

Chapter 3. Basic Topology of \mathbf{R}

Theorem. The open interval $(0, 1) = \{x \in \mathbf{R} : 0 < x < 1\}$ is uncountable.

Proof.

We proceed by contradiction and assume that there exists a function $f : \mathbf{N} \rightarrow (0, 1)$ that is $1-1$ and onto. $1-1$ implies that distinct elements have distinct images. Onto implies that every element in the co-domain has atleast one pre-image. For each $m \in \mathbf{N}$, $f(m)$ is a real number between 0 and 1, and we represent it using the decimal notation

$$f(m) = .a_{m1}a_{m2}a_{m3}a_{m4}a_{m5} \dots$$

What is meant here is that for each $m, n \in \mathbf{N}$, a_{mn} is the digit from the set $\{0, 1, 2, 3, \dots, 9\}$ that represents the n th digit in the decimal expansion of m th real number, $f(m)$. The $1-1$ correspondence between \mathbf{N} and $(0, 1)$ can be summarized in the doubly indexed array:

\mathbf{N}		$(0, 1)$								
1	\longleftrightarrow	$f(1)$	=	. a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	\dots
2	\longleftrightarrow	$f(2)$	=	. a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	\dots
3	\longleftrightarrow	$f(3)$	=	. a_{31}	a_{32}	a_{33}	a_{34}	a_{35}	a_{36}	\dots
4	\longleftrightarrow	$f(4)$	=	. a_{41}	a_{42}	a_{43}	a_{44}	a_{45}	a_{46}	\dots
5	\longleftrightarrow	$f(5)$	=	. a_{51}	a_{52}	a_{53}	a_{54}	a_{55}	a_{56}	\dots
6	\longleftrightarrow	$f(6)$	=	. a_{61}	a_{62}	a_{63}	a_{64}	a_{65}	a_{66}	\dots
\vdots		\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots

The key assumption about this correspondence is that **every** real number in $(0, 1)$ is assumed to appear somewhere on this list.

Now for the pearl of the argument. Define a real number $x \in (0, 1)$ with the decimal expansion $x = .b_1b_2b_3b_4 \dots$ using the rule

$$b_n = \begin{cases} 2 & \text{if } a_{nn} \neq 2 \\ 3 & \text{if } a_{nn} = 2 \end{cases}$$

Now, the real number $x = .b_1b_2b_3b_4 \dots$ cannot be $f(1)$, simply because its first digit b_1 differs from the first digit a_{11} of $f(1)$. Similarly, the second digit b_2 differs from the second digit a_{22} of $f(2)$. In general, the n th digit of x differs from the n th digit of $f(n)$. So, we have constructed a

real number x that is not in the set $\{f(1), f(2), f(3), \dots, f(n)\}$. But, this is a contradiction. Hence, our initial assumption is false. The set of real numbers in $(0, 1)$ are uncountable.

Exercise. [Abbott, 1.6.4] Let S be the set consisting of all sequences of 0s and 1s. Observe that S is not a particular sequence, but rather a large set whose elements are sequences, namely:

$$S = \{(a_1, a_2, a_3, \dots) : a_n = 0 \text{ or } 1\}$$

As an example, the binary sequence $(1, 0, 1, 0, 1, 0, 1, 0, \dots)$ is an element of S as is the sequence $(1, 1, 1, 1, 1, 1, \dots)$. Give a rigorous argument showing that S is uncountable.

Proof.

Suppose that S - the set of all possible binary strings of infinite length is countable. Then, we can define a bijection $f: \mathbf{N} \rightarrow S$ between the natural numbers and S . For each $m \in \mathbf{N}$, $f(m)$ is a binary string in S . Let us enlist the first few elements of this correspondence.

N		(0, 1)								
1	\longleftrightarrow	$f(1)$	=	.	a₁₁	a_{12}	a_{13}	a_{14}	a_{15}	a_{16} ...
2	\longleftrightarrow	$f(2)$	=	.	a_{21}	a₂₂	a_{23}	a_{24}	a_{25}	a_{26} ...
3	\longleftrightarrow	$f(3)$	=	.	a_{31}	a_{32}	a₃₃	a_{34}	a_{35}	a_{36} ...
4	\longleftrightarrow	$f(4)$	=	.	a_{41}	a_{42}	a_{43}	a₄₄	a_{45}	a_{46} ...
5	\longleftrightarrow	$f(5)$	=	.	a_{51}	a_{52}	a_{53}	a_{54}	a₅₅	a_{56} ...
6	\longleftrightarrow	$f(6)$	=	.	a_{61}	a_{62}	a_{63}	a_{64}	a_{65}	a₆₆ ...
\vdots		\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots

Define a binary sequence $x = (b_1, b_2, b_3, b_4, \dots)$ such that

$$b_i = \begin{cases} 1 & \text{if } a_{ii} = 0 \\ 0 & \text{if } a_{ii} = 1 \end{cases}$$

Thus, the binary sequence $x = (b_1, b_2, b_3, b_4, \dots)$ has atleast one bit that differs from all of the elements in S . Consequently, $x \notin S$. This is a contradiction, as S is supposed to contain all binary strings. Hence, S is not countable.

3.1 Discussion: The Cantor Set.

What follows is a fascinating mathematical construction, due to Georg Cantor, which is extremely useful for extending the horizons of our intuition about the nature of subsets of the real line. Cantor's name has already appeared in the first chapter in our discussion of uncountable sets. Indeed, Cantor's proof that \mathbf{R} is uncountable occupies another spot on the short list of the most significant contributions towards the understanding of the mathematical infinite. In the words of the mathematician David Hilbert, "No one shall expel us from the paradise that Cantor has created for us."

Let C_0 be the closed interval $[0, 1]$ and define C_1 to be the set that results when the open middle one third is removed that is,

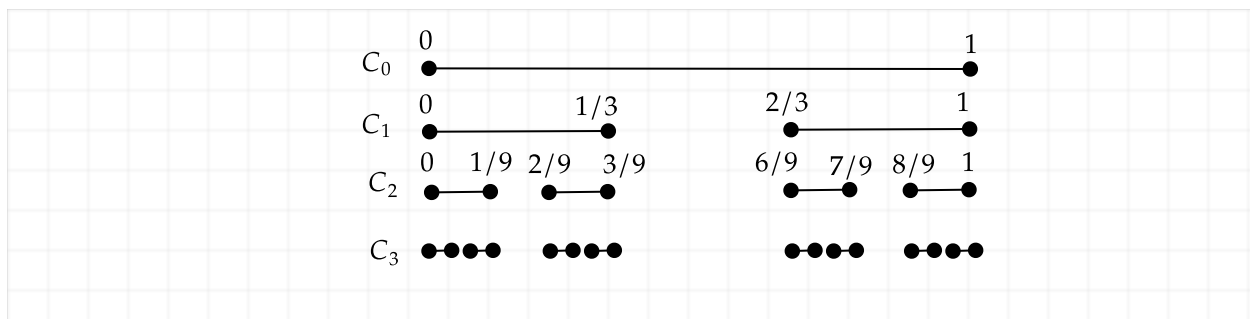
$$C_1 = C_0 \setminus \left(\frac{1}{3}, \frac{2}{3}\right) = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$$

Now, construct C_2 in a similar way by removing the open middle third of each of the two components of C_1 :

$$C_2 = \left(\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{3}{9}\right]\right) \cup \left(\left[\frac{6}{9}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]\right)$$

If we continue this process inductively then for each $n = 0, 1, 2, \dots$ we get a set of C_n consisting of 2^n closed intervals each having length $1/3^n$. Finally, we define the Cantor set C to be the intersection

$$C = \bigcap_{n=0}^{\infty} C_n$$



Defining the Cantor set

It may be useful to understand C as the remainder of the interval $[0, 1]$ after the iterative process of removing open middle one thirds is taken to infinity.

$$C = [0, 1] - \left[\left(\frac{1}{3}, \frac{2}{3} \right) \cup \left(\frac{1}{9}, \frac{2}{9} \right) \cup \left(\frac{7}{9}, \frac{8}{9} \right) \cup \dots \right]$$

There is some initial doubt whether anything remains at all, but notice that because we are always removing open middle one thirds, then for every $n \in \mathbf{N}$, $0 \in C_n$ and hence $0 \in C$. The same argument shows that $1 \in C$. In fact, if y is the endpoint of some closed interval of some particular set C_n , then it is also an endpoint of one of the intervals of C_{n+1} . Because at each stage, the endpoints are never removed, it follows that $y \in C_n$ for all n . Thus, C at least contains the endpoints of all of the intervals that make up each of the sets C_n .

Is there anything else? Is C countable? Does C contain any intervals? Any irrational numbers? These are difficult questions at the moment. All of the endpoints mentioned earlier are rational numbers (they have the form $m/3^n$), which means that if it is true that C consists of only these endpoints, then C would be a subset of \mathbf{Q} and hence countable. We shall see about this. There is some strong evidence that not much is left in C if we consider the total length of the intervals removed. To form C_1 , an open interval of length $1/3$ was taken out. In the second step, we removed two intervals of length $1/9$ and to construct C_n , we removed 2^{n-1} middle thirds of length $1/3^n$. There is some logic, then to defining the length of C to be 1 minus the total

$$\frac{1}{3} + 2\left(\frac{1}{9}\right) + 4\left(\frac{1}{27}\right) + \dots + 2^{n-1}\left(\frac{1}{3^n}\right) + \dots = \frac{\frac{1}{3}}{1 - \frac{2}{3}} = 1$$

The Cantor set has *zero length*.

To this point, the information we have collected suggests a mental picture of C as relatively small. For these reasons, the set C is often referred to as Cantor dust. But, there are some strong counterarguments that imply a very different picture. First C is actually *uncountable*, with cardinality equal to the cardinality of \mathbf{R} . One slightly intuitive but convincing way to see this is to create a 1 – 1 correspondence between C and the sequences of the form $(a_n)_{n=1}^{\infty}$, where $a_n = 0$ or 1. For each $c \in C$, set $a_1 = 0$ if c falls in the left-hand component and set $a_1 = 1$ if c falls in the right hand component. Having established where in C_1 , the point c is located, there are now two possible components of C_2 that might contain c . This time, we set $a_2 = 0$ or 1 depending on whether c falls in the left or right half of these two components of C_2 . Continuing in this way, we come to see that every element $c \in C$ yields a sequence (a_1, a_2, a_3, \dots) of zeroes and ones (a binary string), that acts as a set of directions for how to locate c within C . Because the set of sequences of zeroes and ones is uncountable, we must conclude that C is uncountable as well.

What does this imply? In the first place, because the end points of the approximating sets C_n form a countable set, we are forced to accept the fact that not only are there other points in C , but there are uncountably many of them. From the point of view of cardinality, C is quite large - as large as \mathbf{R} , in fact. This should be contrasted with the fact that from the point of view of length, C measures the same size as a single point. We conclude this discussion with a demonstration that from the point of view of *dimension*, C strangely falls somewhere in between.

There is a sensible agreement that a point has dimension zero, a line segment has dimension one, a square has dimension two, and a cube has dimension three. Without attempting a formal definition of dimension (of which there are several) we can nevertheless get a sense of how one might be defined by observing how the dimension affects the result of magnifying each particular set by a factor of 3. (The reason for the choice of 3 will become clear when we turn our attention back to the Cantor set). A single point undergoes no change at all, whereas a line segment triples in length. For the square, magnifying each length by a factor of 3 results in a larger square that contains 9 copies of the original square. Finally, the magnified cube yields a cube that contains 27 copies of the original cube within its volume. Notice that, in each case, to compute the size of the new set, the dimension appears as the exponent of the magnification factor.

	<i>dim</i>	$\times 3$	new copies
point	0	\rightarrow	$1 = 3^0$
segment	1	\rightarrow	$3 = 3^1$
square	2	\rightarrow	$9 = 3^2$
cube	3	\rightarrow	$27 = 3^3$

Dimension of C ,

Now, apply this transformation to the Cantor set. The set $C_0 = [0, 1]$ becomes the interval $[0, 3]$. Deleting the middle one-third leaves $[0, 1] \cup [2, 3]$, which is where we started in the original construction except that we now stand to produce an additional copy of C in the interval $[2, 3]$. Magnifying the Cantor set by a factor of 3 yields two copies of the original set. Thus, if x is the dimension of C , we must have $3^x = 2$, or $x = \ln 2 / \ln 3 = 0.631$.

The notion of a non-integer or fractional dimension is the impetus behind the term fractal, coined in 1975 by Benoit Mandelbrot to describe a class of sets whose intricate structures have much in common with the Cantor set. Cantor's construction is over hundred years old and for us represents an invaluable testing ground for upcoming theorems and conjectures about the often elusive nature of subsets of the real line.

3.2 Open and Closed Sets.

Given $a \in \mathbf{R}$ and $\epsilon > 0$, recall that the ϵ -neighbourhood of a is the set

$$V_\epsilon(a) = \{x \in \mathbf{R} : |x - a| < \epsilon\}$$

In other words, $V_\epsilon(a)$ is the open interval $(a - \epsilon, a + \epsilon)$ centered at a with radius ϵ .

Definition. A set $O \subseteq \mathbf{R}$ is open if for all points $a \in O$ there exists an ϵ -neighbourhood $V_\epsilon(a) \subseteq O$.

Example.

(i) Perhaps, the simplest example of an open set is \mathbf{R} itself. Given an arbitrary element $a \in \mathbf{R}$, we are free to pick any ϵ -neighbourhood we like and it will always be true that $V_\epsilon(a) \subseteq \mathbf{R}$. It is also the case that the logical structure of the definition requires us to classify the empty set \emptyset as an open subset of the real line.

(ii) For a more useful collection of examples, consider the open interval

$$(c, d) = \{x \in \mathbf{R} : c < x < d\}$$

To see that (c, d) is open in the same sense just defined, let $x \in (c, d)$ be arbitrary. If we take $\epsilon = \min\{x - c, d - x\}$, then it follows that $V_\epsilon(x) \subseteq (c, d)$. It is important to see where this argument breaks down if the interval includes either one of its endpoints.

The union of open intervals is another example of an open set. This observation leads to the next result.

Theorem. (i) The union of an arbitrary collection of open sets is open.

(ii) The intersection of a countably finite collection of open sets is open.

Proof.

To prove (i), we let $\{O_\lambda : \lambda \in \Lambda\}$ be a collection of open sets and $O = \bigcup_{\lambda \in \Lambda} O_\lambda$. Here, Λ could be a countably infinite or an uncountable set. Let a be an arbitrary element of O . In order to show that O is open, the definition insists that we produce an ϵ -neighbourhood of a completely contained in O . But, $a \in O$ implies that a is an element of at least one particular $O_{\lambda'}$. Because we are assuming that $O_{\lambda'}$ is open, by definition we can assert, that there exists $V_\epsilon(a) \subseteq O_{\lambda'}$.

The fact that $O_{\lambda'} \subseteq O$ allows us to conclude that $V_\epsilon(a) \subseteq O$ for all $a \in O$. This completes the proof of (i).

For (ii), let $\{O_1, O_2, \dots, O_N\}$ be a finite collection of open sets. Now, if $a \in \bigcap_{k=1}^N O_k$, then a is an element of each of the open sets. By the definition of an open set, we know that, for each, $1 \leq k \leq N$, there exists a $V_{\epsilon_k}(a) \subseteq O_k$. We are in search of a single neighbourhood of a that is contained in every O_k , so the trick is to take the smallest one. Letting

$\epsilon = \min\{\epsilon_1, \epsilon_2, \dots, \epsilon_N\}$, it follows that $V_\epsilon(a) \subseteq V_{\epsilon_k}(a)$ for all k , and hence $V_\epsilon(a) \subseteq \bigcap_{k=1}^N O_k$ as desired.

Closed Sets.

Definition. A point x is the **limit point** of a set A , if every ϵ -neighbourhood $V_\epsilon(x)$ intersects the set A at some point other than x .

Limit points are also referred to as cluster points or accumulation points, but the phrase, " x is a limit point of A " has the advantage of explicitly reminding us that x is quite literally the limit of the sequence in A .

Theorem. A point x is a limit point of a set A if and only if $x = \lim a_n$ for some sequence (a_n) contained in A , satisfying $a_n \neq x$ for all $n \in \mathbf{N}$.

Proof.

(\implies) Assume that x is a limit point of A . In order to produce a sequence (a_n) converging to x , we are going to consider the particular ϵ -neighbourhoods obtained using $\epsilon = 1/n$. By definition, every neighborhood of x intersects A in some point other than x . This means that, for each $n \in \mathbf{N}$, we are justified in picking a point

$$a_n \in V_{1/n}(x) \cap A$$

with the stipulation that $a_n \neq x$. It should not be too difficult to see why $(a_n) \rightarrow x$. Given an arbitrary $\epsilon > 0$, choose N such that $1/N < \epsilon$. Then for all $n \geq N$, we have $|a_n - x| < \epsilon$.

(\impliedby) For the reverse implication, we assume that $\lim a_n = x$ where $a_n \in A$ but $a_n \neq x$ and let $V_\epsilon(x)$ be an arbitrary neighbourhood of x . The definition of convergence assures us that there exists a term a_N in the sequence satisfying $a_N \in V_\epsilon(x)$. So, every ϵ -neighbourhood, $V_\epsilon(x)$ intersects A in some element other than x . Hence, $x = \lim a_n$. QED.

The restriction that $a_n \neq x$ in the above theorem deserves a comment. Given a point $a \in A$, it is always the case that a is the limit of a sequence in A , if we are allowed to consider the constant sequence (a, a, a, \dots) . There will be occasions where we will want to avoid this somewhat uninteresting situation, so it is important to have vocabulary that can distinguish limit points of a set from isolated points.

Definition. A point $a \in A$ is an isolated point of A if it is not a limit point of A .

As a word of caution, we need to be a little careful about how we understand the relationship between these concepts. Whereas an isolated point is always an element of the relevant set A , it is quite possible for a limit point of A not to belong to A . As an example, consider the endpoint of an open interval. This situation is the subject of the next important definition.

Definition. A set $F \subseteq \mathbf{R}$ is closed if it contains its limit points.

The adjective "closed" appears in several other mathematical contexts and is usually employed to mean that an operation on the elements of a given set does not take us outside of the set. In linear algebra, for example, a vector space is a set that is closed under vector addition and scalar multiplication. In analysis, the operation we are concerned with is the limiting operation. Topologically speaking, a closed set is the one where convergent sequences within the set have limits that are also in the set.

Theorem. A set $F \subseteq \mathbf{R}$ is closed if and only if every Cauchy sequence contained in F has a limit that is also an element of F .

Proof. Exercise 3.2.5.

Example. (i) Consider

$$A = \left\{ \frac{1}{n} : n \in \mathbf{N} \right\}$$

Let's show that each point of A is isolated. Given $1/n \in A$, choose $\epsilon = \frac{1}{n} - \frac{1}{n+1}$. (Note that,

$\frac{1}{n-1} - \frac{1}{n}$ is larger than $\frac{1}{n} - \frac{1}{n+1}$). Then,

$$V_\epsilon(1/n) \cap A = \frac{1}{n}$$

It follows from the definition that $1/n$ is not a limit point and so is isolated. Although all of the points of A are isolated, the set does have one limit point, namely 0. This is because every ϵ -neighbourhood of centered at zero, no matter how small, is going to contain points of A . Because $0 \notin A$, A is not closed. The set $F = A \cup \{0\}$ is an example of a closed set and is called the closure of A . (The closure of a set is discussed in a moment).

(ii) Let's prove that a closed interval

$$[c, d] = \{x \in \mathbf{R} : c \leq x \leq d\}$$

is a closed set using the definition. If x is a limit point of $[c, d]$, then by the previous theorem there exists a sequence (x_n) contained in $[c, d]$ such that $(x_n) \rightarrow x$ and $x_n \neq x$. We need to prove that x belongs to $[c, d]$.

The key to this argument is contained in the Order Limit Theorem, which summarizes the relationship between inequalities and the corresponding limiting process. Because, $c \leq x_n \leq d$, it follows from the the Order Limit Theorem, that $c \leq x \leq d$. Thus, $[c, d]$ is closed.

(iii) Consider the set $\mathbf{Q} \subseteq \mathbf{R}$ of rational numbers. An extremely important property of \mathbf{Q} is that the set of all limit points of \mathbf{Q} is actually all of \mathbf{R} . To see why this is so, recall the theorem, which is referred to as the density property of \mathbf{Q} in \mathbf{R} . \mathbf{Q} is dense in \mathbf{R} implies that \mathbf{Q} sits inside of \mathbf{R} . Between any two real numbers $a, b \in \mathbf{R}$, you can always find a rational number r satisfying, $a < r < b$.

Let $y \in \mathbf{R}$ be arbitrary and consider any neighbourhood $(y - \epsilon, y + \epsilon)$. The density theorem allows us to conclude that there exists a rational number $r \neq y$ that falls in this neighbourhood. Thus, y is a limit point of \mathbf{Q} .

The density property of \mathbf{Q} can now be reformulated in the following way.

Theorem.(Density of \mathbf{Q} in \mathbf{R}). For every $y \in \mathbf{R}$, there exists a sequence of the rational numbers that converges to y .

Proof.

From the above discussion, we know that for an arbitrary $y \in \mathbf{R}$, every ϵ -neighbourhood of y intersects \mathbf{Q} in some point other than y . y is a limit point of \mathbf{Q} . By the theorem on limit points

of sets, there exists a sequence $(x_n) \subseteq \mathbf{Q}$, such that $(x_n) \rightarrow y$, and $x_n \neq y$ for all n .

The same argument can also be used to show that every real number is the limit of a sequence of irrational numbers. Although interesting, part of the allure of the rational numbers is that, in addition to being dense in \mathbf{R} , they are countable. As we will see, this tangible aspect of \mathbf{Q} makes it an extremely useful set, both for proving theorems and for producing interesting counterexamples.

Closure.

Definition. Given a set $A \subseteq \mathbf{R}$, let L be the set of all limit points of A . The closure A is defined to $cl(A) = A \cup L$.

In example 3.2.9 (i), we saw that if $A = \{1/n : n \in \mathbf{N}\}$, then the closure of $A = cl(A) = A \cup \{0\}$. Example 3.2.9 (iii) verifies that $cl(\mathbf{Q}) = \mathbf{R}$. If A is an open interval, then $cl(A) = [a, b]$. If A is a closed interval, then $cl(A) = A$. It is not for lack of imagination that in each of these examples $cl(A)$ is always a closed set.

Theorem. For any $A \subseteq \mathbf{R}$, the closure $cl(A)$ is a closed set and is the smallest closed set containing A .

Proof.

If L is the set of limit points of A , then it is immediately clear that $cl(A)$ contains the limit points of A . There is still something more to prove, however because taking the union of L with A could potentially produce some new limit points of $cl(A)$. In exercise 3.2.7, we outline the argument that this does not happen.

Now, any closed set containing A must contain L as well. This shows that $cl(A) = A \cup L$ is the smallest such closed set containing A .

Complements.

The mathematical notions of open and closed are not antonyms the way they are in standard English. If a set is not open, that does not imply it must be closed. Many sets such as the half-open interval $(c, d] = \{x \in \mathbf{R} : c < x \leq d\}$ are neither open nor closed. The sets \mathbf{R} and \emptyset are simultaneously open and closed, although, thankfully, these are the only ones with this disorientating property. (Exercises 3.2.13). There is, however, an important relationship between open and closed sets. Recall that the complement of a set $A \subseteq \mathbf{R}$ is defined to be the set:

$$A^C = \{x \in \mathbf{R} : x \notin A\}$$

Theorem. A set O is open if and only if O^C is closed. Likewise, a set F is closed if and only if F^C is open.

Proof.

(\implies) Given an open set $O \subseteq \mathbf{R}$, let's first prove that O^C is a closed set. To prove that O^C is a closed, we need to show that it contains all of its limit points. If x is a limit point of O^C , then every ϵ -neighbourhood of x contains some point of O^C other than x . But that is enough to conclude that x cannot be in O , because if $x \in O$, then every open interval $(x - \epsilon, x + \epsilon)$ is contained in O . Consequently, $x \in O^C$. Thus, O^C is a closed set.

(\impliedby) For the converse statement, we assume that O^C is closed and argue that O is open. Thus, given an arbitrary point $x \in O$, we must produce an ϵ -neighbourhood $V_\epsilon(x) \subseteq O$. Because, $x \notin O^C$, we can be sure that x is not a limit point of O^C . If x is not a limit point of O^C , there exists an $\epsilon > 0$, such that $(x - \epsilon, x + \epsilon) \not\subseteq O^C$, or equivalently $(x - \epsilon, x + \epsilon) \subseteq O$. This is precisely what we needed to show.

Theorem. (i) The union of a countably finite collection of closed sets is closed.
(ii) The intersection of an arbitrary collection of closed sets is closed.

Proof.

De Morgan's Laws state that for any collection of sets $\{E_\lambda : \lambda \in \Lambda\}$ it is true that

$$\left(\bigcup_{\lambda \in \Lambda} E_\lambda \right)^C = \bigcap_{\lambda \in \Lambda} E_\lambda^C \quad \text{and} \quad \left(\bigcap_{\lambda \in \Lambda} E_\lambda \right)^C = \bigcup_{\lambda \in \Lambda} E_\lambda^C$$

The result follows directly from these statements and Theorem 3.2.3. The details are requested in Exercise 3.2.9.

Exercises.

1. [Abbot, 3.2.1]