Connected Sets

Sébastien Boisgérault, MINES ParisTech, under CC BY-NC-SA 4.0

January 2, 2018

Contents

Introduction

We characterize the subsets of the plane that are "in one piece". Two slightly different mathematical properties can play this role: *path-connectedness*, whose definition is quite elementary, and *connectedness*, a slightly weaker – and arguably more convoluted – property, but also a more robust and powerful one. The difference matters only when one deals with "pathological" sets; for "well-behaved" sets – and that includes all open sets – the two properties are equivalent.

In this document, we use the word "set" to mean "subset of the complex plane" because this is what we need most of the time. However, the theory still works if we interpret "set" as "subset of a given normed vector space" instead; the only adaption that is required is the replacement of open disks by open balls.

Path-Connected/Connected Sets

Definition – Path-Connected Set. A set A is *path-connected* if any pair of points of A can be *joined* by a path of A:

$$\forall\, (w, z) \in A^2,\ \exists\, \gamma \in C^0([0,1], A),\ \gamma(0) = w \ \text{ and } \ \gamma(1) = z.$$

Definition – Dilation. A set B is a *dilation* of a set A if it is the union of a collection of non-empty open disks whose centers are the points of A:

$$B = \bigcup_{a \in A} D(a, r_a) \ \text{ and } \ \forall\, a \in A,\, r_a > 0.$$

Remark – Non-Uniformity of Dilations. We borrowed the word *dilation* from mathematical morphology, but our use of the word is not completely standard. The dilation of a set A by the non-empty open disk $D(0, r)$ would be classically defined as

$$B = A + D(0, r) = \{a + b \mid a \in A,\, b \in D(0, r)\} = \bigcup_{a \in A} D(a, r).$$

By contrast, the definition that we use allows *non-uniform* dilations: the radius of the disks may change with their centers.

Definition – Connected Set. A set is *connected* if all its dilations are path-connected. A set which is not connected is *disconnected*.

Theorem – Path-Connected/Connected Set. Every path-connected set is connected. Conversely, every open connected set is path-connected.

Proof. Let A be a path-connected set and $B = \cup_{a \in A} D_a$ be a dilation of A. For any points w and z in B, there are points a and b in A such that $w \in D_a$ and $z \in D_b$. There is a path that joins w and a in D_a, a path that joins a and b in A and a path that joins b and z in D_b. The concatenation of these paths joins w and z in B, hence A is connected.

Conversely, let A be an open connected set. For any $a \in A$, the distance r_a between a and the complement of A – which is a closed set – is positive, hence the disk $D_a = D(a, r_a)$ is a non-empty subset of A and $A = \cup_{a \in A} D_a$. The set A is one of its dilations, hence it is path-connected. ■

Corollary – Open Connected Sets. An open set is connected if and only if it is path-connected.

Set Operations

Many properties of connected sets are similar to properties of path-connected sets, so many statements exist in two variants. For example:

Theorem – Union of Sets With a Non-Empty Intersection. if $\mathcal{A}$ is a collection of path-connected/connected sets whose intersection $\cap\mathcal{A}$ is non-empty, then the union $\cup\mathcal{A}$ is path-connected/connected.

Proof. For path-connected sets: let a and b in $\cup\mathcal{A}$. There are some sets A and B in $\mathcal{A}$ such that $a \in A$ and $b \in B$. The intersection $\cap\mathcal{A}$ is included in $A \cap B$, hence $A \cap B$ is not empty; let $c \in A \cap B$. There is a path of A that joins a and c and a path of B that joins c and b; their concatenation joins a and b in $\cup\mathcal{A}$. Hence, this set is path-connected.

For connected sets: let $\cup_{a\in\cup\mathcal{A}} D_a$ be a dilation of $\cup\mathcal{A}$. We have

$$\bigcup_{a\in\cup\mathcal{A}} D_a = \bigcup_{A\in\mathcal{A}} \cup_{a\in A} D_a.$$

For any $A \in \mathcal{A}$, the set $\cup_{a\in A} D_a$ is a dilation of A, hence it is path-connected; the inclusion $A \subset \bigcup_{a\in A} D_a$ provides

$$\cap\mathcal{A} = \bigcap_{A\in\mathcal{A}} A \subset \bigcap_{A\in\mathcal{A}} \cup_{a\in A} D_a,$$

hence the intersection of all $\cup_{a\in A} D_a$ over $A \in \mathcal{A}$ is not empty. We may therefore apply the result of the theorem for path-connected sets to the collection $\{\cup_{a\in A} D_a \mid A \in \mathcal{A}\}$. Our arbitrary dilation of $\cup\mathcal{A}$ is path-connected, hence $\cup\mathcal{A}$ is connected. ■

Theorem – Disjoint Union of Open Sets. If A and B are two non-empty open sets such that $A \cap B = \varnothing$, then $A \cup B$ is not path-connected/connected.

Proof. Assume that γ is a path of $A \cup B$ that joins a point $a \in A$ and a point $b \in B$. Consider the function

$$\phi : t \in [0, 1] \mapsto d(\gamma(t), A) - d(\gamma(t), B).$$

If $z = \gamma(t) \in A$, for example when $t = 0$, $d(z, A) = 0$ and as A is open and $A \cap B = \varnothing$, $d(z, B) > 0$, hence $\phi(t) < 0$. Otherwise, for example when $t = 1$, $z = \gamma(t) \in B$, $d(z, B) = 0$ and as B is open and $A \cap B = \varnothing$, $d(z, A) > 0$, hence $\phi(t) > 0$. But the function ϕ is also continuous; the intermediate value theorem asserts the existence of a $t \in]0, 1[$ such that $\phi(t) = 0$, which is a contradiction. Hence no such path γ can exist and $A \cup B$ is not path-connected; as $A \cup B$ is open, it is not connected either. ■

Connected sets also have some interesting properties that are not shared by all path-connected sets; for example:

Theorem – Closure of Connected Sets. The closure of a connected set is connected.

Proof. Let A be a connected set and let $\cup_{b\in B} D_b$ be a dilation of its closure $B = \overline{A}$. For any $b \in B$, let r_b be the distance between b and the complement of this dilation. We have

$$\cup_{b\in B} D_b = \cup_{b\in B} D(b, r_b).$$

Consider the dilation $\cup_{a\in A} D(a, r_a)$ of A. It is a clearly a subset of the dilation of B; actually, we can prove that both sets are equal. Assume that z belongs to the dilation of B: there is a $b \in B$ such that $|z-b| < r_b$. As B is the closure of A, there is a point $a \in A$ such that $|a-b| < (r_b - |z-b|)/2$; we have

$$|z-a| \leq |z-b| + |a-b| < r_b - |a-b| \leq r_a,$$

hence the point z also belongs to the dilation of A. As the dilation of A is path-connected, so is the dilation of B: B is connected. ■

The equivalent statement is false for some path-connected sets. Actually, we may leverage this difference to build a connected set which is not path-connected:

Example – The Topologist's Sine Curve. Consider

$$A = \{(x, \sin 1/x) \mid x \in]0,1]\}.$$

This set is path-connected – as the image by a continuous function of a path-connected set – hence its closure

$$\overline{A} = A \cup \{(0,y) \mid y \in [-1,+1]\}$$

is connected; however, it is not path-connected.

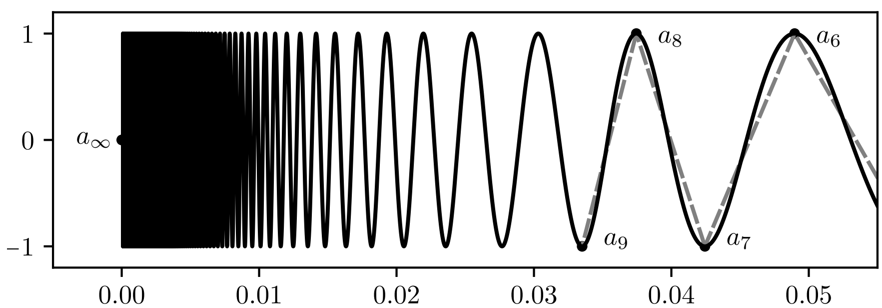

Figure 1: The Topologist's Sine Curve.

Assume on the contrary that $\gamma : [0,1] \to \mathbb{C}$ is a path of $\overline{A}$ that joins the points $a_0 = (2/\pi, 1)$ and $a_\infty = (0,0)$; it has to go through every point

$$a_n = (x_n, y_n),\ n \in \mathbb{N} \ \text{ where } \left| \begin{array}{lcl} x_n & = & 1/((n+1/2)\pi) \\ y_n & = & \sin 1/x_n = (-1)^n \end{array} \right.$$

in this specific order. Indeed, given some $t_n \in [0,1[$ such that $\gamma(t_n) = a_n$, we have $\mathrm{Re}(\gamma(t_n)) = x_n$. As $\mathrm{Re}(\gamma(1)) = \mathrm{Re}(a_\infty) = 0$, by continuity of $t \in [0,1] \mapsto \mathrm{Re}(\gamma(t))$, there is a $t_{n+1} \in]t_n, 1[$ such that $\mathrm{Re}(\gamma(t_{n+1})) = x_{n+1}$. Since for any $x > 0$, there is a unique real number y such that $(x,y) \in A$, this yields $\gamma(t_{n+1}) = a_{n+1}$. Now, since the sequence t_n is increasing and bounded from above, necessarily $|t_{n+1} - t_n| \to 0$ when $n \to +\infty$. But on the other hand, for any $n \in \mathbb{N}$,

$$\begin{aligned} |\gamma(t_{n+1}) - \gamma(t_n)| &= |a_{n+1} - a_n| \\ &\geq |y_{n+1} - y_n| \\ &= 2 \end{aligned}$$

Hence the function γ, despite being continuous and defined on the compact set $[0, 1]$ cannot be uniformly continuous, which is a contradiction.

Components

We define two concepts of components based respectively on path-connectedness and connectedness.

Definition – Component. A *(path-connected/connected) component* of a non-empty set A is a subset of A which is path-connected/connected and maximal with respect to inclusion among such sets – that is, included in no other path-connected/connected subset of A.

Theorem – Partition into Components. The path-connected/connected components of a non-empty set A are a partition of A: they are a collection of non-empty and pairwise disjoint subsets of A whose union is A.

Proof. The proof is identical for path-connected and connected components. Let $a \in A$. Consider the collection $\mathcal{A}_a$ of all connected subsets of A that contain the point a. The set $A_a = \cup \mathcal{A}_a$ is connected. By construction, the set A_a is maximal: it is a component of A. As every component of A is maximal, it contains at least one point $a \in A$: it is therefore non-empty and equal to A_a. Hence the union of all components of A is $\cup_{a \in A} A_a = A$. Finally, if two such components A_a and A_b have a non-empty intersection $c \in A$, the set $A_a \cup A_b$ is connected and contains A_a and A_b, therefore $A_a = A_b$. ■

Corollary – Connectedness & Components. A non-empty set is path-connected/connected if and only if it has a single path-connected/connected component.

Proof. If a set is path-connected/connected, it is one of its components, because it is clearly connected and maximal. As the components form a partition of the set, it is the only component. Conversely, if there is a unique component, again because the components form a partition of the set, it is the set itself, which is therefore path-connected/connected. ■

Theorem – Components of Open Sets. The partitions of a non-empty open set into path-connected components and connected components are identical. All such components are open.

Proof. Let A be an open set and let B be a path-connected component of A. For any $b \in B$, there is a non-empty open disk D centered on b which is included in A. The disk D is a path-connected subset of A that contains a; it is therefore included in the unique maximal path-connected subset of A that contains a: the set B. Therefore, B is open.

The path-connected components of A are open and path-connected, hence they are also connected. They are also maximal among the connected sets of A: a connected component of A contains a path-connected component of A if it contains a point of it; if it were to contain more than one path-connected component, it would be the union several disjoint open sets and hence could not be connected. ∎

Locally Constant Functions

Definition – Locally Constant Function. A function f defined on a set A is locally constant if for any a in A, there is a non-empty open disk D centered on a such that f is constant on $A \cap D$:

$$\forall\, a \in A,\ \exists\, \epsilon > 0,\ \forall\, b \in A,\ |b - a| < \epsilon \Rightarrow f(b) = f(a).$$

Theorem – Locally Constant Functions & Connected Sets. A set A is connected if and only if every locally constant function defined on A is constant.

Proof. Let f be a locally constant function defined on A. Let $a \in A$ and $B = \{b \in A \mid f(b) = f(a)\}$. Assume that f is not constant, that is, that $C = A \setminus B$ is non-empty. As f is locally constant, the distance between any point b of B and the set C is positive; we may define $D_b = D(b, r_b)$ where $r_b = d(b, C)/2 > 0$. We may perform a similar construction for any point c of C and define a disk $D_c = D(c, r_c)$ with $r_c = d(c, B)/2 > 0$. By construction, the sets $\cup_{b \in B} D_b$ and $\cup_{c \in C} D_c$ are non-empty, open and disjoints, hence the dilation $\cup_{a \in A} D_a$ of A is not path-connected. Therefore, A is not connected.

Conversely, if A is not connected, let $\cup_{a \in A} D_a$ be a dilation of A which is not path-connected. It has multiple (path-connected) components; let B be one of them and C be the union of all the others. Since every component is non-empty and open, B and C are both non-empty and open. Then, the function f defined by $f(z) = 1$ if $z \in B$ and $f(z) = 0$ if $z \in C$ is locally constant; however, it is not constant. ∎