# Vector Calculus

Jacob Denson

November 1, 2014

# Chapter 1

# Prelude

On a map, we identify a unique position by a set of coordinates, a lattitude and longitude. Together, the set of all these coordinates form a two dimensional vector space of coordinates which can describe the surface of the earth. These coordinates allow us to make precise quantitative notions of the geometry of space. Rene Descartes ingenious discovery of this method, known today as analytic geometry, revolutionized the fields of physics and mathematics when it was discovered in 1637. It cannot be argued that the fields of the Infinitismal Calculus and Linear algebra were directly influenced by analytic geometries ideas. The subject of this book is an attempt to reconnect these fields that share a common origin.

Human intuition suggest that the spatial properties of the real numbers have obvious analogues in three dimensional space. The aim of the subject of vector calculus is to make the intuition rigorous – that is, to generalize the notions of calculus on the number line to finite dimensional vector spaces over the real numbers. We establish vector calculus based on a prerequisite knowledge of this calculus, as well as an elementary knowledge of linear algebra.

We begin with a familiar definition, that is one of the most important definition of the book.

**Definition 1.** *For any positive interger $n$, the n-dimensional vector space $\mathbf{R}^n$ is defined to be the set of all $n$ tuples where each element of the tuple is a real number. In the language of set theory, a concise description is*

$$\mathbf{R}^n = \{(v_1, v_2, \ldots, v_n) : \forall i : v_i \in \mathbf{R}\}$$

*We also call $\mathbf{R}^n$ the cartesian space, and in this context we call a vector in $\mathbf{R}^n$ a point. In general, for a vector $v \in \mathbf{R}^n$ we denote by $v_i$ the i'th coordinate of the tuple. Think of a vector as a collection of various magnitudes, and a point as the specific location in a space.*

In physics, vectors in cartesian space are used to identify relative differences between various quantities describing objects. Given two vectors $v$ and $w$, we

can then identify a third vector $v + w$, which is identified as the object whose relative differences to the object which $v$ describes is $w$, and vice versa for $w$ to $v$. Of course, if $v$ and $w$ are identified with their cartesian coordinates, $v + w$'s coordinates are just the sum of each corresponding coordinate in $v$ and $w$. This leads to the definition of the summand operation in a vector space $\mathbf{R}^n$.

**Definition 2.** *For two vectors $v$ and $w$, we define the sum of the two vectors,*

$$v + w = (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n)$$
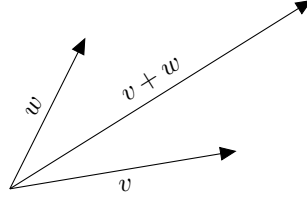


In addition to summing vectors of relative magnitudes, a physicist 'scales' a vector by a number $\lambda$ by identifying a new object whose relative magnitude is scaled propertionally by that number $\lambda$.

**Definition 3.** *Given a vector $v$ and scalar $\lambda$, define by $\lambda v$ the equation*

$$\lambda v = (\lambda v_1, \lambda v_2, \ldots, \lambda v_n)$$

Given these definitions, the set $\mathbf{R}^n$ becomes an $n$ dimensional vector space over the real number field, hence the defined name. The canonical algebraic basis is the set of vectors $e_1, e_2, \ldots, e_n$, defined in the following form:

$$e_i = \underbrace{(0, \ldots, 0, 1, 0, \ldots, 0)}_{\text{Where 1 is in the i'th coordinate}}$$

In this text we shall assume algebraic knowledge of finite dimensional vector spaces. Assuming this elementary understanding, we will extend our understanding to obtain the analytical properties which have an affinity with the methods of the infinitismal calculus.

# Chapter 2

# Analytical properties of $\mathbf{R}^n$

The most basic analytic property is distance, with which we can define almost every other property in the book. In general, arbitrary vector spaces have no notion of distance, thus we must make use of the specific properties of $\mathbf{R}^n$ to define distance on the space. What is the property that allows us to add distance to $\mathbf{R}^n$? The key observation is the following non-rigorous argument, which motivates the definition which preludes the argument:

**Definition 4.** *The* **length***, or* **euclidean norm** *(shortened to just the norm), of a point $v$ in a cartesian space $\mathbf{R}^n$, is denoted $\|v\|$, and defined by*

$$\|v\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

*where the square root is the positive root of the sum of squares. The* **distance** *between two vectors $v$ and $w$ is $\|v-w\|$. Thus the length of a vector $v$ is precisely its distance from the zero vector.*

*Proof.* Let $v$ be an arbitrary point in euclidean space. Suppose the space is one dimensional, which is precisely the real numbers. Then we have that the definition of $\|v\|$ is exactly $|v|$. This intuitively is the 'length' of a number in $\mathbf{R}$. If $v$ is defined in $\mathbf{R}^2$, $\|v\|$ follows from pythagoras' theorem. That is, the length of a the hypotenuse of a triangle is the square root of the sum of the squares of the two sides. Here the hypotenuse is the vector $v$, and the sides are the components of the vector: $v_1$ and $v_2$.

Now suppose by induction that this makes sense in $n-1$ dimensions. Then we can extend the distance to $n$ dimensions in the following manner. Take an arbitrary vector $v$. Then we can consider this vector as the hypotenuse of a triangle in two dimensions. where the first side is $(v_1, v_2, \ldots, v_{n-1}, 0)$ and the second is $(0, 0, \ldots, v_n)$. We can consider the first side of the triangle to be in $n-1$ dimensional space, and the second to be in one dimension. The length of the first side of the triangle in $n-1$ dimensional space is, by induction,

$$\sqrt{v_1^2 + v_2^2 + \cdots + v_{n-1}^2}$$

and the second's length is $\sqrt{v_n^2}$. Then, by Pythagoras' theorem, we obtain that the length of the whole vector is

$$\sqrt{\left(\sqrt{v_1^2 + v_2^2 + \cdots + v_{n-1}^2}\right)^2 + \left(\sqrt{v_n^2}\right)^2} = \sqrt{v_1^2 + v_2^2 + \cdots + v_{n-1}^2 + v_n^2}$$

Thus the notion of a norm makes sense in $n$ dimensions. By induction, we obtain the definition in all dimensions. $\square$

The proof above is inrigorous, relying on ideas that rely on knowledge of the real world. In mathematics this strategy is imprecise and considered unreliable, hence the proof makes little sense. It is included only to provide a reason why we define axiomatically the length of a point in this way. Though it is possible to connect plane geometry to the algebra of vector spaces precisely, the ideas used are quite involved, hence it is left for another time.

We require the length of a vector to be real valued as we need to be able to compare the length of two numbers. If we were to use the complex numbers, we would lose this property of order. Fortunately, our definition of length, defined via the sum of squares, is non-negative, and thus has a square root in the real numbers, so our operation is well-defined.

Some immediate and geometrically obvious properties now result from the definition, showing our definition connects well to the notion of length in the real world.

**Lemma 2.1.** *For every vector $v \in \mathbf{R}^n$, the length of a vector is greater than or equal to zero, and is length zero only when the vector itself is the zero vector.*

*Proof.* Our definition takes the square root that is non-negative in $\mathbf{R}$, hence our norm must be non-negative. If, for some vector $v \in \mathbf{R}^n$, $\|v\| = 0$, then $v_1^2 + \cdots + v_n^2 = 0$, and as this is the sum of squares, which all have non-negative values, we must have, for all $v_i$, $v_i^2 = 0$. This is true if and only if for all $v_i = 0$, hence $v = 0$. The converse is trivial. $\square$
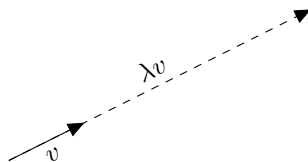
**Lemma 2.2.** *For every vector $v \in \mathbf{R}^n$, and every scalar $\lambda \in \mathbf{R}$, $\|\lambda v\| = |\lambda| \, \|v\|$*

*Proof.* Because then $\lambda v = (\lambda v_1, \lambda v_2, \ldots, \lambda v_n)$, and thus it follows that

$$\|\lambda v\| = \sqrt{(\lambda v_1)^2 + (\lambda v_2)^2 + \cdots + (\lambda v_n)^2}$$
$$= \sqrt{\lambda^2(v_1^2 + \cdots + v_n^2)}$$
$$= \sqrt{\lambda^2}\sqrt{v_1^2 + \cdots + v_n^2}$$
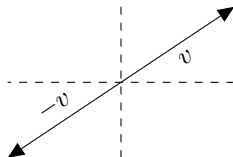$$= |\lambda| \, \|v\|$$

$\square$

Intuitively, if we scale a vector by the number, we should scale the length by the same value proportionally.

**Corollary 2.3.** *For every vector $v$, $\|v\| = \|-v\|$*

*Proof.* The proof is trivial, and left as an exercise. □

The corollary means that, if we mirror a vector about the x and y axis, the size of a vector stays the same.



**Definition 5.** *Given two vectors $v$ and $w$ in the same vector space $\mathbf{R}^n$, we define the **inner product** of the two vectors $v$ and $w$ to be $\langle v, w \rangle$, to be defined*

$$\langle v, w \rangle = v_1 w_1 + v_2 w_2 + \ldots v_n w_n = \sum_{k=1}^{n} v_k w_k$$

What is the inner product? We need to develop the definition of inner products in order to provide a satisfying answer.

**Lemma 2.4** (Symmetry)**.** *For two vectors $v$ and $w$,*

$$\langle v, w \rangle = \langle w, v \rangle$$

*Proof.* The main idea of the proof involves a simple use of commutativity of real numbers.

$$\langle v, w \rangle = \sum_{k=1}^{n} v_k w_k$$
$$= \sum_{k=1}^{n} w_k v_k$$
$$= \langle w, v \rangle$$

□

**Lemma 2.5** (Linearity)**.** *For two vectors $v$ and $w$, and a scalar $\lambda$,*

$$\langle \lambda v, w \rangle = \langle v, \lambda w \rangle = \lambda \langle v, w \rangle$$

5

*Proof.* We get the proof by simply factoring out the scalar $\lambda$ from the equation.

$$\langle \lambda v, w \rangle = \sum_{k=1}^{n} v_k(\lambda w_k)$$
$$= \lambda \sum_{k=1}^{n} v_k w_k{}'$$
$$= \lambda \langle v, w \rangle$$

We obtain another case by noting $\langle v, \lambda w \rangle = \langle \lambda w, v \rangle = \lambda \langle w, v \rangle = \lambda \langle v, w \rangle$. $\quad\square$

First we define another geometric property, angles. The definition results because of the correspondence with the definition of angles $\mathbf{R}^2$: plane geometry.

**Definition 6.** *The* **angle** $\theta$ *between two vectors $v$ and $w$ is the unique angle between 0 and $\pi$ such that*

$$\cos(\theta) = \frac{\langle v, w \rangle}{\|v\|\|w\|}$$

*This definition corresponds to the usual definition in plane geometry.*

*Proof.* Take two arbitrary vectors $v$ and $w$. If we identify $v$ with a line $OA$, and $w$ with a line $OB$, each line with a length equal to the magnitude of each vector, then we obtain the triangle $OAB$ has sides corresponding to $v$, $w$, and $v - w$. The angle $\theta$ between $v$ and $w$ is thus the angle between $OA$ and $OB$. The law of cosines applied to the vectors tells us that

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\|\|w\|\cos(\theta)$$

Noting that for any vector $u$, $\|u\|^2 = \langle u, u \rangle$, we obtain that

$$\|v - w\|^2 = \langle v - w, v - w \rangle$$
$$= \langle v, v \rangle - 2\langle v, w \rangle + \langle w, w \rangle$$
$$= \|v\|^2 - 2\langle v, w \rangle + \|w\|^2$$

Rearranging the previous equation and substituting our new value of $\|v - w\|^2$, we conclude that $\theta$ is the angle defined by the equation

$$\cos(\theta) = \frac{\|v\|^2 + \|w\|^2 - \|v - w\|^2}{2\|v\|\|w\|}$$
$$= \frac{\|v\|^2 + \|w\|^2 - \|v\|^2 + 2\langle v, w \rangle - \|w\|^2}{2\|v\|\|w\|}$$
$$= \frac{\langle v, w \rangle}{\|v\|\|w\|}$$

$$\square$$

Is this angle well defined for all vectors? For any angle $\theta$ that we consider,

$$-1 \leq \cos(\theta) \leq 1$$

Thus the angle between two vectors in $\mathbf{R}^n$ is only well defined if $|\langle v, w \rangle| / \|v\| \|w\|| \leq 1$. The following theorem shows this holds for any vectors in $\mathbf{R}\, R^n$.

**Theorem 2.6** (The Cauchy-Schwarz-Bunyakovsky-Hölder Inequality)**.** *For any two vectors $v, w \in \mathbf{R}^n$, $|\langle v, w \rangle| \leq \|v\| \|w\|$, hence $|\langle v, w \rangle| / \|v\| \|w\| \leq 1$, and the angle between the two vectors is well defined. In addition, $\langle v, w \rangle = \|v\| \|w\|$ if and only if $v$ and $w$ are linearly dependent.*

*Proof.* If $v$ and $w$ are linearly dependent, equality holds by the following calculation. Let $\lambda$ be the scalar such that $w = \lambda v$.

$$\begin{aligned}
\langle v, w \rangle &= \langle v, \lambda v \rangle \\
&= \lambda \langle v, v \rangle \\
&= \lambda \|v\|^2 \\
&= \|v\| \|\lambda v\|
\end{aligned}$$

If $v$ and $w$ are not linearly dependent, $v - \lambda w \neq 0$ for any scalar $\lambda$ (this is precisely the converse of linear dependence). It then follows that $\|v - \lambda w\|^2 > 0$ for all $\lambda$. Expand what this value means, by algebraic manipulations, obtaining that

$$\begin{aligned}
\|v - \lambda w\|^2 &= \langle v - \lambda w, v - \lambda w \rangle \\
&= \|v\|^2 - 2\lambda \langle v, w \rangle + \lambda^2 \|w\|^2
\end{aligned}$$

This can be considered a quadratic function of $\lambda$ with no real solutions. Hence the discriminant is negative. That is,

$$(2\langle v, w \rangle)^2 - 4\|v\|^2 \|w\|^2 < 0$$

Rearranging the equation, we obtain that

$$\langle v, w \rangle^2 < \|v\|^2 \|w\|^2$$

Hence

$$|\langle v, w \rangle| < \underbrace{|\|v\| \|w\||}_{\text{As both are positive}} = \|v\| \|w\|$$

$\square$

**Corollary 2.7.** *The angle between two linearly independant vectors is of magnitude 0.*

*Proof.* When two vectors $v$ and $w$ are linearly independant, from the above inequality, we know for the angle $\theta$ between them,

$$\cos(\theta) = \langle v, w \rangle / \|v\| \|w\| = \|v\| \|w\| / \|v\| \|w\| = 1$$

hence $\theta = 0$. $\square$

**Corollary 2.8.** *Two vectors are at right angles or orthogonal to one another, if and only if the inner product between them is 0.*

*Proof.* Let $v$ and $w$ be two vectors such that the angle $\theta$ between them is a right angle. It follows that $\cos(\theta) = 0$. By definition of the angle, we then know that $\langle v, w \rangle / \|v\|\|w\| = 0$, hence $\langle v, w \rangle = 0$. Conversely, if $\langle v, w \rangle = 0$, we know that $\cos(\theta) = 0$, which happens if and only if $\theta$ is a right angle. $\square$

Now we can intuitively explain the inner product. Take two vectors $v$ and $w$. Project vector $v$ onto vector $w$. What is projection? Precisely, take the vector $u \in \operatorname{span}(w)$ such that $v - u$ is orthogonal to $w$. Scale $w$ by the length of the projection $u$, letting $w' = \|u\|w$. Then $\langle v, w \rangle = \|w'\| = \|u\|\|w\|$. This follows as if $v - u$ is at a right angle to $w$, $\langle v - u, w \rangle = 0$. As $u \in \operatorname{span}(w)$, $u = \lambda w$ for some scalar $\lambda \in \mathbf{R}$. Then by calculation,

$$\langle v - u, w \rangle = \langle v - \lambda w, w \rangle$$
$$= \langle v, w \rangle - \lambda \langle w, w \rangle$$
$$= \langle v, w \rangle - \lambda \|w\|^2$$

Hence $\lambda = \langle v, w \rangle / \|w\|^2$. It follows that $\|u\| = \lambda \|w\| = \langle v, w \rangle / \|w\|$, so $\|u\|\|w\| = \langle v, w \rangle$.
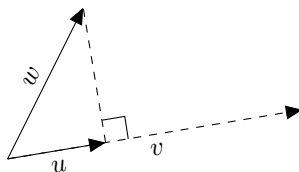


If you understood the above paragraph, you should see that the Cauchy Schwarz inequality is then intuitively true. The length of the projection of one vector is always less than or equal to the vector itself. Since the inner product is the multiplication of the length of this projection by another vector, it is obvious that this is less than the original vector's length multiplied by the other vector.

The most important vector inequality results naturally from the Cauchy-Scharz-Inequality. We know it as the Triangle-Inequality, as should become clear once the statement of the proof is understood.
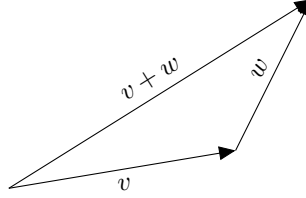
**Theorem 2.9** (The Triangle Inequality). *For any vectors $v$ and $w$,*

$$\|v + w\| \leq \|v\| + \|w\|$$

*Proof.* Let $v$ and $w$ be arbitrary vectors. We prove the statement by a simple calculation.

$$\|v + w\|^2 = \langle v + w, v + w \rangle$$
$$= \|v\|^2 + 2\langle v, w \rangle + \|w\|^2$$
$$\underbrace{\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2}_{\text{By Cauchy-Schwarz}}$$
$$= (\|v\| + \|w\|)^2$$

Hence $\|\|v + w\|\| \leq \|\|v\| + \|w\|\|$, but as both are always non-negative, we obtain the inequality needed. $\square$



The Triangle inequality states that, if we want to go from a point $a$ to a point $b$. The direct route is always less distance than some other route.

**Corollary 2.10.** *For any vectors $v$ and $w$, and for any third vector $u$, we have that $\|v - w\| \leq \|v - u\| + \|u - w\|$.*

*Proof.* This follows as $\|v - w\| = \|(v - u) + (u - w)\|$, which by the triangle inequality, is less than or equal to $\|v - u\| + \|u - w\|$. $\square$

**Corollary 2.11.** *For any vector $v = (v_1, v_2, \ldots, v_n)$ in $\mathbf{R}^n$,*

$$\|v\| \leq \sum_{k=1}^{n} |v_k|$$

*Proof.* $\|v\| = \|\sum_{k=1}^{n} v_k e_k\| \leq \sum_{k=1}^{n} \|v_k e_k\| = \sum_{k=1}^{n} |v_k|$ $\square$

**Lemma 2.12.** *For any vector $v$ with a coordinate $v_i$,*

$$|v_i| \leq \|v\|$$

*Proof.* If $x \leq y$, $\sqrt{x} \leq \sqrt{y}$. As $v_i^2 \leq \sum_{k=1}^{n} v_k^2$ ($v_i$ is in the sum), we know that

$$\sqrt{v_i^2} \leq \sqrt{\sum_{k=1}^{n} v_k^2}$$

and hence $|v_i| \leq \|v\|$. $\square$

These definitions justify the geometric properties of vector spaces $\mathbf{R}^n$. Withit, we can analyse almost all of Euclid analytically. However, we are severely limited by specifying only equalities and inequalities, with which we can only analyze finite sets of points in detail. To extend our notions to the precedence that calculus requires, we require precise notions of geometry to infinite sets. We call these properties developed topological properties.

# Chapter 3

# Topology & Sequences in $\mathbf{R}^n$

In order to justify the analytical properties in the last chapter, we used our intuition of polyhedra such as triangles to justify definition. It thus makes sense that we can define infinite properties in turns of shapes which cannot be defined by finitely many staight lines, the most basic of which is a circle.

**Definition 7.** *Given a point $x$ in $\mathbf{R}^n$, and a positive real number $r$, define the* **open ball** *centred at $x$ with radius $r$, denoted $B(x, r)$, as the set*

$$B(x, r) = \{y \in \mathbf{R}^n : \|x - y\| < r\}$$

*The* **closed ball***, denoted by $\overline{B}(x, r)$ is defined by*

$$\overline{B}(x, r) = \{y \in \mathbf{R}^n : \|x - y\| \leq r\}$$

*The* **circle***, denoted by $S(x, r)$, is defined by*

$$S(x, r) = \{y \in \mathbf{R}^n : \|x - y\| = r\}$$

*The* **punctured ball***, denoted by $\mathring{B}(x, r)$, is the set $B(x, r) - \{x\}$*

Calculus in $\mathbf{R}$ starts with defining properties of sequences. In $\mathbf{R}^n$, this is no different.

**Definition 8.** *Let $(a_i)$ be a sequence of points in $\mathbf{R}^n$. We say that $(a_i)$ converges to a point $a$ in $\mathbf{R}^n$, written $a_i \to a$, or $\lim_{i \to \infty} a_i = a$, if any of the following equivalent statements hold.*

1. *Every ball centered at the point $a$ contains a tail of the sequence $(a_i)$.*

2. *The sequence defined by $\|a_i - a\|$ converges to 0 in the real numbers.*

3. *Every coordinate sequence $([a_i]_k)$ of $a_i$ converges to $a_k$.*

*If a sequence does not converge to any point in $\mathbf{R}^n$, we say the sequence diverges.*

*Proof.* We prove multiple implications that map out a web of equivalences of the definitions. We leave it to the reader to show the proof provides all the implications needed.

- (1) $\implies$ (2): Let $(a_i)$ be a sequence such that every ball $B(a, r)$ contains a tail $(a_i)_{i \geq k}$ for some $k$. Consider the sequence $\|a_i - a\|$. To show this converges to 0, we must use the calculus of the real numbers. Let $\varepsilon > 0$. Then, by considering $B(a, \varepsilon)$, we gain a tail such that $\|a_i - a\| < \varepsilon$, which directly implies that the limit converges.

- (2) $\implies$ (1): Let $(a_i)$ be a sequence such that $\|a_i - a\| \to 0$. Then, for every $\varepsilon > 0$, there is a tail for some integer $k$ such that for any value in $(a_i)_{k \geq i}$, $\|a_i - a\| < \varepsilon$. This means precisely that the tail $(a_i)_{k \geq i}$ is contained in $B(a, \varepsilon)$. As this statement holds for every $\varepsilon$, it holds for any open ball centered at $a$, and thus we obtain (1).

- (2) $\implies$ (3): Let $a_i$ be a sequence such that $\|a_i - a\| \to 0$. By lemma (2.12), we know that for any coordinate $x_k$, $|x_k| \leq \|x\|$. Thus it follows that for any coordinate $(a_i)_k$,

$$|(a_i)_k - a_k| \leq \|a_i - a\|$$

As the second sequence dominates the first, and are both positive sequences, $|(a_i)_k - a_k| \to 0$, which means precisely that $(a_i)_k \to a_k$. This works for arbitrary coordinates, so we obtain (3).

- (3) $\implies$ (2): By Corollary (2.11), $\|x\| \leq \sum_{k=1}^n |x_k|$ for any $x \in \mathbf{R}^n$. We use the same strategy as in the last paragraph. Suppose $|a_{i_k} - a_k| \to 0$ for every $k$. We then know that the finite sum of sequences $(\sum_{k=1}^n |(a_i)_k|)$ converges to 0 also. But this is a dominating sequence of $\|a_k - a\|$, hence $\|a_k - a\| \to 0$, and we obtain (2).

$\square$

Some elementary results from real-valued calculus result from the fact that the coordinates must converge.

**Lemma 3.1.** *If $(v_i)$ and $(w_i)$ are two sequences in $\mathbf{R}^n$ such that $v_i \to v$, $w_i \to w$, then the sequence $(v_i + w_i)$ converges to $v + w$, the sequence $(\lambda v_i)$ converges to $\lambda v$, $(\langle v_i, w_i \rangle) \to \langle v, w \rangle$, and $(\|v_i\|) \to \|v\|$.*

*Proof.* For then the addition and scalar multiplication of every coordinate sequence converges to the coorresponding coordinate needed by application of the calculus of $\mathbf{R}$. $\square$

Now we have stated the meaning of sequences, we can define one of the first relatively deep theorems of vector calculus, an extension of the Bolzano Weirstraß theorem for $\mathbf{R}$. We assume the result in $\mathbf{R}$ to prove our theorem.

**Theorem 3.2** (The Bolzano Weirstraß Theorem in $\mathbf{R}^k$). *Every sequence in $R^n$ contains a convergent subsequence.*

*Proof.* Consider a sequence $(a_i)$ in $R_n$. Define a new sequence in $\mathbf{R}$ $((a_i)_1)$ by taking the first coordinate of every point in the sequence. By the Bolzano Weirstraß theorem in $\mathbf{R}$, we know that there is a convergent subsequence $((a_{n_i})_k)$, a sequence such that $|(a_{n_i})| \to a_1$ for some real value $a_1$, thus we have a sequence $(a_{n_i})$ that converges in the first coordinate. Now, suppose we have a sequence $(a_i)$ that converges in the first through $n-1$'th coordinate. Consider the sequence $|(a_i)_n|$. We then have a subseqence that converges also in the $n$'th coordinate. It follows that we can define a subsequence that converges in every coordinate and thus converges in $\mathbf{R}^n$. $\square$

**Theorem 3.3** (Cauchy's Theorem in $\mathbf{R}^k$). *If a sequence $(a_i)$ satisfies 'Cauchy's Criterion', then it converges. Cauchy Criterion is that, for any $\varepsilon$, there exists a positive integer $N$ such that for all elements $a$ and $b$ in the tail $(a_i)_{i \geq N}$,*

$$\|a - b\| < \varepsilon$$

*Proof.* Suppose $(a_i)$ is a sequence satisfying the property. Let $\varepsilon$ be arbitrary, with corresponding tail $(a_i)_{i \geq N}$. Since $|x_i| \leq \|x\|$ for any coordinate $x_i$ of a vector $x$ by lemma (2.12), we have that for any vectors $a$ and $b$ in the tail, $|a_k - b_k| \leq \|a - b\| < \varepsilon$. Hence by Cauchy's theorem in $\mathbf{R}$, the coordinates converge. It follows that the entire vector sequence converges. $\square$

**Definition 9.** *A point $x$ is a limit or accumulation point of a set $A$ if there exists a sequence $(a_i)$ such that every element is in the set $A$, and $a_i \to x$. Equivalently, a point $x$ is a limit point if every ball around $x$ contains some point in $A$.*

**Definition 10.** *Given a set $A \subset \mathbf{R}^n$, the* **closure** *is defined to be the set*

$$\overline{A} = \{x \in \mathbf{R}^n : x \text{ is an accumulation point of } A\}$$

*We say a set is* **closed** *if $\overline{A} = A$.*

**Lemma 3.4.** *For any set $A$, $\overline{A}$ is closed*

*Proof.* Proving $\overline{\overline{A}} = \overline{A}$ is equivilent to showing every accumulation point of $\overline{A}$ is an accumulation point of $A$. Let $a$ be an accumulation point of $\overline{A}$, so that there exists a sequence $(a_i)$ with every element $a_i$ in $\overline{A}$ such that $a_i \to a$. Define a new sequence $(a_i')$ in $A$ as follows. Let $a_k'$ be an arbitrary element such that $\|a_k - a_k'\| < 1/k$. This is always possible as $a_k$ is a limit point of $A$, and thus every ball around $a_k$ contains some point in $A$. We claim $a_i' \to a$. Let $\varepsilon > 0$ be arbitrary. Let $M$ be the integer such that $(a_i)_{i \geq M}$ is contained in $B(a, \varepsilon/2)$. Consider the tail $(a_i')_{i \geq \max(M, 1/2\varepsilon)}$. Then, for any $a_i'$ in the tail, $\|a_i' - a_i\| \leq \varepsilon/2$, and since $a_i$ is in the specified by $M$, $\|a_i - a\| < \varepsilon/2$ we use the triangle inequality to conclude that

$$\|a_i' - a\| \leq \|a_i' - a_i\| + \|a_i - a\| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$$

Hence $a_i' \to a$, and thus $a$ is a limit point of $A$. It follows that $\overline{A}$ is closed. $\square$

**Exercise 1.** $\overline{\mathbf{Q}} = \mathbf{R}$

**Exercise 2.** *Every closed ball is closed (hence the name makes sense)*

**Definition 11.** *A point $x$ is on the **boundary** of a set $A$ if $x$ is a limit point of $A$ and $A^c$. The set of all boundary points of a set $A$ is denoted $\partial A$. The interior of $A$ is $A^\circ = A - \delta A$.*

**Exercise 3.** $\partial B(c, r) = S(c, r)$

**Exercise 4.** $\partial A = \partial A^c$

**Exercise 5.** $\overline{A} = A \cup \partial A$, *hence a closed set is precisely one that contains its boundary.*

**Definition 12.** *A set $A$ is **open** if $\partial A$ is disjoint from $A$, which is exactly that $A^\circ = A$.*

**Lemma 3.5** (The Open Set Test). *A set $A$ is open if and only if for every point $a \in A$, there is a ball $B(a, r)$ which is a subset of $A$.*

*Proof.* We prove by contraposition. Let $A$ be an arbitrary set. Suppose there is a point $a \in A$ such that every ball $B(a, r)$ contains points in $A^c$. Then $a \in \partial A$, as it is a limit point of $A^c$. We conclude that $A$ is not open, as it contains parts of its boundary. The converse is just the proof in reverse. $\square$

**Theorem 3.6.** *A set $A$ is open if and only if $A^c$ is closed*

*Proof.* If $A^c$ is closed, $\partial A^c \subset A^c$. In exercise (4), it was proved that $\partial A^c = \partial A$, thus $\partial A \subset A^c$, and hence $\partial A \cap A = \emptyset$. If $A$ is open, $\partial A \cap A = \emptyset$, hence $\partial A \cap A^c = \partial A$ so that $\partial A \subset A^c$. Using exercise (4) again, it follows that $\partial A^c \subset A^c$, so that $A^c$ is closed by exercise (5). $\square$

In Mathematics, many groups of objects also have something called a 'dual' set, a group such that almost every theorem of the first group has a corresponding theorem with the second. The dual of a closed set is an open set, and thus we will see many theorems about closed sets have immediate corollaries about open sets, and vice versa.

**Theorem 3.7.** *Let $\mathcal{J}$ be an arbitrary index set, and $(A_j)_{j \in \mathcal{J}}$ a set of open sets. Then $\bigcup_{j \in \mathcal{J}} A_j$ is open.*

*Proof.* We prove by the open set test. Let $a$ be an arbitrary element in $\bigcup_{j \in \mathcal{J}} A_j$. Then there is some specific $A_k$ for which $a \in A_k$, and since this set is open, there is some ball $B(a, r)$ such that $B(a, r) \subset A_k$. As $A_k \subset \bigcup_{j \in \mathcal{J}} A_j$, the same ball must be contained in the union. Thus the union is open. $\square$

**Corollary 3.8.** *If $(A_j)_{j \in \mathcal{J}}$ is a set of closed sets, then $\bigcap_{j \in \mathcal{J}}$ is closed.*

*Proof.* For then $(A_j^c)_{j \in \mathcal{J}}$ is a family of open sets, and

$$\bigcap_{j \in \mathcal{J}} A_j = \left( \bigcup_{j \in \mathcal{J}} A_j^c \right)^c$$

$\square$

**Theorem 3.9.** *If $(A_1, A_2, \ldots, A_n)$ is a finite collection of open sets, then $\bigcap_{k=1}^n A_k$ is open.*

*Proof.* Let $a$ be in $\bigcap_{k=1}^n A_k$. Then $a$ is in every set $A_k$, and for each $A_k$ there is a ball $B(a, r_k)$ contained in $A_k$ as $A_k$ is open. Then, since $B(a, \min(r_1, r_2, \ldots, r_k))$ is a subset of every ball in $A_k$, it is contained in the intersection. Thus the intersection is open. $\square$

**Corollary 3.10.** *The finite union of closed sets is closed.*

It is not in general true that the arbitrary union of open sets is open, and the intersection of closed sets is closed. Take the set of $(B(0, r))_{r \in \mathbf{R}}$. Each of these sets is open, but the intersection is a single point 0, and is not open.

**Definition 13.** *A set $A$ is compact if every sequence in $A$ contains a convergent subsequence that converges to a point in $A$.*

**Definition 14.** *A set $A$ is bounded if, for some point $x$, there is a radius $r$ such that $A \subset B(x, r)$.*

We should specify that $A$ is bounded at the point $x$, but the point is arbitrary, by the lemma below.

**Lemma 3.11.** *A set $A$ in $\mathbf{R}^n$ which is bounded at some point $x$ in $\mathbf{R}^n$ is bounded at every point*

*Proof.* Let $A$ be a set which is bounded at a point $x$ in $\mathbf{R}^n$. Then there is a radius $r$ such that $A \subset B(x, r)$. Let $y$ be an arbitrary point. Take a new radius $r + \|v - w\|$, and consider the ball $B(y, r + \|v - w\|)$. Let $z$ be an arbitrary point in $B(x, r)$. Then $\|x - z\| < r$. By corollary (1.7), $\|y - z\| \le \|y - x\| + \|x - z\| = \|v - w\| + r$. Hence $y \in B(y, r + \|y - x\|)$, and thus $B(x, r) \subset B(y, r + \|y - x\|)$. By the transitive property of subsets, $A \subset B(y, r + \|y - r\|)$. It follows that $A$ is bounded at $y$, for any $y \in \mathbf{R}^n$. $\square$

**Theorem 3.12** (The Heine-Borel theorem (part 1) ). *A set is compact if and only if it is closed and bounded*

*Proof.* Let $A$ be a set that is compact. It is closed because any sequence in $A$ must converge in $A$ (the subsequence contains a convergent subsequence that must converge in $A$). Suppose a set $B$ is unbounded. Then we form a sequence $(b_i)$ that has no convergent subsequence as follows. Let $b_1$ be arbitrary. Given $(b_1, \ldots, b_{n-1})$, define $b_n$ to be a point such that $\|b_n - b_i\| > n$ for all $i$. This must be possible, as otherwise the set is bounded. No subsequence of $(b_i)$ can possibly converge by construction. Thus $B$ cannot be compact, showing this for all sets as $B$ was arbitrary, and then by contraposition, $A$ must be bounded. $\square$

**Definition 15.** *Let $A$ be a set. An open cover of $A$ is a collection $(A_j)_{j \in \mathcal{J}}$ of open sets such that $A$ is a subset of $\bigcup_{j \in \mathcal{J}} A_j$.*

An open cover does not need to be finite or even countable. However, some familiar sets have a property that we may always select finite amounts of a cover to cover the entire set. This shown below by the remaining part of the Heine Borel Theorem, one of the jewels of mathematical analysis.

**Theorem 3.13** (Heine-Borel Theorem (part 2)). *A set $A$ is compact if and only if every open cover of $A$ contains a finite subcover.*

*Proof.* Suppose $A$ is a set such that every open cover contains a finite subcover. Then $A$ is bounded, as the collection $\{B(0, r)\}_{r \in \mathbf{R}}$ forms an open cover of $A$, and thus must contain a finite subcover, in other words a minimum ball that contains $A$. We prove that $A^c$ is open by a similar strategy to above. Let $a$ be an arbitrary element in $A^c$. Consider the set of closed ball complements $\{(\overline{B}(a, r))^c\}_{r \in \mathbf{R}}$. The set of all these forms an open cover of $A$, and thus must contain a finite subcover. We can then take a ball that is the complement of the smallest radius complement in that set, and this ball centered at $a$ becomes a subset of $A^c$. Thus $A^c$ is open, so $A$ is closed. As $A$ is closed and bounded, $A$ is compact, by the first part of the Heine-Borel theorem.

Suppose $A$ is a compact subset of $\mathbf{R}^n$, and hence closed and bounded. As it is bounded, $A$ is contained in a ball. Every ball is contained in a cube, denoted $Q_0$. Let $l_0$ be the length of the diagonal of the cube. Suppose that there is a cover $\mathcal{C}$ of $A$ with no finite subcover. Divide $Q_0$ into $2^n$ subcubes. One of these must not have a finite subcover. Denote this cube $Q_1$. Continue defining these cubes by this method to form a chain $Q_0 \subset Q_1 \subset \dots$. The diagonal of cube $Q_k$ is $l_k = l_1/2^k$. Each cube $Q_k$ is non-empty, hence we may pick some $q_k$ from the cube to form a sequence $(q_k)$. We claim $q_k$ converges to some point $q$, proving the claim by Cauchy's criterion. Given $\varepsilon > 0$, pick an integer $M$ such that $M \geq \lg(l_1/\varepsilon)$. Then the tail $(q_k)_{k \geq M}$ is contained in the cube $Q_M$, whose diagonal $l_M = l_1/2_M$. As $M \geq \lg(l_1/\varepsilon)$, $2^M \geq l_1/\varepsilon$, hence $l_M = l_1/2^M \leq l_1\varepsilon/l_1 = \varepsilon$. It follows that for any points $q'$ and $q''$ in the sequence, $\|q' - q'\| \leq \varepsilon$, as the diagonal is the longest distance between any two points in the square. Thus the sequence converges to some point $q$. There is some open set $C$ in $\mathcal{C}$ such that $q \subset C$. Take the ball $B(q, r)$ such that this ball is contained in $C$. Then there is some square $Q_n$ which is contained in this ball, yet it contains no finite subcover. By contradiction, this cover $\mathcal{C}$ could not have existed. $\square$

The statement of a compact set in terms of open covers is often taken to be the definition of a compact set in most textbooks. We chose our original definition as it is most intuitive.

An example of a non-compact set is $(-1, 1)$. Take the cover $(-r, r)_{r \in (0, 1)}$. Then there is no finite subcover that covers $(-1, 1)$.

We have defined almost too many sets to remember. We need two more types to finish off the topological properties of $\mathbf{R}^n$.

16

Connectedness is an easy thing to see. A circle is connected, two separate circles are not. Like many intuitive concepts, connectedness becomes a very difficult concept to formalize.

**Definition 16.** *A set $A$ is* **disconnected** *if it can be partitioned into two sets $A_1$ and $A_2$, and there are two disjoint open sets $U_1$ and $U_2$ such that $A_1$ is contained in $U_1$ and $A_2$ is contained in $U_2$. A set is* **connected** *if it is not possible to disconnect it.*

Intuitively, a set is convex if, for any two points, the line between those two points remains in the set. How do we formalize this? Well, for any point $c$ between two points $a$ and $b$, $c = \lambda a + (1 - \lambda)b$ for some value $\lambda \in [0, 1]$. This motivates the following definition.

**Definition 17.** *A set $A$ is convex, if, for any two points $a$ and $b$ in the set, and for any scalar $\lambda \in [0, 1]$ $\lambda a + (1 - \lambda)b \in A$.*

What is nice about this definition is it involves no notion of distance. We can consider this for all vector spaces over the real numbers.

**Exercise 6.** *Any ball is convex*

**Theorem 3.14.** *Any convex set is connected*

*Proof.* We prove by contraposition. Suppose a set $A$ is disconnected. Then $A = A_1 \cup A_2$ for two disjoint sets $A_1$ and $A_2$, where there are open sets $U_1$ and $U_2$ such that $A_1 \subset U_1$, $A_2 \subset U_2$. Take two points $a_1 \in A_1$, $a_2 \in A_2$. We claim there is a point on the line between $a_1$ and $a_2$ that is not contained in $A$. Take the supremum of the set $\{\lambda \in [0, 1] : \lambda a_1 + (1 - \lambda)a_2 \in A_1\}$, and denote it $\lambda'$. We claim $x = \lambda' a_1 + (1 - \lambda')a_2$ cannot be an element of $A_1$. If it was, it is contained in an open set $U_1$, and hence there is a ball $B(x, r)$ which is contained in $U_1$, and hence not in $A_2$. Take the value $\lambda = min(\lambda' + r/2\|a_1\|, \lambda + r/2\|a_2\|, 1)$. Then,

$$\|\lambda a_1 + (1 - \lambda)a_2 - x\| = \|(\lambda - \lambda')a_1 + (\lambda' - \lambda)a_2\|$$
$$\leq |\lambda - \lambda'|\|a_1\| + |\lambda' - \lambda|\|a_2\|$$
$$\leq |r/2\|a_1\|\|\|a_1\| + |r/2\|a_2\|\|\|a_2\|$$
$$= r$$

Thus $\lambda'$ is not the supremum, and by contradiction. It cannot be in $A_1$. For similar reasons, it also cannot be in $A_2$, hence it is not in $A$ and thus the set is not convex. $\qquad \square$

**Corollary 3.15.** *Balls are connected*

*Proof.* We leave a proof as an exercise $\qquad \square$

We introduce a final topological notion before we can start studying the familiar notions of calculus.

**Definition 18.** *Given a subset $A$ of $\mathbf{R}^n$, we say $c$ is an accumulation or cluster point if $c \in \overline{A - \{c\}}$. This is equivalent to the notion that, for any radius $r$, $\mathring{B}(c, r) \cap A \neq \emptyset$.*

# Chapter 4

# Functions and Continuity

You should have enough experience in mathematics to know the formal definition of a function. In this book, we deal with functions between real valued vector spaces, that is, functions $f : D \to \mathbf{R}^m$, such that $D \subset \mathbf{R}^n$. We can consider this function to be a function of $n$ variables, mapped to $m$ variables. This is really just the composition of $m$ functions $f_i : \mathbf{R}^n \to \mathbf{R}$, defined by the equation

$$f(v) = (f_1(v), f_2(v), \ldots, f_m(v))$$

**Definition 19.** *Consider a function $f : D \to \mathbf{R}^m$, where $D \subset \mathbf{R}^n$ and $c$ be an accumulation point of $D$. We say that $f$ approaches a vector $w \in \mathbf{R}^m$, and we write $\lim_{v \to c} = w$, if any one of the equivalent notions is defined.*

1. *For every $\varepsilon > 0$, there exists a number $\delta > 0$ such that, for any vector $v \in D$, $0 < \|v - c\| < \delta$, $\|w - f(v)\| < \varepsilon$.*

2. *For every open ball $B(w, \varepsilon)$, there exists a punctured ball $\mathring{B}(c, \delta)$ such that $f(\mathring{B}(c, \delta)) \subset B(w, \varepsilon)$.*

3. *For every sequence $(v_i)$ in $D - \{c\}$ such that $v_i \to c$, $f(v_i) \to w$.*

*Proof.* The equivalence of (1) and (2) is obvious, found by expanding the definitions of (2). We prove the equivalence of (1) and (2) to (3) by proving the corresponding implications.

(1) $\implies$ (3). Suppose $\lim_{v \to c} f(x) = w$ in the sense of the first definition. Take a sequence $(a_i)$ such that $y_i \to c$. Fix any $\varepsilon > 0$. There is some $\delta$ such that if $0 < \|v - c\| < \delta$, $\|f(v) - w\| < \epsilon$. There is some integer $M$ such that the tail $(a_i)_{i \geq M}$ is contained in $B(c, \delta)$. But then the tail $(f(a_i))_{i \geq M}$ is contained in the ball $B(w, \varepsilon)$, so that we get (3).

(3) $\implies$ (1). We prove by contraposition. Suppose $\lim_{v \to c} f(x) \neq w$. Then there is $\varepsilon > 0$ such that, for any $\delta$, there is $v$ such that $0 < \|v - c\| < \delta$ but $\|f(v) - w\| \geq \varepsilon$. Define a sequence $(v_i)$ such that $\|v_i - c\| < 1/i$, but $\|f(v_i) - w\| \geq \varepsilon$. Then $v_i \to c$, but $f(v_i) \not\to w$. By contraposition, we obtain that (3) implies (1). $\square$

**Corollary 4.1.** *The limit of a function is unique.*

*Proof.* Let $f$ be a function that converges at an accumulation point $c$ to $v$ and $w$. Then for every sequence $(a_i)$ such that $a_i \to c$, $a_i \to v$ and $a_i \to w$, hence $v = w$. But this only happens if there is a sequence with this property. This follows as $c$ is an accumulation point. $\square$

**Corollary 4.2.** *If $\lim_{v \to c} f(v) \to l$ and $\lim_{v \to c} f(v) \to m$, then*

$$\lim_{v \to c} (f \pm g)(x) = l \pm w$$

$$\lim_{v \to c} \langle f, g \rangle = \langle l, w \rangle$$

$$\lim_{v \to c} \|f\| = \|l\|$$

*Proof.* The proof follows by connection of a function to sequences; the argument is left to the reader. $\square$

**Definition 20.** *Suppose $D \subset \mathbf{R}^n$, and we have a function $f : D \to \mathbf{R}^m$. Then, for a point $c \in D$, we say that $f$ is continuous at $c$ if*

$$\lim_{v \to c} f(v) = f(c)$$

*We say that, for a subset $C \subset D$, $f$ is continuous on $C$ if $f$ is continuous at every point $c \in C$.*

Intuitively, continuity means we can draw the function without taking pen off paper. The limit of a function is precisely the point that make the function continuous. They also have many more interesting properties.

**Lemma 4.3.** *For any sequence $(v_i)$, such that $v_i \to v$ and any function $f$ continuous at $v$*

$$\lim_{i \to \infty} f(v_i) = f(\lim_{i \to \infty} v_i)$$

*Proof.* The proof follows from the definition of limits established previously. $\square$

**Lemma 4.4.** *If $f : D \to \mathbf{R}^m$ and $g : E \to \mathbf{R}^m$ are continuous at a point $c$, where $D$ and $E$ are subsets of $\mathbf{R}^n$ and $h : F \to \mathbf{R}^t$, where $F$ is a subset of $\mathbf{R}^m$ that is continuous at $g(c)$:*

1. *$f \pm g$ is continuous at $c$.*

2. *$\langle f, g \rangle$ is continuous at $c$.*

3. *$\|f\|$ is continuous at $c$.*

4. *Every component function of $f$ is continuous at $c$.*

5. *$g \circ f$ is continuous at $c$.*

**Theorem 4.5.** *A function $f : D \to \mathbf{R}^m$ where $D \subset R^n$ is continuous on $D$ if and only if, for every open set $C$ in $\mathbf{R}^m$, $f^{-1}(C)$ is open.*

*Proof.* Let $f$ be as above. We prove the statement for open balls, from which the entire theorem follows. Take a ball $B(x, r)$ for some point $x$ and some radius $r$. Then $f^{-1}(B(x, r))$ is defined to be the set

$$\{a \in D : \|f(a) - x\| < r\}$$

Since $f$ is continuous, there is $\delta$ such that $B(a, \delta) \subset f^{-1}(B(f(a), r - \|f(a) - x\|))$. This set is a subset of $B(x, r)$, as if $\|y - f(a)\| < r - \|f(a) - x\|$, $\|y - x\| \leq \|y - f(a)\| + \|x - f(a)\| < r - \|f(a) - x\| + \|f(a) - x\| = r$. It follows that the set $f^{-1}(B(x, r))$ is open.

Suppose for any open set $C$ in $\mathbf{R}^m$, $f^{-1}(C)$ is open. Then for any $\varepsilon > 0$, and for any point $x$, $B(f(x), \varepsilon)$ is open, hence $f^{-1}(B(f(x), \varepsilon))$ is open. As $x \in f^{-1}(B(f(x), \varepsilon))$, we know there is a ball $B(x, \delta)$ such that $f(B(x, \delta)) \subset B(f(x), \varepsilon)$. Thus we get continuity. $\square$

**Corollary 4.6.** *A function is continuous if and only if, for every closed set $C$ in $\mathbf{R}^m$, $f^{-1}(C)$ is closed.*

We obtain some simple practical applications from this theorem.

**Lemma 4.7.** *The set of solutions to the equation $\cos(x^2 + y^2) > 1/2$ is open.*

*Proof.* Let $f(x, y) = \cos(x^2 + y^2)$. This function is continuous as it is the composition of continuous functions. If $(x, y)$ is a solution to the inequality, this means exactly that $f(x, y) > 1/2$, which is true if and only if $f(x, y) \in f^{-1}((1/2, \infty))$. As this set is open, the inverse of that set is open. $\square$