Connected sets

G. C. Boore
website: https://gcboore.com
email: art@gcboore.com

October 6, 2015

Contents

1 Introduction

One definition of the *Mandelbrot set* is

$$M = \{c : c \in \mathbb{C},\ J(f_c) \text{ is connected}\},$$

where $f_c : \mathbb{C} \to \mathbb{C}$ is the quadratic map defined by $f_c(z) = z^2 + c$ and $J(f_c)$ is the associated Julia set.

In 1982 Douady and Hubbard [3] proved that the Mandelbrot set itself is a connected set (there is another proof in [1, Section 9.10]). Clearly being connected is an important property for a set to have, so in these notes we try and explain what it means to say that a set is connected. The result is the briefest of introductions to topology, starting from scratch. Hopefully anyone who hasn't studied any topology at all will be able to get some idea of what's involved although some mathematical experience is needed. There is also the question as to whether or not the Mandelbrot set is locally connected, and we will try and make clear what this means too, and why it is an important open problem. Our guide to most of the definitions and results is Sutherland's book [10], where many other areas of topology are covered.

2 Open sets

Any open set has something of the infinite about it. In some cases like the complex plane, $\mathbb{C}$, which is an open set, this is fairly obvious as we can travel off to infinity in all directions. For the open unit disc $D = \{z : z \in \mathbb{C}, |z| < 1\}$ the infinite becomes apparent as soon as we realise that D is much the same as $\mathbb{C}$ under the one-to-one and onto map $f(z) = z/(1 - |z|)$. We prove D is topologically equivalent to $\mathbb{C}$ in Lemma 2.8.

The definitions of the various types of connectedness that we will meet require only the ideas of open sets and continuous functions. This makes topological spaces rather than metric spaces the most general setting for the definitions. However topological spaces can always be derived from the open sets in metric spaces, so we start by defining an open set in the context of a general metric space before giving the definition of a topological space in Subsection 2.2.

2.1 Metric spaces

A *metric space* (X, d) is a non-empty set X together with a *metric* (a distance function) $d : X \times X \to \mathbb{R}$ which satisfies the following axioms for all $x, y, z \in X$

(M1) $d(x, y) \geqslant 0$; $d(x, y) = 0$ if and only if $x = y$,
(M2) $d(x, y) = d(y, x)$,
(M3) $d(x, z) \leqslant d(x, y) + d(y, z)$.

The third axiom (M3) is called the triangle inequality. For a metric space (X, d) if $A \subset X$ is non-empty then (A, d_A) can be considered as a *metric subspace* where d_A is the restriction of d to $A \times A$. Clearly the axioms for a metric hold for d_A, for $x, y, z \in A$, because they hold for d.

The *open ball* of centre x and radius $r > 0$, is defined as

$$S(x, r) = \{y : y \in X,\ d(x, y) < r\}.$$

A set $A \subset X$ is *open* if and only if for every $x \in A$ there exists $r > 0$ such that $S(x, r) \subset A$.

We use phrases like "open in X" or "open subset of X" and the notation $S_X(x, r)$ whenever we wish to emphasise the parent space when this may not be obvious from the context. This is because whether or not a set is open depends on the parent space being referred to. We also use phrases like "open with respect to the metric d" or "d-open" when we wish to point out the metric being used. For open balls we may write $S(d)(x, r)$ because changing the metric changes open balls as we now describe. (The open unit disc in $\mathbb{C}$ with the Euclidean metric we can write as $D = S_{\mathbb{C}}(|\ |)(0, 1)$.)

Consider the following three metrics which are taken from [10, Examples 2.2.3]. These metrics are defined on n-dimensional Euclidean space, which is the Cartesian n-fold product of $\mathbb{R}$

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ times}} = \{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R},\ 1 \leqslant i \leqslant n\},$$

and are defined, for $x, y \in \mathbb{R}^n$, as

$$\begin{aligned} d_1(x,y) &= \sum_{i=1}^{n} |x_i - y_i| \,, \\ d_2(x,y) &= |x-y| = \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{\frac{1}{2}} , \\ d_\infty(x,y) &= \max_{1 \leqslant i \leqslant n} \{|x_i - y_i|\} \,. \end{aligned} \tag{2.1}$$

The first metric $d_1(x, y)$ is known as the taxicab metric as it measures the distance between points in $\mathbb{R}^2$ as lengths of the sides of right-angled triangles, so it measures the distance a taxi may travel in the grid-like street plan of a modern city. We call $d_2(x, y)$ the *Euclidean metric* and use the notation $|x - y|$ for it. For the complex plane $\mathbb{C}$, which is the same set of points as $\mathbb{R}^2$, the Euclidean metric is the same as the modulus function. It is easy to see that (M1) and (M2) hold but for (M3) we need to make use of the Cauchy-Schwartz inequality (for a proof see [10, Example 2.2.1]). Let $a, b \in \mathbb{R}^n$ then the Cauchy-Schwartz Inequality states that

$$\left(\sum_{i=1}^{n} a_i b_i \right)^2 \leqslant \left(\sum_{i=1}^{n} a_i^2 \right) \left(\sum_{i=1}^{n} b_i^2 \right). \tag{2.2}$$

Let $x, y, z \in \mathbb{R}^n$, then to show (M3) holds for the Euclidean metric we need to show $|x - z| \leqslant |x - y| + |y - z|$. This is the following inequality, after squaring both sides and putting $a_i = x_i - y_i$, $b_i = y_i - z_i$, so that $a_i + b_i = x_i - z_i$,

$$\sum_{i=1}^{n} (a_i + b_i)^2 = \sum_{i=1}^{n} (a_i^2 + 2a_i b_i + b_i^2) \leqslant \sum_{i=1}^{n} a_i^2 + 2 \left(\sum_{i=1}^{n} a_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} b_i^2 \right)^{\frac{1}{2}} + \sum_{i=1}^{n} b_i^2,$$

which holds by (2.2). To prove (M3) holds for $d_1(x, y)$ and $d_\infty(x, y)$ is now straightforward as they are defined in terms of the Euclidean metric for which we can now use the triangle inequality.

We introduced these metrics in order to show how different metrics produce different open balls. This is illustrated in Figure 2.1 where we can see the different shapes of $S(x, r)$, in $\mathbb{R}^2$ (or $\mathbb{C}$), for each of the metrics (a) $d_1(x, y)$, (b) $d_2(x, y) = |x - y|$ and (c) $d_\infty(x, y)$. It is clear from the definitions that for any $x, y \in \mathbb{R}^n$, $d_1(x, y) \geqslant d_2(x, y) \geqslant d_\infty(x, y)$ and it follows that $S(d_1)(x, r) \subset S(d_2)(x, r) \subset S(d_\infty)(x, r)$ which is also evident from Figure 2.1. For future reference we collect together some inequalities between the three metrics in the next lemma.

Lemma 2.1. *For $x, y \in \mathbb{R}^n$, and the metrics defined in* (2.1)*,*

$$d_1(x,y) \geqslant d_2(x,y) \geqslant d_\infty(x,y) \geqslant \frac{1}{\sqrt{n}} d_2(x,y) \geqslant \frac{1}{n} d_1(x,y).$$

Proof. The first two inequalities follow from the definitions. For the third clearly $nd_\infty(x, y)^2 \geqslant d_2(x, y)^2$. For the last inequality $nd_2(x, y)^2 \geqslant d_1(x, y)^2$ follows from the Cauchy-Schwartz Inequality, since putting $b_i = 1$, for $1 \leqslant i \leqslant n$, in (2.2) gives $n \sum_{i=1}^{n} a_i^2 \geqslant \left(\sum_{i=1}^{n} a_i\right)^2$. □

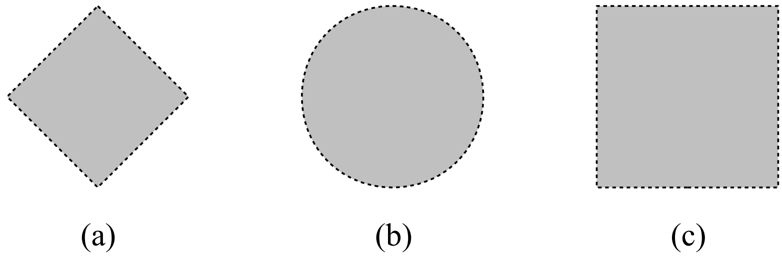

Figure 2.1: The open balls (a) $S(d_1)(x,r)$, (b) $S(d_2)(x,r)$ and (c) $S(d_\infty)(x,r)$, in $\mathbb{R}^2$.

We call $S(x,r)$ an open ball so now we prove that it is actually an open set in the next lemma which is [10, Lemma 2.3.7].

Lemma 2.2. *Let $S(x,r)$ be an open ball in a metric space (X,d). Then $S(x,r)$ is open in X.*

Proof. Let $y \in S(x,r)$. Put $s = r - d(x,y) > 0$, then for $z \in S(y,s)$, using (M3), $d(x,z) \leqslant d(x,y) + d(y,z) < d(x,y) + r - d(x,y) = r$ and so $z \in S(x,r)$. This proves $S(y,s) \subset S(x,r)$. □

For $A, B \subset X$ the *set difference* or *complement of B with respect to A* is defined as $A \setminus B = \{x : x \in A, x \notin B\}$.

A set $A \subset X$ is *closed in X* if and only if $X \setminus A$ is open.

In general open sets and closed sets are quite special and most sets are in fact neither open nor closed as the next example shows. This example also emphasises the importance of the parent space. (We don't provide a picture of it as it's impossible to draw one).

Example 2.3. *Let $(\mathbb{C}, |\ |)$ be the metric space which is the complex plane together with the Euclidean metric and let $X = P \cup Q \subset \mathbb{C}$, where P is the closed unit disc centred at -1 without the point at 0, Q is the closed unit disc centred at 1 also without the point at 0, and $P \cap Q = \emptyset$. Specifically*

$$P = \{z \in \mathbb{C} : \ |z+1| \leqslant 1\} \setminus \{0\}\,,$$
$$Q = \{z \in \mathbb{C} : \ |z-1| \leqslant 1\} \setminus \{0\}\,.$$

Then P and Q are neither open nor closed in $\mathbb{C}$. However P and Q are both open and closed in X.

Consider an open ball centred at $-2 \in P$, then $S_{\mathbb{C}}(-2,r) \not\subset P$ for any value of $r > 0$, so P is not open in $\mathbb{C}$. Similarly $0 \in \mathbb{C} \setminus P$ and $S_{\mathbb{C}}(0,r) \not\subset \mathbb{C} \setminus P$ for any $r > 0$, so $\mathbb{C} \setminus P$ is not open in $\mathbb{C}$. This means P is not closed in $\mathbb{C}$ either. A similar argument also works for Q.

To show P is open in X we need to consider $(X, |\ |)$ as a metric subspace of $(\mathbb{C}, |\ |)$. (Strictly speaking $(X, |\ |) = (X, |\ |_X)$ where $|\ |_X$ is the restriction of $|\ |$ to X, but the notation is complicated enough already). We need to show that for any

$z \in P$ there exists $r > 0$ such that $S_X(z,r) \subset P$. For any point $z \in P$ there is always a positive distance to the y-axis (the imaginary axis in $\mathbb{C}$), so we can always make r small enough so that $S_X(z,r) \cap Q = \emptyset$ which implies $S_X(z,r) \subset P$ since $P \cap Q = \emptyset$. This means P is open in X. Similarly Q is open in X. As $Q = X \setminus P$ and $P = X \setminus Q$ it follows that P and Q are also both closed in X.

The definitions of the different types of connectedness that follow only need the ideas of open sets and continuous functions and this makes topological spaces, rather than metric spaces, the most general setting for the definitions.

2.2 Topological spaces

A *topological space* $(X, \mathcal{T}_X)$ consists of a non-empty set X, together with a fixed collection $\mathcal{T}_X$ of subsets of X satisfying

(T1) $\emptyset,\ X \in \mathcal{T}_X$,

(T2) the intersection of any two sets in $\mathcal{T}_X$ is in $\mathcal{T}_X$,

(T3) the union of any collection of sets in $\mathcal{T}_X$ is in $\mathcal{T}_X$.

In this general setting to say $A \subset X$ is an open set means only that $A \in \mathcal{T}_X$ with no other special meaning given to the word open and $\mathcal{T}_X$ is known as a *topology* for X. To give one example, if we put $\mathcal{T}_X = \mathcal{P}(X)$, the set of all subsets of X, then the axioms (T1), (T2) and (T3) hold and $\mathcal{T}_X$ is a topology known as the *discrete* topology for X.

We now show that any metric space (X, d) can always be regarded as a topological space.

Lemma 2.4. *Let (X, d) be a metric space. Then*

(a) $\emptyset,\ X$ *are open sets,*

(b) *the intersection of any two open sets is open,*

(c) *the union of any collection of open sets is open.*

Proof. (a) Vacuously $\emptyset$ is an open set, and clearly X is open since $x \in X$ implies $S(x,r) \subset X$ for any $r > 0$.

(b) If $A \cap B = \emptyset$ then $A \cap B$ is open by part (a), so we may assume $A \cap B \neq \emptyset$. Let $x \in A \cap B$. Since A and B are open there exist $r, s > 0$ such that $S(x,r) \subset A$ and $S(x,s) \subset B$. It follows that $S(x, \min(r,s)) \subset A \cap B$ which proves $A \cap B$ is open.

(c) Let $A \neq \emptyset$ be the union of any collection of open sets in X and let $x \in A$. It follows that $x \in B$ for some open set B, so there exists $r > 0$ such that $S(x,r) \subset B \subset A$. This proves A is open. □

De Morgan's laws state that

$$X \setminus \bigcup_{\alpha \in I} A_\alpha = \bigcap_{\alpha \in I} (X \setminus A_\alpha) \text{ and } X \setminus \bigcap_{\alpha \in I} A_\alpha = \bigcup_{\alpha \in I} (X \setminus A_\alpha), \tag{2.3}$$

and prove the following corollary to Lemma 2.4 which shows that a topological space can also be defined in terms of its closed sets.

Corollary 2.5. *Let* (X, d) *be a metric space. Then*

(a) $\emptyset$, X *are closed sets,*

(b) *the union of any two closed sets is closed,*

(c) *the intersection of any collection of closed sets is closed.*

Lemma 2.4 shows that the axioms (T1), (T2), and (T3) hold for the open sets of any metric space. This means we can always consider any metric space (X, d) as a topological space $(X, \mathcal{T}_X)$ where $\mathcal{T}_X$ is the metric topology defined as

$$\mathcal{T}_X = \{U : U \subset X,\ U \text{ is open in } X\}.$$

If it is not clear from the context then the notation $\mathcal{T}_X(d)$ can be used to make clear which metric is being referred to with

$$\mathcal{T}_X(d) = \{U : U \subset X,\ U \text{ is } d\text{-open in } X\}.$$

Now it makes sense to talk about $\mathbb{C}$ with the Euclidean topology, meaning the topological space $(\mathbb{C}, \mathcal{T}_\mathbb{C})$ where $\mathcal{T}_\mathbb{C}$ is the set of all open sets of $\mathbb{C}$ with respect to the metric $|\ |$. As we have already seen in Example 2.3, $\mathcal{T}_\mathbb{C} \subsetneqq \mathcal{P}(\mathbb{C})$.

Whilst all metric spaces can be regarded as topological spaces not all topological spaces are metrizable. That is, given a topological space we can't always turn it into a metric space, see [10, Example 3.1.6].

An important part of a topological space are its topological subspaces, see [10, Definition 3.4.1].

Let $(Y, \mathcal{T}_Y)$ be a topological space and $X \subset Y$ a non-empty subset, then the *subspace topology* $\mathcal{T}_X$ is defined as

$$\mathcal{T}_X = \{U \cap X : U \in \mathcal{T}_Y\},$$

and $(X, \mathcal{T}_X)$ is called a *topological subspace* of $(Y, \mathcal{T}_Y)$.

The axioms (T1), (T2), and (T3) hold for $\mathcal{T}_X$.

We now consider Example 2.3 again but this time in a topological setting (we will see in Lemma 4.3 of Section 4 that $(X, \mathcal{T}_X)$ is disconnected). We have already shown that P and Q are neither open nor closed in the metric space $(\mathbb{C}, |\ |)$, so by definition P and Q are neither open nor closed in the Euclidean topology $\mathcal{T}_\mathbb{C}$. We now consider the subspace topology $\mathcal{T}_X$, where $X = P \cup Q$. Let $U = \{z :\ z = x + iy \in \mathbb{C},\ x < 0\}$ and $V = \{z :\ z = x + iy \in \mathbb{C},\ x > 0\}$, then $U,\ V \in \mathcal{T}_\mathbb{C}$ and consequently $U \cap X = P = X \setminus Q \in \mathcal{T}_X$ and $V \cap X = Q = X \setminus P \in \mathcal{T}_X$. So P and Q are both open and closed in the subspace topology. This agrees with what we found considering $(X, |\ |)$ as a metric subspace of $(\mathbb{C}, |\ |)$.

Whether we consider metric subspaces on their own or as topological subspaces, the definitions produce the same results for open and closed sets. Of course the definitions have been designed to agree in this respect, see [10, Exercise 3.9.13].

We now briefly describe continuous functions in a topological setting. For a map $f : X \to Y$ suppose $A \subset X$ then we use the notation $f(A)$ for the *image of* A, with $f(A) = \{f(a) : a \in A\} \subset Y$. For $B \subset Y$ we use the notation $f^{-1}(B)$ for the *preimage of* B, with $f^{-1}(B) = \{x : \ x \in X, \ f(x) \in B\} \subset X$. We note the following for $P, Q \subset X$ and $R, S \subset Y$,

(a) $f(P) \subset R$ if and only if $P \subset f^{-1}(R)$,
(b) $f^{-1}(R \cup S) = f^{-1}(R) \cup f^{-1}(S)$,
(c) $f^{-1}(R \cap S) = f^{-1}(R) \cap f^{-1}(S)$,
(d) $f(P \cup Q) = f(P) \cup f(Q)$,
(e) $f(P \cap Q) \subset f(P) \cap f(Q)$.

To see we don't have equality for (e) let $f : \mathbb{R} \to \mathbb{R}$ be the map $f(x) = (x-1)^2$ then if $P = [0,1]$ and $Q = [1,2]$ we have $f(P \cap Q) = f(\{1\}) = \{0\} \subsetneqq [0,1] = f(P) \cap f(Q)$.

As another example where we don't have equality for (e), let $g : [0,1] \to [0,1]$ be the doubling map defined as follows

$$g(x) = \begin{cases} 2x & \text{if } x \in \left[0, \frac{1}{2}\right], \\ 2x - 1 & \text{if } x \in \left(\frac{1}{2}, 1\right]. \end{cases}$$

Let $P = \left[0, \frac{1}{2}\right]$ and $Q = \left[\frac{1}{2}, 1\right]$ then $g(P \cap Q) = g(\{1/2\}) = \{1\} \subsetneqq (0,1] = g(P) \cap g(Q)$.

It is important to note that if f is injective then we do have equality for (e). Comparing (c) and (e), preimages are better behaved than images under set intersection in general.

In Definitions 3 and 4 that follow preimages rather than images are used in the definition of continuity. To see why this is the case consider the function $f : \mathbb{R} \to \mathbb{R}$ defined as $f(x) = 0$. Let $U \subset \mathbb{R}$ be open then $f(U) = \{0\}$ so the images of open sets are closed for f. This is not the case for the preimages of open sets. If $0 \notin U$ then $f^{-1}(U) = \emptyset$ which is open and if $0 \in U$ then $f^{-1}(U) = \mathbb{R}$ which is also open, so preimages of open sets are open for f and so, by Definitions 3 and 4, f is a continuous function.

Before giving the general definition of a continuous function between topological spaces in Definition 4, we first give three different but equivalent definitions of a continuous function between metric spaces, as it's useful to see them all in one place, see [10, Definitions 2.1.3, 2.3.6, and Proposition 2.3.13]. In the first two definitions $\delta = \delta(\varepsilon, a)$, that is δ depends on both ε and a.

1. The (ε, δ)-definition. Let (X, d_1) and (Y, d_2) be metric spaces. A map $f : X \to Y$ is *continuous at* $a \in X$ if and only if given any $\varepsilon > 0$, there exists $\delta > 0$ such that $d_1(x, a) < \delta$ implies $d_2(f(x), f(a)) < \varepsilon$.

2. The open ball definition. Let (X, d_1) and (Y, d_2) be metric spaces. A map $f : X \to Y$ is *continuous at* $a \in X$ if and only if given any $S(d_2)(f(a), \varepsilon)$, there exists $S(d_1)(a, \delta)$ such that $f(S(d_1)(a, \delta)) \subset S(d_2)(f(a), \varepsilon)$.

We say f *is continuous* when f is continuous at all points of X.

3. The open set definition. Let (X, d_1) and (Y, d_2) be metric spaces. A map $f : X \to Y$ is *continuous* if and only if U is d_2-open in Y implies $f^{-1}(U)$ is d_1-open in X.

Definition 3 generalises to a topological setting.

4. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. A map $f : X \to Y$ is *continuous* if and only if $U \in \mathcal{T}_Y$ implies $f^{-1}(U) \in \mathcal{T}_X$.

Such f is said to be $(\mathcal{T}_X, \mathcal{T}_Y)$-continuous.

The next lemma is useful and states what we expect to be the case.

Lemma 2.6. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and suppose $f : X \to Y$ is a continuous map. Let $(A, \mathcal{T}_A)$ be a topological subspace of $(Y, \mathcal{T}_Y)$ such that $f(X) \subset A \subset Y$, then $f : X \to A$ is also continuous.*

Proof. Let $U \in \mathcal{T}_A$ then, from the definition of a continuous function, we need to show that $f^{-1}(U) \in \mathcal{T}_X$. As $U \in \mathcal{T}_A$, $U = V \cap A$ for some $V \in \mathcal{T}_Y$ with $f^{-1}(V) \in \mathcal{T}_X$. It is also the case that $X \subset f^{-1}(A) \subset X$ so $f^{-1}(A) = X$. Therefore $f^{-1}(U) = f^{-1}(V \cap A) = f^{-1}(V) \cap f^{-1}(A) = f^{-1}(V) \cap X \in \mathcal{T}_X$ by (T1) and (T2). □

Given two topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ a map $f : X \to Y$ is a *homeomorphism* if and only if f is a bijection (one-to-one and onto) such that f is $(\mathcal{T}_X, \mathcal{T}_Y)$-continuous and f^{-1} is $(\mathcal{T}_Y, \mathcal{T}_X)$-continuous.

If f is a homeomorphism between $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ then it follows that $U \in \mathcal{T}_X$ if and only if $f(U) \in \mathcal{T}_Y$. So the two spaces must have the same open set structure and they are said to be *homeomorphic* or *topologically equivalent*.

It is natural to ask whether continuity in both directions is really necessary for a *homeomorphism*. One way to show that this is the case is with examples involving compact sets which we now briefly discuss.

Let $(X, \mathcal{T}_X)$ be a topological space. For $A \subset X$ any collection of open sets in X, $\{A_\alpha : \alpha \in I\}$, such that $A \subset \bigcup_{\alpha \in I} A_\alpha$, is called an *open cover* of A and $\{A_\alpha : \alpha \in I\}$ contains a *finite subcover* if $A \subset \bigcup_{\beta \in J} A_\beta$, where $J \subset I$ is a finite indexing set. A set $A \subset X$ is *compact* in X if and only if every open cover of A contains a finite subcover.

Suppose $(X, \mathcal{T}_X)$ is a topological subspace of $(Y, \mathcal{T}_Y)$ then it follows from the definitions that if X is compact in Y then it is also compact in X, that is $(X, \mathcal{T}_X)$ is a compact topological space. Compactness is inherited by subspaces. It also follows from the definitions that if X is compact in X then it is also compact in Y.

In n-dimensional Euclidean space $(\mathbb{R}^n, |\ |)$, a set is compact if and only if it is closed and bounded. This is the Heine-Borel Theorem [10, Theorem 5.7.1]. (Let (X, d) be a metric space then $A \subset X$ is *bounded* if there exists $K > 0$ such that $d(x, y) \leqslant K$ for all $x, y \in A$.) So for example $[0, 1]$ is a compact subset of $\mathbb{R}$ as it is closed and bounded. The Mandelbrot set is compact (for a proof see [1, Section 9.10]) and so closed and bounded. (In general if a metric space is closed and bounded this is not enough to imply compactness. For example $((0, 1), |\ |)$ is closed (in itself) and bounded but it is not compact. With more definitions and work it can be shown that a metric space (X, d) is compact if and only if it is complete and totally bounded, see [10, Propositions 7.2.9, 9.2.4 and Theorem 10.1.8].)

To illustrate the definitions we show that the half-open interval $[0, 1)$ is not compact. Consider $[0, 1)$ with the Euclidean subspace topology $\mathcal{T}_{[0,1)}$ as a subspace of $(\mathbb{R}, \mathcal{T}_\mathbb{R})$. As $(-1, 1 - \frac{1}{n}) \in \mathcal{T}_\mathbb{R}$, it follows from the definition of a subspace topology that $[0, 1 - \frac{1}{n}) = (-1, 1 - \frac{1}{n}) \cap [0, 1) \in \mathcal{T}_{[0,1)}$. Clearly $[0, 1) \subset \bigcup_{n \in \mathbb{N}} [0, 1 - \frac{1}{n})$ where $\left\{[0, 1 - \frac{1}{n}) : n \in \mathbb{N}\right\}$ is an open cover in $\mathcal{T}_{[0,1)}$. However it doesn't contain a finite subcover. If it did then there exists a maximum m such that $[0, 1 - \frac{1}{m})$ contains all other intervals in a finite cover, but $[0, 1 - \frac{1}{m})$ doesn't cover $[0, 1)$. So $([0, 1), \mathcal{T}_{[0,1)})$

is not compact. Since compactness is inherited by subspaces $[0,1)$ is not compact in $\mathbb{R}$ or $\mathbb{R}^2$ either.

We now return to our discussion of homeomorphisms. A property that is preserved under homeomorphisms is called a *topological invariant*. The property of being compact is a topological invariant since the continuous image of a compact set is also compact, see [10, Proposition 5.5.1]. So if two spaces are homeomorphic and one of them is compact then so must be the other. This suggests a strategy for providing examples where a map f is a continuous bijection but f^{-1} is not continuous. It is enough to find a continuous bijection $f : X \to Y$, where X is not compact but Y is, then f^{-1} can't be continuous (as this would imply X is compact). Consider $[0,1)$ and the unit circle C as subsets of $\mathbb{C}$, we have seen that $[0,1)$ is not compact in $\mathbb{C}$, whereas C is closed and bounded in $\mathbb{C}$ and so compact. Now $f : [0,1) \to C$, with $f(z) = e^{2\pi i z}$ is a continuous bijection but f^{-1} is not continuous.

In Lemma 2.7 that follows we consider a condition on metrics which ensures that two metric spaces (X, d_1) and (X, d_2) with the same underlying set X are homeomorphic. In fact we show more than topological equivalence, because we show that the open sets are actually identical, that is we show $\mathcal{T}_X(d_1) = \mathcal{T}_X(d_2)$ which is the same thing as saying that the identity map is a homeomorphism. The condition on the metrics required is known as Lipschitz equivalence.

Let (X, d_1) and (X, d_2) be two metric spaces with the same underlying set X. The metrics d_1 and d_2 are *Lipschitz equivalent* if there exist constants $h, K > 0$ such that for any $x, y \in X$

$$hd_2(x,y) \leqslant d_1(x,y) \leqslant Kd_2(x,y). \tag{2.4}$$

Lemma 2.7. *Let (X, d_1) and (X, d_2) be two metric spaces with Lipschitz equivalent metrics.*

Then

(a) *Every d_2-open ball contains a d_1-open ball and vice versa.*

(b) $\mathcal{T}_X(d_1) = \mathcal{T}_X(d_2)$.

(c) *The identity map $I : X \to X$ is a homeomorphism.*

Proof. (a) Let $x \in X$ and $r > 0$. First we show $S(d_1)(x, rh) \subset S(d_2)(x, r)$. Let $y \in S(d_1)(x, rh)$ then $d_1(x,y) < rh$ and by (2.4), $d_2(x,y) \leqslant (1/h)d_1(x,y) < r$ so $y \in S(d_2)(x, r)$.

Secondly we show $S(d_2)(x, r/K) \subset S(d_1)(x, r)$. Let $y \in S(d_2)(x, r/K)$ then $d_2(x,y) < r/K$ and by (2.4), $d_1(x,y) \leqslant Kd_2(x,y) < r$ so $y \in S(d_1)(x, r)$.

Working from the definitions it is clear that statements (a), (b) and (c) are equivalent. ☐

It is not the case that if $\mathcal{T}_X(d_1) = \mathcal{T}_X(d_2)$ then the metrics are Lipschitz equivalent. As an example consider $(\mathbb{R}^2, d_1)$ and $(\mathbb{R}^2, d_2)$ with $d_1(x,y) = \min\{1, d_2(x,y)\}$ and $d_2(x,y)$ the Euclidean metric.

Lemma 2.1 shows that the three different metrics defined in (2.1) are all Lipschitz equivalent and so the metric topologies are identical by Lemma 2.7. That is, for the metric spaces $(\mathbb{R}^n, d_1)$, $(\mathbb{R}^n, d_2)$ and $(\mathbb{R}^n, d_\infty)$, where the metrics are as defined in

(2.1) the open sets are identical with $\mathcal{T}_{\mathbb{R}^n}(d_1) = \mathcal{T}_{\mathbb{R}^n}(d_2) = \mathcal{T}_{\mathbb{R}^n}(d_\infty)$. So even though the open balls are very different in shape, as shown in Figure 2.1 for $\mathbb{R}^2$, the open sets are the same for these spaces. For problems dealing with continuity this gives some freedom in the choice of metric.

We now give a proof (see [8]) that the open unit disc and the complex plane are topologically equivalent.

Lemma 2.8. *The metric spaces* $(D, |\ |)$ *and* $(\mathbb{C}, |\ |)$ *are homeomorphic.*

Proof. Let $f : \mathbb{C} \to D$ be defined by $f(z) = z/(1 + |z|)$. Suppose $f(z) = f(w)$ then $|z|(1 + |w|) = |w|(1 + |z|)$ so $|z| = |w|$, $z = w$ and f is one-to-one. A calculation shows the inverse of f is $f^{-1}(z) = z/(1 - |z|)$ so for any $w \in D, f^{-1}(w) \in \mathbb{C}$ and f is onto. It remains to show f and f^{-1} are continuous and we will use the (ε, δ)-definition.

(a) f is continuous.

Let $z, a \in \mathbb{C}$ and let

$$B = z\,|a| - a\,|z| = \left(\frac{|z| + |a|}{2}\right)(z - a) - \left(\frac{z + a}{2}\right)(|z| - |a|)\,. \tag{2.5}$$

Then

$$\begin{aligned} |f(z) - f(a)| &= \left|\frac{z}{1 + |z|} - \frac{a}{1 + |a|}\right| = \left|\frac{z - a + z\,|a| - a\,|z|}{(1 + |z|)(1 + |a|)}\right| \\ &= \frac{|z - a + B|}{1 + |z| + |a| + |z|\,|a|} \end{aligned} \tag{2.6}$$

and using (2.5), (M3), and the backwards form of (M3) which is $||z| - |a|| \leqslant |z - a|$, we obtain

$$\begin{aligned} |z - a + B| &= \left|\left(1 + \frac{|z| + |a|}{2}\right)(z - a) - \left(\frac{z + a}{2}\right)(|z| - |a|)\right| \\ &\leqslant (1 + |z| + |a|)\,|z - a|\,. \end{aligned}$$

Using this inequality in (2.6) gives

$$\begin{aligned} |f(z) - f(a)| &= \frac{|z - a + B|}{1 + |z| + |a| + |z|\,|a|} \\ &\leqslant \left(\frac{1 + |z| + |a|}{1 + |z| + |a| + |z|\,|a|}\right)|z - a| \\ &\leqslant |z - a|\,. \end{aligned}$$

Given any $\varepsilon > 0$, putting $\delta = \varepsilon$, we have shown that $|z - a| < \delta$ implies $|f(z) - f(a)| < \varepsilon$ and since this holds for any $a \in \mathbb{C}$, f is continuous.

(b) f^{-1} is continuous.

Let $g(z) = f^{-1}(z) = z/(1 - |z|)$ and let $z, a \in D$, then

$$\begin{aligned} |g(z) - g(a)| &= \left|\frac{z}{1 - |z|} - \frac{a}{1 - |a|}\right| = \left|\frac{z - a + a\,|z| - z\,|a|}{(1 - |z|)(1 - |a|)}\right| \\ &= \frac{|z - a - B|}{(1 - |z|)(1 - |a|)} \end{aligned} \tag{2.7}$$

and using (2.5), (M3), and the backwards form of (M3), we obtain

$$\begin{aligned}|z-a-B| &= \left|\left(1-\left(\frac{|z|+|a|}{2}\right)\right)(z-a)+\left(\frac{z+a}{2}\right)(|z|-|a|)\right| \\ &\leqslant (1+|z|+|a|)\,|z-a|\,,\end{aligned}$$

and so (2.7) can be written as

$$|g(z)-g(a)| \leqslant \left(\frac{1+|z|+|a|}{(1-|z|)(1-|a|)}\right)|z-a|\,. \tag{2.8}$$

Now $a \in D$ so $|a| < 1$ and $S(a,(1-|a|)/2) \subset D$. Let $z \in S(a,(1-|a|)/2)$ then $|z| = |z-a+a| \leqslant |z-a|+|a| < (1-|a|)/2+|a| = (1+|a|)/2$. It follows that $1-|z| > 1-(1+|a|)/2 = (1-|a|)/2$ and so $1/(1-|z|) < 2/(1-|a|)$. As $z \in D$, $|z| < 1$, and (2.8) becomes

$$|g(z)-g(a)| < \left(\frac{6}{(1-|a|)^2}\right)|z-a|\,.$$

Given any $\varepsilon > 0$, putting $\delta = \varepsilon(1-|a|)^2/6$, we have shown that $|z-a| < \delta$ implies $|g(z)-g(a)| < \varepsilon$ and since this holds for any $a \in \mathbb{C}$, $g = f^{-1}$ is continuous. □

3 Closure

In this section $(X, \mathcal{T}_X)$ is a topological space so to say A is open means only $A \in \mathcal{T}_X$. Any set can be turned into a closed set by taking the closure of it as we show in Lemma 3.1(a).

The definition of a closed set in a topological space corresponds to the one for metric spaces given in Section 2.

A set $A \subset X$ is *closed in* X if and only if $X \setminus A \in \mathcal{T}_X$.

The *interior* of a set $A \subset X$, denoted by $Int(A)$, is the largest open set contained in A, that is

$$Int(A) = \bigcup_{G \in \mathcal{G}} G,$$

where $\mathcal{G} = \{G : G \subset A,\ G \text{ is open }\}$.

The *closure* of a set $A \subset X$, denoted by $Cl(A)$, is the smallest closed set containing A, that is

$$Cl(A) = \bigcap_{F \in \mathcal{F}} F,$$

where $\mathcal{F} = \{F : A \subset F \subset X,\ F \text{ is closed}\}$. As before we may write $Cl_X(A)$ to emphasise the parent space when it is not clear from the context.

A point $x \in X$ is a *limit point of* $A \subset X$ if and only if every open set U, for which $x \in U$, is such that $U \cap (A \setminus \{x\}) \neq \emptyset$. We use A' for the *set of limit points* of A. (Note: for metric spaces it is enough that $S(x,r) \cap (A \setminus \{x\}) \neq \emptyset$, for every $r > 0$).

The closure of a set can also be defined in terms of its limit points as the next lemma shows.

Lemma 3.1. *Let $A \subset X$ then*

(a) *$Cl(A)$ is closed,*
(b) *$Cl(A) = A \cup A'$,*
(c) *A is closed if and only if $A = Cl(A)$.*
(d) *$Int(A)$ is open.*
(e) *A is open if and only if $A = Int(A)$.*

Proof. (a) From the definition of closure and (2.3)

$$X \setminus Cl(A) = X \setminus \bigcap_{F \in \mathcal{F}} F = \bigcup_{F \in \mathcal{F}} (X \setminus F),$$

where $\mathcal{F} = \{F : A \subset F \subset X,\ F \text{ is closed}\}$. This is a union of open sets which is open by (T3). This proves $Cl(A)$ is closed.

(b) First we prove $Cl(A) \subset A \cup A'$. From the definition of closure $A \subset Cl(A)$ so we need only consider $x \in Cl(A) \setminus A$ and we need to show $x \in A'$. For a contradiction suppose $x \notin A'$, then there exists an open set U such that $x \in U$ and $U \cap (A \setminus \{x\}) = \emptyset$. As $x \notin A$, $U \cap A = \emptyset$ and so $X \setminus U$ is a closed set with $A \subset Cl(A) \subset X \setminus U$. However $x \notin X \setminus U$ which means $x \notin Cl(A)$ and is the required contradiction.

Now we show $A \cup A' \subset Cl(A)$. As before we need only consider $x \in A' \setminus A$ so for a contradiction suppose $x \notin Cl(A)$. From the definition of closure there must exist at least one closed set F, $A \subset F \subset X$, such that $x \notin F$. For such F, $x \in X \setminus F$ which is an open set with $(X \setminus F) \cap (A \setminus \{x\}) = \emptyset$ so $x \notin A'$ which is a contradiction.

(c) Suppose A is closed, it follows from the definition of closure that $A \subset Cl(A)$ and $Cl(A) \subset A$ (as $A \in \mathcal{F}$) so $A = Cl(A)$.

If $A = Cl(A)$ then A is closed by part (a).

(d) Any union of open sets is open by (T3).

(e) Suppose A is open, it follows from the definition of the interior that $A \subset Int(A)$ (as $A \in \mathcal{G}$) and $Int(A) \subset A$ (as $G \subset A$ for each $G \in \mathcal{G}$), so $A = Int(A)$.

If $A = Int(A)$ then from the definition of $Int(A)$, A is a union of open sets which is open. □

We conclude this section with the definitions of dense and nowhere dense sets, see [10, Definitions 3.7.20, 3.7.21].

Let $(X, \mathcal{T}_X)$ be a topological space then a subset $A \subset X$ is *(everywhere) dense* in X if and only if $Cl(A) = X$.

The set of rationals $\mathbb{Q}$ is a dense subset of $\mathbb{R}$. Since the open interval $(x, x + \varepsilon)$ contains rational numbers for any $x \in \mathbb{R}$ and any $\varepsilon > 0$, it follows from the definition of a limit point that $x \in \mathbb{Q}'$, that is x is a limit point of $\mathbb{Q}$. This means $Cl(\mathbb{Q}) = \mathbb{Q} \cup \mathbb{Q}' = \mathbb{R}$.

Let $(X, \mathcal{T}_X)$ be a topological space then a subset $A \subset X$ is *nowhere dense* in X if and only if $Int(Cl(A)) = \emptyset$.

The integers $\mathbb{Z}$ are nowhere dense in $\mathbb{R}$. This follows since $Cl(\mathbb{Z}) = \mathbb{Z}$, so $\mathbb{Z}$ is closed and clearly $Int(\mathbb{Z}) = \emptyset$. Lemma 3.3 shows that $\mathbb{R} \setminus \mathbb{Z}$ is dense in $\mathbb{R}$.

The next lemma is useful in the proof of Lemma 3.3 and provides another convenient characterization of the closure of a set.

Lemma 3.2. *Let $(X, \mathcal{T}_X)$ be a topological space with $A \subset X$ and let $x \in X$.*

Then $x \in Cl(A)$ if and only if every open set $U \subset X$, with $x \in U$, is such that $U \cap A \neq \emptyset$.

Proof. From Lemma 3.1(b), $Cl(A) = A \cup A'$. Suppose $U \subset X$ is an open set with $x \in U$. If $x \in A$ then $U \cap A \neq \emptyset$. From the definition of a limit point, if $x \in (A' \setminus A)$ then $U \cap (A \setminus \{x\}) = U \cap A \neq \emptyset$. This proves that if $x \in Cl(A)$ then $U \cap A \neq \emptyset$.

For the converse suppose every open set $U \subset X$, with $x \in U$, is such that $U \cap A \neq \emptyset$. If $x \in A \subset Cl(A)$ there is nothing to prove. Suppose then that $x \notin A$, then $U \cap A = U \cap (A \setminus \{x\}) \neq \emptyset$ and $x \in A' \subset Cl(A)$. □

The next lemma is [10, Proposition 3.7.28].

Lemma 3.3. *Let $(X, \mathcal{T}_X)$ be a topological space then a subset $A \subset X$ is nowhere dense in X if and only if $X \setminus Cl(A)$ is dense in X.*

Proof. Let $x \in X$. Then $Int(Cl(A)) = \emptyset$ if and only if every open set $U \subset X$, with $x \in U$, is such that $U \cap (X \setminus Cl(A)) \neq \emptyset$ if and only if $x \in Cl(X \setminus Cl(A))$. The last if and only if follows by Lemma 3.2 and so $Int(Cl(A)) = \emptyset$ if and only if $X = Cl(X \setminus Cl(A))$. □

4 Connected and totally disconnected sets

Here is the definition of a connected set in a topological setting, see [10, Definition 6.2.2, Proposition 6.2.3].

A topological space $(X, \mathcal{T}_X)$ is *disconnected* if and only if there exist non-empty disjoint open subsets $A, B \subset X$ such that $X = A \cup B$. We call such sets $\{A, B\}$ a *partition* of X.

$(X, \mathcal{T}_X)$ is *connected* if it is not disconnected.

An important aspect of this definition is that a set is disconnected with respect to its own topology, not to the topology of a parent space. So for $X \subset Y$ with $\mathcal{T}_X$ the subspace topology, $(X, \mathcal{T}_X)$ is disconnected if and only if there exist non-empty disjoint subsets, $A, B \in \mathcal{T}_X$, such that $X = A \cup B$. A subspace $(X, \mathcal{T}_X)$ is *connected* if it is not disconnected.

The topological subspace $(X, \mathcal{T}_X)$ of Example 2.3 is disconnected as $\{P, Q\}$ partitions X.

The next lemma shows that the continuous image of a connected set is connected. Connectedness is a topological invariant.

Lemma 4.1. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and suppose $f : X \to Y$ is continuous. If X is connected then so is $f(X)$.*

Proof. Suppose $f(X)$ is disconnected with $\{A, B\}$ a partition. Then $f(X) = A \cup B$, $A \cap B = \emptyset$ with $A, B \in \mathcal{T}_{f(X)}$. Now $f : X \to f(X)$ is continuous, by Lemma 2.6, so $X = f^{-1}(A) \cup f^{-1}(B)$ with $f^{-1}(A), f^{-1}(B) \in \mathcal{T}_X$ and $\emptyset = f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$. That is $\{f^{-1}(A), f^{-1}(B)\}$ partitions X. □

The next lemma is useful and follows from the definitions.

Lemma 4.2. *Let $(Y, \mathcal{T}_Y)$ be a topological space and $(X, \mathcal{T}_X)$ a disconnected topological subspace with $\{A, B\}$ a partition of X.*

Then there exist $U, V \in \mathcal{T}_Y$ such that
(a) $A = U \cap A$, $B = V \cap B$,
(b) $A \cap V = B \cap U = \emptyset$.

Proof. From the definition of a subspace topology, as $A, B \in \mathcal{T}_X$ there exist $U, V \in \mathcal{T}_Y$ such that $A = U \cap X$ and $B = V \cap X$. First we prove (b).

(b) Since $X = A \cup B$ and $A \cap B = \emptyset$, we obtain $\emptyset = A \cap B = A \cap (V \cap X) = A \cap (V \cap (A \cup B)) = A \cap V$ and similarly $B \cap U = \emptyset$.

(a) As $X = A \cup B$ and $A \cap B = \emptyset$, it follows that $A = U \cap X = U \cap (A \cup B) = U \cap A$ using (b), and similarly $B = V \cap B$. □

The next lemma [10, Corollary 6.2.4] can be used to prove that $X = P \cup Q$ in Example 2.3 is disconnected as P and Q are both open and closed in X.

Lemma 4.3. *A topological space $(X, \mathcal{T}_X)$ is connected if and only if the only subsets of X that are both open and closed are $\emptyset$ and X.*

Proof. First suppose $\emptyset$ and X are the only subsets of X that are both open and closed. Assume $(X, \mathcal{T}_X)$ is disconnected then $X = A \cup B$, for non-empty $A, B \in \mathcal{T}_X$, $A \cap B = \emptyset$ with $X \setminus A = B$ and $X \setminus B = A$, so A and B are both open and closed in X with $\{A, B\} \neq \{\emptyset, X\}$. This contradiction proves $(X, \mathcal{T}_X)$ is connected.

For the converse let $(X, \mathcal{T}_X)$ be connected. Suppose there exists $A \subset X$, $A \neq \emptyset$, $A \neq X$, and that A is both open and closed. Then $X = A \cup (X \setminus A)$ where A and $X \setminus A$ are non-empty open and disjoint and so $(X, \mathcal{T}_X)$ is disconnected. This contradiction proves that the only subsets of X that are both open and closed are $\emptyset$ and X. □

We now relate the partition of a disconnected subspace to its parent space.

Lemma 4.4. *Let $(Y, \mathcal{T}_Y)$ be a topological space and $(X, \mathcal{T}_X)$ a disconnected topological subspace with $\{A, B\}$ a partition of X.*

Then
(a) $Cl_Y(A) \cap B = A \cap Cl_Y(B) = \emptyset$,
(b) *If X is closed in Y then so are A and B,*
(c) *If X is open in Y then so are A and B.*

Proof. (a) $Cl_Y(A) \cap B = \emptyset$.

There exists $V \in \mathcal{T}_Y$ such that $B = V \cap X \in \mathcal{T}_X$ and $A \cap V = \emptyset$, by Lemma 4.2(b). It follows that $A \subset Y \setminus V$, where $Y \setminus V$ is closed. For a contradiction suppose $Cl_Y(A) \cap B \neq \emptyset$, then there exists $x \in Cl_Y(A) \cap B$. As $x \in B = V \cap X$, $x \in V$. From the definition of closure in Section 3, $x \in F$ where F is any closed set in Y with $A \subset F$, so $x \in Y \setminus V$. This is a contradiction.

(b) *A is closed in Y.*

As X is closed in Y, $X = Cl_Y(X)$ by Lemma 3.1(c). From the definition of closure $A \subset X$ implies $Cl_Y(A) \subset Cl_Y(X)$ and so $Cl_Y(A) \subset X = A \cup B$. By part(a) $Cl_Y(A) \subset A$ which means $A = Cl_Y(A)$ and A is closed in Y by Lemma 3.1(c).

(c) *A is open in Y.*

By Lemma 3.1(a), $Cl_Y(B)$ is closed in Y, so $W = Y \setminus Cl_Y(B)$ is open with $W \cap B = \emptyset$. By part(a) $A \subset W$. Now $W \cap X = W \cap (A \cup B) = W \cap A = A$. As $W \cap X \in \mathcal{T}_Y$ by (T2), A is open in Y. □

We can see Lemma 4.3(a) holds for the partition $\{P, Q\}$ of Example 2.3 as clearly $Cl_{\mathbb{C}}(P) = P \cup \{0\}$ so $Cl_{\mathbb{C}}(P) \cap Q = \emptyset$. As $X = P \cup Q$ is neither open nor closed in $\mathbb{C}$, parts (b) and (c) don't apply.

We have written out the proof of the next lemma in full just to show that the definition of a subspace topology carries through as expected.

Lemma 4.5. *Let $(X, \mathcal{T}_X)$ be a connected subspace of a topological space $(Y, \mathcal{T}_Y)$ and suppose $X \subset K \subset Y$ where $(K, \mathcal{T}_K)$ is disconnected with partition $\{A, B\}$. Then either $X \subset A$ or $X \subset B$.*

Proof. By Lemma 4.2 there exist $U, V \in \mathcal{T}_Y$ such that $A = U \cap A$, $B = V \cap B$ and $A \cap V = B \cap U = \emptyset$.

First we show $X \cap A = X \cap U$. Clearly $X \cap A = X \cap (U \cap A) \subset X \cap U$. Suppose $x \in X \cap U$ with $x \notin A$. It follows that $x \in X \cap B$, but $B \cap U = \emptyset$ and this contradiction implies $X \cap U \subset X \cap (U \cap A) = X \cap A$. In the same way $X \cap B = X \cap V$.

If both $X \cap A$ and $X \cap B$ are non-empty, then $X = (X \cap A) \cup (X \cap B) = (X \cap U) \cup (X \cap V)$ and X is disconnected. It follows that one of $X \cap A$ and $X \cap B$ is empty and so either $X \subset A$ or $X \subset B$. □

The next lemma is [10, Proposition 6.2.18].

Lemma 4.6. *Let $(Y, \mathcal{T}_Y)$ be a topological space and $(X, \mathcal{T}_X)$ a connected topological subspace. If $X \subset K \subset Cl_Y(X)$ then $(K, \mathcal{T}_K)$ is connected.*

Proof. For a contradiction assume $(K, \mathcal{T}_K)$ is disconnected with $\{A, B\}$ a partition of K. By Lemma 4.5 it follows that either $X \subset A$ or $X \subset B$. Without loss of generality we may assume $X \cap B = \emptyset$ with $X \subset A$ and so from the definition of closure $Cl_Y(X) \subset Cl_Y(A)$. As B is non-empty and $B \subset K \subset Cl_Y(X) \subset Cl_Y(A)$ it follows that $B \cap Cl_Y(A) \neq \emptyset$ which is impossible by Lemma 4.4(a). □

The next lemma, which is [10, Proposition 6.2.15], says nothing about intersections being connected. Consider a triangle, let two sides be the set A and the third B, then A and B are connected with non-empty intersection, so $A \cup B$ (the whole triangle) is connected. However $A \cap B$ consists of two distinct points and is disconnected. Also consider three sets consisting of the three sides of a triangle, the pairwise intersections consist of single points, and the intersection of all three is empty but their union is still connected.

Lemma 4.7. *Let $\{X_\alpha : \alpha \in I\}$ be a collection of connected subspaces of a topological space $(Y, \mathcal{T}_Y)$ for some indexing set I, and suppose for all $\alpha, \beta \in I$, $X_\alpha \cap X_\beta \neq \emptyset$. Then $X = \bigcup_{\alpha \in I} X_\alpha$ is a connected subspace of $(Y, \mathcal{T}_Y)$.*

Proof. Suppose X is disconnected and $\{A, B\}$ is a partition for X. Now $X_\alpha \subset X = A \cup B$. By Lemma 4.5, either $X_\alpha \subset A$ or $X_\alpha \subset B$. Suppose there exist $X_\alpha \subset A$ and $X_\beta \subset B$ then $X_\alpha \cap X_\beta = \emptyset$ which can't happen either. It follows that one of A or B is empty which means $\{A, B\}$ is not a partition for X. This contradiction proves X is connected. □

If not explicitly stated the Euclidean topology $(\mathbb{R}^n, \mathcal{T}_{\mathbb{R}^n})$ is always assumed from now on.

We now consider the connectedness of subsets of the real line. It can be shown that X is a connected subspace of $\mathbb{R}$ if and only if X is an interval, see [10, Theorems 6.2.7, 6.2.8], but we aim to prove only that the interval $[0, 1] \subset \mathbb{R}$ is connected.

We use the conventional notation for a *real interval*, so for $a, b \in \mathbb{R}$, $(a, b] = \{x : x \in \mathbb{R}, a < x \leqslant b\}$ and so on.

The completeness axiom for $\mathbb{R}$ ensures that any non-empty set of real numbers, $A \subset \mathbb{R}$, that is bounded above, has a *least upper bound* or *supremum* which we denote by $\sup A$. Specifically for a non-empty set $A \subset \mathbb{R}$, that is bounded above, there exists $\sup A \in \mathbb{R}$, such that

$$\begin{aligned} &\text{(i)} \quad a \leqslant \sup A, \text{ for all } a \in A, \\ \text{and } &\text{(ii)} \quad \text{if } b \text{ is an upper bound for } A, \text{ then} \sup A \leqslant b. \end{aligned}$$

The *greatest lower bound* or *infimum* is defined in a similar way and is denoted by $\inf A$.

If $A \subset \mathbb{R}$ with $\sup A \notin A$ then it follows from the definitions that $\sup A$ is a limit point of A and similarly for $\inf A$. This means that it is always the case that $\inf A, \sup A \in Cl_{\mathbb{R}}(A)$ by Lemma 3.1(b). We use this fact in the proof of the next theorem.

Theorem 4.8. *The interval* $[0, 1] \subset \mathbb{R}$ *is connected.*

Proof. For a contradiction suppose $[0, 1]$ is disconnected with $\{A, B\}$ a partition and, relabelling if necessary, let $0 \in A$.

(a) *Either* $\inf B \in A$ *or* $\inf B \in B$.

This follows since $\inf B \in Cl_{\mathbb{R}}(B) \subset Cl_{\mathbb{R}}([0, 1]) = [0, 1]$. The last equality follows by Lemma 3.1(c) as $[0, 1]$ is closed.

(b) $\inf B \in Cl_{\mathbb{R}}(A)$ *and* $\inf B \in Cl_{\mathbb{R}}(B)$.

We only need to show $\inf B \in Cl_{\mathbb{R}}(A)$. If $\inf B = 0$ then $Cl_{\mathbb{R}}(B) \cap A = \{0\}$, this contradicts Lemma 4.4(a) and so $0 < \inf B$. As $[0, 1] = A \cup B$ and $A \cap B = \emptyset$, it must be the case that $[0, \inf B) \subset A$ and so $\inf B \in Cl_{\mathbb{R}}([0, \inf B)) \subset Cl_{\mathbb{R}}(A)$.

Statements (a) and (b) contradict Lemma 4.4(a), so $[0, 1]$ is connected. ☐

This proof doesn't carry through for $\mathbb{Q} \cap [0, 1]$ because $\inf B$ may not be in $\mathbb{Q} \cap [0, 1]$, so statement (a) may not hold. In fact $\mathbb{Q} \cap [0, 1]$ is totally disconnected as we see below.

As described in [10, Section 6.5] a topological space can always be expressed as a union of connected components using an equivalence relation. Let $(X, \mathcal{T}_X)$ be a topological space then we can define an equivalence relation $\sim$ on X as follows. For $x, y \in X$,

$$x \sim y \textit{ if and only if } x, y \in C \textit{ for some connected subspace } C \subset X.$$

Clearly $x \sim x$ (reflexive) and $x \sim y = y \sim x$ (symmetric). To see $x \sim y$, $y \sim z$ implies $x \sim z$ (transitive), suppose $x, y \in C$ and $y, z \in C'$ for some connected subspaces C, C'. As $y \in C \cap C'$, $C \cap C' \neq \emptyset$. It follows that $x, z \in C \cup C'$ which is connected by Lemma 4.7. The equivalence relation $\sim$ partitions X either into one component in which case X is connected or else it partitions X into non-empty mutually disjoint equivalence classes which we call the connected components of X. These components can be defined explicitly as the largest connected subspaces of X. Let $x \in X$, then $C_x \subset X$, the *largest connected component containing* $\{x\}$, is defined as

$$C_x = \bigcup_{C \in \mathcal{C}} C \tag{4.1}$$

where $\mathcal{C} = \{C : x \in C \subset X, C \text{ is connected}\}$. Now $\{x\} \in \mathcal{C}$ so $x \in C_x$ and C_x is non-empty. As $x \in C$ for all $C \in \mathcal{C}$, C_x is connected by Lemma 4.7. Also by Lemma 4.6, $Cl_X(C_x)$ is connected so $Cl_X(C_x) = C_x$ and C_x is closed by Lemma 3.1(c).

To summarise, any topological space $(X, \mathcal{T}_X)$ can be expressed as the union of one or more non-empty, mutually disjoint, closed, connected components.

When all the components are single points $\{x\}$ then $(X, \mathcal{T}_X)$ is *totally disconnected*.

In Example 2.3, $X = P \cup Q$ is disconnected where P and Q are two closed (in X) connected components.

The set of rational numbers $\mathbb{Q}$ (as a subspace of $\mathbb{R}$) is totally disconnected. To see this consider $(\mathbb{Q}, \mathcal{T}_\mathbb{Q})$ as a subspace of $(\mathbb{R}, \mathcal{T}_\mathbb{R})$ with $\mathcal{T}_\mathbb{R}$ the Euclidean topology. Suppose there exists a connected component $C \subset \mathbb{Q}$ with $p, q \in C$, and $p < q$. Now consider $(C, \mathcal{T}_C)$ as a subspace of $(\mathbb{R}, \mathcal{T}_\mathbb{R})$ and let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ with $p < \alpha < q$. Let $U = \{x : \ x \in \mathbb{R}, \ x < \alpha\}$ and $V = \{x : \ x \in \mathbb{R}, \ x > \alpha\}$, then $\{C \cap U, C \cap V\}$ is a partition of C and C is not connected. It follows that all components of $\mathbb{Q}$ consist of just one point and so $\mathbb{Q}$ is totally disconnected. (If we consider $(\mathbb{Q}, \mathcal{T}_\mathbb{Q})$ on its own with $\mathcal{T}_\mathbb{Q}$ the Euclidean topology then we would still have to construct $\mathbb{R}$ in order to show it is totally disconnected).

The quadratic Julia set $J(f_c)$ for $c \notin M$ is a totally disconnected set and for $c \in M$ it is a connected set. In fact it can be shown that any $J(f_c)$ is a non-empty, uncountable, compact set, with no isolated points, see [1], [6] and [7] for more information. Having no isolated points means every point of $J(f_c)$ is a limit point. So even though $J(f_c)$ may be totally disconnected, no point is out on its own, cut off from all the others by any finite distance. Douady and Hubbard proved in [3] that the Mandelbrot set M is also a connected set.

In general for any rational map of degree $n \geqslant 2$, the Julia set J is uncountably infinite with no isolated points and either $J = \mathbb{C} \cup \{\infty\}$ or $Int(J) = \emptyset$, see [1, Theorems 4.2.1, 4.2.3, 4.2.4]. Also J is either connected or else it has uncountably many connected components, see [7, Corollaries 4.14, 4.15] and also [1, Theorem 5.7.1].

5 Path-connected sets

In this section we define a path-connected set, a path-connected set is connected but the converse is not true.

Let $(X, \mathcal{T}_X)$ be a topological space. For $a, b \in X$ a *path from a to b in X* is a continuous function $f : [0, 1] \to X$ such that $f(0) = a$ and $f(1) = b$. We say that *a and b are joined* by the path f, see [10, Definition 6.3.1].

A topological space $(X, \mathcal{T}_X)$ is *path-connected* if any two points in X can be joined by a path in X, see [10, Definition 6.3.2].

We state the next lemma without a proof that h is continuous, see [10, Lemma 6.3.3].

Lemma 5.1. *Let $(X, \mathcal{T}_X)$ be a topological space. For $a, b, c \in X$, if f is a path from a to b and g is a path from b to c, then we can construct a path h from a to c, by putting,*

$$h(x) = \begin{cases} f(2x) & \text{if } x \in \left[0, \frac{1}{2}\right], \\ g(2x-1) & \text{if } x \in \left[\frac{1}{2}, 1\right]. \end{cases}$$

The next lemma is [10, Proposition 6.4.1].

Lemma 5.2. *A path-connected topological space is connected.*

Proof. Let $(X, \mathcal{T}_X)$ be path-connected but disconnected with partition $\{A, B\}$. As A and B are non-empty there exist $a \in A$ and $b \in B$. Let f be a path that joins a to b. By Lemma 4.1, $f([0,1])$ is a connected subspace of X since $[0,1]$ with the Euclidean topology is connected by Theorem 4.8. Using Lemma 4.5 either $f([0,1]) \subset A$ or $f([0,1]) \subset B$ neither of which is possible as $f(0) = a \in A$ and $f(1) = b \in B$. This contradiction implies $(X, \mathcal{T}_X)$ is connected. □

The next lemma shows that the continuous image of a path-connected set is path-connected. Path-connectedness is a topological invariant.

Lemma 5.3. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and suppose $f : X \to Y$ is continuous. If X is path-connected then so is $f(X)$.*

Proof. First of all $f : X \to f(X)$ is continuous by Lemma 2.6. Let $c, d \in f(X)$ then there exist $a, b \in X$ such that $f(a) = c$ and $f(b) = d$. Let g be a path from a to b in X, then $f \circ g$ is a path from c to d in $f(X)$ (the composition of continuous functions is continuous). Therefore $f(X)$ is path-connected. □

The next lemma is the analogue of Lemma 4.7 but for path-connected subspaces.

Lemma 5.4. *Let $\{X_\alpha : \alpha \in I\}$ be a collection of path-connected subspaces of a topological space $(Y, \mathcal{T}_Y)$ for some indexing set I, and suppose for all $\alpha, \beta \in I$, $X_\alpha \cap X_\beta \neq \emptyset$. Then $X = \bigcup_{\alpha \in I} X_\alpha$ is a path-connected subspace of $(Y, \mathcal{T}_Y)$.*

Proof. Let $a, b \in X$, then $a \in X_\alpha$ and $b \in X_\beta$ for some $\alpha, \beta \in I$ and let $c \in X_\alpha \cap X_\beta$. As X_α and X_β are path-connected there are paths, f from a to c and g from c to b. It follows that there is a path h from a to b by Lemma 5.1. □

Let $(X, \mathcal{T}_X)$ be a topological space, much as we did for connected components, we can define an equivalence relation $\sim$ on X as follows.

For $x, y \in X$ let $x \sim y$ if and only if there is a path in X from x to y.

Clearly $x \sim x$ (reflexive) and $x \sim y = y \sim x$ (symmetric). For $x \sim y$, $y \sim z$ implies $x \sim z$ (transitive) follows by Lemma 5.1. We use the notation P_x for the equivalence class containing $x \in X$ where

$$P_x = \{y : \ y \in X, \text{ there is a path in } X \text{ from } x \text{ to } y\}, \tag{5.1}$$

and call P_x the *path-connected component of x in X*.

From the definition of the connected component C_x in (4.1) it follows that $P_x \subset C_x$ since P_x is connected by Lemma 5.2. This means that any connected component C_x can be expressed as a union of non-empty, mutually disjoint, path-connected components. Unlike connected components, which are closed, these path-connected components may not be closed in C_x. For an example see the discussion of the topologist's sine at the end of this section.

The key word in the next lemma is the word open, this is [10, Proposition 6.4.2].

Lemma 5.5. *Any non-empty connected open subset $U \subset \mathbb{R}^n$ is path-connected.*

Proof. Let $x \in U$ and let $P_x \subset U$ be as defined in (5.1) with $U = X$.

(a) *P_x is open in U.*

Let $y \in P_x$ then there is a path f from x to y in P_x. As $y \in U$ there exists $r > 0$ such that $S(y, r) \subset U$, as U is open in $\mathbb{R}^n$. Let $z \in S(y, r)$, then because the open ball $S(y, r)$ is a convex set there is a path g in $S(y, r)$ from y to z. By Lemma 5.1 there exists a path h in U from x to z so $z \in P_x$. This means $S(y, r) \subset P_x$ and P_x is open in $\mathbb{R}^n$ and hence in U.

(b) *P_x is closed in U.*

Let $y \in U \setminus P_x$, as U is open there exists $r > 0$ such that $S(y, r) \subset U$. Suppose $S(y, r) \not\subset U \setminus P_x$, then there exists $z \in S(y, r)$ such that $z \in P_x$ so there is a path from x to z. But $z \in S(y, r)$ means there is a path from z to y, so by Lemma 5.1 there is a path from x to y and $y \in P_x$. This contradiction means $S(y, r) \subset U \setminus P_x$ and $U \setminus P_x$ is open in $\mathbb{R}^n$ and hence in U. This proves P_x is closed in U.

The fact that U is connected and P_x is both open and closed in U means P_x is either $\emptyset$ or U by Lemma 4.3 and, since P_x is non-empty, $P_x = U$. That is U is path-connected. □

Returning to Example 2.3, suppose we restore the point $\{0\}$ to P and Q, so that they are closed unit discs centred at -1 and 1 respectively, with $P \cap Q = \{0\}$. Inuitively $X = P \cup Q$ is all of one piece and the easiest way to prove it is connected is to show it is path-connected. As P and Q are convex they are path-connected. So we need only consider the situation with $w, z \in X$, $w \in P$ and $z \in Q$. There are paths from w to 0 in P and from 0 to z in Q so there is a path from w to z in X and X is path-connected. Lemma 5.2 proves X is connected. Removing 0 disconnects X.

To end this section we consider an example of a connected set that is not path-connected. This is the well-known topologist's sine and can be defined as the set $K = G \cup J \subset \mathbb{R}^2$ where $G = \big\{(x, \sin(\frac{1}{x})) : 0 < x \leqslant 1\big\}$ is the graph of $y = \sin(\frac{1}{x})$ for $0 < x \leqslant 1$ and $J = \{(0, y) : y \in [-1, 1]\}$ is the closed interval on the y-axis joining the points $(0, -1)$ to $(0, 1)$, see [10, Figure 6.1] and [2, Figure 16, page 97] for illustrations. This is actually a compact topologist's sine, because the topologist's sine shown in [10, Figure 6.1] is not compact as there G is defined for $0 < x$. For a detailed proof that K is connected but not path-connected see [10, Examples 6.2.19, 6.3.5]. As G is the continuous image of $(0, 1]$, a connected interval, G is connected by Lemma 4.1. The proof that K is connected depends on showing that $Cl_{\mathbb{R}^2}(G) = G \cup J = K$, for then $G \subset K \subset Cl_{\mathbb{R}^2}(G)$, and since G is connected so is K by Lemma 4.6. Clearly G and J are disjoint path-connected components of K. Intuitively we may expect K not to be path-connected (travelling along G towards the y-axis we are never going to get there) but the proof requires quite a lot of work, see [10, Example 6.3.5].

As we pointed out earlier, connected components are closed. This is not the case for path-connected components and K provides an example of a connected set consisting of two disjoint path-connected components J and G with J closed in K and G open in K. We summarize the situation in $\mathbb{R}^2$ (with the Euclidean topology) and then in K as follows.

1. In $\mathbb{R}^2$: K and J are closed, G is neither open nor closed.

Clearly G is not open and J is closed in $\mathbb{R}^2$. As $Cl_{\mathbb{R}^2}(G) = G \cup J$, see [10, Example 6.2.19], $K = Cl_{\mathbb{R}^2}(K)$ so K is closed by Lemma 3.1(c). Similarly $Cl_{\mathbb{R}^2}(G) \neq G$ and G is not closed by Lemma 3.1(c).

2. In K: J is closed and not open, G is open and not closed.

Formally, as $K \setminus J = K \cap (\mathbb{R}^2 \setminus J) \in \mathcal{T}_K$ from the definition of a subspace topology, J is closed in K. As $K \setminus G = J$ is closed, G is open in K. If G were closed in K then $\{G, J\}$ would be a partition and K would be disconnected, so G is not closed. (Alternatively $Cl_K(G) = G \cup J$ so $Cl_K(G) \neq G$.) As G is not closed J is not open.

6 Simply connected sets

Here is a formal definition of a simply connected set from [9, Section 10.38]. A path $f : [0, 1] \to X$ is *closed* if $f(0) = f(1)$.

Let $(X, \mathcal{T}_X)$ be a topological space and let f_0 and f_1 be closed paths in X then f_0 and f_1 are X-*homotopic* if there is a continuous mapping $H : [0, 1] \times [0, 1] \to X$, such that

$$H(s, 0) = f_0(s), \quad H(s, 1) = f_1(s), \quad H(0, t) = H(0, t). \tag{6.1}$$

Now put $f_t(s) = H(s, t)$, then (6.1) defines a one-parameter family of closed curves f_t in X which connects f_0 and f_1. Here connects means we can continuously deform f_0 into f_1 (via intermediate paths f_t) whilst staying within X.

A *constant path* f is such that $f([0, 1]) = \{x\}$ for some single point $x \in X$. If f_0 is X-homotopic to a constant path f_1 then f_0 is *null-homotopic* in X.

A connected topological space $(X, \mathcal{T}_X)$ is *simply connected* if and only if every closed path in X is null-homotopic.

Roughly speaking X is simply connected if any closed path can be continuously deformed to a point, so X has no holes. As the definition is based on paths, which are continuous functions, the property of being simply connected is a topological invariant (preserved under homeomorphisms). This means, for example, that a sphere is not homeomorphic to a torus as a torus is connected but it is not simply connected.

We use the notation Σ for the Riemann sphere which can be defined as the complex plane together with an added abstract point $\{\infty\}$ with $\Sigma = \mathbb{C} \cup \{\infty\}$. It can also be regarded as the surface of a unit sphere S in $\mathbb{R}^3$ centred at the origin $(0, 0, 0)$ with $\mathbb{C} = \{(x, y, z) : (x, y, z) \in \mathbb{R}^3, \ z = 0\}$, considered as a plane in $\mathbb{R}^3$. Under stereographic projection $\mathbb{C}$ is homeomorphic to $S \setminus (0, 0, 1)$ where (0,0,1) is the pole at the top of the sphere, which is identified with $\{\infty\}$, see [1, Section 2.1] and [9, Section 13.1]. It is usual to write $A \subset \Sigma$, for a subset $A \subset \mathbb{C}$, when what is really meant is the embedding of A in Σ under stereographic projection. Also (Σ, σ) is a metric space where σ is the restriction of the Euclidean metric in $\mathbb{R}^3$ to Σ, σ is known as the chordal metric, see [9, Section 13.1] for more information.

In [9, Theorem 13.11] Rudin proves that nine different definitions of a simply connected set are equivalent, we state just three of them in the next theorem. For a non-empty open connected set $U \subset \mathbb{C}$ we have already seen in Lemma 5.5 that U is path-connected.

Theorem 6.1. *Let $U \subset \mathbb{C}$ be a non-empty connected set which is open in $\mathbb{C}$.*

Then the following three statements are equivalent

(a) *U is simply connected,*

(b) *U is homeomorphic to the open unit disc D,*

(c) *$\Sigma \setminus U$ is connected.*

We now briefly discuss the three statements of this theorem referring the reader to [9, Theorem 13.11] for a proof.

To get an idea as to why the Riemann sphere makes an appearance consider $U = \mathbb{C} \setminus Cl(D)$, then U is open and connected but not simply connected since a closed path encircling the closed unit disc $Cl(D)$ cannot be deformed to a point whilst staying in $\mathbb{C}$. However $\mathbb{C} \setminus U = Cl(D)$ is connected. Using the Riemann sphere this is not the case as $\Sigma \setminus U = Cl(D) \cup \{\infty\}$ which is not connected.

Some of the implications are much more difficult to prove than others. For example to prove (b) implies (a) is relatively straightforward as the property of being simply connected is a topological invariant and D is simply connected. However (a) implies (b) requires the Riemann Mapping Theorem for $U \neq \mathbb{C}$. (In Lemma 2.8 we proved (b) for $U = \mathbb{C}$).

To prove (a) implies (c) is also more difficult than it looks. For a contradiction suppose $X = \Sigma \setminus U$ is disconnected with $\{A, B\}$ a partition. As X is closed we know from Lemma 4.4(b) that A and B are closed, so they are disjoint compact subsets of Σ. This means there must be a strictly positive minimum distance between them, however to go on and prove that U cannot then be simply connected requires more work.

Finally we point out that Theorem 6.1 also holds if we consider $U \subset \Sigma$, $U \neq \Sigma$, as a non-empty connected set which is open in Σ. We can always rotate U on Σ if necessary so that $U \subset \Sigma \setminus \{\infty\}$, then U is homeomorphic under stereographic projection to $U \subset \mathbb{C}$.

7 Locally connected and locally path-connected sets

We now give the definition of a locally connected set and a locally path-connected set, basing these definitions on [10, Definition 6.5.5] and [7, pages 182-185].

A topological space $(X, \mathcal{T}_X)$ is *locally connected* if and only if given $x \in U \subset X$, where U is open in X, there exists a connected open set V such that $x \in V \subset U$.

A topological space $(X, \mathcal{T}_X)$ is *locally path-connected* if and only if given $x \in U \subset X$, where U is open in X, there exists a path-connected open set V such that $x \in V \subset U$.

That any topological space which is locally path-connected is locally connected follows from the definitions using Lemma 5.2. There is a proof in [7, Lemma 17.17] of the next lemma (which suggests there are examples of topological spaces that are locally connected but not locally path-connected).

Lemma 7.1. *Any compact locally connected metric space (X, d) is locally path-connected.*

This means for the Mandelbrot set M, which is compact, local connectedness is equivalent to local path-connectedness. As M is connected the next lemma shows that if M is locally connected (locally path-connected) then it is path-connected.

Lemma 7.2. *Let $(X, \mathcal{T}_X)$ be a connected topological space which is also locally path-connected.*

Then $(X, \mathcal{T}_X)$ is path-connected.

Proof. Let $x \in X$, and let P_x be the corresponding path-component as defined in (5.1).

(a) *P_x is open in X.*

Let $y \in P_x$ then $y \in X$ and as X is open there exists a path-connected open set V_y, with $y \in V_y$. For any $z \in V_y$, as there is a path from x to y in P_x and a path from y to z in V_y, there is a path from x to z in X, by Lemma 5.1, so $z \in P_x$, and $V_y \subset P_x$. This means P_x is a union of open sets and so is open in X by (T3).

(b) *P_x is closed in X.*

If $X \setminus P_x = \emptyset$ then $P_x = X$ and we are done. So now we consider X to be the disjoint union of path-connected components, one of which is P_x. It follows that $X \setminus P_x$ is a disjoint union of path-connected components which are open by part(a), so $X \setminus P_x$ is open by (T3), and P_x is closed.

Because X is connected the only subsets of X that are both open and closed are $\emptyset$ and X, by Lemma 4.3. As $x \in P_x$, $X = P_x$ and so X is path-connected. □

Here are just a few sets which illustrate some combinations of the different definitions of connectedness which we have encountered. (Locally connected is equivalent to locally path-connected in these examples).

1. *In Example 2.3, X is locally connected but not connected or path-connected.*

2. *A quadratic Julia set that is totally disconnected is not locally connected.*

3. *The topologist's sine K is connected, but not path-connected and not locally connected.*

To see that K is not locally connected we can argue as follows. If K is locally path-connected then it must be path-connected by Lemma 7.2. So K is not locally path-connected. As K is closed and bounded it is compact. It follows by Lemma 7.1 that K is not locally connected. Therefore K is not locally path-connected or locally connected.

4. *The comb space C is path-connected but not locally connected.*

The comb space, $(C, \mathcal{T}_C)$, for $C \subset \mathbb{R}^2$ with $\mathcal{T}_C$ the subspace topology, is defined by $C = J \cup N \cup M$ where J is the closed interval on the y-axis of the topologist's sine with

$$\begin{aligned} J &= \{(0, y) : y \in [-1, 1]\}, \\ N &= \bigcup_{n \in \mathbb{N}} \{(1/n, y) : y \in [-1, 1]\}, \\ M &= \{(x, -1) : x \in [0, 1]\}, \end{aligned}$$

see [2, Figure 16, page 97] for an illustration. As C is compact, not locally path-connected implies not locally connected by Lemma 7.1. To see that C is not locally path-connected consider the point $(0, 0) \in C$ and let $U = S_C((0, 0), \frac{1}{2})$, then $S_C((0, 0), r)$ is not path-connected for all $0 < r \leqslant \frac{1}{2}$ so there can't exist a path-connected open set V with $(0, 0) \in V \subset U$. Therefore C is not locally path-connected or locally connected.

5. *A Sierpiński curve is a non-empty, compact, connected, locally connected and nowhere dense subset of the complex plane, with the property that any two complementary domains are bounded by simple closed curves that are disjoint.*

An interesting class of Julia sets are Sierpiński curve Julia sets. Sierpiński curve Julia sets can be produced by some rational maps of the form $f(z) = z^2 + c/z^2$ and they are homeomorphic to the Sierpiński carpet. This provides a link between some repellers (Julia sets) and the attractor of an iterated function system (the Sierpiński carpet). More information on Sierpiński curve Julia sets can be found in Devaney and Look [4].

8 The Mandelbrot set and local connectivity

In this concluding section we discuss briefly what it might mean if the Mandelbrot set M is or is not locally connected. (What follows is based on the articles [2] and [5] where more information may be found.)

As discussed in [2, Section 8], if M is not locally connected then it is possible its boundary may have some of the properties of the topologist's sine K or the comb C. Let R be the open right half-plane $R = \{(x, y) : \ (x, y) \in \mathbb{R}^2, \ x > 0\}$, then $K, C \subset R \cup J$ where $J = \{(0, y) : y \in [-1, 1]\}$ is the closed interval on the y-axis of the topologist's sine and the comb. Let $I = int(J)$, so $I = \{(0, y) : y \in (-1, 1)\}$ is an open interval on the y-axis with $I \subset K$ and $I \subset C$. It follows that there is no path in $(R \setminus K) \cup I$ or in $(R \setminus C) \cup I$ that joins a point in $R \setminus K$ or $R \setminus C$ to a point in I. That is points in I are not reachable by paths that start from external points that lie in $(R \setminus K)$ or $(R \setminus C)$. If M is not locally connected and its boundary is similar to K then it may not be path-connected, if its boundary is like C then it may be path-connected. In either case we would expect there to be points on its boundary which are not accessible by paths starting from external points.

If M is locally connected then this is enough to imply that M is path-connected by Lemma 7.2. Now we consider one other consequence if M is locally connected. From Douady and Hubbard's paper [3] they were able to construct an (analytic) homeomorphism

$$\Psi : \Sigma \setminus Cl(D) \to \Sigma \setminus M.$$

If $f : \Sigma \to \Sigma$ is inversion in the unit circle defined as $f(z) = 1/z$, then

$$\Psi \circ f : D \to \Sigma \setminus M.$$

is a homeomorphism from D, the open unit disc (on Σ) to $\Sigma \setminus M$. It follows by Theorem 6.1 that $\Sigma \setminus M$ is simply connected and hence M is connected. (As $\Sigma \setminus M$ is connected, M is also simply connected). There is a theorem of Carathéodory, see [7, Theorem 17.14], which states that if M is locally connected then the map $\Psi \circ f$ can be extended to a continuous map

$$\widetilde{\Psi} \circ f : Cl(D) \to Cl(\Sigma \setminus M).$$

The formal definition of the boundary of a set is as follows.

Let $(X, \mathcal{T}_X)$ be a topological space and let $H \subset X$ then the *boundary of* H is $b(H) = Cl(H) \cap Cl(X \setminus H)$, [10, Definition 3.7.31].

As M is closed $b(M) = M \cap Cl(\Sigma \setminus M)$, which means $Cl(\Sigma \setminus M) = Cl(\Sigma \setminus M) \cap (\Sigma \setminus M \cup M) = \Sigma \setminus M \cup b(M)$. Similarly $Cl(D) = D \cup b(D)$ where $b(D)$ is just the unit circle.

It follows that if M is locally connected then there would be a continuous map from $D \cup b(D)$ to $\Sigma \setminus M \cup b(M)$. Specifically if g is a path that joins $g(0) \in D$ to a point $g(1) \in b(D)$ on the boundary of the unit disc, then $\widetilde{\Psi} \circ f \circ g$ is a path that joins the point $(\widetilde{\Psi} \circ f \circ g)(0) \in \Sigma \setminus M$ to the point $(\widetilde{\Psi} \circ f \circ g)(1) \in b(M)$. This means that any point on $b(M)$ would be accessible by a path from any point in $\Sigma \setminus M$. It would also mean that $b(M)$ could be parametrized using the unit circle.

To summarize:

(a) Locally connected is equivalent to locally path-connected for M by Lemmas 5.2 and 7.1.

(b) If M is not locally connected then there may be points on the boundary of M which are not accessible by paths starting from external points. In this case M may or may not be path-connected.

(c) If M is locally connected then M is path-connected by Lemma 7.2 and all the points on its boundary would be accessible by paths starting from any external points. Also the boundary could be parametrized using the unit circle.

Knowing whether or not the Mandelbrot set is locally connected would provide important information for the understanding of its geometry

References

[1] A. F. Beardon, *Iteration of Rational Functions*, Springer-Verlag, New York, (2000).

[2] B. Branner, The Mandelbrot set, *Chaos and Fractals: The Mathematics Behind the Computer Graphics*, Ed. R. L. Devaney and L. Keen, Proc. Symp. App. Math., **39** (1989), 75-105.

[3] A. Douady and J. H. Hubbard, *Itérations des polynômes quadratiques complexes*, C. R. Acad. Sc. Paris, t.294 (1982), 123-126.

[4] R. L. Devaney and D. M. Look, *A criterion for Sierpiński curve Julia sets*, Topology Proceedings, **30** (2006), 163-279.

[5] J. Ewing, *Can we see the Mandelbrot set?*, The College Math. J., **26** (1995), No.2, 90-99.

[6] K. J. Falconer, *Fractal Geometry, Mathematical Foundations and Applications*, John Wiley, Chichester, 2nd Ed. 2003.

[7] J. Milnor, *Dynamics in One Complex Variable*, Princeton University Press, Princeton and Oxford, (2006).

[8] V. Nestoridis and N. Papadatos, *An extension of the disc algebra, II*, Complex Variables and Elliptic Equations, **59**:7 (2014), 1003-1015.

[9] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, Boston, (1987).

[10] W. A. Sutherland, *Introduction to Metric and Topological Spaces*, Oxford University Press, Oxford, (2003).