, similarly, the second zero must be the zero vector, because the product of a scalar time a vector is a vector, so “$0, u$” is a vector.
Theorem 6. Let $V$ be a vector space over $\mathbb{R}$. Then, if $r$ is a scalar and $u$ is a vector,
\[ r \cdot u = 0 \implies (r = 0 \text{ or } u = 0). \]

Proof. You do this proof.

4 Subtraction of vectors

Now that we know how to add vectors and what the zero vector is, we can start subtracting vectors. We want to define what “$u - v$” means if $u$ and $v$ are vectors. Clearly, $u - v$ must be the vector $w$ such that $u = v + w$. So, in order to be able to talk about $u - v$ and give it a name, we have to prove that this $w$ exists and is unique.

Theorem 7. Let $V$ be a vector space over $\mathbb{R}$ and let $u, v$ be vectors in $V$. Then there exists a unique vector $w$ such that
\[ u = v + w. \] (2.20)

Proof. The fact that $w$ exists is just the subtraction law.

Now we have to prove uniqueness. Suppose $w_1, w_2$ are two vectors that satisfy
\[ u = v + w_1 \quad \text{and} \quad u = v + w_2. \]

Then
\[ v + w_1 = v + w_2. \]

By the commutative law of addition:
\[ w_1 + v = w_2 + v. \]

It then follows from the cancellation law that
\[ w_1 = w_2. \]

So we have proved that any two vectors $w$ that satisfy (2.20) must be equal. So $w$ is unique, and this completes our proof. \( \text{Q.E.D.} \)
Now that we have proved that the difference of two vectors $u - v$ exists and is unique, we can give it a name. And—surprise!—we will call it $u - v$.

**Definition 16.** Let $V$ be a vector space over $\mathbb{R}$ and let $u, v$ be vectors in $V$. Let $w$ be the unique vector $w$ such that $u = v + w$. Then $w$ is called the difference of $u$ and $v$, or the vector $u$ minus $v$, and is denoted by the expression $u - v$.

Finally, we need to know what we mean by the “negative” of a vector $u$. We have two choices: we could define $-u$ to be $0 - u$, or we could define $-u$ to be $(-1)u$. These are two different definitions. And it turns out that they both lead to the same vector. But, since we have to make a choice, I will choose the first one, and then I will ask you to prove that the second one would yield the same vector.

**Definition 17.** Let $V$ be a vector space over $\mathbb{R}$ and let $u$ be a vector in $V$. Then the negative of $u$ is the vector $-u$ defined by $-u = 0 - u$.

Then, as I said before, this gives the same result as if we had chosen the other definition.

**Theorem 8.** Let $V$ be a vector space over $\mathbb{R}$ and let $u$ be a vector in $V$. Then $(-1)u = -u$.

**Exercise 4.** Prove Theorem 8

5 “Points” vs. “Vectors”

In our study of geometry by means of linear algebra, we are going to be talking about “points” and “vectors”. For the time being, until further notice, “points” and “vectors” are the same thing. It turns out that the correct setting for geometry is that of **affine spaces**, not vector spaces, and **in affine spaces “points” and “vectors” are not the same thing**. But this more sophisticated analysis will be discussed later. For the time being, then, “points” are the same thing as “vectors”.

---

**Important notational convention.** If $A, B$ are two points of the vector space $V$, then the expression “$A\hat{\rightarrow}B$” will denote the vector $B - A$. 

6 Some geometry, at last!

I strongly recommend that you read the articles “Affine Geometry” and “Play-far’s Axiom” in Wikipedia.

Now that we have taken the necessary but very boring first steps in linear algebra\(^{10}\) let us start doing some geometry.

As you probably remember for previous math courses, the first thing one studies in geometry is lines, segments, triangles, planes, and so on. And one either postulates or proves basic properties such as

1. Given two different points \(p, q\), there exists a unique line that goes through \(p\) and \(q\).

2. If \(L\) is a line and \(P\) is a point of \(V\) then there exists a unique line through \(P\) which is parallel\(^{11}\) to \(L\).

3. If \(A, B, C, D\) are four distinct points, and the vectors \(\overrightarrow{AB}\) and \(\overrightarrow{CD}\) are equal, then the vectors \(\overrightarrow{AC}\) and \(\overrightarrow{BD}\) are equal.

4. The three medians of a triangle meet in one point.

So let us show how one can do geometry in our vector space setting, by proving the four statements I have just listed.

Naturally, we cannot prove anything about “lines”, “parallel lines”, “triangles”, and “medians”, unless we first define what those words mean. So we are going to start by defining what a “line” is.

7 Lines

We define the concept of “line” in two steps:

\(^{10}\)”Linear Algebra” is the study of vector spaces.

\(^{11}\)If you think that the meaning of two lines being “parallel” is that they do not meet, then this assertion is false in three-dimensional space, because given a line \(L\) in three-dimensional space and a point \(P\) not on \(L\), there exist lots of lines through \(P\) that do not meet \(L\). (The assertion is true in two dimensions, though, as long as \(P\) does not lie on \(L\).) But the correct definition of “parallel” is different, and with the correct definition the statement is true in all dimensions.
1. First, we will define “parametrized line”;

and then

II. we will define “line”.

**Definition 18.** Let $V$ be a vector space over $\mathbb{R}$. A parametrized line in $V$ is a function\(^{12}\) $\lambda : \mathbb{R} \to V$ such that, for a point\(^{13}\) $p \in V$ and a vector $v \in V$ such that $v \neq 0$, the following holds:

\begin{equation}
\lambda(t) = p + tv \quad \text{for every } t \in \mathbb{R}.
\end{equation}

The point $p$ is the time zero point, or starting point, of $\lambda$, and the vector $v$ is the velocity vector of $\lambda$. $\square$

**Definition 19.** Suppose that $\lambda$ is a parametrized line and $Q$ is a point of $V$. We say that $Q$ lies on $\lambda$, or $Q$ belongs to $\lambda$, or $\lambda$ goes through $Q$, if $Q = \lambda(t)$ for some $t \in \mathbb{R}$. If $\lambda(t) = Q$, we say that $\lambda$ goes through $Q$ at time $t$.

We then define the set of points of $\lambda$, or the carrier of $\lambda$, to be the set of points $Q$ such that $\lambda$ goes through $Q$.

So, in formal language:

\[
\text{Set}(\lambda) = \left\{ Q \in V : (\exists t \in \mathbb{R})Q = \lambda(t) \right\}.
\]

$\square$

**Definition 20.** Let $V$ be a vector space over $\mathbb{R}$. Then

1. A line in $V$ is a subset $L$ of $V$ such that

\[
L = \text{Set}(\lambda) \quad \text{for some parametrized line } \lambda.
\]

2. If $L$ is a line, then any parametrized line $\lambda$ such that $L = \text{Set}(\lambda)$ is said to be a parametrization of $L$.

3. If $P$ is a point of $V$, and $L$ is a line in $V$, we say that $L$ goes through $P$, or that $P$ lies on $L$, if $P$ belongs to $L$. $\square$

\(^{12}\)“$\lambda$” os the Greek letter “lambda”. Do not forget that if you want to do mathematics you have to know the Greek letters.

\(^{13}\)At this moment, and until further notice, “point” means the same as “vector”. Later, when we study affine spaces, we will change our point of view, and “point” will not mean exactly the same thing as “vector.”
In other words: a “line” is the set of all points that lie on some parametrized line. Two parametrized lines can be different but give rise to the same line.

Remark 6. Isn’t it weird that, instead of just saying “$P$ belongs to $L$”, we also have those two other ways of saying the same thing, namely, “$L$ goes through $P$” and “$P$ lies on $L$”?

The answer is that geometry is a very old subject (about 4,000 years old), but set theory is very new (about 100 years old). People got used to talking about “points” and “lines”, and lines “going through points” without thinking that a line is a set. (Even today, when we say that “the train from Boston to Washington goes through New York”, we don’t think that the train trajectory is a set of points, and New York is one of those points.) Now that we have sets, it is very convenient to think of a line as a set of points, but in geometry we all go on using the old terminology, and talk about lines going through points and points lying on lines. □

8 The line through two distinct points

Theorem 9. Let $V$ be a vector space over $\mathbb{R}$, and let $P, Q$ be two points of $V$ such that $P \neq Q$. Then there exists a unique line $L$ that goes through $P$ and $Q$.

Proof. First, we let $v = Q - P$. Define a function $\lambda : \mathbb{R} \rightarrow V$ by letting

$$\lambda(t) = P + tv \quad \text{for} \quad t \in \mathbb{R}.$$  

Then $\lambda$ is a parametrized line according to our definition. (Definition 18 requires that $v \neq 0$, but this is a consequence of our assumption that $Q \neq P$.)

Furthermore, $\lambda$ goes through $P$, because

$$\lambda(0) = P,$$

and $\lambda$ goes through $Q$, because

$$\lambda(1) = P + v = P + (Q - P) = Q.$$

Let $L$ be the set $\text{Set}(\lambda)$. Then $L$ is a line according to our definition, and $P$ and $O$ belong to $L$. So we have found a line $L$ that goes through $P$ and $Q$, and this proves the existence of such a line.
Now we have to prove uniqueness. For this purpose, we have to show that if \( \hat{L} \) is any line that goes through \( P \) and \( Q \), then \( \hat{L} = L \).

So let \( \hat{L} \) be a line that goes through both \( P \) and \( Q \). We want to prove that \( \hat{L} = L \). And we are going to prove this the way we prove that two sets are equal: we will prove that \( L \subseteq \hat{L} \) and that \( \hat{L} \subseteq L \).

First, let us observe that, since \( \hat{L} \) is a line, we have \( \hat{L} = \text{Set}(\hat{\lambda}) \) for some parametrized line \( \hat{\lambda} \). And we can pick a point \( \hat{P} \in V \) and a vector \( \hat{v} \in V \) such that

\[ \hat{v} \neq 0 \]

and

\[ \hat{\lambda}(t) = \hat{P} + t\hat{v} \quad \text{for every} \quad t \in \mathbb{R}. \]

Since \( P \) lies on \( \hat{L} \), we may pick a time \( t_P \in \mathbb{R} \) such that

\[ \hat{\lambda}(t_P) = P. \]

And, since \( Q \) lies on \( \hat{L} \), we may pick a time \( t_Q \in \mathbb{R} \) such that

\[ \hat{\lambda}(t_Q) = Q. \]

Then

\[ t_Q \neq t_P, \]

because \( Q \neq P \). (If \( t_Q \) was equal to \( t_P \), then \( \hat{\lambda}(t_Q) \) would be equal to \( \hat{\lambda}(t_P) \), so \( Q \) would be equal to \( P \).)

If \( t \in \mathbb{R} \), then we have

\[
\begin{align*}
\lambda(t) &= P + tv \\
&= P + t(Q - P) \\
&= P + tQ - tP \\
&= (1 - t)P + tQ \\
&= (1 - t)\hat{\lambda}(t_P) + t\hat{\lambda}(t_Q) \\
&= (1 - t)(\hat{P} + t_P\hat{v}) + t(\hat{P} + t_Q\hat{v}) \\
&= (1 - t)\hat{P} + (1 - t)t_P\hat{v} + t\hat{P} + tt_Q\hat{v} \\
&= \hat{P} + \left((1 - t)t_P + tt_Q\right)\hat{v} \\
&= \hat{\lambda}\left((1 - t)t_P + tt_Q\right),
\end{align*}
\]
so
\[\lambda(t) = \hat{\lambda}(1 - t) \rho t + t \eta t \] for every \( t \in \mathbb{R} \).

It follows that if \( p \) is any point of \( L \), then \( t \in \hat{L} \), because we may pick \( t \in \mathbb{R} \) such that \( \lambda(t) = p \), and then \( \hat{\lambda}(1 - t) \rho t + t \eta t \) = \( p \), so \( p \in \text{Set}(\hat{\lambda}) \), i.e., \( p \in \hat{L} \). Hence \( L \subseteq \hat{L} \).

We now prove that \( \hat{L} \subseteq L \). Let \( p \) be an arbitrary member of \( \hat{L} \). Then we may pick a real number \( s \) such that
\[ p = \hat{\lambda}(s) .\]

Let \( t \in \mathbb{R} \) be such that \( (1 - t) \rho t + t \eta t = s \). (Precisely, \( t \) is the solution of \( (1 - t) \rho t + t \eta t = s \), i.e., of \( \rho t + t(\eta t - \rho t) = s \), so
\[ t = \frac{s - \rho t}{\eta t - \rho t} .\]

Notice that the right-hand side of (2.23) is well defined, because \( \eta t \neq \rho t \).

Then
\[ \hat{\lambda}(s) = \lambda(t) ,\]

i.e., \( \hat{\lambda} = \lambda(t) \). This shows that \( \hat{\lambda} \in \text{Set}(\lambda) \), i.e., \( \hat{\lambda} \in L \). So we have shown that \( \hat{L} \subseteq L \).

Since \( L \subseteq \hat{L} \) and \( \hat{L} \subseteq L \), we have proved that \( \hat{L} = L \), and this completes our proof of uniqueness.

\textbf{Q.E.D.}

\textbf{Important notational convention.} If \( A, B \) are two points of the vector space \( V \), such that \( A \neq B \), then the expression “\( AB \)” will denote the unique line that goes through \( A \) and \( B \).

9 \ Parallelograms

The following theorem gives information about parallelograms.
Theorem 10. Let $B$ be a vector space over $\mathbb{R}$. If $A, B, C, D$ are four distinct points of $V$, and the vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are equal, then the vectors $\overrightarrow{AC}$ and $\overrightarrow{BD}$ are equal.

Proof. We are assuming that $\overrightarrow{AB} = \overrightarrow{CD}$, i.e., that $B - A = D - C$.

Then $B - A + C = D$, so $B + C = D + A$. Hence $B + C - A = D$, so $C - A = D - B$. Hence $\overrightarrow{AC} = \overrightarrow{BD}$. This completes our proof. Q.E.D.

10 The three medians of a triangle

Definition 21. Let $V$ be a vector space over $\mathbb{R}$, and let $A, B$ be points of $V$. The midpoint of the segment from $A$ to $B$ is the point $\text{midpt}(A, B)$ given by

$$\text{midpt}(A, B) \overset{\text{def}}{=} A + \frac{1}{2} \overrightarrow{AB}.$$

It then follows that

$$\text{midpt}(A, B) = \frac{1}{2}(A + B),$$

because

$$\text{midpt}(A, B) = A + \frac{1}{2} \overrightarrow{AB}$$
$$= A + \frac{1}{2}(B - A)$$
$$= A + \frac{1}{2}B - \frac{1}{2}A$$
$$= \frac{1}{2}(A + B).$$
If we have three points $A, B, C$, we can talk about the “triangle” with vertices $A$, $B$, and $C$. (We do not need to give a precise definition of “triangle” at this point, so we won’t.)

**Definition 22.** Three points $A, B, C$ in a vector space $V$ are **non collinear** if they do not lie on a line. That is, $A, B$ and $C$ are non collinear if there does not exist a line $L$ such that $A \in L$, $b \in L$ and $C \in L$. \[ \square \]

**Lemma 1.** Suppose $A, B, C$ are non collinear points. Then

(1) $A \neq B$, $A \neq C$, and $B \neq C$,

(2) the midpoints of the lines $AB$, $BC$, and $AC$ (which are well defined because $A \neq B$, $A \neq C$, and $B \neq C$), satisfy

(2.24) \quad \text{midpt}(A, B) \neq C \\
(2.25) \quad \text{midpt}(A, C) \neq B \\
(2.26) \quad \text{midpt}(B, C) \neq A$.

**Proof.** If $A = B$, then the line $BC$ (or any line containing $B$, if $B = C$) would also contain $A$, so $A, B, C$ would not be non collinear. So $A \neq B$. The inequalities $A \neq C$ and $B \neq C$ follow in a similar way.

The midpoint midpt($A, B$) obviously lies on the line $AB$. So if $C$ was equal to midpt($A, B$) it would follow that $C$ lies on the line $AB$, and then $A, B, C$ would not be non collinear. Hence (2.24) follows. Formulas (2.25) and (2.26) are proved in a similar way. \[ \text{Q.E.D.} \]

If $A, B, C$ are non collinear points, then let us define

(2.27) \quad P = \text{midpt}(A, B), \\
(2.28) \quad Q = \text{midpt}(B, C), \\
(2.29) \quad R = \text{midpt}(A, C).

Lemma 1 tells us that $A \neq Q$, $B \neq R$, and $C \neq P$. Hence the three lines $AQ$, $BR$, $CP$ are well defined.

**Definition 23.** The medians of the triangle $ABC$ are the lines $AQ$, $BR$, $CP$, where $P, Q, R$ are the points defined by (2.24), (2.25), (2.26). \[ \square \]
Theorem 11. The three medians of a triangle $ABC$ meet at one point. That is, there exists a point $O$ that belongs to the three lines $AQ$, $BR$, and $CP$.

In other words,

\[ AQ \cap BR \cap RP \neq \emptyset, \tag{2.30} \]

where $P, Q, R$ are the points defined by Formulas (2.27), (2.28), (2.29).

The three medians of a triangle meet at a point

**Proof.** Let

\[ O = \frac{1}{3}(A + B + C). \]

Then

\[
O = \frac{1}{3}A + \frac{1}{3}(B + C) \\
= \frac{1}{3}A + \frac{2}{3}\left(\frac{1}{2}(B + C)\right)
\]
\[= \frac{1}{3}A + \frac{2}{3}\text{midpt}(B, C)\]
\[= \frac{1}{3}A + \frac{2}{3}Q\]
\[= \left(1 - \frac{2}{3}\right)A + \frac{2}{3}Q\]
\[= A - \frac{2}{3}A + \frac{2}{3}Q\]
\[= A + \frac{2}{3}(Q - A)\]
\[= A + \frac{2}{3}A\text{Q} .\]

Let \(\lambda_A : \mathbb{R} \mapsto V\) be the function given by
\[\lambda_A(t) = A + t(Q - A) \quad \text{for} \quad t \in \mathbb{R} .\]

(That is, \(\lambda_A(t) = A + t\overrightarrow{AQ}\) for \(\in \mathbb{R}\).)

Then \(\lambda_A\) is a parametrized line, which goes through \(A\) and \(Q\). And we have shown that \(O = A + \frac{2}{3}\overrightarrow{AQ}\). So \(O = \lambda_A(\frac{2}{3})\). Hence \(O \in \text{Set}(\lambda_A)\), and \(\text{Set}(\lambda_A) = AQ\). So \(O\) belongs to the line \(AQ\).

A similar argument proves that \(O\) belongs to the lines \(BR\) and \(CP\). Hence \(O \in AB \cap BR \cap CP\), so \(AB \cap BR \cap CP \neq \emptyset\), completing our proof. Q.E.D.

**Remark 7.** As you have seen, proving Theorem 11 is very easy, *if you know a formula for the point \(O\); you just plug in the formula and show that the point \(O\) belongs to all three lines.*

This is a perfectly valid proof, because in order to prove that a set is nonempty, all we need is to exhibit a member of the set, and it does not matter how we figured out which formula was the right one to try. We could have found the formula by “solving the equations”, or by “having a hunch”. Or maybe a voice from heaven told us. That’s immaterial: if we exhibit a member of a set, then that proves that the set is nonempty.

Some students may feel that this is “cheating”, because “we didn’t really find \(O\)”. Surely, if you want to do a proof like this on your own, you will have to figure out a formula for the solution; but once you have figured it out, there is no need to tell me how you got it: all that matters is that the expression you try works. \(\square\)
11 Linear independence of two vectors

Definition 24. Let $V$ be a vector space over $\mathbb{R}$. Two vectors $u, v$ in $V$ are linearly independent if neither one is a scalar multiple of the other, that is, if there does not exist a scalar $r$ such that $u = rv$, and there does not exist a scalar $r$ such that $v = ru$. □

Definition 25. Let $V$ be a vector space over $\mathbb{R}$. Let $u, v$ be two vectors in $V$.

1. a vector $w \in V$ is a linear combination of $u$ and $v$ if there exist scalars $a, b$ such that $w = au + bv$.

2. The linear span of $u$ and $v$ is the set $\text{Span}(u, v)$ whose members are all the linear combinations of $u$ and $v$. That is

$$\text{Span}(u, v) \overset{\text{def}}{=} \left\{ w \in V : (\exists a \in \mathbb{R})(\exists b \in \mathbb{R})w = au + bv \right\}. \quad \square$$

Exercise 5. Prove the following statement: If $V$ is a vector space over $\mathbb{R}$, and $A, B, C$ are three distinct\textsuperscript{14} points of $V$, then $A, B, C$ are non collinear if and only if the vectors $u = \overrightarrow{AB}$ and $v = \overrightarrow{AC}$ are linearly independent. □

Exercise 6. Prove the following statement: If $V$ is a vector space over $\mathbb{R}$, and $u, v, w, x$ are four vectors in $V$, such that $u$ and $v$ are linearly independent, then $\text{Span}(u, v) = \text{Span}(w, x)$ if and only if there exist scalars $a, b, c, d$ such that

$$w = au + bv,$$
$$x = cu + dv,$$

and

$$ad \neq bc.$$\textsuperscript{14}that is, such that $A \neq B$, $A \neq C$, and $B \neq C$
12 Planes

Definition 26. Let $V$ be a vector space over $\mathbb{R}$.

1. A parametrized plane in $V$ is a function $\lambda : \mathbb{R}^2 \to V$ such that, for a point $A$ of $V$ and two linearly independent vectors $u, v$ in $V$, we have
   \[ \lambda(t, s) = A + tu + sv \quad \text{for every} \quad (t, s) \in \mathbb{R}^2. \]

2. If $\lambda$ is a parametrized plane, and $P$ is a point of $V$, then we say that $\lambda$ goes through $P$ if there exist parameter values $t \in \mathbb{R}$, $s \in \mathbb{R}$, such that
   \[ \lambda(t, s) = P. \]

3. If $\lambda$ is a parametrized plane, then the expression
   \[ \text{Set}(\lambda) \]
   denotes the set of all points $P$ of $V$ such that $\lambda$ goes through $P$. That is,
   \[ \text{Set}(\lambda) = \left\{ P \in V : (\exists t \in \mathbb{R})(\exists s \in \mathbb{R})\lambda(t, s) = P \right\}. \]

4. A plane in $V$ is a subset $P$ of $V$ such that
   \[ P = \text{Set}(\lambda) \quad \text{for some parametrized plane} \quad \lambda. \]

5. If $P$ is a plane in $V$ and $A$ is a point of $V$, we say that $A$ lies on $P$, or that $P$ goes through $A$, if $A$ belongs to $P$.

6. If $P$ is a plane in $V$, and $\lambda$ is a parametrized plane such that $\text{Set}(\lambda) = P$, then $\lambda$ is called a parametrization of $P$. \hfill \Box

Theorem 12. Let $V$ be a vector space over $\mathbb{R}$, and let $A, B, C$ be three non collinear points of $V$. Then there exists a unique plane $P$ in $V$ that goes through $A$, $B$ and $C$.

Proof. We will first we construct a plane $P$ that goes through $A$, $B$ and $C$, thus proving existence, and then we will prove that any plane $\hat{P}$ that goes through $A$, $B$ and $C$ must coincide with $P$. 

To construct $\mathbf{P}$, we define

$$u = \overrightarrow{AB},$$
$$v = \overrightarrow{AC}.$$

We then define a function $\lambda : \mathbb{R}^2 \mapsto \mathbb{R}$ by letting

$$\lambda(t, s) = A + tu + sv \quad \text{for} \quad t \in \mathbb{R}, \; s \in \mathbb{R}.$$

We would like to assert that $\lambda$ is a parametrized plane, but to be able to say that we have to know that

$(\#)$ $u$ and $v$ are linearly independent.

So let us prove $(\#)$ first.

Suppose $u$ and $v$ are not linearly independent. Then either

(a) $u = rv$ for some $r \in \mathbb{R},$

or

(b) $v = ru$ for some $r \in \mathbb{R}.$

Suppose (a) holds, and write $u = rv, \; r \in \mathbb{R}.$ Let $L$ be the line $AC.$ (This line is well defined because $A \neq C.$) Then, if we define $\gamma : \mathbb{R} \mapsto V$ by letting

$$\gamma(t) = A + tv \quad \text{for} \quad t \in \mathbb{R},$$

it is clear that $\gamma$ is a parametrized line, and the line $\text{Set}(\gamma)$ is precisely $L$, because $\gamma(0) = A$ and $\gamma(1) = A + v = A + (C - A) = C,$ so $A$ and $C$ both belong to $\text{Set}(\gamma)$.

Now, $B = A + u = A + rv,$ so $\gamma(r) = B.$ Hence $B$ belongs to $\text{Set}(\gamma),$ so $B \in L.$ It follows that $A, \; B$ and $C$ belong to $L,$ and this contradicts the fact that $A, \; B$ and $C$ are non collinear.

A similar contradiction is obtained if (b) holds. So the assumption that $u$ and $v$ are not linearly independent leads to a contradiction, and we have proved $(\#)$.

Now that we know that $u$ and $v$ are linearly independent, we are able to assert that $\lambda$ is a parametrized plane. Let $\mathbf{P}$ be the plane $\text{Set}(\lambda).$
Clearly,

\[
\begin{align*}
\lambda(0, 0) &= A, \\
\lambda(1, 0) &= A + u = A + (B - A) = B \\
\lambda(0, 1) &= A + v = A + (C - A) = C.
\end{align*}
\]

So \(A, B\) and \(C\) belong to \(P\).

We have thus proved the existence of a plane \(P\) such that \(A, B\) and \(C\) belong to \(P\), by actually constructing one such plane.

Next, we have to prove that if \(Q\) is any plane that goes through \(A, B\) and \(C\), then \(Q = P\).

So let \(Q\) is a plane that goes through \(A, B\) and \(C\). Let \(\mu : \mathbb{R}^2 \mapsto V\) be a parametrization of \(Q\). Let \(A_*, u_*, v_*\) be such that

\[
\mu(t, s) = A_* + t u_* + s v_* \quad \text{for} \quad t \in \mathbb{R}, \; s \in \mathbb{R},
\]

and \(u_*\) and \(v_*\) are linearly independent.

Since \(A, B\) and \(C\) belong to the plane \(Q\), we may pick pairs \((t_A, s_A) \in \mathbb{R}^2\), \((t_B, s_B) \in \mathbb{R}^2\), \((t_C, s_C) \in \mathbb{R}^2\), such that

\[
\begin{align*}
A &= \mu(t_A, s_A), \\
B &= \mu(t_B, s_B), \\
C &= \mu(t_C, s_C).
\end{align*}
\]

So we have

\[
\begin{align*}
(2.31) \quad A &= A_* + t_A u_* + s_A v_* , \\
(2.32) \quad B &= A_* + t_B u_* + s_B v_* , \\
(2.33) \quad C &= A_* + t_C u_* + s_C v_* .
\end{align*}
\]

If \(R\) is any point of \(P\), then we may write

\[
R = \lambda(t, s) = A + tu + sv
\]
for some scalars \( t, s \). Then

\[
R = \lambda(t, s) = A + tu + sv
\]

\[
R = A + t(B - A) + s(C - A)
\]

\[
= (1 - t - s)A + tB + sC
\]

\[
= (1 - t - s)\left( A_\tau + t_A u_\tau + s_A v_\tau \right) + tB + sC
\]

\[
\mu(\tau, \sigma)
\]

where

\[
\tau = t_A + t(t_B - t_A) + s(s_B - s_A),
\]

\[
\sigma = t_A + t(t_C - t_A) + s(s_C - s_A).
\]

So \( R \) belongs to \( Q \). Hence we have shown that every point of \( P \) belongs to \( Q \), and then \( P \subseteq Q \).

We now want to prove that, conversely, \( Q \subseteq P \). Let \( R \) be an arbitrary point of \( Q \). Then we may pick scalars \( \tau, \sigma \), such that

\[
R = \mu(\tau, \sigma).
\]

If we could find scalars \( t, s \) such that (2.34) and (2.35) hold, then it will follow from our previous calculation that

\[
\lambda(t, s) = \mu(\tau, \sigma) = R,
\]

and this will prove that \( R \) belongs to \( \text{Set}(\lambda) \), so \( R \in P \).

So all we need to do is prove that

\[
\text{(\%)} \text{ For every pair of scalars } \tau, \sigma, \text{ there exists a pair } (t, s) \text{ such that (2.34), (2.35) hold.}
\]
In other words, we have to prove that Equations (2.34), (2.35) can be solved so as to find, for every pair \((\tau, \sigma)\), a pair \((t, s)\) for which the equations hold.

This is an elementary Linear Algebra problem. We can rewrite the system (2.34), (2.35) as

\[
\begin{align*}
(2.36) & \quad t(t_B - t_A) + s(s_B - s_A) = \tau - t_A, \\
(2.37) & \quad t(t_C - t_A) + s(s_C - s_A) = \sigma - t_A.
\end{align*}
\]

And this, in turn, can be written in matrix form as

\[
Mx = y,
\]

where \(M\) is the \(2 \times 2\) matrix given by

\[
M = \begin{bmatrix}
t_B - t_A & s_B - s_A \\
t_C - t_A & s_C - s_A
\end{bmatrix},
\]

\(x\) is the column vector given by

\[
x = \begin{bmatrix}
t \\
s
\end{bmatrix},
\]

and \(y\) is the column vector given by

\[
y = \begin{bmatrix}
\tau - t_A \\
\sigma - s_A
\end{bmatrix}.
\]

We need to prove that for every \(y \in \mathbb{R}^2\) there exists \(x \in \mathbb{R}^2\) such that \(Mx = y\). And we know from Linear algebra that this happens if and only if

\[
\det M \neq 0,
\]

where “\(\det M\)” stands for “determinant of \(M\).”

We also know from Linear Algebra that the determinant of a square matrix is nonzero if and only if the rows of the matrix are linearly independent. In our case, it will follow that \(\det M \neq 0\) if we prove that

\((\%)\) The rows of \(M\) are linearly independent.

To prove \((\%)\), we assume that the rows of \(M\) are linearly dependent, and try to derive a contradiction. Let \(\rho_1, \rho_2\) be the rows of \(M\). Since \(\rho_1\) and \(\rho_2\) are linearly dependent, one of them must be a multiple of the other one. So either
(1) $\rho_1$ is a multiple of $\rho_2$,

or

(2) $\rho_2$ is a multiple of $\rho_1$.

Consider the case when (1) holds. Write $\rho_1 = r\rho_2$, for some real number $r$.

we then have

\[
\begin{align*}
t_B - t_A &= r(t_C - t_A), \\
s_B - s_A &= r(s_C - s_A).
\end{align*}
\]

Since

\[
\begin{align*}
A &= A_* + t_A u_* + s_A v_*, \\
B &= A_* + t_B u_* + s_B v_*, \\
C &= A_* + t_C u_* + s_C v_*, \\
u &= B - A, \\
v &= C - A,
\end{align*}
\]

we have

\[
\begin{align*}
u &= (t_B - t_A)u_* + (s_B - s_A)v_* \\
&= r(t_C - t_A)u_* + r(s_C - s_A)v_* \\
&= r((t_C - t_A)u_* + (s_C - s_A)v_*) \\
&= r(C - A) \\
&= rv,
\end{align*}
\]

so $u$ is a multiple of $v$, and we have reached a contradiction, because the vectors $u, v$ are supposed to be linearly independent.

In exactly the same way, we reach a contradiction if we assume that (2) holds.

So we have proved that the rows of $M$ are linearly independent, and this completes our proof. \[Q.E.D.\]
3 Homework assignment No. 3, due on Thu. Feb. 11

PROBLEM 1. Prove Theorems 5 and 6, on Page 20.
PROBLEM 2. Prove Theorem 8 on Page 22.
PROBLEM 3. Exercise 5 on Page 32.

Problem 5. In this problem we discuss how to move the origin of a vector space $V$, and put it at any point $A$ of $V$ other than the “true” origin $0_V$.

Let $V$ be a vector space over $\mathbb{R}$. Fix a point $A$ of $V$.

We define on the set $V$ a new addition operation, called $+_A$ (“addition with origin $A$”), and a new scalar multiplication operation, called $\cdot_A$ (“scalar multiplication with origin $A$”), as follows:

1. If $P \in V$ and $Q \in V$, we let
   \[ P +_A Q = P + Q - A. \] (3.38)

2. If $P \in V$ and $r \in \mathbb{R}$, we let
   \[ r \cdot_A P = r(P - A) + A, \] (3.39)
   (so that $r \cdot_A P = rP - (1 - r)A$).

We let $V_A$ be the set $V$ equipped with these new operations of addition and multiplication. Prove that

1. $V_A$ is a vector space over $\mathbb{R}$.
2. The origin of $V_A$ is $A$.