Jordan's Proof of the Jordan Curve Theorem

Thomas C. Hales*

University of Pittsburgh

Abstract. This article defends Jordan's original proof of the Jordan curve theorem.

The celebrated theorem of Jordan states that every simple closed curve in the plane separates the complement into two connected nonempty sets: an interior region and an exterior. In 1905, O. Veblen declared that this theorem is "justly regarded as a most important step in the direction of a perfectly rigorous mathematics" [13]. I dedicate this article to A. Trybulec, for moving us much further "in the direction of a perfectly rigorous mathematics."

1 Introduction

Critics have been unsparing in their condemnation of Jordan's original proof. According to Courant and Robbins, "The proof given by Jordan was neither short nor simple, and the surprise was even greater when it turned out that Jordan's proof was invalid and that considerable effort was necessary to fill the gaps in his reasoning" [2]. A web page maintained by a topologist calls Jordan's proof "completely wrong." Morris Kline writes that "Jordan himself and many distinguished mathematicians gave incorrect proofs of the theorem. The first rigorous proof is due to Veblen" [8]. A different Kline remarks that "Jordan's argument did not suffice even for the case of a polygon" [7].

Dissatisfaction with Jordan's proof originated early. In 1905, Veblen complained that Jordan's proof "is unsatisfactory to many mathematicians. It assumes the theorem without proof in the important special case of a simple polygon and of the argument from that point on, one must admit at least that all details are not given" [13]. Several years later, Osgood credits Jordan with the theorem only under the assumption of its correctness for polygons, and further warns that Jordan's proof contains assumptions.[1]

Nearly every modern citation that I have found agrees that the first correct proof is due to Veblen in 1905 [13]. See, for example, [9, p. 205].

* This research has been supported by NSF grant 0503447.

[1] "Es sei noch auf die Untersuchungen von C. Jordan verwiesen, wo der Satz, unter Annahme seiner Richtigkeit für Polygone, allgemein für Jordansche Kurven begründet wird. Jordan beweist hiermit mehr als die Funktionentheorie gebraucht; dagegen macht er Voraussetzungen, welche diese Theorie streng begründet wissen will" [11].

ISBN 978-83-7431-128-1 ISSN 0860–150X

My initial purpose in reading Jordan was to locate the error. I had completed a formal proof of the Jordan curve theorem in January 2005 and wanted to mention Jordan's error in the introduction to that paper [3]. In view of the heavy criticism of Jordan's proof, I was surprised when I sat down to read his proof to find nothing objectionable about it. Since then, I have contacted a number of the authors who have criticized Jordan, and each case the author has admitted to having no direct knowledge of an error in Jordan's proof. It seems that there is no one still alive with a direct knowledge of the error.

The early criticisms from Veblen and Osgood are in fact rather harmless. True, Jordan did not write out a proof for polygons, but then again, the proof for polygons is widely regarded as completely trivial. For example, the book *What is Mathematics* "presupposes only knowledge that a good high school course could impart," and yet it presents two different proofs of the polygon case of the Jordan curve theorem. The second of these proofs is sketched in eight lines, with "the details of this proof is left as an exercise" [2]. If Veblen and Osgood had stronger evidence to discredit Jordan, why was their only explicit objection such a trivial one?

We wonder whether Kline's objection that "Jordan's argument did not suffice even for the case of a polygon" might merely be a reiteration of Veblen's objection that the case of polygons was omitted. These authors are correct in stating that Jordan stated the polygon version of the Jordan curve theorem without proof. However, a careful analysis of his proof (which we provide below) shows that Jordan does not make essential use of the Jordan curve theorem for polygons. Rather, he relies instead on the considerably weaker statement that there is a well-defined parity function that counts the number of times a ray crosses a polygon.

I have found one supporter of Jordan's proof. A 1996 paper gives a nonstandard proof along the lines of Jordan's original article [6]. It does not make sense for an essentially flawed approach to shed its defects when translated into another language, any more than it would for pulp to become great literature in translation. Puzzled, I contacted Reeken, one of the authors of the nonstandard proof. He replied that "Jordan's proof is essentially correct… Jordan's proof does not present the details in a satisfactory way. But the idea is right and with some polishing the proof would be impeccable" [12].

At the same time, Veblen's proof has suffered with the passage of time. His proof was part of his larger project to axiomatize *analysis situs* as an isolated field of mathematics. The model for this project was Hilbert's axiomatization of the foundations of geometry in 1899 [4]. This work precedes the rise of set theory as an axiomatic discipline. (Zermelo's first paper on the axioms of set theory appeared three years after Veblen's proof, in 1908.) Veblen's system of axioms was later abandoned when R. L. Moore showed in 1915 that his axioms describe nothing but the ordinary Euclidean plane. According to one account, the "results published in 1915 by Moore were rather devastating" for Veblen's line of research [10]. Thus, after a century, the entire framework of Veblen's proof is largely forgotten.

In view of the fundamental importance of the Jordan curve theorem to geometry, I present Jordan's proof anew. I have brought the terminology and language up to date without changing any essential ideas. In this way, I hope to preserve all of Jordan's major ideas, while avoiding its minor shortcomings. By presenting his ideas once again, we revive an elegant argument that has been unfairly condemned. Ultimately, if history has filled the gaps so completely that it becomes a serious challenge for us to discern them, then this speaks all the more forcibly in favor of Jordan.

There is another reason to take particular interest in Jordan's proof. This reason is the proof of the isoperimetric inequality for a general rectifiable Jordan curve J. The proof of the isoperimetric inequality for polygons is quite simple. To deduce the general version of the isoperimetric inequality from the special case of polygons, one must construct a polygon approximation to J whose length is no greater than that of J, and whose enclosed area exceeds that of J by no more than ϵ (for any $\epsilon > 0$). This is precisely what Jordan constructs in his proof. Hence, the isoperimetric inequality comes as a corollary to Jordan's proof. As far as I am aware, this corollary of Jordan's proof has gone unnoticed for over a century.

2 All About Polygons

2.1 Basic Definitions

Let d be the standard metric on $\mathbb{R}^2$.

Definition 1. *A simple closed curve J, also called a Jordan curve, is the image of a continuous one-to-one function from $\mathbb{R}/\mathbb{Z}$ to $\mathbb{R}^2$. We assume that each curve comes with a fixed parametrization $\phi_J : \mathbb{R}/\mathbb{Z} \to J$. We call $t \in \mathbb{R}/\mathbb{Z}$ the time parameter. By abuse of notation, we write $J(t) \in \mathbb{R}^2$ instead of $\phi_J(t)$, using the same notation for the function ϕ_J and its image J.*

We say that I is a (short) *interval* in $\mathbb{R}/\mathbb{Z}$ if there is an interval $[t, t']$ in $\mathbb{R}$ with $0 < t' - t < 1/2$ such that I is the image of $[t, t']$ under the canonical projection to $\mathbb{R}/\mathbb{Z}$. The upper bound on $t' - t$ ensures that an interval is uniquely determined in $\mathbb{R}/\mathbb{Z}$ by its endpoints.

Remark 1. We adopt the following useful convention, when two distinct parameter values $t, t' \in \mathbb{R}/\mathbb{Z}$ are used in a symmetrical manner, we swap them as necessary and identify them with real representatives such that $0 < t' - t \leq 1/2$.

Definition 2. *A polygon is a Jordan curve that is a subset of a finite union of lines. A polygonal path is a continuous function $P : [0, 1] \to \mathbb{R}^2$ that is a subset of a finite union of lines. It is a polygonal arc, if it is $1 - 1$.*

2.2 Parity Function for Polygons

The Jordan curve theorem for polygons is well known. We will only need a weak form, essentially saying that the complement of a polygon has at least two connected components. For this, it is enough to construct a non-constant locally constant function (the parity function).

Lemma 1. *Let P be a polygon. There exists a locally constant function on the complement of P in $\mathbb{R}^2$ that takes two distinct values.*

Proof. If the vertical line through $p \in \mathbb{R}^2 \setminus P$ does not contain any vertices of P, define its parity to be the parity of the number of intersections of P with downward directed vertical ray starting at p.

This function extends to a function on $\mathbb{R}^2 \setminus P$. In fact, if the downward vertical ray through p intersects P at a vertex, any sufficiently small horizontal displacement of the ray will not pass through a vertex of P.

The parity is independent of the small displacement. In fact, let V be the semi-infinite vertical strip between the two rays. The intersection of P with V consists of a finite number of disjoint polygonal arcs, with endpoints along the bounding rays. Since each polygonal arc has two endpoints, the total number of intersections of P with the two vertical rays is even. Thus, the parity is independent of the small displacement.

The fact that the parity is independent of small displacements is also used to establish the local constancy of the parity function.

The parity function takes two distinct values, as is seen by considering values along a ray that intersects P.

In reference to polygons, terms such as 'interior' or 'inside' mean odd parity and terms such as 'exterior' or 'outside' mean even parity. By using these terms, we do not assume the Jordan curve theorem for polygons, which makes the stronger assertion that the interior and exterior sets are each connected.

Now that we have established the existence of a parity function, we can compute the parity in degenerate situations where the vertical ray passes through a vertex of the polygon.

There are various related properties of the parity function on the complement of a polygon. The following statements about polygonal paths are easily established by the same method of 'displacement' that is used in the preceding lemma.

- We can compute the parity with respect to a ray in any direction through a point p and get the same value for the parity. It is not necessary to use the downward directed rays.
- Any point in the unbounded component has even parity.
- If a polygonal arc L from p to q crosses the polygon P transversally (meaning that L and P do not meet at a vertex of either one) m times, then p and q have the same parity with respect to P if and only if m is even.

2.3 Parity Function for Arcs

Let A be a polygonal arc in the plane with endpoints p and q. Let R_p and R_q be the vertical rays emanating from p and q directed upwards. Then by arguments similar to those used for a polygon, we see that there is a well-defined locally constant parity function $x \mapsto \pi_A(x)$ on

$$\mathbb{R}^2 \setminus (A \cup R_p \cup R_q)$$

that counts the crossing parity of A with a downward directed ray emanating from a point x.

If L is a line that does not meet p or q, then the parity of $\pi_A(x)$ is independent of x for all $x \in L$ with sufficiently large second coordinate. Thus, we may speak of the crossing parity between a line and a polygonal arc A, provided the line does not pass through the endpoints of of A.

If B is a second polygonal arc with endpoints p and q such that $A \cup B$ forms a simple closed curve, then the conditions $\pi_A(x) = \pi_B(x)$ and $\pi_A(x) \neq \pi_B(x)$ extend from

$$\mathbb{R}^2 \setminus (A \cup B \cup R_p \cup R_q)$$

to locally constant conditions on

$$\mathbb{R}^2 \setminus (A \cup B).$$

We have $\pi_A(x) = \pi_B(x)$ if and only if x has even parity with respect to the polygon $A \cup B$.

If A is a polygonal arc joining p and q, and L is a vertical line that does not meet p or q, then the line L meets A an odd number of times if and only if L separates p and q.

2.4 Adding a Polygonal Arc to a Polygon

Let A, B, and C be polygonal arcs that have the same endpoints p, q, but do not meet except at the endpoints. We then have polygons $P(A, B)$, $P(B, C)$, and $P(A, C)$ formed by pairs of arcs. Every point (other than the endpoints) of A (resp. B, C) has the same parity with respect to $P(B, C)$ (resp. $P(A, C)$, $P(A, B)$). So we may speak of the parity of any one of A, B, C with respect to the polygon formed by the other two.

Write π_A, π_B, π_C for the functions of $x \in \mathbb{R}^2 \setminus (A \cup B \cup C \cup R_p \cup R_q)$ giving the parities of crossing with A, B, and C with a downward directed ray starting from a point x.

Lemma 2. *Let x be a point of the plane in the complement of $A \cup B \cup C$ that lies outside $P(A, C)$. Then x lies outside $P(A, B)$ if and only if it lies outside $P(B, C)$.*

Proof. To say that x lies outside $P(A, C)$ means more precisely that $\pi_A = \pi_C$, and so forth. Then the conclusion follows directly from the following boolean tautology:

$$(\pi_A = \pi_C) \Rightarrow [(\pi_A = \pi_B) \Leftrightarrow (\pi_B = \pi_C)].$$

Lemma 3. *Exactly, one of A, B, C has odd parity with respect to the polygon formed by the other two.*

Proof. Assume first that C has odd parity for $P(A, B)$. We show that A has even parity for $P(B, C)$. Starting from a point in the plane where all parities are equal, approach $A \cup B \cup C$. We first meet a point of A or B. If we meet A first, the result is immediate. Say we meet B first. This means that $\pi_A = \pi_C$ along B. Also, we are given that $\pi_A \neq \pi_B$ along C. Analyzing a small disk around a point where A, B, and C meet, we see that these conditions give $\pi_B = \pi_C$ along A.

In the other direction, we show that if C has even parity for $P(A, B)$, and B has even parity for $P(A, C)$, then A has odd parity for $P(B, C)$. This is established in a similar manner, by examining a small disk around a point where A, B, and C meet.

2.5 Joining a Polygon Top to Bottom

In the following lemma, we join the "top" segment of a polygon to the "bottom" part by a linear segment in the interior region.

Lemma 4. *Let P be a polygon. Let L be a vertical line in the plane such that there exists a point $p \in P$ to the left side of L and a point q to the right side of L. Let A and B be the two polygonal arcs with endpoints at p and q (so that $P = A \cup B$, and $\{p, q\} = A \cap B$). Then there is a segment $L' \subset L$, that (except for its two endpoints on P) lies in an interior region formed by the polygon and such that one endpoint of L' lies on A and the other on B.*

We have seen from Lemma 1 that there is a well-defined parity function for crossings of L with P. That is, parities are well-defined even if L passes through a vertex of P.

Proof. As we move along L, we intersect one of the branches of the curve (A or B) with a certain parity, then the other branch with a certain parity, then the first again, and so forth. We write this sequence of parities as

$$a_1, b_1, a_2, b_2, \ldots$$

(where we swap A and B if necessary to make a_1 first in this sequence). Since A runs from p to q which are on opposite sides of L, the sum of the parities a_i is odd. Similarly, the sum of the parities b_i is odd. This means that we can split the sequence of parities after the first odd term. For example, if a_1 and b_1 are even, but a_2 is odd, we split the sequence into a_1, b_1, a_2 and $b_2, a_3, \ldots$. Geometrically, let L' be the segment of L that connects the last crossing of the first group (say, a_1, b_1, a_2) with the first crossing of the second group (say $b_2, a_3, \ldots$). It is clear that L' has the desired properties.

2.6 Interior of Perturbed Polygons

This subsection shows that an interior point of a polygon is also an interior point to a second polygon obtained by perturbing the first.

Lemma 5. *Let P be a (time-parameterized) polygon. Let p lie in the complement of P. Let D be the distance from P to p. Let P' be a (time-parameterized) polygon such that for all t,*

$$d(P(t), P'(t)) < D.$$

Then the parity of p with respect to P equals the parity of p with respect to P'.

Proof. By perturbing p to a nearby point (which we continue to call p) in such a way to preserve the inequality of the hypothesis, we may assume that the horizontal line L_H through p meets P in finitely many points. Let these points be

$$q_1 = P(t_1), \quad q_2 = P(t_2), \ldots$$

listed in cyclic order around P. The polygonal arc $P_i \subset P$ from q_i to q_{i+1} meets the vertical line L_V through p with parity π_i. Let $q'_i = P'(t_i)$, P'_i, π'_i be the corresponding quantities for P'.

Each of the points q_i is at distance greater than D to the left or to the right of p. Since P' approximates P within D, each of the points q'_i is to the left or right of L_V, according to whether q_i is to the left or right. Thus, the parity π_i, which is the number of left-right crossings of L_V, is equal to π'_i.

Since q_i and q_{i+1} are consecutive points on L_H, all of the crossings of L_V and P_i lie distance $> D$ above L_H or all lie distance $> D$ below L_H. Since P' approximates P within D, all the crossings of $P'_i \cap L_V$ are above or below L_H according to what happens to P_i.

Since $\pi_i = \pi'_i$, and all of the crossings of L_V are above or below in the same for both P_i and P'_i, for each i, it follows that the parity of p with respect to P is the same as the parity of p with respect to P'.

3 Tubes

In this section we construct a system of tubes around the edges of a polygon P. Let $r > 0$.

Let e be an edge of P. On both sides of e, construct an edge parallel to e at distance r and of the same length as e. Then cap both ends with a semicircle. This is T_e. This set is precisely the set of points that have distance r from e. Let T be the union of all the tubes, for all the edges of P.

We will use these tubes only for generic values of r. This means that through any point x in the plane, there are at most two tubes that pass through x and the intersection of these two tubes is transverse (an intersection of two curves, each a semicircle or line segment). Furthermore, for generic r, there exist only finitely many points x that meet more than one tube. (Call these points *jump points.*) The generic points are dense in the positive reals.

Let U be a component of the complement of T with the property that it contains a point at distance greater than r from P. We claim that every point of U has distance greater than r from P. Otherwise, there is a point $p \in U$ that lies in the convex hull of some T_e. (When we speak of the interior of a tube, we mean the set of points in the convex hull of T_e but not in T_e itself; there is no parity function for tubes.) If p is interior to the tube, then the entire component is interior to the tube, contrary to the assumption that some point of U has distance greater than r from P.

If q lies on the boundary of U, then it lies on some tube T_e and its distance to e is exactly r. Its distance to P is at least r, so this means that the closest point to q on P is a point on e. If q is not a jump point, then this closest point on P is unique. If q is a jump point, then q has distance r from exactly two points of P.

The segment from a point on the boundary of U to its closest point on P does not meet P, except at that closest point.

Since U is connected and does not meet P, the parity function with respect to P is constant on the connected set U.

The boundary of U is formed by a finite collection of circular arcs and line segments. We may assume that these arcs and segments extend as far as possible along the boundary of U, so that each circular arc and segment is terminated on both ends by a jump point (or by an endpoint of a tube's semicircle).

Remark 2. Let p and q be two distinct points having distance exactly r from P. Let $\bar{p}$ and $\bar{q}$ be points on P at distance r from p and q, respectively. We claim that the line segments $p\bar{p}$ and $q\bar{q}$ do not meet (except at $\bar{p}$, when $\bar{p} = \bar{q}$). In fact, since the distance from p (resp. q) to P is exactly r, we see that the distances from p to $\bar{q}$ and from q to $\bar{p}$ are at least r. Thus, the perpendicular bisector of the segment $\bar{p}\bar{q}$ separates $p\bar{p}$ from $q\bar{q}$.

3.1 Separating Polygons

We continue in the same context, with a polygon P, a generic r, a union of tubes T, and a component U of the complement of T that has a point of distance greater than r from P. Let V also be such a component.

Lemma 6. *Assume that the parity function takes the same value on U and V. Let p and q lie on the boundaries of U and V respectively. Assume that $p \neq q$. Suppose that there exists $m \in P$ such that its distance to p is r, and its distance to q is also r. Then one of the following two options holds:*

- *$U = V$ and p and q are connected by a circular arc in the boundary of U.*
- *There exists a polygonal arc C of length at most $2r$ with endpoints on P and not otherwise meeting P, such that p and q have different parities with respect to each of the two polygons formed by C and P (that is, polygons $P(A, C)$ and $P(B, C)$ of Section 2.4).*

(Note that in the second option, every point of C has distance at most r from P. In particular, it cannot meet a component such as U whose points have distance greater than r from P.)

Proof. The point m lies on the edge e and p and q lie on the tube T_e, for some e.

If p is not on the tube's semicircles, then there are exactly two points on T_e at distance r from m. These points are p and q. The line segment pq meets e transversally at m, and does not otherwise meet P. This shows that p and q have opposite parities with respect to P. This is contrary to the assumption that U and V have the same parity.

Thus, p and q both lie on the same semicircle. We can follow the contour of the semicircle from p to q, staying in the boundary of the same component U, unless there is a jump point s between them on the semicircle. If the first case of the conclusion of the lemma fails to hold, we will use the jump point s to show the second case of the conclusion holds. Let m and m' be the points of P closest to s. The polygonal arc from m' to s to m then joins P to P, has length $2r$, and separates p from q (except in the degenerate case when $s = p$ or $s = q$). In the degenerate case when $s = p$ or $s = q$, we need to take a small polygonal perturbation of this path – still of length at most $2r$ – to avoid passing directly through s. This can be done, since r is generic, and the two tubes passing through s have overlapping interiors.

3.2 Single Component

We continue in the context of the preceding subsection, with a polygon P, a generic r, a union of tubes T, and a component U of the complement of T that has a point of distance greater than r from P. Let V also be such a component.

The next lemma is a key result. It gives a set of conditions that are sufficient to ensure that two components U and V of the same parity are equal. The idea is quite simple. We trace out the boundary of U. As we do this, we trace out time parameters t of the closest points on P to these boundary points. There is a jump in time parameter at every jump point of the boundary of U. However, the hypotheses of the lemma ensure that these time jumps are small. The time jumps are so small that they cannot contain a time parameter t coming from a boundary point of U or V. In this way, we find that the boundary of U "seizes essential control of all the time parameters." Consequently, the time parameter of a boundary point of V can then be matched with a time parameter of a boundary point of U. This allows us to show that U and V have a boundary point in common. For a generic tube size r, this implies that the components U and V are actually equal.

In my view, the biggest (and rather harmless) omission in Jordan's original proof was in failing to state the details about why the "boundary of U seizes essential control of the time parameters." (This is Case 3 in the proof provided below.)

Lemma 7. *Let U, V, r, T, P be as above. Let $R > 0$. Assume U and V have the same parity with respect to P. Assume that V contains a point that has distance greater than $\max(r, R)$ from P. We make the following additional three assumptions:*

1. *There is an interior point of P that has distance greater than R from P.*
2. *U contains a point that has distance greater than $\max(r, R)$ from P.*
3. *For every polygonal arc C with endpoints $(P(t_1), P(t_2))$ on P, not otherwise meeting P, of the same parity as U, and of length at most $2r$, the time parameters t_1, t_2 satisfy $t_1 < t_2$ and $t_2 - t_1 < 1/2$ (after interchanging t_1 and t_2 if necessary). Moreover, if we let A be the image of $[t_1, t_2]$ under P, then the polygon $P(A, C)$ formed by A and C lies in a disk of diameter R.*

Then under these assumptions, $U = V$.

Proof. Let X be the set of $t \in \mathbb{R}/\mathbb{Z}$ such that there is a point p on the boundary of U such that $P(t)$ is a closest point on P to p. Each line segment in the boundary of U contributes an interval to X. Each circular arc on the boundary contributes a point to X. The set X is closed. Each point on the boundary of U that is not a jump point gives a well-defined point on X. Each jump point gives precisely two points t_1 and t_2 on X. The points $P(t_1)$ and $P(t_2)$ both have distance r from the jump point, and are thus connected by a polygonal arc C satisfying the third assumption of the lemma. Thus, we may assume that $t_1 < t_2$ and $t_2 - t_1 < 1/2$. Let $Y \subset \mathbb{R}/\mathbb{Z}$ be the union of these intervals (t_1, t_2) as we run over the jump points on the boundary of U.

Let q lie on the boundary of V, and let $P(t_q)$ be a point on P closest to q. We consider three cases $t_q \in X$, $t_q \in Y$, and $t_q \notin X \cup Y$.

Case 1: Assume that $t_q \in X$. By the definition of X, there is a point p on the boundary of U that also gives $t_q \in X$. By Lemma 6, there are two possibilities. One is that q is joined to p by a circular arc in the boundary of U. In particular, q lies in the boundary of both U and V. Since r is generic, this implies that the components U and V are equal. The other possibility is that there is a polygonal arc C of length at most $2r$ joining P to P and separating p from q. By parity arguments, one of V or U lies inside $P(A, C)$, hence inside a disk of diameter R. This is impossible, since both U and V have points farther than R from P. Thus, $t_q \notin X$.

Case 2: Assume that $t_q \in Y$. To dismiss this case, we show more generally that for any q' in the boundary of $V' \in \{U, V\}$, and such that $P(t_0)$ a closest point to q', we have $t_0 \notin Y$. In particular, X and Y are disjoint. Assume to the contrary, that $t_0 \in Y$. This means that $t_0 \in (t_1, t_2)$, where $P(t_1), P(t_2)$ are the closest points on P to a jump point p on the boundary of U. The two line segments $pP(t_i)$ form a polygonal arc C joining P to P of length $2r$. The component V', and the polygonal arc C have the same parity with respect to P.

We have $P(t_0) \in A$, where A is the image of $[t_1, t_2]$ under P. Let B be the other branch of P from $P(t_1)$ to $P(t_2)$, so $P = A \cup B$. The assumptions on the diameters of U, V and $P(A, B)$ prevent V' from lying in the interior of $P(A, C)$. By Lemma 2, V' has the same parity with respect to $P(B, C)$ as it does with respect to $P = P(A, B)$.

The line segment from q' to $P(t_0) \in A$ does not meet C (by Remark 2). This line connects V' to A without crossing $P(B, C)$. Thus, A (excluding endpoints) has the same parity with respect to $P(B, C)$ as V'. We conclude that the parity of A with respect to $P(B, C)$ is the parity of C with respect to $P(A, B)$. We can now apply By Lemma 3, which forces the parity of B to be odd with respect to $P(A, C)$. Thus, A, B, and $P = A \cup B$ are all contained in a disk of radius R. This is inconsistent with the first of the list of three assumptions of the Lemma.

Case 3: Finally, consider the case $t_q \notin X \cup Y$. This implies that $X \cup Y \neq \mathbb{R}/\mathbb{Z}$. The set $X \cup Y$ is constructed as a union of closed intervals, with endpoints in X. This implies that there is $t' \in X$ such that $P(t')$ is a closest point to a jump point p' in the boundary of U and such that there exists an open interval in $\mathbb{R}/\mathbb{Z}$ that has endpoint t' and is disjoint from $X \cup Y$. We consider two cases, depending on whether t' is an isolated point in X, or the endpoint of a non-trivial closed interval.

If t' is an isolated point in X, then it comes from a circular arc of the boundary of U. Along the boundary of U, both endpoints of this circular arc are jump points, and these two jump points give time parameters $t'', t''' \in X$ with $(t'', t'), (t', t''') \subset Y$. We may assume that $t'' < t' < t'''$. (For example, we cannot have $t'' < t''' < t'$, because Y and X are disjoint by Case 2, so $t''' \notin (t'', t')$.) We have $(t'', t''') \subset X \cup Y$. This is contrary to the hypothesis that t' lies on the boundary between $X \cup Y$ and its complement.

If t' is the endpoint of an interval I in X, then it comes from a line segment in the boundary of U. Since t' is a jump point, there exists t'' such that $P(t'')$ is a second point on P closest to that boundary point of U. Again, by Case 2, t'' must lie on the opposite side of t' from I. The sets $(t', t'') \subset Y$ and I cover a neighborhood of t', contrary to the hypothesis that t' lies on the boundary between $X \cup Y$ and its complement.

4 Polygon Approximation

We are now at an advanced stage of the proof, and yet everything so far has been about polygons. To proceed further, we need to prove that every Jordan curve can be approximated by polygons. That is the purpose of this section.

4.1 Uniform Continuity

To avoid a proliferation of deltas and epsilons, we introduce a special notation, c and c', for two of the deltas. The uniform continuity of a Jordan curve J can be expressed by the existence of a delta (which we call c) (depending on J) such that

$$\forall \epsilon\ t\ t'.\ \epsilon > 0 \wedge d(t, t') < c(\epsilon) \;\Rightarrow\; d(J(t), J(t')) < \epsilon.$$

By redefining $c(\epsilon)$ to be even smaller if necessary, we may assume that for all $\epsilon > 0$, we have $c(\epsilon) < 1/2$. J is a homeomorphism from its domain to its image. The uniform continuity of J^{-1} can be expressed by the existence of a delta c' such that

$$\forall \epsilon\ t\ t'.\ \epsilon > 0 \wedge d(J(t), J(t')) < c'(\epsilon) \;\Rightarrow\; d(t, t') < \epsilon.$$

The following lemma tells us how to ensure that an arc of a Jordan curve lies near $J(t)$ by controlling the location of its endpoints.

Lemma 8. *Let J be a Jordan curve. Let $\epsilon > 0$, and let t, t' be given that satisfy*

$$d(J(t), J(t')) < c'(c(\epsilon)).$$

Then, (adopting our conventions that $c(\epsilon) < 1/2$, $t < t'$, and $t' - t \leq 1/2$) for all $t'' \in [t, t']$, we have

$$d(J(t), J(t'')) < \epsilon.$$

Proof. Combine the uniform continuity of J and J^{-1}.

4.2 Approximation

Lemma 9. *For every (time-parameterized) Jordan curve J and $\epsilon > 0$, there is a time-parameterized polygon P such that $d(J(t), P(t)) < \epsilon$ for all t.*

Proof. Conceptually, the proof is very simple. Form a polygonal path P_N by joining $J(i/N)$ to $J((i+1)/N)$ (with each separate line segment parameterized with constant speed) for some large N and $i = 0, \ldots, N-1$. This path needn't be simple, but there is a subset of the image of P_N that avoids self-intersections and approximates J to within ϵ.

First we note that for any positive ϵ, there is N sufficiently large, so that the polygon P_N approximates J to within ϵ at each time t.

Using the conventions for parameter values t, t', we see that if $d(J(t), J(t')) < c'(c(\epsilon))$, then $d(J(t), J(t'')) < \epsilon$ for all $t'' \in [t, t']$. More strongly, for $\epsilon > 0$, we can find N such that for any double point $P_N(t) = P_N(t')$, we have $d(P_N(t), J(t'')) < \epsilon$ for $t'' \in [t, t']$. That is, the curve J stays within ϵ of the double point $P_N(t)$ throughout that interval. We excise the closed loop of P_N on $[t, t']$ making the path stop at the double point $P_N(t)$ during the time period $[t, t']$. The resulting map remains within distance ϵ of J at all times.

To make a consistent excision of all the self-intersections, from the finite collection of all such intervals, and pick one that is not contained as a subinterval of any other. Excise it, then repeat the process from the beginning. By excising an interval at each stage that is nested in no other, the iterative process picks out a disjoint collection of intervals to excise. The process terminates because the number of self-intersections drops with each iteration.

The resulting subset of $\mathbb{R}^2$ is the desired polygon, but poorly parameterized: a path that stops for a moment each time it reaches an excised interval is not simple. However, this is a simple matter to fix. Near each excised interval, reparameterize so that instead of remaining constant the values progress in a strictly monotone manner, changing the parameter values by such a small perturbation that the ϵ-approximation still holds. This completes the proof.

5 Constructing an Interior Point

Let J be a Jordan curve. The purpose of this section is to construct a point θ_0 that lies in the interior region of every sufficiently close polygonal approximation to J.

5.1 Constants β and w

In this section, we attach positive real numbers β and w to a Jordan curve J, as well as two special points p and q. These will be used later in this section.

Given a Jordan curve J, we define $w = w_J > 0$ to be the width of the narrowest vertical strip containing J. Fix this strip.

Pick p on J along the left edge of the strip and q on J along the right edge of the strip. There are distinct arcs A and B in J that connect p to q.

If we pick coordinates so that the y-axis is the left end of the strip, we can find a constant $\beta > 0$ such that for every point $a \in A$ and $b \in B$ whose x-coordinates lie in $[w/3, 2w/3]$, we have $d(a, b) \geq \beta$. Making β smaller if necessary, we can assume that

$$w/3 > \beta.$$

5.2 Constructing θ_0 and θ_∞

Lemma 10. *Let J be a Jordan curve. Attach data β, w, p, q, A, B to J as above. Pick ϵ so that $0 < \epsilon < \beta/2$. Let P be an ϵ-approximation to J. Set $p' = P(t)$ and $q' = P(t')$, where $p = J(t)$ and $q = J(t')$. Let A' and B' be the polygonal arcs in the image of P, from p' to q' (corresponding to A and B for J). Let L be a vertical line bisecting the vertical strip fitting J. Let L' be the segment on L determined by Lemma 4 (applied to P, L, p', q', A', B'). Then some point θ_0 along L' has distance at least $\beta/2 - \epsilon$ from A' and B'.*

Proof. Each point along L' has distance greater than $w/6 > \beta/2$ from points on J outside the vertical strip between $[w/3, 2w/3]$ to the right of p. Let U_A (resp. U_B) be the subset of the plane consisting of points at distance less than $\beta/2$ from A (resp. B). U_A and U_B are open. Both have nonempty intersection with L', because each contains an endpoint of L' by construction of Lemma 4. Their intersection along L' is empty, because such a point would put A at distance less than β from B within the vertical strip $[w/3, 2w/3]$. By the connectedness of L', we conclude that there is some point θ_0 of L' that is not in $U_A \cup U_B$. It has distance at least $\beta/2$ from A and B, so distance at least $\beta/2 - \epsilon$ from A' and B'.

Let θ_0 be the point that the lemma shows to exist (for some small ϵ). By Lemma 5, it lies in the interior of P' for all sufficiently close approximations P' of J.

Let θ_∞ be a point "at infinity", that is, any point far away from "all the action" in the proof. It will then be in the exterior region of P' for all sufficiently close approximations P' of J.

We let α be an index that runs over the set $\{0, \infty\}$, and write θ_α for the corresponding points.

6 Constructing the Interior and Exterior

We are now ready to give the main argument in the proof of the Jordan curve theorem. Let J be a Jordan curve. Let θ_α be the points constructed in the previous section.

For each $n \in \mathbb{N}$, and $\alpha \in \{0, \infty\}$, we will construct nonempty connected open regions U_α^n as the connected component of θ_α in the complement of some finite union T^n of tubes. The regions will have the following key properties:

1. $U_0^n \cap U_\infty^n = \emptyset$
2. $U_\alpha^n \cap J = \emptyset$

3. $U_\alpha^n \subset U_\alpha^{n+1}$
4. $\mathbb{R}^2 \setminus J \subset \bigcup_{n,\alpha} U_\alpha^n$.

Theorem 1. *(Jordan curve theorem) Let J be a Jordan curve. The complement of the image of J is the union of two disjoint nonempty open connected sets.*

In the words of Jordan, *"toute courbe [fermée] continue [et sans point multiple] divise le plan en deux régions, l'une extérieure, l'autre intérieure, cette dernière ne pouvant se réduire à zéro"* [5, p. 99].

Lemma 11. *If regions U_α^n can be constructed satisfying these four properties, then the Jordan curve theorem holds.*

Proof. Set $V_\alpha = \bigcup\{U_\alpha^n \mid n\}$. We claim that V_0 and V_∞ are the desired disjoint nonempty open connected sets. First, V_α is open because it is the union of open sets. It is nonempty, because it contains the point θ_α. The set V_α is connected because it is the union of connected sets containing the common point θ_α. The sets lie in the complement of J by Property (2), and give the full complement by Property (4). Finally, V_0 and V_∞ are disjoint by Properties (1) and (3).

We pick a sequence of positive real numbers r_n tending to 0. Pick r_0 small enough for the construction of the point θ_0 to work. Then pick

$$r_{n+1} < r_n/5.$$

Take P^n to be a $r_n/4$-approximation to J given by Lemma 9, and T^n to be the union of tubes around P^n for a generic parameter $r'_n \in (r_n/2, r_n)$. Take U_α^n to be the component of the complement of T_α^n complement containing θ_α.

To complete the proof of the Jordan curve theorem, it is enough to prove four lemmas, showing that Properties (1)–(4) hold for these choices.

6.1 Verifying the Four Properties

Lemma 12. *Property (1) holds for these choices. That is, $U_0^n \cap U_\infty^n = \emptyset$.*

Proof. Points of U_0^n and U_∞^n have different parities with respect to the polygons P^n.

Lemma 13. *Property (2) holds for these choices. That is, $U_\alpha^n \cap J = \emptyset$.*

Proof. U_α^n is the θ_α-component of $\mathbb{R}^2 \setminus T^n$. Every point of U_α^n has distance greater than $r'_n > r_n/2$ from P^n (as in Section 3). However, every point of J lies within distance $r_n/4 < r_n/2$ from P^n.

Lemma 14. *Property (3) holds for these choices. That is, $U_\alpha^n \subset U_\alpha^{n+1}$.*

Proof. We claim that every point of the boundary of U_α^{n+1} lies strictly within distance r'_n from P^n. In fact, to bound this distance, go from a boundary point to P^{n+1}, then to J, then to P^n, which have respective distances bounded by

$$r'_{n+1} + r_{n+1}/4 + r_n/4 < 5r_{n+1}/4 + r_n/4 < r_n/4 + r_n/4 < r'_n$$

by the recursion inequality for r_n. However, the points of the boundary of U_α^n have distance exactly r'_n from P^n. So $U_\alpha^n \subset U_\alpha^{n+1}$.

Lemma 15. *Property (4) holds for these choices. That is,* $\mathbb{R}^2 \setminus J \subset \bigcup_{n,\alpha} U^n_\alpha$*. More precisely, for any* $\epsilon > 0$*, there exists an* n *such that* U^n_α *contains all points* x *of* $\mathbb{R}^2 \setminus J$ *whose distance from* J *is at least* ϵ *and whose parity with respect to* P^n *is the same as that of* θ_α*.*

Proof. Fix α. Pick $\epsilon > 0$. We may assume that ϵ is less than the distance between θ_β and J, for $\beta = 0, \infty$. We pick n large enough so that the following conditions hold. (These conditions are listed in matching order with the hypotheses of Lemma 7, with R instantiated to ϵ.)

- The interior point θ_0 of P^n has distance greater than ϵ from P^n. (This can be arranged for large n because the distance of θ_0 to J is greater than ϵ.)
- The point $\theta_\alpha \in U^n_\alpha$ has distance greater than $\max(r'_n, \epsilon)$ from P^n. (For example, take n large enough that $r'_n < \epsilon$, then use the facts that the distance from θ_α to J is greater than ϵ, and that P^n approximates J.)
- Every pair of points $P^n(t_1), P^n(t_2)$ on P^n whose separation is at most $2r'_n$ satisfies $t_1 < t_2$ and $t_2 - t_1 < 1/2$ (after interchanging t_1 and t_2 if necessary). (This is just the uniform continuity properties of J^{-1} from Section 4.1, combined with the fact that P^n gives a sequence of approximations to J.) Moreover, for every polygonal arc C with endpoints $P^n(t_1), P^n(t_2)$, of the same parity as U^n_α, and of length at most $2r'_n$, the image A of P^n on $[t_1, t_2]$ and C lie in a disk of diameter ϵ. (To get this condition, apply Lemma 8 applied to J, and polygonal arcs of length less than some δ, to get a closed curve inside a disk of diameter $\epsilon' < \epsilon$. Then sufficiently close approximations P^n will yield data inside a disk of diameter ϵ.)

Fix n satisfying these conditions. Let V be the connected component in $\mathbb{R}^2 \setminus T^n$ of an element x as given in the statement. Applying Lemma 7 to U and V, we see that $U = V$. So $x \in U$.

7 The Isoperimetric Inequality

We briefly sketch a proof of the isoperimetric inequality, based on Jordan's proof of the Jordan Curve Theorem.

Corollary 1. *Let* J *be a Jordan curve of finite one-dimensional Hausdorff measure* b*. Let* a *be the area of the interior region. Then*

$$4\pi a \leq b^2.$$

Proof. It is enough to prove the weaker inequality for all $\epsilon > 0$:

$$4\pi(a - \epsilon) \leq b^2.$$

A lower bound on the length of J is the length of the piecewise linear curve obtained by choosing a finite list of points S sequentially along J and joining each consecutive pair of points of S by a line segment. Each polygon approximation P^n to J is obtained by removing a finite number of closed loops from such a piecewise linear curve for some $S = S_n$. (See Section 4.2.) Thus, the length of each polygon approximation P^n to J is no greater than the length of J. Hence the length b_n of each P^n is at most b. The regions U_0^n lie in the interior of P^n and J.

Let $\epsilon' = \epsilon/(\pi b)$. Cover J with finitely many disks whose radii r_i are less than ϵ' and that sum to less than b. (The sum of the diameters can be brought arbitrarily close to b. The sum of the radii can be brought arbitrarily closed to $b/2$.) Using Lemma 15, pick n large enough that the U_0^n and the disks cover the interior of J. Let a_n be the area of the interior of P^n. Then

$$a \le a_n + \sum \pi r_i^2 < a_n + \pi\epsilon' \sum r_i \le a_n + \epsilon.$$

The isoperimetric inequality now follows from the isoperimetric inequality for polygons:

$$4\pi(a - \epsilon) \le 4\pi a_n \le b_n^2 \le b^2.$$

The isoperimetric inequality usually includes the statement that circles are the only rectifiable Jordan curve for which equality is obtained. Again, this is not difficult to prove, once the inequality is known. The deepest part of the proof of the isoperimetric inequality is the existence of a suitable polygon approximation, as provided by Jordan.

References

1. J. W. Alexander, A Proof and Extension of the Jordan-Brouwer Separation Theorem, Trans. AMS, 23(4), 1922, 333-349.
2. R. Courant and H. Robbins, What is Mathematics, second edition, Oxford, 1996.
3. T. Hales, The Jordan Curve theorem, both formally and informally, to appear in the Amer. Math. Monthly.
4. D. Hilbert, Foundations of Geometry, Open Court, 1971.
5. C. Jordan, Cours D'Analyse de l'École Polytechnique, Paris, second edition, 1893.
6. V. Kanovei and M. Reeken, A nonstandard proof of the Jordan curve theorem, preprint 1996.
7. J. R. Kline, What is the Jordan Curve Theorem? Amer. Math Monthly, 49, no. 5, 281–286.
8. M. Kline, Mathematical thought from ancient to modern times. Vol. 3. Second edition. The Clarendon Press, Oxford University Press, New York, 1990.
9. M. H. A. Newman, Elements of the Topology of Plane Sets of Points, Dover, 1992.
10. D. Reginald Traylor, Creative Teaching: The Heritage of R. L. Moore http://at.yorku.ca/i/a/a/b/21.htm
11. W. F. Osgood, Lehrbuch der Funktionentheorie, Teubner, 1912.
12. M. Reeken, private communication, 21 Feb 2005.
13. O. Veblen, Theory on Plane Curves in Non-Metrical Analysis Situs, Trans. AMS, 6, No. 1 (1905), 83–98.