RECTANGLES, CURVES, AND KLEIN BOTTLES

RICHARD EVAN SCHWARTZ

ABSTRACT. In this article I will survey some results about inscribing triangles and quadrilaterals in Jordan curves. I will focus on the recent result of Josh Greene and Andrew Lobb, which says that for any smooth embedded loop C and any aspect ratio λ, there are four points in C which make the vertices of a rectangle of aspect ratio λ.

1. INTRODUCTION

This article starts with the question of picking out four special points on a curve in the plane and ends with a discussion of Shevchishin's theorem that you cannot embed a Klein bottle in $\boldsymbol{R}^4$, four-dimensional Euclidean space, if it is Lagrangian. I will explain below what this means.

The notorious Toeplitz conjecture, which goes all the way back to 1911, asks whether any Jordan curve contains four points which make the vertices of a square. (The edges of the square might intersect the curve in a messy way.) Such a collection of points is called an *inscribed square*. Figure 1 shows an example of a red square inscribed in a hexagon. The Toeplitz conjecture is also popularly known as the *square peg conjecture*.
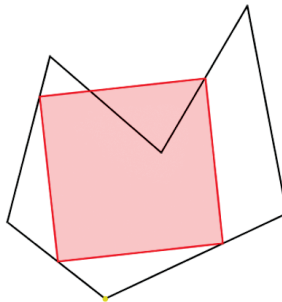

FIGURE 1. A square inscribed in a hexagon

It is important to remember that a Jordan curve is any *continuous* loop in the plane. If you put reasonable constraints on the curve, then the result has been known for a long time. In 1913, A. Emch [E] proved the result for convex curves. In 1929, L. G. Shnirelman [Shn] proved the result for sufficiently smooth curves. In general, the problem is that the curve could be unreasonable. Perhaps you start

with a polygon and then make little changes to the curve on small scales. Then you go in with a microscope and make even smaller changes, and so on.

The square peg problem has a long and sprawling history. See, for instance, [AA], [ACFSST], [CDM], [E], [FG], [H1], [H2], [Jer], [Mak1], [Mak2], [Ma1], [Ma2], [M], [N], [NW], [S1], [Shn], [St], [Ta], [Tv], [Va]. B. Matschke's paper [Ma1] gives a survey of what had been known up to 2014, and I. Pak's book [P] has an even more recent survey.

One can relax the question and ask about inscribed rectangles. The first general result along these lines, due to H. Vaughan, is that every Jordan curve (no matter how wild) has an inscribed rectangle. Meyerson's paper [M] cites a lecture that Vaughan gave about the proof at University of Illinois at Urbana-Champaign in 1977. My own involvement in this business is that I proved in [S2] that every point on a Jordan curve, with at most four exceptions, is the vertex of an inscribed rectangle.

As an aside, M. Meyerson [M] proved in 1980 that all but at most two points of any Jordan curve are vertices of inscribed equilateral triangles. This kind of result is not known for any other shape of triangle (e.g., right-angled isosceles), though M. Neilson [N] shows that a dense set of points in any Jordan curve are vertices of inscribed triangles of any desired shape.

Just like triangles, a rectangle has a shape to it, namely its *aspect ratio*, the ratio of its length to its width. One can ask whether every Jordan curve has an inscribed rectangle of any given aspect ratio. In 2018, C. Hugelmeyer made the first progress on this problem, showing in [H1] that every smooth embedded loop has an inscribed rectangle of aspect ratio $\sqrt{3}$. He later showed the following result in [H2]: For any smooth embedded loop, at least one third of the aspect ratios (as measured in a natural way) arise as aspect ratios of inscribed rectangles. This result involved a clever conversion of the problem into a question about certain Moebius bands intersecting in $\boldsymbol{R}^4$. Roughly speaking, Hugelmeyer constructs a continuous one-parameter family of embedded Moebius bands, all having the same boundary. Using topological methods and a bit of measure theory, he then shows that at least one third of the pairs have to intersect away from their boundaries.

This year (as of this writing, 2020), Josh Greene and Andrew Lobb [GL] made a breakthrough on the aspect ratio problem for inscribed rectangles. They proved that any smooth Jordan curve has inscribed rectangles of every aspect ratio. Their proof builds in Hugelmeyer's idea, and it considers a related one-parameter family of Moebius bands embedded in $\boldsymbol{R}^4$. The added twist is that they use the additional structure of $\boldsymbol{R}^4$ coming from its identification with $\boldsymbol{C}^2$, the space of pairs of complex numbers, and this allows them to bring in tools from symplectic geometry. They then use symplectic methods to show that *every pair* of Moebius bands in this family must intersect each other away from the common boundary. This breakthrough was the subject of a recent article by Hartnett in *Quanta Magazine* [Q].

Where do the Klein bottles come from? Well, if two Moebius bands meet along a common boundary, then their union is a Klein bottle with a kind of seam along the boundary. Suitably smoothing out this seam, you wind up with a Klein bottle. As I will explain, Greene and Lobb arrange for both the embeddings and the smoothing to be compatible with symplectic geometry, and the result is that the Klein bottle has the special property of being Lagrangian.

In this article, I will give an account of some of my favorite results in this area and then focus on the Greene–Lobb result. More honestly, I will give an account of the results whose proofs I actually understand well enough to give a nice explanation. The reader should know that my taste is partly dictated by my ignorance of the wider field. I am probably omitting a lot of beautiful material just by accident.

Here is an outline of the paper. In the brief §2, I will say a few words about Jordan curves. In §3 I will sketch proofs of Meyerson's theorem and of the square peg result for Jordan curves which are locally graphs of functions, as well as a few other related results. The material in §3 is not needed for the Greene–Lobb result. In §4 I will explain the ideas behind the Greene–Lobb result, using some of the symplectic geometry as a black box. At the end of §4 I will discuss an even stronger recent result of Greene and Lobb.

2. Jordan curves

2.1. **Basic definition.** A *Jordan curve* is the image $J = f(S^1)$ of a continuous and one-to-one map $f : S^1 \to \mathbf{R}^2$. Here S^1 is the unit circle. The famous *Jordan curve theorem* says that $\mathbf{R}^2 - J$ has two components, one bounded and one unbounded. The bounded one is called the *inside* and the unbounded one is called the *outside*. There are many proofs of the Jordan curve theorem; see, e.g., [Tv].

The case for polygons is fairly elementary: Color the points of $\mathbf{R}^2 - J$ black or white according to whether a generic ray emanating from the point intersects J an odd or an even number of times. (The argument given in §3.4 below justifies the claim that this parity does not depend on the line.) These black and white regions turn out to be the inside and the outside regions.

If you want to avoid using the Jordan curve theorem, which in general is rather tricky to prove, let me suggest an alternate definition. Say that a *special Jordan curve* is the image $h(S^1)$ where $h : \mathbf{R}^2 \to \mathbf{R}^2$ is a homeomorphism—i.e., a bijection which is continuous and whose inverse is continuous. In this case, J automatically inherits the topological properties of S^1, such as having an inside and an outside. The Jordan curve theorem and the two-dimensional Shoenflies theorem together say that every Jordan curve is special.

2.2. **Polygonal approximation.** For every Jordan curve J there is a sequence $\{J_n\}$ of (embedded) polygons such that $d(J_n, J) \to 0$ as $n \to \infty$. Here, $d(J_n, J)$ is the infimal value of ϵ so that every point of J is within ϵ of J_n and vice versa. This metric is called the *Hausdorff metric*. One way to get this result is to apply a sledgehammer from complex analysis. Take the Riemann map from the unit disk to one side of J in the plane. Then look at the images of circles of radius $1-\epsilon$ as $\epsilon \to 0$. This gives a family of smooth approximating curves which one can approximate by polygons in an easy way.

Now I'll describe a more elementary method based on the one in [Tv].[1] The method uses a square grid. The finer the grid, the better the approximation.

Let $J = J_0$ be the image of a map $f_0 : S^1 \to \mathbf{R}^2$. We think of S^1 as $\mathbf{R}/\mathbf{Z}$ so that it makes sense to speak of a linear map on an arc of S^1. Place J inside a square grid that is fine enough so that for each square Q, the set $f_0^{-1}(Q)$ is contained in

---

[1]After reading Tverberg's construction carefully, I think that he might need to add the maximality assumption to his selection of squares, as I do.

an arc $A_{Q,0} \subset S^1$ that is (say) less than $1/3$ the length of S^1. Note that f_0 maps the endpoints of $A_{Q,0}$ into the boundary of Q.

Call Q *maximal* if there is no other square R such that $A_{Q,0} \subset A_{R,0}$. Pick a maximal square Q. We create $f_1 : S^1 \to \mathbf{R}^2$ by setting $f_1 = f_0$ on $S^1 - A_{Q,0}$ and making f_1 linear on $A_{Q,0}$. The image $J_1 = f_1(S^1)$ is still embedded because this new curve intersects Q only in a straight line segment and otherwise is a subset of J_0. For any other square R we have $A_{R,1} \subset A_{R,0}$. In particular, there is no R such that $A_{Q,0} \subset A_{R,1}$.

We repeat the construction, producing maps $f_2, f_3, \dots$ and Jordan curves $J_2, J_3, \dots$ until we reach some $J' = J_k$ which intersects each square in a line segment. In particular, J' is an embedded polygon. Each modification we make only affects a small arc and, moreover, once we make a modification the maximality assumption guarantees that it is not erased at a later time. So, J' remains close to J.

There is one additional case where we use this kind of polygonal approximation. Following Meyerson, say that a *triod* is the union of three continuous arcs joined at a single point, like the letter Y, and otherwise disjoint from each other. The triod is *polygonal* if it is a finite union of line segments. The same kind of polygonal approximation idea shows that an arbitrary triod can be approximated by polygonal triods.

3. Some results about inscribed shapes

3.1. Triangles with an arbitrary shape.
Let Δ be a triangle. Say that another triangle T has the same *shape* as Δ if there is an orientation preserving similarity which maps Δ to T. Such a map is the composition of a rotation, a dilation, and a translation. In this section I will prove that every point of every differentiable Jordan curve is the vertex of an inscribed triangle of any given shape. This result is, in a sense, the triangular analogue of the Greene–Lobb result.

Let J be a differentiable curve, and let p_0 be a point on J. Let p_t be a parametrization of J so that as t ranges from 0 to 1 the point p_t moves all the way around J, say counterclockwise. For each choice of $t \in (0,1)$ there is a unique point q_t so that the points (p_0, p_t, q_t) are vertices of a triangle T_t which has the same shape as Δ. In Figure 2, I have drawn T_t in red when t is near 0 and in blue when t is near 1.
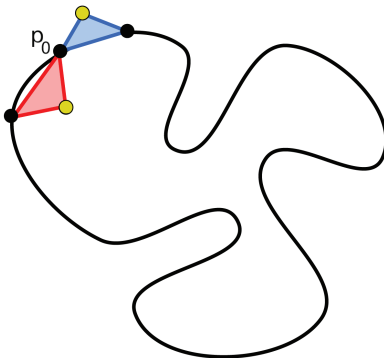

FIGURE 2. A family of triangles

Notice that in Figure 2 the point q_t, in yellow, starts out on the inside of J when t is near 0 and switches to the outside when t is near 1. Since q_t varies continuously, there must be some time $u \in (0, 1)$ for which $p_u \in J$. But then all the vertices of T_u lie in J. This completes the proof.

The proof gives a bit more: Any point of differentiability on any Jordan curve is the vertex of triangles of arbitrary shape.

3.2. **Meyerson's theorem.** Meyerson's theorem says that all but at most two points of an arbitrary Jordan curve J are vertices of inscribed equilateral triangle. In this section I will give a proof that is complete modulo details of polygonal approximation. The proof here is somewhat like Meyerson's proof and, of course, is based on his ideas, but it relies more heavily on polygonal approximation to make the analysis simpler.

Lemma 3.1. *Suppose that $p_0 \in J$ is some point, and there exist two other points p_1', p_2' in the region bounded by J such that $p_0 p_1' p_2'$ is an equilateral triangle. Then J has an inscribed equilateral triangle with vertex p_0.*

Proof. Let B and U denote the bounded and unbounded components of $\mathbf{R}^2 - J$. Let ρ be the 60-degree rotation about p_0 such that $R(p_1') = p_2'$. Consider extending the rays $p_0 p_1'$ and $p_0 p_2'$ outward until they first hit J at points p_1, p_2. Without loss of generality $p_0 p_1$ is not longer than $p_0 p_2$. Hence $R(p_1) \in J \cup B$. Let q_1 be a point of J maximally far from p_0. We have $R(q_1) \in J \cup U$. So, by continuity there is some $r_1 \in J$ such that $R(r_1) \in J$. Our equilateral triangle has vertices $p_0, r_1, R(r_1)$. ♠

Recall that a triod is a continuous version of the letter Y. Call the triod *good* if there is an equilateral triangle inscribed in the triod having one end of the triod as vertex. Otherwise call it *bad*. Call such triangles *end-inscribed triangles*. The key observation is that any three vertices of J are the endpoints of a triod that stays entirely in the region bounded by J. This is easy to see if J is a special Jordan curve. Just take one for the round disk and map it over.

Suppose for the moment that all triods are good. Choose any $a, b, c \in J$ and take a triod staying entirely inside J and having a, b, c as endpoints. Since this triod is good, there is an equilateral triangle inscribed in it having one of a, b, c as vertex, say a. But then the previous lemma applies to this triangle and shows that J has an inscribed equilateral triangle with a as vertex. So, to prove Meyerson's theorem we just have to show that all triods are good.

We will prove that all triods are good in three steps: polygonal triods, end-straight triods, general triods. The polygonal case really shows the meat of the argument. The other cases just amount to fooling around with approximations and limits.

Lemma 3.2. *A polygonal triod is good.*

Proof. Assume not, for the sake of contradiction. Let A denote the union of the first two legs of T. Let a be the endpoint of T not in A. For any $x \in T$, let A_x denote the result of rotating A by 60 degrees clockwise about x. When $x \in T - \partial A$, we have $\partial A \cap A_x = \emptyset$ and $A \cap \partial A_x = \emptyset$. Otherwise we would get the desired triangle. This means that the mod 2 intersection number I_x between A and A_x is well-defined and constant for all $x \in T - \partial A$.

Let b be an endpoint of A. The two arcs A and A_b make a 60-degree angle at b. If they intersect anywhere besides b we are done. So, assume that they intersect

nowhere else. By compactness, A_x and A cross exactly once, at x, for x sufficiently close to b. Hence $I_x = 1$ for all $x \in T - \partial A$. In particular, $I_a = 1$. But then we have an inscribed equilateral triangle with vertex a. ♠

A triod is *end straight* if the triad is polygonal sufficiently near the ends.

Lemma 3.3. *An end-straight triod is good.*

Proof. Let T be end straight. We can approximate T by a sequence $\{T_n\}$ of polygonal triods having the same final segments. By the previous lemma, T_n has an end-inscribed equilateral triangle Δ_n. Not all points of Δ_n can be on the same final segment of T_n because then Δ_n would consist of three collinear points. Note also that $T_n \to T$ and T is embedded. Combining these two observations, we see that there is a uniform positive lower bound to the size of Δ_n. Hence we can take a limit and find the end-inscribed equilateral triangle on T. ♠

Lemma 3.4. *An arbitrary triod is good.*

Proof. Now let T be an arbitrary triod, with ends a, b, c. For any large integer n, move out along the triple point of T until you reach the first point that is exactly $1/n$ from a. Call this point a'. Likewise define b', c'. Let T_n be the triod obtained by adding the segments aa', bb', cc' and erasing the arcs of T which join a to a', etc. If n is large enough, all points of $T' - (aa')$ are further than $1/n$ from a, etc.

By construction T_n is end straight. Let Δ_n be an end-inscribed triangle on T_n. Note that Δ_n, being equilateral, cannot have a as a vertex, and another vertex on aa'. So, either Δ_n is inscribed in T, and we're done, or else (after relabeling) Δ_n has a as a vertex and one point in bb'. Letting $n \to \infty$, we get an inscribed equilateral triangle with both a and b as vertices. ♠

3.3. **Squares inscribed in local graphs.** Say that J is a *local graph* if J has a finite covering by (open) rectangles $R_1, \ldots, R_n$ such that for each i the intersection $J \cap R_i$ is the graph of a function $y = f_i(x)$ after the picture has been rotated so that R_i has its sides parallel to the coordinate axes. In this section I will show that any local graph has an inscribed square. First, we can approximate J by polygons $\{J_n\}$ which all have the graph property with respect to the same set of rectangles.

If some J has an inscribed square Q, then the vertices of Q inherit two cyclic orders, one from the inclusion in Q and one from their inclusion in J. We call Q *gracefully inscribed* in J if these two orders coincide. Below I will sketch a proof that every polygon has a gracefully inscribed square.

Let Q_n be a square gracefully inscribed in the polygon J_n. After passing to a subsequence, we reduce to two cases. Either there is a positive lower bound δ to the diameters of Q_n or else there is a single point $p \in J$ such that $Q_n \to p$. In the first case we can take a limit on a subsequence and find a square of sidelength at least δ inscribed in J. Let us rule out the second case.

Rotating and scaling, we can assume that the limit point is the origin and that J and J_n intersect a rectangle R which contains the origin in its interior in such a way that $J_n \cap R$ is the graph of a function. See Figure 3. This is meant to hold for all n. The cyclic order on the vertices of Q_n imposed by J_n goes from left to right, as indicated in Figure 3. However, this order cannot coincide with the cyclic order on the vertices imposed by Q_n. This is a contradiction. So, the second case cannot occur.
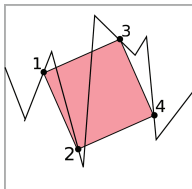
FIGURE 3. A contradiction at a small scale

This proof suggests a stronger version of the square peg conjecture that has been discussed quite often in connection with this problem. Say that a polygon P is *wide* if the bounded component of $\mathbf{R}^2 - P$ contains a disk of radius 1.

Conjecture 3.5 (Big peg conjecture). *There is some $\epsilon_0 > 0$ with the following property. Every wide polygon has an inscribed square of sidelength at least ϵ_0.*

The *big peg conjecture* and polygonal approximation immediately imply the original square peg conjecture. The idea is that you first dilate your Jordan curve J so that it has an approximation by wide polygons $\{J_n\}$. Each polygon J_n has a big inscribed square Q_n. The limit $\{Q_n\}$ must converge, on a subsequence, to some square inscribed in J. The lower bound on the size prevents these squares from shrinking to points. The big peg conjecture is quite seductive because it only involves polygons.

3.4. **A warmup problem.** Before getting to the existence of gracefully inscribed squares, let's consider a warmup problem that captures many features of the argument we give below. The argument we give is one of the steps in the proof of the polygonal Jordan curve theorem.

Let's prove that a generic horizontal line intersects a generic polygon X an even number of times. Here, a *generic polygon* means one having no pair of vertices on the same horizontal line. A *generic horizontal line* (with respect to X) is one which does not contain a vertex of X. Let L_1 be a generic line. Start with a line L_0 lying entirely below X. Let L_t be the family of horizontal lines which sweeps upward. Call a parameter t *critical* if L_t contains a vertex of X and otherwise *ordinary*.

At the ordinary parameters, the intersection points vary continuously and so their number does not change. There are only finitely many critical parameters, and at each critical parameter there is only one intersection point that lies at a vertex. As we wiggle the line up or down near a critical parameter, we see that near the critical parameter there are only three things that can happen. Figure 4 shows two of them, and the third possibility is like the first one but turned upside down.
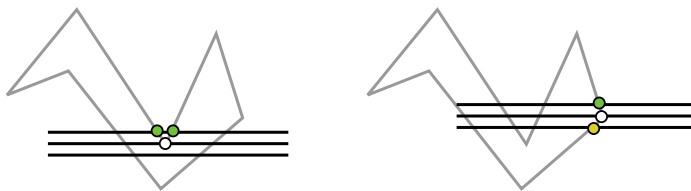

FIGURE 4. The behavior of intersection points

In all cases, the parity of the number of intersections does not change. So, L_1 intersects X an even number of times. An examination of the argument shows that all we really needed was that X does not have any horizontal sides. We made X even more generic just so as to deal with the critical intersections one at a time.

3.5. **Existence of gracefully inscribed squares.** We show that a generic polygon has an odd number of gracefully inscribed squares. Here *generic* means that no two sides of the polygon are parallel, no three sides lie in lines having a triple intersection, no three vertices lie in a square, and so on. The reasons for using generic polygons are similar to the reasons in the warmup problem. For instance, if a polygon has two long parallel sides close together, it will have infinitely many inscribed squares.

There are a variety of proofs that a generic polygon has an odd number of inscribed squares; see [Shn], [St], or [P, Theorem 23.11]. These arguments do not specifically ask for gracefully inscribed squares, but the variational proof—at least the one I sketch below—works when we restrict our attention to gracefully inscribed squares.

Suppose P_1 is a generic polygon. Start with some easy-to-understand polygon P_0 having the same number of sides as P_1 and having an odd number of inscribed squares. For instance, P_0 could be a slight perturbation of a subdivision of an obtuse triangle. Now consider a continuous family P_t of polygons that interpolates between between P_0 and P_1. You cannot necessarily make all the polygons in the family *completely generic*. For instance, you may not be able to avoid some edges becoming parallel along the way. However, if the edges of P_t are very short, then a square inscribed in P_t can have at most two vertices inscribed in this union of parallel edges. Also, you can make coincidences like parallel edges happen one at a time.

Say that a vertex of an inscribed square is *critical* if it is a vertex of P_t, and otherwise ordinary. Call the square critical if it has a critical vertex, and call the parameter t critical if there is an associated critical square. We can make the family generic enough so that there are only finitely many critical parameters, and at each critical parameter there is only one critical square, and this critical square has only one critical vertex. Moreover, we can make all the ordinary vertices vary continuously with the parameter. The continuity property is a *local* one: it only involves a statement about how a polygon interacts with at most four lines, and it can be settled by a direct algebraic calculation as in [S1] or [S2]. So, the square count changes only when we pass through a critical parameter.

What happens when we pass through a critical parameter? Notice that the question is also local: at most five lines are involved. Figure 5 shows a typical picture. The black edges at the intersection point are edges of the polygon, and the red and blue segments are meant to depict the lines containing these edges.

We can think of the square in question as two copies of the same square, one red and one blue. The red (respectively blue) square is inscribed in the union of the black and red (respectively blue) lines. As the parameter varies, the red square and the blue square separate from each other in a continuous way. We can translate the picture so that the red and blue lines always intersect at the origin. If the family is generic enough, the critical vertices, both from the red square and the blue square, vary monotonically through the origin.
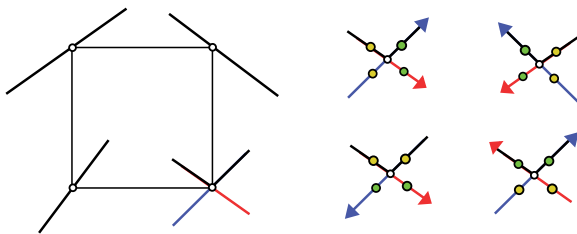
FIGURE 5. The behavior of intersection points

We have illustrated this with the figures at the right. The yellow intersection points are the locations of the critical vertices just before we reach the critical parameter, and the green points indicate the positions just after. Figure 5 shows the four possibilities for the directions that the points can move as a function of the varying parameter. The directions indicate the motion as a function of the parameter. The red and blue squares may or may not be inscribed in the polygon; it depends on whether the critical vertices lie on the segments of the polygon or on the red or blue segments extending them. In the bottom left case at right in Figure 5, both squares are inscribed in the polygon before the critical parameter and neither are inscribed after. In the top left case, one polygon is inscribed in the polygon before and one is inscribed after. The other two cases have similar treatments. Thus, the parity of the number of inscribed squares does not change as we pass through a critical point.

Notice that if the squares are gracefuly inscribed before, they are gracefully inscribed after. So, the parity of the number of graceful squares does not change either. Finally, the square inscribed in a (subdivided, perturbed) obtuse triangle is gracefully inscribed. Our initial polygon has an odd number of gracefully inscribed squares and therefore so does the final one.

3.6. **Existence of an inscribed rectangle.** In section 4 I will explain Vaughan's proof that every Jordan curve has an inscribed rectangle. For the reader who really wants an elementary explanation, Vaughan's proof leaves a bit to be desired: It requires some graduate level algebraic topology. Let me sketch a different proof based on the material in my papers [S1] and [S2]. This proof is more complicated but avoids algebraic topology.

It turns out that generically the space of rectangles inscribed in four lines is a one-dimensional manifold. Rather beautifully, the set of centers of these rectangles generically forms a hyperbola; see [S1]. This fact, coupled with the kind of analysis done in section 3.5, shows that the space of rectangles inscribed in a generic polygon is a one-dimensional manifold. Some of the components are loops and some of the other components are arcs. The endpoints of the arc components correspond to rectangles of aspect ratio 0 or ∞. (Working with labeled rectangles, we can tell the difference.) It turns out that a component of the manifold in question contains an even number of gracefully inscribed squares unless it has one of two properties:

- It is a loop connecting a gracefuly inscribed square $ABCD$ to the same inscribed square $BCDA$ with its vertices rotated. Call this a *rotator*. Every vertex of the polygon is the vertex of some rectangle in the rotator.

- It is an arc whose one end corresponds to rectangles of aspect ratio near 0 and whose other end corresponds to rectangles of aspect ratio near ∞. Call this a *sweepout*. All but at most four vertices of the polygon are vertices of a rectangle in a sweepout.

Given that there is an odd number of graceful squares, the generic polygon has an odd number of rotators and sweepouts combined; hence, at the very least it has one or the other. The existence of a sweepout would establish that the polygon has inscribed rectangles of every aspect ratio, but I could not rule out the existence of rotators. But, in either case, if we have sequence of polygons $\{J_n\}$ approximating a Jordan curve, we can extract from either a sweepout or a rotator a uniformly large rectangle R_n (in the sense of its minimum side length) inscribed in R_n. (I'll say more about this below.) The limit $\lim R_n$ will be a nontrivial rectangle inscribed in J. In [S2] I soup up this argument to show that all but at most four points of J are vertices of inscribed rectangles.

Now I will say more about finding the big rectangle R_n. Consider a set S_n of (say) 100 distinct points on J_n. We choose S_n so that $\{S_n\}$ converges to a set S of 100 distinct points on J. Each rectangle R inscribed in J_n cuts J_n off into four arcs, $\alpha, \beta, \gamma, \delta$ going in order. Let $I(R)$ denote the number of points in $\alpha \cup \gamma$ minus the number of points in $\beta \cup \delta$. By a rough form of continuity, we can always find some rectangle R_n, either in a rotator or a sweepout, having $|I(R_n)| < 10$. Then some pair of adjacent arcs cut off by R_n, say α_n, β_n, are such that that $S_n \cap \alpha_n$ and $S_n \cap \beta_n$ each have at least ten points. By construction, no side of R_n can shrink to a point. I have deliberately used more points than strictly necessary so as to avoid needing a careful count.

4. Existence of inscribed rectangles

4.1. Vaughan's theorem.
Let me first explain Vaughan's argument that every Jordan curve has an inscribed rectangle. This result was the inspiration for the work of Hugelmeyer and of Greene and Lobb. The argument is not really elementary because it relies on the fact from algebraic topology that there is no continuous embedding of the Klein bottle into $\mathbf{R}^3$.

First of all, the Klein bottle is the quotient obtained from gluing the sides of a square together according to the identifications shown on the left hand side of Figure 6. The right hand side of Figure 6 shows a decomposition of the Klein bottle into two Moebius bands, joined together along their common boundary.
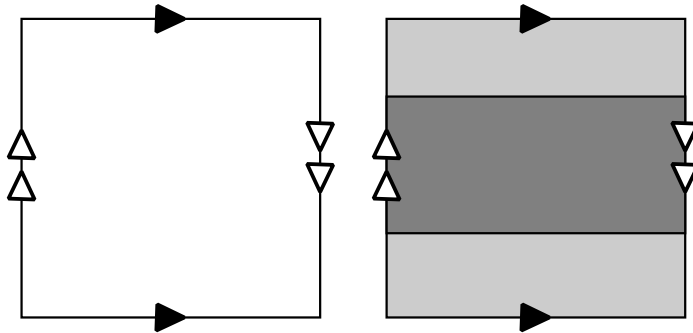

FIGURE 6. A Klein bottle in $\mathbf{R}^3$ with a seam

Theorem 4.1. *There is no continuous embedding of the Klein bottle into $\mathbf{R}^3$.*

Sketch of the proof. If there were an embedding of a Klein bottle into $\mathbf{R}^3$, there would also be an embedding into S^3, the 3-sphere. Let $K \subset S^3$ be the image of this embedding. It follows from Alexander duality that

$$\mathbf{Z}/2 \simeq H^2(K) \simeq \widetilde{H}_0(S^3 - K) \simeq \mathbf{Z}^r,$$

for some r. This is a contradiction. Here $\widetilde{H}_0$ is reduced singular homology with $\mathbf{Z}$ coefficients and H^2 is Čech cohomology with $\mathbf{Z}$ coefficients. This result is discussed in Hatcher's book on algebraic topology [Hat, p. 256]. ♠

Given a Jordan curve J, let S denote the set of unordered and unequal pairs of points in J. The space S is an open Moebius band. There are various ways to see this. This topological statement works the same for any Jordan curve, so we might as well consider the unit circle, and we also might as well consider it as a subset of the real projective plane. Every unequal unordered pair of points in the circle determines a unique point in $\mathbf{RP}^2$: Take the two tangent lines to the circle at these points and intersect them. This map identifies the space S with the complement of the closed unit disk in the projective plane, and this is an open Moebius band.

Vaughan defines a map $\phi : S \to \mathbf{R}^3$ by the formula

(1) $$\phi(a,b) = \left(\frac{a+b}{2}, |a-b| \right).$$

Geometrically, ϕ maps the ordered pair to a point encoding the midpoint of the segment $\overline{ab}$ and its length. If $\phi(a_1, b_1) = \phi(a_2, b_2)$, it means that the corresponding segments have the same length and meet at their midpoint. This gives an inscribed rectangle. So, we just have to prove that ϕ is not one-to-one.

We argue by contradiction. Notice that the image $\phi(S)$ lies in the upper half-space, and $\phi(\partial S)$ lies in the XY-plane. Let ρ denote reflection in the XY-plane. The union

$$K = \phi(S) \cup \phi(\partial S) \cup \rho \circ \phi(S)$$

consists of two Moebius bands meeting along their boundary and thus is a Klein bottle. The three pieces separately are disjoint, and if ϕ is one-to-one, then K is embedded. This contradicts the nonexistence of an embedded Klein bottle.

4.2. **The Greene–Lobb theorem.** Inspired by Hugelmeyer's papers, [H1] and [H2], Greene and Lobb use a construction that is similar to the one given in the proof of Vaughan's theorem. They also consider the space S of unordered distinct points on J. They define the map $\phi : S \to \mathbf{C}^2$ using the map

(2) $$f(a,b) = \left(\frac{a+b}{2}, \frac{(a-b)^2}{2\sqrt{2}|a-b|} \right).$$

One can extract from $f(a,b)$ the location of the center of $\overline{ab}$, the length of $\overline{ab}$, and the slope of $\overline{ab}$. Notice also that when a and b come together, the formula gives $f(a,b) = f(a,a) = (a,0)$. They also introduce the map

(3) $$R_\phi(z,w) = (z, e^{i\phi}(w)).$$

If $f(a_1, b_1) = R_{2\phi} \circ f(a_2, b_2)$, it means that the segments $\overline{a_1 b_1}$ and $\overline{a_2 b_2}$ have the same midpoint and the same length. Also, one of them is a rotation of the

other through an angle of ϕ. Therefore, an intersection like this corresponds to an inscribed rectangle whose diagonals make an angle of ϕ. Let

$$M_\phi = R_{2\phi} \circ f(S).$$

The set M_0 is an embedded Moebius band because one can recover a and b from $f(a, b)$. The set M_ϕ is just a rotation of M_0. These two Moebius bands limit on a common boundary, namely $f(\partial S)$. The union

$$K_\phi = M_0 \cup f(\partial S) \cup M_\phi$$

is a Klein bottle, and it is embedded unless M_0 and M_ϕ intersect away from their common boundary.

The upshot of the discussion above is that if K_ϕ is not embedded, then J has an inscribed rectangle whose diagonals make an angle of ϕ with each other. To prove that J has an inscribed rectangle of every aspect ratio, it suffices to prove that K_ϕ is never embedded.

It is possible to embed a Klein bottle in $\boldsymbol{R}^4$, but it turns out that after one performs a certain smoothing operation on K_ϕ, the result is a *Lagrangian Klein bottle*. Such a thing cannot be embedded. In the remainder of this chapter I will explain what a Lagrangian Klein bottle is and also how the smoothing operation works.

4.3. **Lagrangian surfaces.** As a prelude to discussing Lagrangian Klein bottles, I need to talk about some symplectic geometry. The standard symplectic structure on $\boldsymbol{R}^4$ can be described entirely in terms of real numbers, but it is nice to describe it in terms of complex numbers. We can naturally identify $\boldsymbol{R}^4$ with the space $\boldsymbol{C}^2$. The map is given by

$$(x_1, y_1, x_2, y_2) \rightarrow (z_1, z_2), \qquad z_j = x_j + iy_j.$$

On $\boldsymbol{C}^2$ there is a natural operation on pairs of vectors $V, W \in \boldsymbol{C}^2$. Writing $V = (V_1, V_2)$ and $W = (W_1, W_2)$, we define

$$(4) \qquad\qquad \langle V, W \rangle = V_1 \overline{W}_1 + V_2 \overline{W}_2.$$

Here $\overline{z}$ denotes the complex conjugate of z. This is known as a *Hermitian inner product*. It is linear in each argument and also $\langle V, W \rangle = \overline{\langle W, V \rangle}$.

It is instructive to write this in real coordinates. Let $V = (a_1 + ia_2, a_3 + ia_4)$ and $W = (b_1 + ib_2, b_3 + ib_4)$. Then

$$\langle V, W \rangle = (a_1 b_1 + a_2 b_2 + a_3 b_3 + a_4 b_4) + i(a_1 b_2 - a_2 b_1 + a_3 b_4 - a_4 b_3).$$

The real part of this expression is the dot product, and the imaginary part is (up to a *rotation*) the standard symplectic form on $\boldsymbol{R}^4$. We will write the imaginary part as ω. So,

$$(5) \qquad\qquad \omega(V, W) = \mathrm{Im}\langle V, W \rangle.$$

We distinguish two special kinds of planes in $\boldsymbol{C}^2$.

- A 2-plane $\Pi \subset \boldsymbol{C}^2$ is called *Lagrangian* if the restriction of Ω to Π is identically 0. An easy example of a Lagrangian plane is $\boldsymbol{R}^2$. The Hermitian inner product of any pair of vectors in $\boldsymbol{R}^2$ is real, so the imaginary part vanishes.

- A 2-plane Π is called *complex* if $i\Pi$ and Π are parallel. In case Π contains the origin, Π is complex if $i\Pi = \Pi$. The plane $\boldsymbol{C}^1 \subset \boldsymbol{C}^2$ is the prototypical complex plane.

A *smooth embedding* is a map $f : \Sigma \to \boldsymbol{R}^4$, where Σ is a surface and f is a smooth and injective map. This means, in particular, that the differential df maps each tangent plane of Σ to a plane in $\boldsymbol{R}^4$. The map f is *Lagrangian* if df maps each tangent plane to a Lagrangian plane. The construction by Greene and Lobb gives a Klein bottle which, if embedded, would be Lagrangian.

It is well known that one can embed the Klein bottle as a surface in $\boldsymbol{R}^4$. The classic approach is to almost embed it in $\boldsymbol{R}^3$ as one of those famous blown-glass models, and then fix it up. These blown-glass models do not quite work, because one of the necks of the bottle crashes through the surface and makes a seam. This is shown in Figure 7.
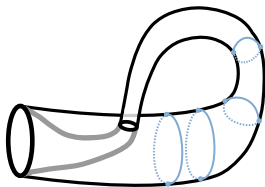

FIGURE 7. A Klein bottle in $\boldsymbol{R}^3$ with a seam

To embed this surface in $\boldsymbol{R}^4$, we add a fourth coordinate to separate out the two parts of the surface which crash into each other. Imagine traveling on the outside of the bottle, starting at the base and moving around. Make the fourth coordinate zero near the base and then (initially moving left to right) gradually increase it as you move around. By the time you wind around to the seam, the neck has a large positive fourth coordinate and is disjoint from the seam. Now gradually decrease the fourth coordinate so that the neck can rejoin with the base in $\boldsymbol{R}^4$.

To be more systematic about this, you could first draw a family of *parallel* loops on the Klein bottle, as suggested by the several pale blue ones in Figure 7. (More formally, this would be called a *foliation* of the Klein bottle by loops.) You could make the fourth coordinate constant on each of the loops and adjust the value on each loops, as above. I mention this because this foliation by loops will come up again below.

Shevchishin's theorem says, in particular, that some tangent plane along the resulting surface must fail to be Lagrangian.

Theorem 4.2 (Shevchishin's theorem). *There does not exist a smooth Lagrangian embedding of a Klein bottle in $\boldsymbol{R}^4$.*

Here is a lower-dimensional analogy. Let X be the plane with the origin removed. Each point $p \in X$ has a special tangent line through it, namely the line through p that is perpendicular to the ray $\overline{0p}$. It is possible to embed the circle in X in many ways, but it is impossible to do so in such a way that its tangent line is never special. The tangent line will be special at each point of the embedded circle that is maximally far from the origin.

There are several proofs of Shevchishin's result. All the proofs rely on some symplectic geometry machinery that is beyond the scope of the article. One of the proofs is in the paper [Ne] by S. Nemirovski.

I won't pretend to understand a proof of Shevchishin's theorem, but let me give you a sense of the depth of the result. I will repeat, as filtered through my own understanding, the sketch of a proof that Helmut Hofer mentioned when I asked him about it after Josh Greene's talk at the Institute of Advanced Study. Any errors in this account are due to my misunderstanding. The space K has a smooth foliation by closed loops, as discussed above. Each of these loops can be filled in by a pseudo-holomorphic curve. This is harder. The Lagrangian condition forces these holomorphic curves to be transverse to K, and then certain technical conditions force them to vary continuously and to be disjoint.

Consider the union of these disks. If we remove one of them, then topologically we have the product of a disk and a segment. When we put in the missing one, we are gluing the ends of this solid tube together. There are essentially two choices for the gluing: orientation preserving and orientation reversing. Now for the punchline: A pseudo-holomorphic curve has a canonical orientation coming from the (almost) complex structure involved in its definition. So the gluing must be orientation preserving and therefore the union must be a solid torus rather than a twisted disk bundle. But then the boundary of a solid torus cannot be a Klein bottle, and this is a contradiction.

I got the impression that the details of this approach have not been worked out. Nemirovski's proof goes through something called Luttinger surgery and a result of Gromov and McDuff on the classification of open symplectic manifolds which are standard (i.e., look like symplectic $\mathbf{R}^4$) at infinity. The Gromov–McDuff result, in turn, involves the kind of pseudo-holomorphic disks mentioned above.

4.4. Lagrangian smoothing.

The classic *bump function* is a smooth function f such that $f(x) = 0$ for $|x| \geq 2$ and $f(x) = 1$ for $|x| \leq 1$. This function is used all over the place in the theory of smooth manifolds. It is also used in the construction below, which is called *Lagrangian smoothing*.

Let's consider a very simple situation first. In $\mathbf{C}^2$ we consider two totally real planes which intersect along a line. The first plane is $\Pi_1 = \mathbf{R}^2$. The second plane also goes through the origin and is spanned by $(1,0)$ and $(0,i)$. Both planes contain the vector $(1,0)$ and so intersect along a line. Let

$$Y = \Pi_1 \cup \Pi_2.$$

Note that Y is the union of two totally real planes. Put another way, Y is the union of two Lagrangian surfaces which meet transversely along a curve which happens to be a line.

Now consider a family of planes X_t that is perpendicular to Π_1 and Π_2. The plane X_t contains the point $(t,0)$ and is spanned by the vectors $(0,1)$ and $(0,i)$. These planes are all complex planes, parellel to the second copy of $\mathbf{C}$, namely $\{0\} \times \mathbf{C}$. Each plane X_t intersects Y in the union Y_t of two perpendicular lines, as shown at left in Figure 8. We can produce a surface by replacing each Y_t by a union Z_t of two smooth curves, as shown at right in Figure 8. We can make Y_t and Z_t agree outside, say, the unit disk. This construction makes use of a bump function.

The union $Z = \bigcup Z_t$ is a Lagrangian surface which agrees with Y outside a small neighborhood of the line. The reason why Z is Lagrangian is that the tangent planes at each point are spanned by vectors of the form $(0,u)$ and $(1,0)$ for some unit complex u. As we have already mentioned, such planes are totally real.

FIGURE 8. A local model for Lagrangian smoothing

This operation is a local model for Lagrangian smoothing. Suppose we have two Lagrangian embedded surfaces which meet along a closed curve. Using a suitable change of coordinates, which comes from a variant of the so-called Darboux theorem, one can arrange that the local picture is just like in the simple model above. One then performs the local surgery described above and produces a union of two disjoint Lagrangian surfaces which agrees with the original union outside a small-as-you-like neighborhood of the original curve of intersection. The details of this coordinate change are worked out in [GL].

4.5. **The construction revisited.** Now let us go back to the main construction. Greene and Lobb have cooked up their map so that M_0 is Lagrangian. There are two main points to the proof. First, f extends to a map from $\boldsymbol{C}^2$ to $\boldsymbol{C}^2$ which preserves the symplectic form ω. So, f maps Lagrangian surfaces to Lagrangian curves. Second, the set of ordered distinct pairs of points in J is a Lagrangian surface inside $\boldsymbol{C}^2$. Indeed, the tangent plane at each point of J is spanned by vectors of the form $(z, 0)$ and $(0, z)$.

The union K_ϕ has a seam along $f(\partial S)$ that looks locally like one quarter of the left side of Figure 8, except that the angle between the two surfaces is 2ϕ rather than $\pi/2$. The left side of Figure 9 shows what we mean.

Simplifying things a bit, what Greene and Lobb do is smooth out the seam by doing *half* of the Lagrangian smoothing discussed above. The result would be a smooth embedded Lagrangian Klein bottle, which is a contradiction. This proves that M_0 and M_ϕ indeed intersect away from their common boundary.

Let me say a few more words about what it means to do half the Lagrangian smoothing. What they do is pass to a double cover, writing f as a composition $f = \sigma \circ \widehat{f}$, where $\widehat{f}$ is a map defined in a way very similar to f and σ is a 2-fold branched covering map from $\widehat{K}_\phi$ to K_ϕ. Here $\widehat{K}_\phi$ is the object like K_ϕ that is constructed using $\widehat{f}$ in place of f. They perform the smoothing of $\widehat{K}_\phi$ in a way that is equivariant with respect to σ, and then they push down the image via σ. Effectively, this does the smoothing as indicated in Figure 9.

FIGURE 9. Another Lagrangian smoothing construction

4.6. **A further development.** In [GL2] Greene and Lobb prove an even stronger result by similar methods. A *cyclic quadrilateral* is one which can be inscribed in a round circle. They prove that any cyclic quadrilateral (up to rotation and scaling) can be inscribed in any smooth Jordan curve. This result immediately implies the rectangle theorem, because a rectangle of any aspect ratio is a cyclic quadrilateral.

The proof of this stronger result proceeds along similar lines. They proceed by contradiction and reduce the problem to a statement that there exists a Lagrangian embedding of a torus with some additional topological constraints—a *minimum Maslov number* of 4. Once again, this time by a theorem of Polterovich and Viterbo, there is no such Lagrangian embedding with this property. This contradiction proves the theorem.

Acknowledgments

About the author

Richard Schwartz is currently Chancellor's Professor in Mathematics at Brown University. He was an Invited Speaker at the 2002 International Congress of Mathematicians (ICM) in Beijing, People's Republic of China, and also an Invited Speaker at the 2022 ICM in St. Petersburg, Russia.

References

[AA]   A. Akopyan and S. Avvakumov, *Any cyclic quadrilateral can be inscribed in any closed convex smooth curve*, Forum Math. Sigma **6** (2018), Paper No. e7, 9, DOI 10.1017/fms.2018.7. MR3810027

[ACFSST]   J. Aslam, S. Chen, F. Frick, S. Saloff-Coste, L. Setiabrata, and H. Thomas, *Splitting loops and necklaces: variants of the square peg problem*, Forum Math. Sigma **8** (2020), Paper No. e5, 16, DOI 10.1017/fms.2019.51. MR4061975

[CDM]   J. Cantarella, E. Denne, and J. McCleary, *Transversality in configuration spaces and the square peg problem*, `arXiv:1402.6174` (2014).

[E]   A. Emch, *Some Properties of Closed Convex Curves in a Plane*, Amer. J. Math. **35** (1913), no. 4, 407–412, DOI 10.2307/2370404. MR1506193

[FG]   P. Feller and M. Golla, *Non-orientable slice surfaces and inscribed rectangles* `arXiv:2003.01590v1` (2020).

[GL]   J. Greene and A. Lobb, *The rectangular peg problem*, Ann. of Math. (2) **194** (2021), no. 2, 509–517, DOI 10.4007/annals.2021.194.2.4. MR4298749

[GL2]   J. Greene and A. Lobb, *Cyclic quadrilaterals and smooth Jordan curves*, `arXiv:2011.05216` (2021)

[Hat]   A. Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge, 2002. MR1867354

[H1]   C. Hugelmeyer, *Every smooth Jordan curve has an inscribed rectangle with aspect ratio equal to $\sqrt{3}$.* `arXiv:1803:07417` (2018)

[H2]   C. Hugelmeyer, *Inscribed rectangles in a smooth Jordan curve attain at least one third of all aspect ratios*, Ann. of Math. (2) **194** (2021), no. 2, 497–508, DOI 10.4007/annals.2021.194.2.3. MR4298748

[Jer]   R. P. Jerrard, *Inscribed squares in plane curves*, Trans. Amer. Math. Soc. **98** (1961), 234–241, DOI 10.2307/1993494. MR120604

[Mak1]   V. V. Makeev, *On quadrangles inscribed in a closed curve* (Russian), Mat. Zametki **57** (1995), no. 1, 129–132, DOI 10.1007/BF02309400; English transl., Math. Notes **57** (1995), no. 1-2, 91–93. MR1339220

[Mak2]   V. V. Makeev, *On quadrangles inscribed in a closed curve and the vertices of the curve* (Russian, with English and Russian summaries), Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **299** (2003), no. Geom. i Topol. 8, 241–251, 331, DOI 10.1007/s10958-005-0412-8; English transl., J. Math. Sci. (N.Y.) **131** (2005), no. 1, 5395–5400. MR2038265

[Ma1]   B. Matschke, *A survey on the square peg problem*, Notices Amer. Math. Soc. **61** (2014), no. 4, 346–352, DOI 10.1090/noti1100. MR3184501

[Ma2]   B. Matschke, *Quadrilaterals inscribed in convex curves*, `arXiv:1801:01945v2` (2018).

[M]   M. D. Meyerson, *Equilateral triangles and continuous curves*, Fund. Math. **110** (1980), no. 1, 1–9, DOI 10.4064/fm-110-1-1-9. MR600575

[N]   M. Neilson, *Triangles inscribed in simple closed curves*, Geom. Dedicata **43** (1992), no. 3, 291–297.

[Ne]   S. Yu. Nemirovskiĭ, *The homology class of a Lagrangian Klein bottle* (Russian, with Russian summary), Izv. Ross. Akad. Nauk Ser. Mat. **73** (2009), no. 4, 37–48, DOI 10.1070/IM2009v073n04ABEH002462; English transl., Izv. Math. **73** (2009), no. 4, 689–698. MR2583965

[NW]   M. J. Nielsen and S. E. Wright, *Rectangles inscribed in symmetric continua*, Geom. Dedicata **56** (1995), no. 3, 285–297, DOI 10.1007/BF01263570. MR1340790

[P]   I. Pak, *Lectures on discrete and polyhedral geometry*, online book, `https://www.math.ucla.edu/~pak/book.htm`.

[Q]   K. Hartnett, *New geometric perspective cracks an old problem about rectangles*, `https://www.quantamagazine.org` (2020).

[S1]   R. E. Schwartz, *Four lines and a rectangle*, J. Experimental Math. (to appear) 2020.

[S2]   R. E. Schwartz, *A trichotomy for rectangles inscribed in Jordan loops*, Geom. Dedicata **208** (2020), 177–196, DOI 10.1007/s10711-020-00516-8. MR4142923

[Shn]   L. G. Šnirelʹman, *On certain geometrical properties of closed curves* (Russian), Uspehi Matem. Nauk **10** (1944), 34–44. MR0012531

[St]   W. Stromquist, *Inscribed squares and square-like quadrilaterals in closed curves*, Mathematika **36** (1989), no. 2, 187–197 (1990), DOI 10.1112/S0025579300013061. MR1045781

[Ta]   T. Tao, *An integration approach to the Toeplitz square peg problem*, Forum Math. Sigma **5** (2017), Paper No. e30, 63, DOI 10.1017/fms.2017.23. MR3731730

[Tv]   H. Tverberg, *A proof of the Jordan curve theorem*, Bull. London Math. Soc. **12** (1980), no. 1, 34–38, DOI 10.1112/blms/12.1.34. MR565480

[Va]   H. Vaughan, *Rectangles and simple closed curves*, lecture, University of Illinois at Urbana-Champaign, 1977.

DEPARTMENT OF MATHEMATICS, BROWN UNIVERSITY, PROVIDENCE, RHODE ISLAND