Vector Calculus

Jacob Denson

November 3, 2014

Chapter 1

Prelude

> Any problem in geometry can easily be reduced to such terms that a knowledge of the lengths of certain straight lines is sufficient for its construction.
>
> Rene Descartes - 1637

On a map, we identify a unique position by a pair of coordinates, a lattitude and a longitude. Via this method, description of the earth's surface is made precise by a simple pair of numbers. We can easily extend this coordinate method to three dimensions, allowing precise, quantitative statements about the geometry of space to be made. Rene Descartes' ingenious discovery of this concise description, known today as analytic geometry, revolutionized the fields of physics and mathematics when it was discovered in 1637. Inspired by Descartes' ideas, the two mathematicians Sir Isaac Newton and Gottfried Leibnitz went on to discover the infinitismal calculus. Meanwhile, the study of analytic geometry was extended to arbitrary dimensions, with the new name of linear algebra. As calculus of a single variable focuses on the infinitismal properties of the analytic geometry of two dimensions, vector calculus realizes these properties in the arbitrary dimensional vector spaces of linear algebra.

Human intuition suggests that the spatial properties of the real numbers have obvious analogues in three dimensional space. The aim of this report is to create rigour to fill in gaps of reasoning that intuition often results in. Assuming the properties of differential and integral calculus in 'one dimensional space', we show how results from $\mathbf{R}$ naturally extend to arbitrary finite dimensional vector spaces with scalars in the real numbers.

We begin with what should be a familiar definition:

Definition 1. *For any positive integer n, the n-dimensional real 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, $\mathbf{R}^n$ is equal to*

$$\{(v_1, v_2, \ldots, v_n) : v_1, v_2, \ldots, v_n \in \mathbf{R}\}$$

Each element in the tuple is called a coordinate.

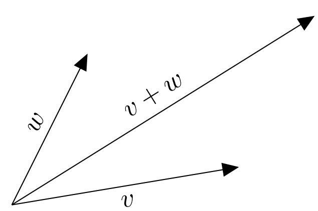
We also call $\mathbf{R}^n$ the Cartesian or Euclidean space, and in this context we call a vector in $\mathbf{R}^n$ a point.

When we call an element of $\mathbf{R}^n$ a vector, think of it as a collection of various magnitudes that an object posses, whereas think of a point as a geometric location in space: the tip of the corresponding vector.

In physics, vectors 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 whose relative difference to w is v. Of course, if v and w are identified with a vector of 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 the 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 λ by identifying a new object where each magnitude of that object is scaled proportionally by λ. We describe this mathematically as follows.

Definition 3. *Given a vector v and scalar λ, define by λv the equation*

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

Given the operations of addition and scalar multiplication, the set $\mathbf{R}^n$ becomes an n dimensional vector space over the field of real numbers. The canonical algebraic basis is the set of vectors $e_1, e_2, \ldots, e_n$, defined in the following form:

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

In this text, algebraic knowledge of $\mathbf{R}^n$ is assumed. We focus our efforts on discovering analytical properties which have an affinity with the methods of the infinitismal calculus. To be facecious, you should know vectors in $\mathbf{R}^n$ by associated equalities, and we shall now learn the interactions of inequalities.

Chapter 2

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

The introduction of numbers as
coordinates is an act of violence.

Hermann Weyl - 1927

Perhaps the most basic geometric property is distance. With it, we can define almost every other property in the book. What is important about distance? It gives a number to quantize a relationship between a pair of objects. 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 a geometric argument which follows the definition below.

Definition 4. *The* **length***, or* **euclidean norm** *of a vector v in $\mathbf{R}^n$, denoted $\|v\|$, is defined as*

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

where the square root is the positive root of the sum of squares. The **distance** *between two points x and y in $\mathbf{R}^n$ is $\|x - y\|$. 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 the hypotenuse of a triangle is the square root of the sum of the squares of the two sides. To establish this, let v be the hypotenuse of a triangle, and let the sides be the components of the vector: v_1 and v_2.

Now suppose that the definition of length has been established 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, the definition makes sense in an arbitrary cartesian space $\mathbf{R}^n$. $\square$

The proof above is not rigorous, relying on the ideas of plane geometry which are only connected to $\mathbf{R}^2$ by an intuitive study of analytic geometry. In mathematics, intuition is considered imprecise, and has resulted in unreliable proofs in the long run. Back when Calculus was invented, this style of proof was the norm, but in the 1800s, when the study of the calculus of vectors became mainstream, it was found that some ideas, thought intuitively right, resulted in deep errors in the foundations of calculus. It was only when Karl Weierstraß formalized the concepts of limits, and Bernhard Riemann created his formal definition of the integral, that these problems were sorted out. Because of the problems that result from the intuitive method, all proofs in modern mathematics must be based on a solid axiomatic footing. The proof above is included only as motivational reason why we axiomatically define the length of a vector as the root of a sum of squares as above.

It is easy to see that the length of an arbitrary vector is defined in $\mathbf{R}$. A sum of squares is always non-negative, and a non-negative number has a unique positive root in $\mathbf{R}$. We obviously need this in any definition of distance. Imagine the havoc that would be caused by a vector which has 'no distance' in comparison to other vectors, or multiple distances!

Some immediate geometrical properties arise from the definition of the euclidean norm above. The statements are obvious to intuitive statements in the real world. This just provides more evidence that our definition fits the intuitive sense of distance.

Lemma 2.1. *For every vector $v \in \mathbf{R}^n$, $\|v\| \geq 0$. The length of v is 0 only when the vector itself is the zero vector.*

Proof. Since our definition takes the unique non-negative square root of a number in $\mathbf{R}$, our norm must be non-negative. If, for some vector $v \in \mathbf{R}^n$, $\|v\| = 0$, then $v_1^2 + \cdots + v_n^2 = 0$. As this is the sum of squares, which all have non-negative values, we must have for all coordinates v_i, $v_i^2 = 0$. This is true if and only if $v_i = 0$ for all coordinates v_i, hence $v = 0$. $\square$

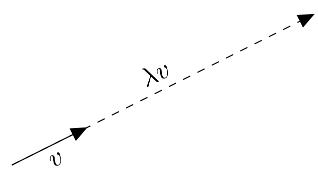
Lemma 2.2. *For every vector $v \in \mathbf{R}^n$, and every scalar $\lambda \in \mathbf{R}$, $\|\lambda v\| = |\lambda|\ \|v\|$.*

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

$$\begin{aligned}
\|\lambda v\| &= \sqrt{(\lambda v_1)^2 + (\lambda v_2)^2 + \cdots + (\lambda v_n)^2} \\
&= \sqrt{\lambda^2 (v_1^2 + \cdots + v_n^2)} \\
&= \sqrt{\lambda^2}\sqrt{v_1^2 + \cdots + v_n^2} \\
&= |\lambda|\ \|v\|
\end{aligned}$$

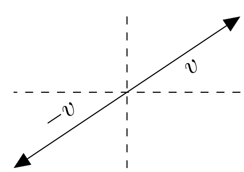
$\square$

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



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

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 vector space $\mathbf{R}^n$, we define the* **inner product** *of the two vectors v and w, denoted $\langle v, w\rangle$, as*

$$v_1 w_1 + v_2 w_2 + \cdots + v_n w_n = \sum_{k=1}^{n} v_k w_k$$

What is the inner product? In order to provide a satisfying answer, we must discover some resultant properties of the definition. For a little motivation of its use, see that for any vector v, $\|v\|^2 = \langle v, v\rangle$.

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

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

Proof. The idea of the proof rests on the commutativity of real numbers.

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

□

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

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

Proof. This proof rests on the distributive property of $\mathbf{R}$.

$$\begin{aligned}
\langle \lambda v, w \rangle &= \sum_{k=1}^{n} v_k (\lambda w_k) \\
&= \lambda \sum_{k=1}^{n} v_k w_k {}^{\text{‘}\prime} \\
&= \lambda \langle v, w \rangle
\end{aligned}$$

We obtain the reverse case by noting that $\langle v, \lambda w \rangle = \langle \lambda w, v \rangle = \lambda \langle w, v \rangle = \lambda \langle v, w \rangle$. □

Lemma 2.6 (Additivity)**.** *For three vectors v, w, and u,*

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

Proof.

$$\begin{aligned}
\langle v + u, w \rangle &= \sum_{k=1}^{n} (v_k + u_k)(w_k) \\
&= \sum_{k=1}^{n} v_k w_k + \sum_{k=1}^{n} u_k w_k \\
&= \langle v, w \rangle + \langle u, w \rangle
\end{aligned}$$

□

To see intuitively what the dot product is, we rigourously define another geometric property, angles. Like with distance, the definition results because of the correspondence with the definition of angles $\mathbf{R}^2$: plane geometry.

Definition 6. *The* **angle** θ *between two vectors* v *and* w *is the unique angle between 0 and* π *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. OAB has sides corresponding to v, w, and $v - w$. The angle θ 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

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

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

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

□

Are angles well defined for all vectors? If either of the vectors are zero, then there magnitude is zero, hence the angle is not well defined (a zero vector has no 'point' to define an angle on). For any angle θ,

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

Thus the angle between two vectors in $\mathbf{R}^n$ is only well defined if

$$\frac{|\langle v, w\rangle|}{\|v\|\|w\|} \leq 1$$

The following theorem shows this holds for any pair of non-zero vectors in $\mathbf{R}^n$.

Theorem 2.7 (The Cauchy-Schwarz-Bunyakovsky-Hölder Inequality)**.** *For any two vectors $v, w \in \mathbf{R}^n$,*

$$|\langle v, w\rangle| \leq \|v\|\|w\|$$

hence

$$\frac{|\langle v, w\rangle|}{\|v\|\|w\|} \leq 1$$

and thus 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 λ be the scalar such that $w = \lambda v$.

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

If v and w are not linearly dependent, $v - \lambda w \neq 0$ for any scalar λ (this is precisely the converse of linear dependence). It then follows that $\|v - \lambda w\|^2 > 0$ for all λ. Expanding what this value means, by algebraic manipulations, we obtain 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 λ with no solutions in $\mathbf{R}$. 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

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

□

Corollary 2.8. *The angle between two linearly independant vectors is of magnitude 0.*

Proof. When two vectors v and w are linearly independant, from the above equality, we know for the angle θ between them,

$$\cos(\theta) = \frac{\langle v, w \rangle}{\|v\|\|w\|} = \frac{\|v\|\|w\|}{\|v\|\|w\|} = 1$$

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

Corollary 2.9. *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 θ between them is a right angle. It follows that $\cos(\theta) = 0$. By definition of the angle, we then know that

$$\frac{\langle v, w \rangle}{\|v\|\|w\|} = 0$$

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 θ is a right angle. $\square$

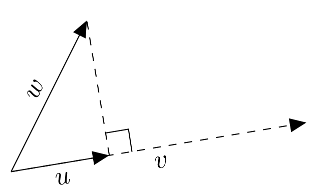
Now we can intuitively explain the inner product. Take two vectors v and w. Project vector v onto vector w. What is projection? Precisely, take the vector $u \in \operatorname{span}(w)$ such that $v - u$ is orthogonal to w. Scale w by the length of the projection u, letting $w' = \|u\|w$. Then $\langle v, w \rangle = \|w'\| = \|u\|\|w\|$. 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,

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

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

$$\|u\| = \lambda \|w\| = \frac{\langle v, w \rangle}{\|w\|}$$

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 a vector is always less than or equal to the length of 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.

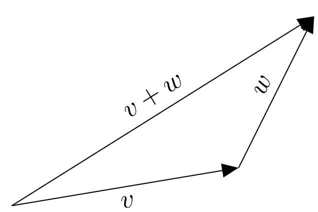
Theorem 2.10 (The Triangle Inequality). *For any vectors v and w,*

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

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

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

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



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 2.11. *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 $\|v - u\| + \|u - w\|$. □

Corollary 2.12. *For any vector $v = (v_1, v_2, \ldots, v_n)$ in $\mathbf{R}^n$,*

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

Proof.

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

□

Lemma 2.13. *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 $|v_i| \leq \|v\|$. □

These definitions justify the geometric properties of vector spaces $\mathbf{R}^n$. With it, we can analyse almost all of the arguments 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 in infinite sets. We call these properties developed topological properties.

Chapter 3

Topology 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*

$$\{y \in \mathbf{R}^n : \|x-y\| < r\}$$

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

$$\{y \in \mathbf{R}^n : \|x-y\| \leq r\}$$

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

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

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

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 ε, 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).

$\square$

We obtain some elementary results from facts from real-valued calculus result from the correspondence of definition (2) of coordinates with convergence of real numbers. We state the facts in $\mathbf{R}$ without proof, and assume the reader can prove the corresponding theorem in $\mathbf{R}^n$ using definition (2).

Theorem 3.1. *If (a_i) and (b_i) are two sequences in $\mathbf{R}$, and c is some fixed constant in $\mathbf{R}$, such that $a_i \to a$ and $b_i \to b$, then*

$$a_i + b_i \to a + b$$

$$\lambda a_i \to \lambda a$$

Corollary 3.2. *If (v_i) and (w_i) are two sequences in $\mathbf{R}^n$ such that $v_i \to v$, $w_i \to w$, then*

$$v_i + w_i \to v + w$$

$$\lambda v_i \to \lambda v$$

$$\|v_i\| = \|v\|$$

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 3.3 (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 $([a_{n_i}]_1)$, a sequence such that $[a_{n_i}]_1 \to a_1$ for some real value a_1. Now, suppose we have a sequence (a_i) that converges in the first through $n-1$'th coordinate. Consider the sequence $([a_i]_n)$. By another application of Bolzano Weirstraß in $\mathbf{R}$, we obtain a subseqence that converges in the n'th coordinate. It follows that we can define a subsequence that converges in every coordinate and thus converges in $\mathbf{R}^n$. □

Like with Bolzano Weirtstraß, it is not too difficult to extend Cauchy's theorem to $\mathbf{R}^k$ as well.

Theorem 3.4 (Cauchy's Theorem in $\mathbf{R}^k$)**.** *If a sequence (a_i) satisfies 'Cauchy's Criterion', then it converges. Cauchy's Criterion is that, for any ε, 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 ε 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. □

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 in the sequence 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 3.5. *For any set A, $\overline{A}$ is closed*

Proof. Proving $\overline{\overline{A}} = \overline{A}$ is equivalent 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 $(a_i')_{i \geq \max(M, 2/\varepsilon)}$. 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 ∂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 ∂A is disjoint from A, so $A^\circ = A$.*

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 $\partial A \cap A = \emptyset$, so A is open. 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.

Lemma 3.7 (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. Thus by contraposition, a set is open if there is a ball $B(a, r)$ for every point a in the set which is contained in the set. The converse follows the same argument strategy, and is left to the reader. □

Theorem 3.8. *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. □

Corollary 3.9. *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

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

□

Theorem 3.10. *If $(A_1, A_2, \dots, A_n)$ 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 $B(a, \min(r_1, r_2, \dots, r_k))$ is a subset of every ball in A_k, it is contained in the intersection. Thus the intersection is open. □

Corollary 3.11. *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 3.12. *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), $\|y - z\| \leq \|y - x\| + \|x - z\| = \|v - w\| + r$. 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 $y \in \mathbf{R}^n$. □

Theorem 3.13 (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 sequence 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. By contraposition, A must be bounded. □

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 3.14 (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, and hence A is closed, 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.

Exercise 6. $(-1,1)$ *is a non-compact set. Find an open cover of* $(-1,1)$ *that has no finite subcover.*

We need two more types of sets to finish off the topology of $\mathbf{R}^n$, the first being the quality of connectedness. 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, in a way that 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.*

Exercise 7. *A set A is disconnected if and only if it can be partitioned into two sets A_1 and A_2 such that $\overline{A_1} \cap A_2 = \emptyset$, and $A_1 \cap \overline{A_2} = \emptyset$.*

Our final property is convexity. 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 8. *Any ball is convex.*

Exercise 9. *A point v in a convex set A is an extreme point if there does not exist points u and w and some scalar $\lambda \in [0, 1]$ such that $v = \lambda u + (1-\lambda)w$. Prove or disprove that this only occurs when v is a limit point of the set.*

Definition 18. *Let $v_1, v_2, \ldots, v_n$ be vectors in $\mathbf{R}^n$, and let $\lambda_1, \ldots, \lambda_n$ be non-negative real numbers which sum to 1. The vector*

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

is said to be a convex combination of the set of vectors v_1 through v_n. The convex hull of these vectors is the set of all convex combinations of the vectors.

Exercise 10. *Prove that the convex hull of a finite set of vectors is convex.*

Theorem 3.15. *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 λ'. 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,

$$\begin{aligned}
\|\lambda a_1 + (1-\lambda)a_2 - x\| &= \|(\lambda - \lambda')a_1 + (\lambda' - \lambda)a_2\| \\
&\leq |\lambda - \lambda'|\|a_1\| + |\lambda' - \lambda\|a_2\| \\
&\leq |r/2\|a_1\|\|\|a_1\| + |r/2\|a_2\|\|\|a_2\| \\
&= r
\end{aligned}$$

Thus λ' 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. ∎

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

Definition 19. *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

After enough mathematics to understand this report, you should know the formal definition of a function. In this book, we deal with functions from a subset of $\mathbf{R}^n$ to $\mathbf{R}^m$. We can consider this function to be a function of n variables, mapped to m variables. Really, this is just the composition of m functions $f_i : \mathrm{dom}(f) \to \mathbf{R}$, defined by the equation

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

Definition 20. *Consider a function f, mapping a subset of $\mathbf{R}^n$ to $\mathbf{R}^m$ and suppose u is an accumulation point of $\mathrm{dom}(f)$. We say that f approaches a vector $w \in \mathbf{R}^m$ as its domain approaches u, and we write $\lim_{v\to u} = w$, 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 $v \in \mathrm{dom}(f)$ such that $0 < \|v - u\| < \delta$, $\|w - f(v)\| < \varepsilon$.*
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) with elements in $\mathrm{dom}(f) - \{c\}$ 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 δ 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 δ, 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). $\square$

As the equivalence of sequences of vectors to sequences of real numbers allowed us to prove theorems, the equivalence of limits of functions to sequences implies many theorems on par with ones you have already seen.

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

Proof. Let f be a function that converges at an accumulation point u 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 u is an accumulation point, so there must be a sequence with this property. $\square$

We leave the rest of these arguments to the reader.

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

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

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

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

Corollary 4.3. *If f maps from a subset of* $\mathbf{R}^n$ *to* $\mathbf{R}^m$*, such that*

$$\lim_{v\to u} f(v) = l$$

and g maps from a subset of $\mathbf{R}^m$ *to* $\mathbf{R}^l$ *such that*

$$\lim_{w\to l} g(w) = m$$

Then

$$\lim_{v\to u} (g \circ f)(v) = m$$

Definition 21. *Suppose we have a function f from a subset of* $\mathbf{R}^n$ *to* $\mathbf{R}^m$*. Then, for a point* $u \in \mathrm{dom}(f)$*, we say that f is continuous at u if, for any* $\varepsilon > 0$*, there is a δ such that, for any vector v such that* $\|v - u\| < \delta$*,* $\|f(v) - f(u)\| < \varepsilon$*. If u is an accumulation point, this is equivalent to the fact that*

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

We say that, for a subset $C \subset \mathrm{dom}(f)$*, 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. If u is not an accumulation point, then the function is continuous at that point, since we may pick a δ such that u is the only element in the ball. Continuous functions have many useful properties. We leave the proofs to the reader as they follow immediately from statements about limits of functions.

Lemma 4.4. *For any sequence (v_i), such that $v_i \to v$ and any function f continuous at v*

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

Lemma 4.5. *Let f and g be functions from subsets of $\mathbf{R}^m$ both continuous at a point u, and h maps from a subset of $\mathbf{R}^n$, that is continuous at $f(c)$:*

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. *$h \circ f$ is continuous at c.*

The following theorem is very important in the field of topology

Theorem 4.6. *A function f from a subset of $\mathbf{R}^n$ to $\mathbf{R}^m$ is continuous on its domain if and only if, for every open set C in $\mathbf{R}^m$, there exists an open set U such that $U \cap \operatorname{dom}(f) = f^{-1}(C)$.*

Proof. Let f be as above, continuous on its domain. We prove the statement for open balls, from which the entire theorem follows as an arbitrary open set is the union of open balls. 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

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

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

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

Corollary 4.7. *A function is continuous if and only if, for every closed set C in $\mathbf{R}^m$, there exists a closed set D such that $D \cap \mathrm{dom}(f) = f^{-1}(C)$.*

Some simple practical applications result from this theorem.

Lemma 4.8. *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. $\square$

The interests of continuous functions in Topology rely in the fact that the functions preserve the topological properties of a space in some way. We list some important applications here.

Theorem 4.9. *Let f be a continuous function, and $A \subset \mathrm{dom}(f)$ a compact set. Then $f(A)$ is compact.*

Proof. Let U be an open cover on $f(A)$. For each open set u in U, there is an open set M such that $f^{-1}(u) = M \cap \mathrm{dom}(f)$. Form the set of M's for each u. This is an open cover of A, and hence contains a finite subcover U'. For each $u \in U'$, take the corresponding open set in $f(A)$. This forms a finite subcover, hence $f(A)$ is compact. $\square$

Corollary 4.10. *If f is a mapping and $\mathrm{dom}(f)$ is compact, then $f(\mathrm{dom}(f))$ is bounded.*

Another topological notions is maintained by a continuous map.

Theorem 4.11. *If f is a contiuous mapping, and A is a separable subset of the range of f, then $f^{-1}(A)$ is separable. Hence by contraposition, if B is a connected subset of $\mathrm{dom}(f)$, $f(B)$ is also connected.*

Proof. Assume A is a non-empty separable subset of the range of f, so $A = A_1 \cup A_2$, for non-empty subsets A_1 and A_2 such that $A_1 \cap \overline{A_2} = \overline{A_1} \cap A_2 = \emptyset$. Consider the two sets $f^{-1}(A_1)$, and $f^{-1}(A_2)$. Then $f^{-1}(A_1 \cup A_2) = f^{-1}(A_1) \cup f^{-1}(A_2)$, and each is non-empty. Since $A_1 \subset \overline{A_1}$, $f^{-1}(A_1) \subset f^{-1}(\overline{A_1})$, and the same for A_2. These sets are closed as their image is closed. It follows that the closure of $f^{-1}(A_1)$ is a subset of $f^{-1}(\overline{A_1})$, as for A_2, and

$$f^{-1}(\overline{A_1}) \cap f^{-1}(A_2) = f^{-1}(\overline{A_1} \cap A_2) = f^{-1}(\emptyset) = \emptyset$$

The proof is similar for A_2, showing that $f^{-1}(A)$ is separable. $\square$