The Cut Locus and the Jordan Curve Theorem

Rich Schwartz

November 19, 2015

1 Introduction

A *Jordan curve* is a subset of $\boldsymbol{R}^2$ which is homeomorphic to the circle, S^1. The famous Jordan Curve Theorem says that, for any Jordan curve Γ, the complement $\boldsymbol{R}^2 - \Gamma$ has two path connected components. Some years ago, Helge Tverberg wrote a paper which gives a very clean elementary proof. See

H. Tverberg, *A Proof of the Jordan Curve Theorem*,
Bulletin of the London Math Society, 1980, pp 34-38.

Tverberg's paper hardly leaves anything to be desired, except that perhaps it is too terse for an undergraduate reader. Probably like a lot of people, I noticed that Tverberg's main technical lemmas could be formulated in terms of the *cut locus* of a polygon. I'm sure that this is what Tverberg had in mind, though he doesn't discuss it explicitly. So, I thought I would explain this point of view in some notes, while following Tverberg's general approach.

For the JCT, the first step is to prove it for polygons. This is easy, and I will just sketch the proof. I'll also just sketch the proof that an arbitrary Jordan curve can be approximated arbitrarily well by a sequence of embedded polygons. Tverberg has a beautiful construction for this, which I'll explain lightly and briefly.

The main difficulty in finishing the proof of the JCT is deducing the general case from a sequence of polygonal cases. This seems obvious at first, but that is where all the work resides.

There are, of course, many other approaches to the JCT. Various topological machines, like the Meyer-Vietoris sequence for reduced homology, kill it like a truck smashing into a butterfly.

2 The Polygonal Jordan Curve Theorem

Here is a quick sketch of this. Let Γ be a polygon in the plane. Given two points $p, q \in \Gamma$, consider a polygonal path α which joins p to q. Assume that α avoids the vertices of Γ and intersects each edge it meets at a definite angle. Let $N(p, q, \alpha)$ be the number of intersection points, mod 2. If you wiggle α you will see that the possible intersection points cancel or are created in pairs. Hence $N(p, q)$ makes sense, independent of the choice of α. Figure 1 shows how this works.

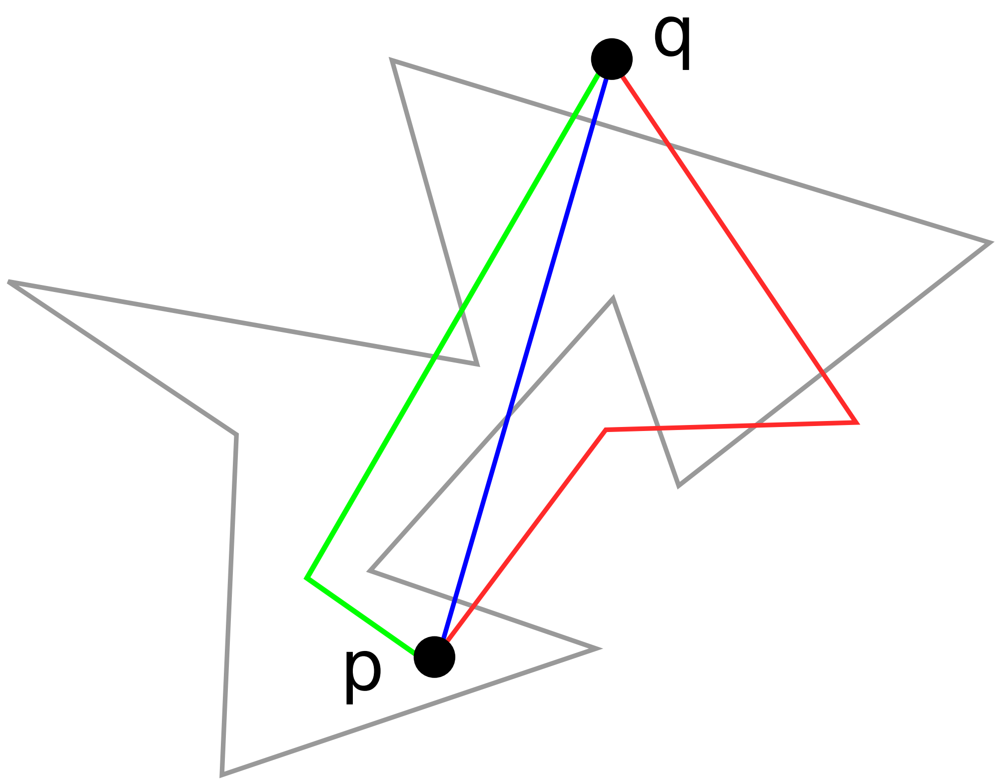

Figure 1: $N(p, q)$ defined

Note that $N(p, q) + N(q, r) = N(p, r)$ mod 2. If $N(p, q) = 1$ then these points cannot be joined by a path which avoids Γ. If $N(p, q) = 0$ then Figure 2 shows how one can reduce the number of intersections of a path α joining p to q. Repeating this trick finitely many times, we get a path connecting p to q which avoids Γ. Hence there are exactly two connected components.

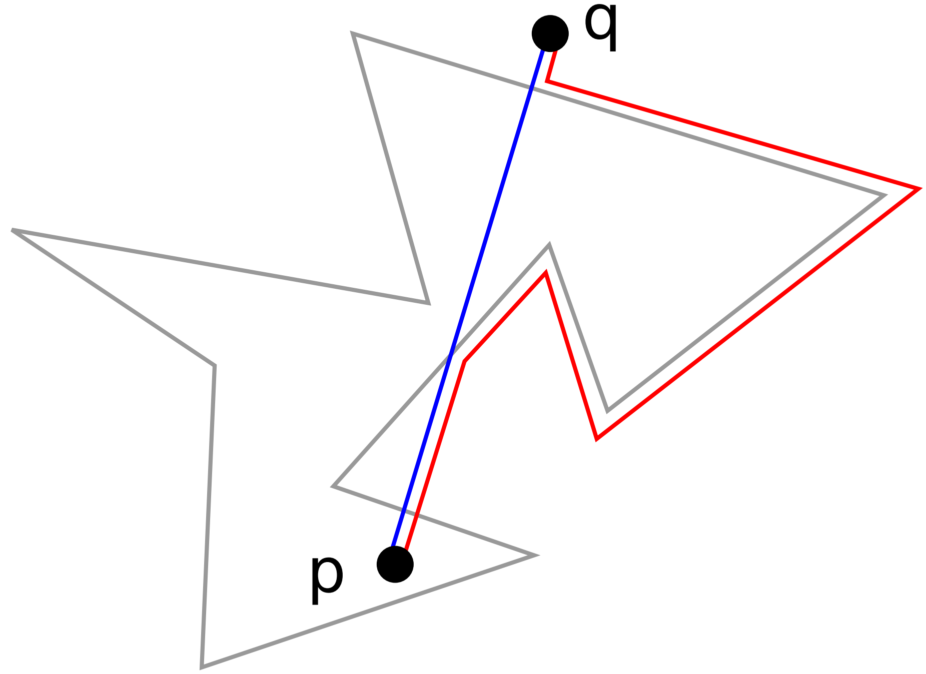

Figure 2: Reducing the number of intersections by 2.

2.1 Polygonal Approximation

The naive approach of just joining together points along the curve will probably not produce an *embedded* approximation - at any scale! Here's Tverberg's trick. Actually, I'll modify the construction to make it more clear that it works. Let Γ be a Jordan curve. Superimpose over Γ a tiling of the plane by squares so that no vertex of a tile square lies on Γ. Color the tiles red and blue, in a checkerboard pattern. Order the red squares. Let R be the first red square. There is some maximal arc of Γ whose endpoints are on edges of R. Replace this arc by a line segment. The result is still embedded. Now do this for the next red square. And so on. When you are done, the approximation is a polygon in each red square and it only enters and exits each blue square once. Now replace all the arcs in blue squares by segments. Figure 3 shows the first half of this, where we modify things according to the red squares. The second half is pretty obvious.

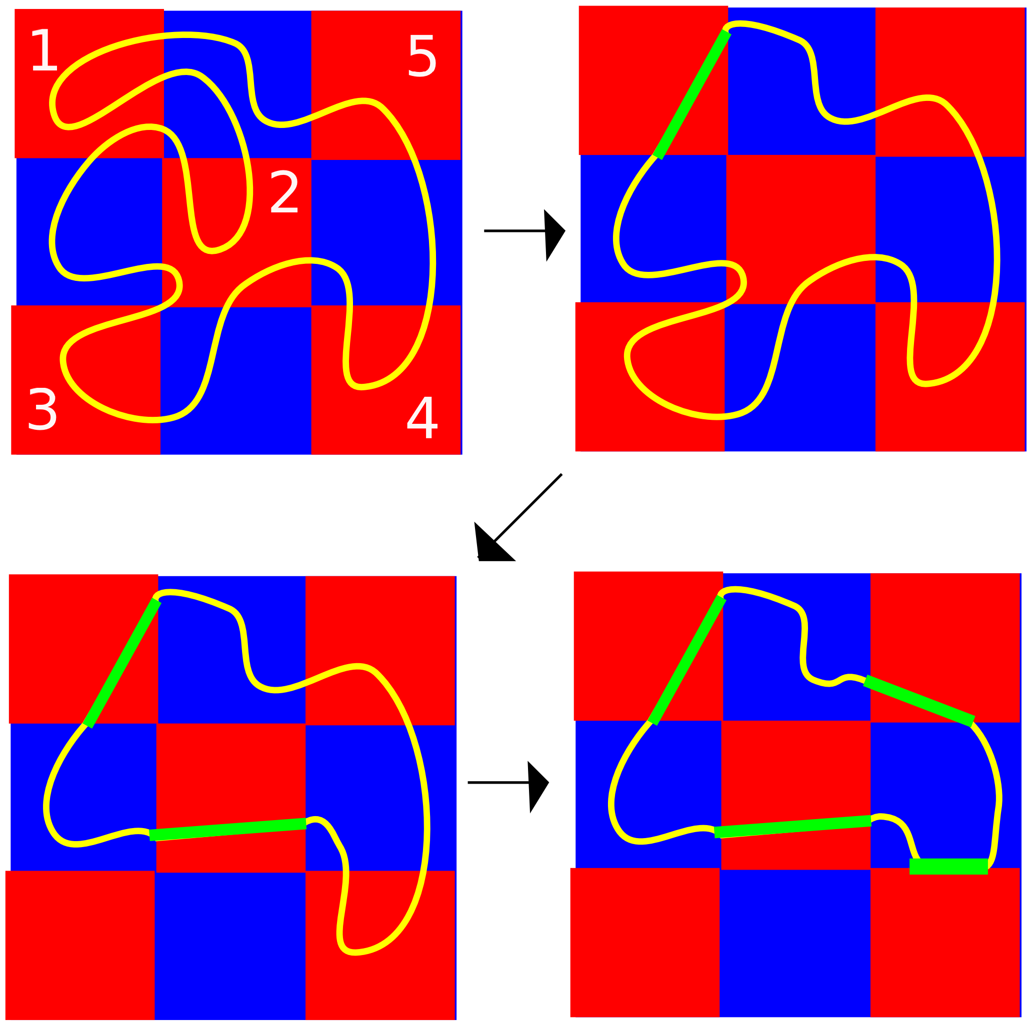

Figure 3: Approximating a Jordan Curve by a polygon

Figure 4 shows how this is done with a very crude mesh. If you use a finer mesh, the approximation is better. There is some amount of fussing to make sure that you can get an arbitrarily good approximation by taking an arbitrarily fine mesh. The point is that, after the red modifications are done, every part of the original curve is close to a part which has replaced it. Then, the blue modifications retain this property.

2.2 The Cut Locus

Let P be region in the plane whose boundary, ∂P, is a polygon. Figure 1 shows a typical example. I'll call P a *polygonal region*.

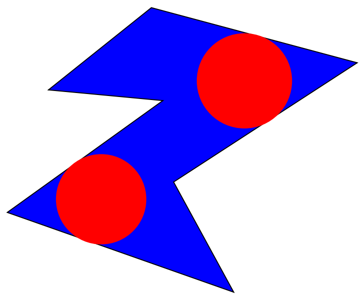

Figure 4: A Polygon with 2 maximal disks drawn

A *maximal disk* in P is a disk $D \subset P$ which contains at least 2 points of ∂P. The *cut locus* is defined to be the union of all the centers of the maximal disks. Figure 2 shows a rough picture of the cut locus of the polygon in Figure 1, as well as some indication of how I drew it. It would be better to have a computer program to draw the thing!

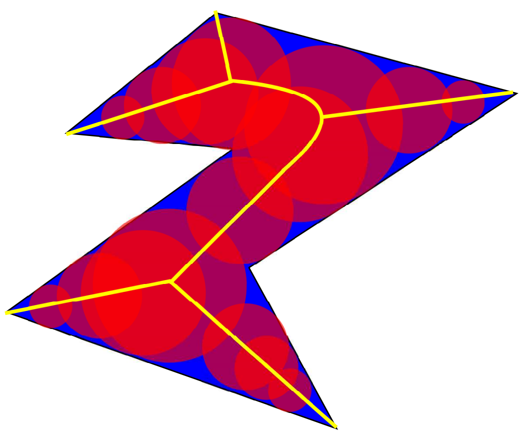

Figure 5: Examples of the cut locus

Here is the main theorem.

Theorem 2.1 *The cut locus of a polygon is a tree whose individual edges are either straight line segments or arcs of parabolas.*

The proof has three parts. The first part is a case-by-case local analysis of the cut locus. The second part is a proof that the cut locus has no loops. The third part is the proof that the cut locus only has one path component.

3 Facts about Circles

Here we prove some facts about circles.

Lemma 3.1 *There is at most one circle through* 3 *distinct points.*

Proof: Any two circles intersect in at most two points. ♠

Lemma 3.2 *Suppose that C and C' are two circles tangent to* 3 *distinct lines. Then there is some $\epsilon > 0$ such that the centers of C and C' are at least ϵ apart.*

Proof: Every pair of intersecting lines defines two angle bisectors. The circles C and C' must be centered at the intersection of a pair of these angle bisectors. But then there are only finitely many choices for the centers. If the centers of C and C' are too close, they must therefore coincide. But then there is only one circle centered at some point and tangent to some line. This would force $C = C'$. ♠

Lemma 3.3 *Suppose that C and C' are two circles tangent to* 2 *distinct lines and containing a third point. Then there is some $\epsilon > 0$ such that the centers of C and C' are at least ϵ apart.*

Proof: Let L_1 and L_2 be the two lines and let p be the point. The set of points equidistant from L_1 and p is a parabola. The set of points equidistant from L_1 and L_2 is the pair of angle bisectors. Hence the centers of C and C' must lie on the intersection of a parabola with a pair of lines. Again there are only finitely many choices. The rest of the proof is the same as in the previous case. ♠

Lemma 3.4 *Suppose that C and C' are two circles tangent to the same line and contining two distinct points not on the line. Then there is some $\epsilon > 0$ such that the centers of C and C' are at least ϵ apart.*

Proof: Same idea as in the previous cases. ♠

4 The Cut Locus is a Graph

Let C denote the cut locus and let $x \in C$ be some point. There is some maximal disk Δ such that x is the center of Δ. Note that x cannot be the center of two maximal disks. The larger one would slop over the boundary. So, Δ is unique. There are finitely many points $p_1, ..., p_m \in \partial P \cap \partial\Delta$. We will consider several possibilities in turn.

We will first consider the case when $m = 2$. There are three possibilities.

Case A: Suppose first that $m = 2$ and p_1, p_2 both belong to interiors of edges e_1 and e_2 in P. Figure 6a shows this case. In this case, x lies on the center of the line L which is the angle bisector of the angle made by the two lines L_1 and L_2 extending e_1 and e_2. All points on this line are equidistant, so the cut locus contains a neighborhood of L about x. Conversely, any point of the cut locus sufficiently close to x must lie on L.

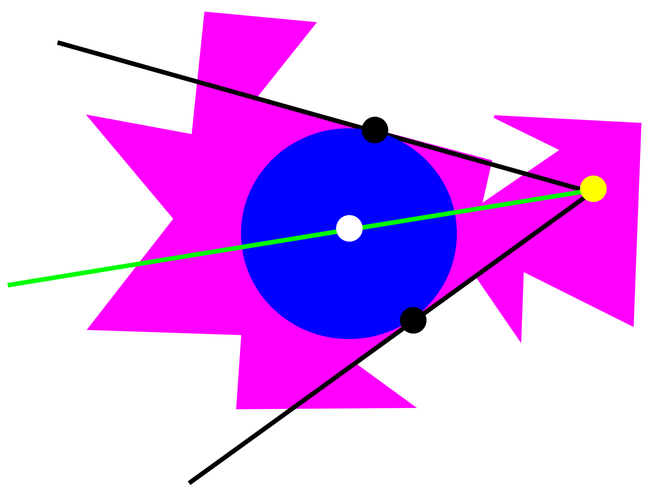

Figure 6a: The edge-edge case

Case B: Suppose that p_1 and p_2 are both vertices of ∂P. Figure 4 shows this case. All the nearby points in C come from circles which contain these same two vertices, and the centers of these circles are equidistant from p_1 and p_2. So, again in this case, the cut locus is a line segment in a neighborhood of x.

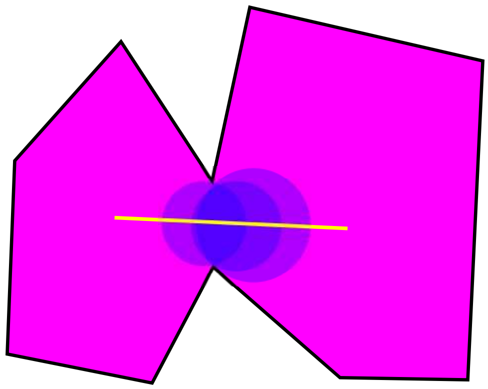

Figure 6b: The vertex-vertex case.

Case C: Suppose that p_1 is a vertex of P and p_2 lies on an edge e_2 of P. In this case, all the maximal disks near x will be equidistant from the line L_2 containing e_2 and p_1. The set of such points is a parabola. Hence, in a neighborhood of x, the cut locus agrees with an arc of a parabola. Figure 6c shows this situation.

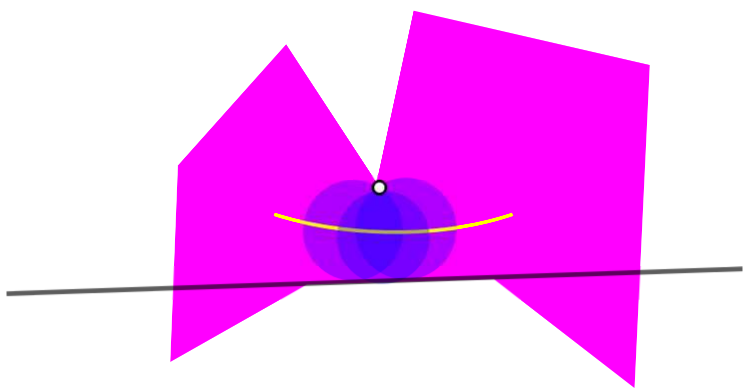

Figure 6c: The edge-vertex case

Consider now the general case. Let $p_1, ..., p_m$ be the points contained $\partial\Delta$, the maximal disk centered at x. There is some k such that $p_1, ..., p_k$ are on the interiors of edges $e_1, ..., e_k$ and the remaining points are vertices of ∂P.

Let Δ' be some maximal disk whose center is very close to x. Then $\partial\Delta' \cap \partial P$ must be contained in the union

$$e_1 \cup ... \cup e_k \cup p_{k+1} \cup ... \cup p_n.$$

But then, from the facts about circles, $\partial\Delta' \cap \partial P$ can only contain two points. So, Δ' fits into one of the special cases already analyzed. This shows that a neighborhood of x in the cut locus is contained in a finite union of arcs of parabolas and line segments.

We want to see that a neighborhood of the cut locus about x is precisely such a union. We are worried about the following situation: Suppose that Δ' and Δ'' are maximal disks centered very close to x, and the centers x' and x'' lie on the same line segment or parabolic arc. Call this segment-or-arc S. We worry that perhaps not all points of S between x' and x'' lie in the cut locus. Let's rule this out.

Consider the case when Δ' and Δ'' are both tangent to edges in our collection. Then any disk Δ''' centered on S between x' and x'' and tangent to the same two edges lies in the convex hull of $\Delta' \cup \Delta''$. Hence, such a disk is contained in P and its center is part of the cut locus. The same remarks apply to the edge-vertex case.

In the vertex-vertex case the disk Δ''' is actually contained in $\Delta' \cup \Delta''$. Have a good look at Figure 4. So again, the center of Δ'' lies in the cut locus. This takes care of our worry.

Our local analysis tells us that C has the structure of a finite graph, whose edges are either line segments or arcs of parabolas. The vertices of the graph are the points where the associated maximal circles contain more than 2 points of ∂P on their boundaries.

5 The Cut Locus has No Loops

Here is a proof that the cut locus has no loops. Each point $p \in P - C$ is the center of a disk Δ_p which touches P at exactly one point. Call this point q. Let s be the segment which connects p to q.

Lemma 5.1 *$s \cap P = \emptyset$.*

Proof: For the sake of contradiction suppose there is some point $p' \in s \cap C$. Let Δ' be the maximal disk centered at p'. Note that s is contained in a diameter of Δ and Δ' simultaneously. Since p' is closer to q than p, we see that $\Delta' \subset \Delta$. Otherwise q would be in the interior of Δ'. But then Δ' only intersects ∂P at q. This contradicts the fact that Δ' is a maximal disk. ♠

Corollary 5.2 *Every point of $P - C$ can be joined to ∂P by a line segment which remains in $P - C$.*

Suppose now that C has a loop L. By the Jordan Curve Theorem for polygons, $\boldsymbol{R}^2 - L$ has two path connected components. One of these components is disjoint from ∂P. Call this component A. From the corollary, every point of A can be joined to ∂P by a line segment which avoids C. In particular, this segment avoids L. But then we have joined two points in different path components of $\boldsymbol{R}^2 - L$. This is a contradiction. Hence, L cannot exist. There are no loops in the cut locus.

6 The Cut Locus is a Tree

We know from the previous section that the cut locus is a finite union of trees. We want to see that the cut locus is just a single tree. For the moment we will consider one tree component of the cut locus and eventually we will see that it must be the only component. Let $T \subset C$ be a tree component.

Say that a *maximal path* of T is a path which connects one vertex of T to another. The vertices of T must be vertices of ∂P. Technically, these points to not belong to T, but rather to the closure of T. Again by the Jordan Curve for Polygons, each maximal path separates P into two components. There is at least one maximal path $S \subset T$ which has the property that the rest of T lies in one of the components of $P - T$. We call such a path a *peripheral path*. Figure 7 shows an example.

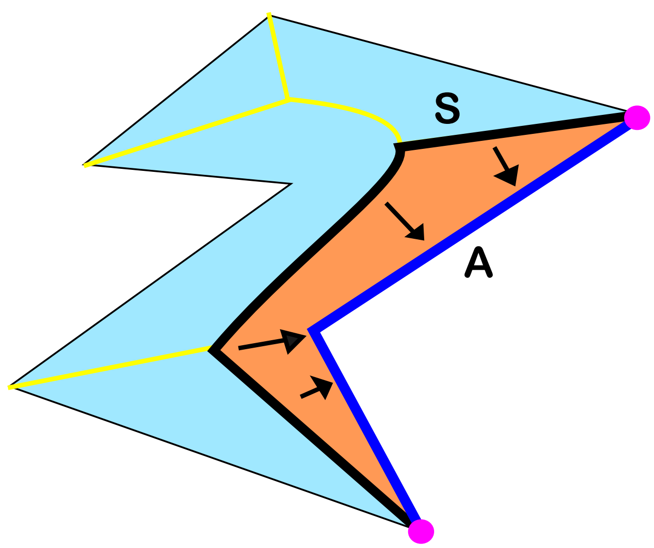

Figure 7: a peripheral path (in black).

We say that the *empty component* is the component of $P - S$ that does not contain $T - S$. The empty component is colored orange in Figure 7. The boundary of the empty component is $S \cup A$, where $A \subset \partial P$ is an arc joining the endpoints of S. The arc A is shown in blue.

Lemma 6.1 *There is a continuous map $\phi : S \to A$ which is the identity on each vertex of $S \cap A$. This map has the property that the line segment joining any $s \in S$ to $\phi(s) \in A$ avoids the whole cut locus.*

Proof: Consider any point $p \in S$. First suppose that p lies in the interior of an edge e of S. Then P is contained in a maximal disk Δ which intersects ∂P in exactly two points. Looking back at our local analysis, we see that the two segments joining p to ∂P must lie on opposite sides of e. Hence, one of the the points of $\Delta \cap \partial P$ lies in A. We define this point to be $\phi(p)$. The line segment joining p to $\phi(p)$ avoids the cut locus for the same reason here that the line segment considered in the previous section does.

Our local analysis above also shows that ϕ is continuous as p varies in the interior of e. The map extends to the closure of e. All that is happening here is that there is some third point of ∂A which lies on the same boundary circle.

Consider two edges e_1 and e_2 of S which meet at a vertex v. Let Δ be the maximal disk centered at v. The fact that S is a peripheral curve guarantees that there is a unique point of $\Delta \cap \partial P$ lying in A. Otherwise some edge of T would emanate from v and stick into the empty component. Hence, the map $\phi|_{e_1}$ and the map $\phi|_{e_2}$ agree on v. Hence ϕ extends to be a continuous map from S to A. ♠

Now we can establish the path connectness. Suppose that the cut locus is the union of two or more trees. Then there will be a pair of trees T_1 and T_2 and a peripheral path $S_1 \in T_1$ such that T_2 lies in the empty component determined by S_1. Figure 8 shows the situation.

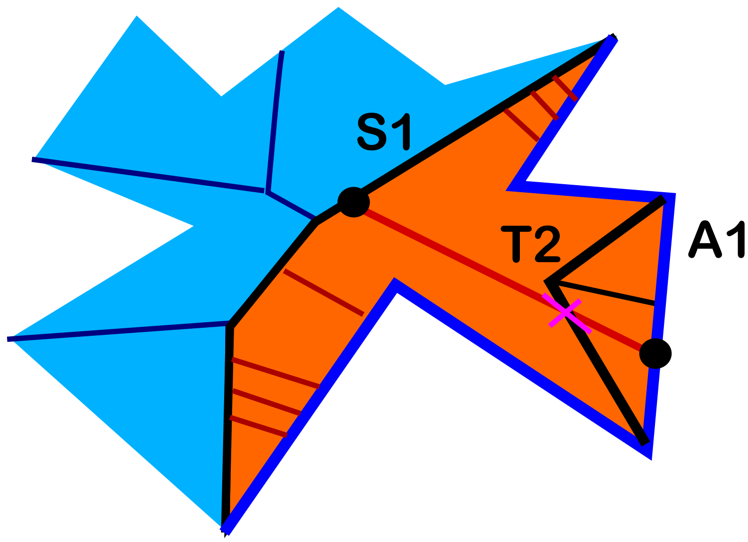

Figure 8: The trees T_1 and T_2.

Consider the union of line segments which join points $s \in S_1$ to the boundary arc A_1 defined above. These segments start out near one vertex and vary continuously until they reach other vertex. In particular, one of these segments must connect a point of S_1 to a point of A_1 which is separated from S_1 by a maximal path of T_2. This is a contradiction.

7 Fat Disks

Suppose we have some homeomorphism from S^1 to ∂P. We assume that S^1 has length 1. This allows us to transfer a measure to ∂P. The measure of any arc A of ∂P is defined to be the length of $h^{-1}(A)$. We call a maximal disk *fat* if it touches ∂P in two points which are at least $2/3$ apart in measure. This is Tverberg's Lemma 3 in his paper on the Jordan Curve Theorem.

We will suppose the result is false and derive a contradiction. If the result is false, we can assign an orientation to each edge of the cut locus. We have the edge point into the arc of smaller measure. Figure 8 shows what we are talking about. Figure 8 is just a schematic picture, where we draw the tree inside the unit disk in any way that is topologically equivalent to the cut locus, and then we join points on the tree to the points on the circle corresponding to the points on the polygon. This gives us a good way to picture the measure of the arcs: they just correspond to arc length on the circle.

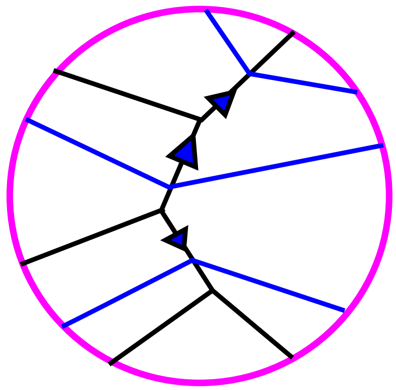

Figure 9: Orienting the edges of the tree.

You get the same orientation for any point on the interior of an edge, because there are supposedly no fat disks.

What happens at a vertex? If there is no arc of ∂P having measure less than $1/2$ which contains all the points, then at least two of them are at least $2/3$ apart. In case there are 3 points, this is pretty obvious. For more than 3 points, you have to draw some pictures of points on the circle to convince yourself of this.

Since there are no fat disks, all the points on ∂P associated to the vertex do not all lie in some arc of measure less than $1/2$. Then, as Figure 10 indicates, there is one ingoing edge and the rest of the edges are outgoing.

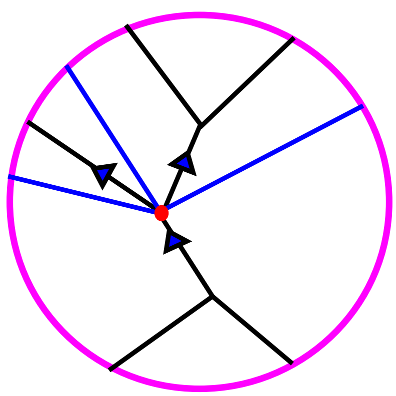

Figure 10: The orientations at a vertex.

Note that our orientation points outward at each of the outer edges of the tree, because these edges limit on a vertex. Figure 11 shows what we are talking about.

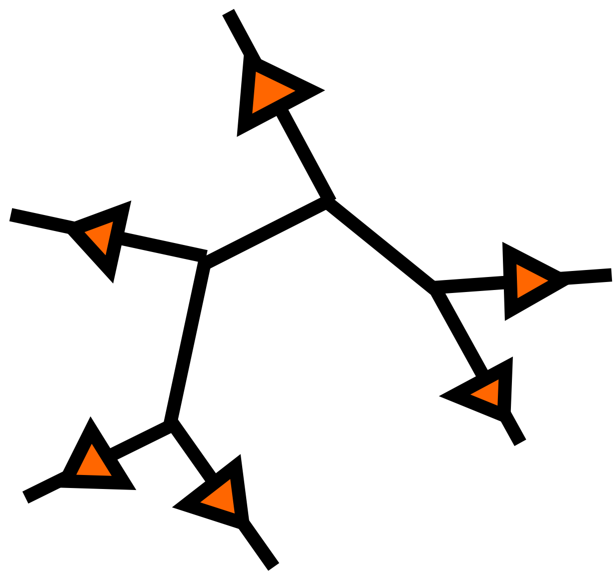

Figure 11: Ourward pointing edges.

So, we have oriented the edges of the tree so that

1. at each vertex there is one inward pointing edge and the rest outward pointing

2. The outer edges are all outward pointing.

Call this a *strangely oriented tree*.

Lemma 7.1 *A strangely oriented tree cannot exist.*

Proof: If T just has one edge then obviously T is not strangely oriented. So, let T be a strangely oriented edges with the fewest number of edges. If we delete one of the outer edges of T we still have a strangely oriented tree, and it has one fewer edge. This is a contradiction. ♠

Since there are no strangely oriented trees, the cut locus must have a fat disk.

8 Bottlenecks

Say that a *bottleneck* of ∂P is a polygonal path of length less than 2 which joins two points of ∂P. A bottleneck separates P into two path components. Now we suppose that P has been scaled up so that it contains points x and y which are at least one unit from ∂P. We prove that we can either connect x to y by a path which stays 1 unit away from ∂P or else there is a bottleneck which separates x from y. This is almost the same thing as Tverberg's Lemma 4 (which uses chords instead of paths) but the role our result plays in the proof of the Jordan Curve Theorem is exactly the same.

First of all, x is contained in some disk Δ which intersects ∂P at some point. If x is already contained in the cut locus, just do nothing at this step. Otherwise, we can continuously enlarge Δ until it touches some second point of ∂P. The notion of the center of this expanding disk gives a path from x to the cut locus which stays at least 1 unit away from ∂P. The same goes for y. In short, we can join x and y to the cut locus by paths which stay at least 1 unit from ∂P. So, without loss of generality, we can assume that x and y belong to the cut locus.

Now, color a point of the cut locus blue if it is at least 1 unit away from ∂P and otherwise red. Note that x and y are both colored blue. If x and y lie in the same blue connected component, then they can be joined by a path which stays at least 1 unit away from ∂P. Otherwise, since the cut locus is a tree, the path in the cut locus from x to y must go through a red point z. Moreover, we can assume that this red point lies in the interior of an edge. The radius of the maximal disk Δ centered at z is less than 1, and two radii of Δ connecting z to ∂P form the bottleneck.

9 The Jordan Separation Theorem

Now I'll explain how these results help prove the Jordan Curve Theorem. First I'll prove the Jordan Separation Lemma.

Suppose $h : S^1 \to \boldsymbol{R}^2$ is a homeomorphism from S^1 into its image $h(S^1)$. We want to show that $h(S^1)$ separates $\boldsymbol{R}^2$ into two connected components. We know the JCT for polygons already. We also have a sequence of maps $\{f_n\}$ such that $P_n = f_n(S^1)$ is an embedded polygon and $\|f_n(x) - h(x)\| < 1/n$ for all x.

Let P_n be the closure of the bounded component of $f_n(S_1)$. By construction, P_n is a polygonal region.

Lemma 9.1 *There is some $\epsilon > 0$ so that P_n contains a disk of radius ϵ.*

Proof: P_n contains a disk Δ_n which touches ∂P_n in two points y_n, y_n such that $f_n^{-1}(x_n)$ and $f_n^{-1}(y_n)$ are at east $2/3$ apart, when distances are measured along S_1. Suppose that the radius of Δ_n tends to 0. Then, passing to a subsequence, we can arrange that $f_n^{-1}(x_n)$ and $f_n^{-1}(y_n)$ both converge to points $x, y \in S^1$. These points are at least $2/3$ apart. In particular, $x \neq y$. But $h(x) = h(y)$. This is a contradiction. ♠

Lemma 9.2 (Jordan Separation) *$h(S^1)$ separates $\boldsymbol{R}^2$ into at least two components.*

Proof: Passing to a subsequence, we can assume that the big disks $\{\Delta_n\}$ guaranteed by the preceding lemma converge to a single disk Δ. But then the center p of Δ is at least $\epsilon/2$ away from ∂P_n for large n.

Suppose that p can be connected to ∞ by a continuous path γ which avoids $h(S_1)$. Since $\boldsymbol{R}^2 - h(S^1)$ is open, γ stays at least $\eta > 0$ from $h(S_1)$. But once n is large, we have $\eta > 1/n$, and γ must always intersect P_n. Butthen γ comes within η of some point of $h(S^1)$. This is a contradiction. ♠

10 The Jordan Curve Theorem

Suppose now that $\boldsymbol{R}^2 - h(S^1)$ has at least 3 path components, two of which are bounded. Choose points p, q in two different bounded components of $\boldsymbol{R}^2 - h(S^1)$.

First suppose that p and q both lie in P_n for all sufficiently large n. Since p and q are in different path components of $h(S^1)$, there there is no way to connect p to q in P_n without coming within $1/n$ of ∂P_n. But then ∂P_n has a bottleneck b_n of length less than $1/n$ which separates p from q. See Figure 12 below. There is an arc $A_n \subset \partial P_n$ so that $A_n \cup b_n$ separates p from q. But the diameter of $A_n \cup b_n$ tends to 0 as n tends to ∞. This contradicts the fact that both points remain uniformly far from P_n.

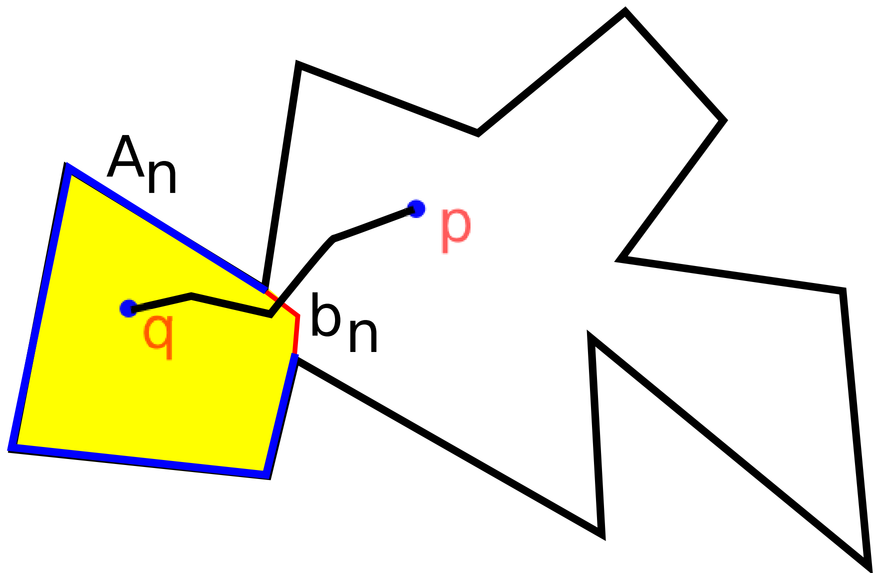

Figure 12: A bottleneck

Suppose that p and q lie outside P_n for all n. We can translate the picture so that some disk centered at the origin lies in P_n for all n. Consider the map

$$F(z) = 1/z. \tag{1}$$

Let $h' = F \circ h$. The map F is a homeomorphism of $\boldsymbol{C} \cup \infty$. The curve $h'(S^1)$ is another Jordan curve. The sets $F(\partial P_n)$ are Jordan curves made from finitely many circular arcs. We can replace these by polygonal approximations P'_n so that $p' = F(p)$ and $q' = F(q)$ lie inside P'_n for all n. In this way, we reduce to the case already handled.