



AN INTEGRATION APPROACH TO THE TOEPLITZ SQUARE PEG PROBLEM

TERENCE TAO

UCLA Department of Mathematics, Los Angeles, CA 90095-1555, USA;
email: tao@math.ucla.edu

Received 22 November 2016; accepted 5 September 2017

Abstract

The ‘square peg problem’ or ‘inscribed square problem’ of Toeplitz asks if every simple closed curve in the plane inscribes a (nondegenerate) square, in the sense that all four vertices of that square lie on the curve. By a variety of arguments of a ‘homological’ nature, it is known that the answer to this question is positive if the curve is sufficiently regular. The regularity hypotheses are needed to rule out the possibility of arbitrarily small squares that are inscribed or almost inscribed on the curve; because of this, these arguments do not appear to be robust enough to handle arbitrarily rough curves. In this paper, we augment the homological approach by introducing certain integrals associated to the curve. This approach is able to give positive answers to the square peg problem in some new cases, for instance if the curve is the union of two Lipschitz graphs $f, g : [t_0, t_1] \rightarrow \mathbb{R}$ that agree at the endpoints, and whose Lipschitz constants are strictly less than one. We also present some simpler variants of the square problem which seem particularly amenable to this integration approach, including a periodic version of the problem that is not subject to the problem of arbitrarily small squares (and remains open even for regular curves), as well as an almost purely combinatorial conjecture regarding the sign patterns of sums $y_1 + y_2 + y_3$ for y_1, y_2, y_3 ranging in finite sets of real numbers.

2010 Mathematics Subject Classification: 55N45

1. Introduction

A subset Γ of the plane \mathbb{R}^2 is said to *inscribe a square* if it contains the four vertices of a square of positive sidelength. Note that despite the terminology, we do not require the solid square with these four vertices to lie in the region enclosed by Γ (in the case that Γ is a closed curve); see Figure 1.

© The Author 2017. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

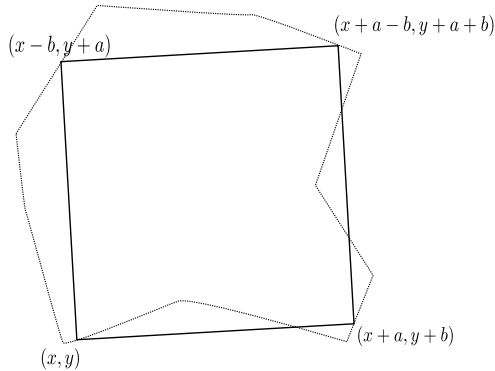


Figure 1. A (simple, closed, polygonal) curve inscribing a square.

The following conjecture of Toeplitz [26] is usually referred to as the *Square Peg Problem* or *Inscribed Square Problem*:

CONJECTURE 1.1 (Square peg problem). [26] *Let $\gamma : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{R}^2$ be a simple closed curve. Then $\gamma(\mathbb{R}/L\mathbb{Z})$ inscribes a square.*

In this paper, a curve $\gamma : I \rightarrow M$ denotes a continuous map from a domain I that is either an interval $[t_0, t_1]$ or a circle $\mathbb{R}/L\mathbb{Z}$ to a manifold M , with the curve being called *closed* in the latter case. Such a curve is called *simple* if γ is injective.

A recent survey on this problem, together with an extensive list of further references may be found in [16]; the brief summary below is drawn from that survey.

Thanks to arguments of a homological nature, the above conjecture is known assuming that the curve γ is sufficiently regular. For instance:

- Toeplitz [26] claimed Conjecture 1.1 for convex curves, but did not publish a proof. This case was partially resolved by Emch [4, 5] and then fully resolved in [3, 28]; this case can also be deduced from the ‘table theorem’ of Fenn [7];
- Hebert [10] gave a proof of Conjecture 1.1 for quadrilaterals. A proof of Conjecture 1.1 for arbitrary polygons was given in [18]; see also [21, 22] for some further ‘discretizations’ of the conjecture. See [19] for some computer-assisted quantitative bounds on one such discretization;
- Emch [6] established Conjecture 1.1 for piecewise analytic curves. An alternate proof of the analytic case was obtained by Jerrard [11];

- Schnirelman [23] (see also [8]) established Conjecture 1.1 for continuously twice differentiable curves (and in fact was also able to treat some curves outside this class). An alternate argument treating continuously twice differentiable curves (obeying an additional technical geometric condition) was obtained by Makeev [14];
- Nielsen and Wright [17] established Conjecture 1.1 for curves symmetric around a line or point;
- Stromquist [25] established Conjecture 1.1 for locally monotone curves. An alternate treatment of this case was given in [27];
- Cantarella *et al.* [2] established Conjecture 1.1 for C^1 curves, and also for bounded curvature curves without cusps;
- Matschke [15] established Conjecture 1.1 for an open dense class of curves (namely, curves that did not contain small trapezoids of a certain form), as well as curves that were contained in annuli in which the ratio between the outer and inner radius is at most $1 + \sqrt{2}$, and which had nontrivial winding number around the centre of the annulus.

One can broadly classify the methods of proof in the above positive results as being ‘homological’ in nature, in that they use algebraic topology tools such as the intersection product in homology, bordism, equivariant obstruction theory, or more elementary parity arguments counting the number of intersections between curves. In fact, many of these proofs establish the stronger claim that (under some suitable genericity and regularity hypotheses) there are an *odd* number of squares with vertices lying on the curve.

It is tempting to attack the general case of Conjecture 1.1 by approximating the curve γ by a curve in one of the above classes, applying an existing result to produce squares with vertices on the approximating curve, and then taking limits somehow (for example, by a compactness argument). However, in the absence of some additional regularity hypotheses on the curve, it could conceivably happen that the approximating inscribed squares degenerate to a point in the limit. Even if the original curve is smooth except for a single singular point, one could imagine that all squares in the approximating curves concentrate into that singular point in the limit. This scenario of squares degenerating to a point is the primary reason for the inability to remove the regularity hypotheses in the known positive results on the problem.

In this paper, we propose to modify the homological approach to the inscribed square problem, by focusing not only on the parity of intersections between various geometric objects associated to the curve γ , but also on bounding certain

integrals associated to these curves. As with previous works, one requires a certain amount of regularity (such as rectifiability, Lipschitz continuity, or piecewise linearity) on the curves in order to initially define these integrals; but the integrals enjoy more stability properties under limits than intersection numbers, and thus may offer a route to establish more cases of Conjecture 1.1. As an instance of this, we give the following positive result towards this conjecture, which appears to be new. For any $I \subset \mathbb{R}$ and any function $f : I \rightarrow \mathbb{R}$, we define the *graph* $\text{Graph}_f : I \rightarrow \mathbb{R}^2$ to be the function $\text{Graph}_f(t) := (t, f(t))$, so in particular $\text{Graph}_f(I) \subset \mathbb{R}^2$ is the set

$$\text{Graph}_f(I) := \{(t, f(t)) : t \in I\}.$$

Such a function f is said to be *C-Lipschitz* for a given $C > 0$ if one has $|f(s) - f(t)| \leq C|s - t|$ for all $s, t \in I$. Similarly if f is defined on a circle $\mathbb{R}/L\mathbb{Z}$ rather than an interval I (using the usual Riemannian metric on $\mathbb{R}/L\mathbb{Z}$).

THEOREM 1.2 (The case of small Lipschitz constant). *Let $[t_0, t_1]$ be an interval, and let $f, g : [t_0, t_1] \rightarrow \mathbb{R}$ be $(1 - \varepsilon)$ -Lipschitz functions for some $\varepsilon > 0$. Suppose also that $f(t_0) = g(t_0)$, $f(t_1) = g(t_1)$, and $f(t) < g(t)$ for all $t_0 < t < t_1$. Then the set*

$$\text{Graph}_f([t_0, t_1]) \cup \text{Graph}_g([t_0, t_1]) \tag{1.1}$$

inscribes a square.

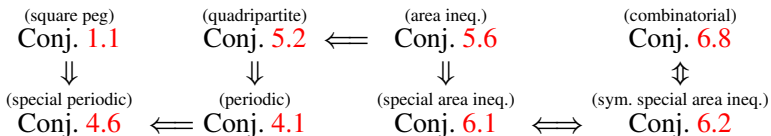
In other words, Conjecture 1.1 holds for curves that traverse two Lipschitz graphs, as long as the Lipschitz constants are strictly less than one. The condition of having Lipschitz constant less than one is superficially similar to the property of being locally monotone, as considered in the references [25, 27] above; however, due to a potentially unbounded amount of oscillation at the endpoints $\text{Graph}_f(t_0) = \text{Graph}_g(t_0)$ and $\text{Graph}_f(t_1) = \text{Graph}_g(t_1)$, the sets in Theorem 1.2 are not necessarily locally monotone at the endpoints, and so the results in [25, 27] do not directly imply Theorem 1.2. Similarly for the other existing positive results on the square peg problem.

We prove Theorem 1.2 in Section 3. A key concept in the proof will be the notion of the (signed) area $\int_\gamma y \, dx$ under a rectifiable curve γ ; see Definition 3.2. This area can be used to construct a ‘conserved integral of motion’ when one traverses the vertices of a continuous family of squares; see Lemma 3.5. Theorem 1.2 will then follow by applying the contraction mapping theorem to create such a continuous family of squares to which the conservation law can be applied, and then invoking the Jordan curve theorem to obtain a contradiction. The hypothesis of f, g having Lipschitz constant less than one is crucially used

to ensure that the curve that one is applying the Jordan curve theorem to is simple; see Proposition 3.8.

Since the initial release of this paper, we have learnt that a very similar method was introduced by Karasev [12] to obtain a partial result on the related problem of Makeev [14] of inscribing a given cyclic quadrilateral in a simple smooth closed curve, and indeed the proof of Theorem 1.2 can be generalized to handle inscribing an equiangular trapezoid if the Lipschitz constants are small enough; see Remark 3.10. We thank Benjamin Matschke for this reference and observation.

Without the hypothesis of small Lipschitz constant, the Jordan curve theorem is no longer available, and we do not know how to adapt the argument to prove Conjecture 1.1 in full generality. However, in later sections of the paper we present some related variants of Conjecture 1.1 which seem easier to attack by this method, including a periodic version (Conjecture 4.1), an area inequality version (Conjecture 5.6), and a combinatorial version (Conjecture 6.8). In contrast to the original square problem, the periodic version remains open even when the curves are piecewise linear (and this case seems to contain most of the essential difficulties of the problem). Conjecture 5.6 is the strongest of the new conjectures in this paper, implying all the other conjectures stated here except for the original Conjecture 1.1 (although it would imply some new cases of that conjecture). Conjecture 6.8 is a simplified version of Conjecture 5.6 that is an almost purely combinatorial claim regarding the sign patterns of triple sums $y_1 + y_2 + y_3$ of numbers y_1, y_2, y_3 drawn from finite sets of real numbers. It seems to be a particularly tractable ‘toy model’ for the original problem, though the author was not able to fully resolve it. The logical chain of dependencies between these conjectures, as well as some more technical variants of these conjectures that will be introduced in later sections, is summarized as follows, in which each conjecture is annotated with a very brief description:



2. Notation

Given a subset E of a vector space V and a shift $h \in V$, we define the translates

$$E + h := \{p + h : p \in E\}.$$

Given two subsets E, F in a metric space (X, d) , we define the distance

$$\text{dist}(E, F) := \inf_{p \in E, q \in F} d(p, q).$$

We use the asymptotic notation $X = O(Y)$ to denote an estimate of the form $|X| \leq CY$ for some implied constant C ; in many cases we will explicitly allow the implied constant C to depend on some of the ambient parameters. If n is an asymptotic parameter going to infinity, we use $X = o(Y)$ to denote an estimate of the form $|X| \leq c(n)Y$ where $c(n)$ is a quantity depending on n (and possibly on other ambient parameters) that goes to zero as $n \rightarrow \infty$ (holding all other parameters fixed).

We will use the language of singular homology (on smooth manifolds) in this paper, thus for instance a *1-chain* in a manifold M is a formal integer linear combination of curves $\gamma : I \rightarrow M$, and a *1-cycle* is a 1-chain σ whose boundary $\partial\sigma$ vanishes. Two k -cycles are *homologous* if they differ by a k -boundary, that is to say the boundary ∂U of a $k + 1$ -cycle. We integrate (continuous) k -forms on (piecewise smooth) k -chains in the usual fashion, for instance if $\sigma = \sum_{i=1}^n c_i \gamma_i$ is a 1-chain that is an integer linear combination of curves γ_i , and θ is a 1-form, then $\int_{\sigma} \theta := \sum_{i=1}^n c_i \int_{\gamma_i} \theta$. See for instance [9] for the formal definitions of these concepts and their basic properties. We will not use particularly advanced facts from singular homology; perhaps the most important fact we will use is the claim that if two (piecewise linear) cycles γ_1, γ_2 in an oriented manifold are homologous and intersect a smooth oriented submanifold V (without boundary) transversely, then their intersections $\gamma_1 \cap V, \gamma_2 \cap V$ are homologous cycles in V . Indeed, if γ_1 and γ_2 differ by the boundary of some cycle σ , then $\gamma_1 \cap V$ and $\gamma_2 \cap V$ differ by the boundary of $\sigma \cap V$ (viewed as a cycle with an appropriate orientation); alternatively, one may use Poincaré duality and the theory of the cup product.

3. Proof of positive result

We now prove Theorem 1.2. It will be convenient to give a name to the space of all squares.

DEFINITION 3.1 (Squares). We define $\text{Squares} \subset (\mathbb{R}^2)^4$ to be the set of all quadruples of vertices of squares in \mathbb{R}^2 traversed in anticlockwise order; more explicitly, we have

$$\text{Squares} := \{((x, y), (x + a, y + b), (x + a - b, y + a + b), (x - b, y + a)) : x, y, a, b \in \mathbb{R}; (a, b) \neq (0, 0)\}.$$

By abuse of notation we refer to elements of Squares as (nondegenerate) squares. Thus we see that a set $\Gamma \subset \mathbb{R}^2$ inscribes a square if and only if Γ^4

intersects Squares. We also form the closure

$$\overline{\text{Squares}} := \{(x, y), (x + a, y + b), (x + a - b, y + a + b), (x - b, y + a) : x, y, a, b \in \mathbb{R}\}$$

in which the sidelength of the square is allowed to degenerate to zero; this is a four-dimensional linear subspace of $(\mathbb{R}^2)^4$. A quadruple $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ of curves $\gamma_1, \gamma_2, \gamma_3, \gamma_4 : [t_0, t_1] \rightarrow \mathbb{R}^2$ is said to *traverse squares* if one has $(\gamma_1(t), \gamma_2(t), \gamma_3(t), \gamma_4(t)) \in \overline{\text{Squares}}$ for all $t \in [t_0, t_1]$; note that we allow the square traversed to degenerate to a point. Equivalently, $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ traverses squares if and only if there exist continuous functions $x, y, a, b : [t_0, t_1] \rightarrow \mathbb{R}^2$ such that

$$\begin{aligned} \gamma_1(t) &= (x(t), y(t)) \\ \gamma_2(t) &= (x(t) + a(t), y(t) + b(t)) \\ \gamma_3(t) &= (x(t) + a(t) - b(t), y(t) + a(t) + b(t)) \\ \gamma_4(t) &= (x(t) - b(t), y(t) + a(t)). \end{aligned}$$

We will also need the notion of the area under a (rectifiable) curve. Recall that a curve $\gamma : [t_0, t_1] \rightarrow \mathbb{R}^2$ is rectifiable if the sums $\sum_{i=0}^{n-1} |\gamma(s_{i+1}) - \gamma(s_i)|$ are bounded for all partitions $t_0 = s_0 < s_1 < \dots < s_n = t_1$. If we write $\gamma(t) = (x(t), y(t))$, this is equivalent to requiring that the functions $x, y : [t_0, t_1] \rightarrow \mathbb{R}^2$ are of bounded variation.

DEFINITION 3.2 (Area under a curve). Let $\gamma : [t_0, t_1] \rightarrow \mathbb{R}^2$ be a rectifiable curve, and write $\gamma(t) = (x(t), y(t))$ for $t \in [t_0, t_1]$. The *area under* γ , denoted by $\int_\gamma y \, dx$, is defined to be the real number

$$\int_\gamma y \, dx := \int_{t_0}^{t_1} y(t) \, dx(t)$$

where the integral on the right-hand side is in the Riemann–Stieltjes sense, that is to say the limit of $\sum_{i=0}^{n-1} y(s_i^*)(x(s_{i+1}) - x(s_i))$ for any partition $t_0 = s_0 < \dots < s_n = t_1$ and $s_i \leq s_i^* \leq s_{i+1}$ as $\max_{0 \leq i \leq n-1} |s_{i+1} - s_i|$ goes to zero.

Of course, if γ is piecewise smooth, this definition of $\int_\gamma y \, dx$ agrees with the usual definition of $\int_\gamma y \, dx$ as the integral of the 1-form $y \, dx$ on γ (now viewed as a 1-chain).

EXAMPLE 3.3. If $f : [t_0, t_1] \rightarrow \mathbb{R}$ is continuous of bounded variation, then the area under the curve Graph_f is just the usual Riemann integral:

$$\int_{\text{Graph}_f} y \, dx = \int_{t_0}^{t_1} f(t) \, dt.$$

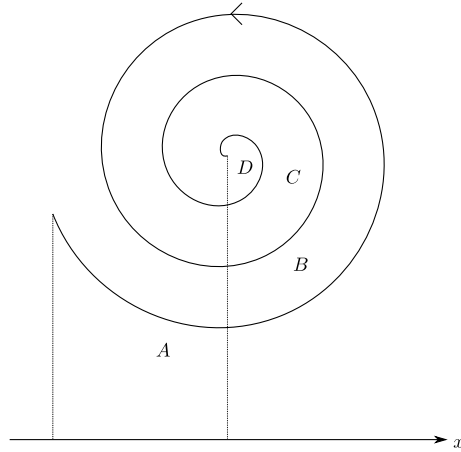


Figure 2. The area under the spiral (with the anticlockwise orientation) is equal to $|A| - |B| - 2|C| - 3|D|$, where $|A|$ is the area of the region A , and similarly for $|B|, |C|, |D|$; the weights $+1, -1, -2, -3$ attached to A, B, C, D here are the winding numbers of the spiral together with the horizontal and vertical line segments used to close up the curve.

In particular, if f is positive, $\int_{\text{Graph}_f} y \, dx$ is the area of the region bounded by Graph_f , the real axis, and the vertical lines $\{t_0, t_1\} \times \mathbb{R}$. If γ is not a graph, the area under γ is more complicated; see Figure 2.

It will be particularly important to understand the area under a closed curve:

LEMMA 3.4. *Let $\gamma : [t_0, t_1] \rightarrow \mathbb{R}^2$ be a simple closed anticlockwise rectifiable curve, and let Ω be the bounded open region enclosed by this curve as per the Jordan curve theorem. Then the area under the curve γ is equal to the negative of the Lebesgue measure $|\Omega|$ of Ω . In particular, this area is nonzero.*

Of course, if γ were clockwise instead of anticlockwise, then the negative sign in the above lemma would be removed; however, it would still be true that the area under this curve is nonzero.

Proof. In the case that γ is a polygonal path, this claim is clear from Stokes' theorem:

$$\int_{\gamma} y \, dx = \int_{\partial\Omega} y \, dx = - \int_{\Omega} dx \wedge dy. \tag{3.1}$$

Now we consider the general case. The strategy is to approximate γ by a polygonal path, apply (3.1), and take limits; but (as with the proof of the Jordan curve theorem) some care must be taken with the limiting argument.

We can normalize $[t_0, t_1] = [0, 1]$. Let $\varepsilon_0 > 0$ be a small parameter (which will eventually be sent to zero). By continuity of γ , there exists $0 < \varepsilon_1 < \varepsilon_0$ such that $|\gamma(t) - \gamma(t')| \leq \varepsilon_0$ whenever $t \bmod 1$ and $t' \bmod 1$ are separated by distance at least ε_1 on the unit circle \mathbb{R}/\mathbb{Z} . By a further application of continuity and the hypothesis that γ is simple, there exists $0 < \varepsilon_2 < \varepsilon_1$ such that $|\gamma(t) - \gamma(t')| \geq \varepsilon_2$ whenever $t \bmod 1$ and $t' \bmod 1$ are separated by distance at least ε_1 on the unit circle \mathbb{R}/\mathbb{Z} . Now let n be a natural number, that we assume to be sufficiently large depending on $\gamma, \varepsilon_0, \varepsilon_1, \varepsilon_2$. Let $\gamma_n : [0, 1] \rightarrow \mathbb{R}^2$ be the piecewise polygonal path formed by joining up the points $\gamma_n(j/n) := \gamma(j/n)$ for $j = 0, \dots, n$ by line segments, thus

$$\gamma_n \left(\frac{j + \theta}{n} \right) := (1 - \theta)\gamma \left(\frac{j}{n} \right) + \theta\gamma \left(\frac{j + 1}{n} \right)$$

for $j = 0, \dots, n - 1$ and $0 \leq \theta \leq 1$. As γ is uniformly continuous, we see for n large enough that

$$|\gamma_n(t) - \gamma(t)| < \frac{\varepsilon_2}{2} \tag{3.2}$$

for all $t \in [0, 1]$. Also, it is clear that the length of γ_n is bounded by the length of the rectifiable curve γ .

The path γ_n is closed, but it need not be simple. However, from (3.2), the triangle inequality, and the construction of ε_2 , we see that a collision $\gamma_n(t) = \gamma_n(t')$ can only occur if $t \bmod 1$ and $t' \bmod 1$ differ by at most ε_1 in the unit circle. In such a case, γ_n can be viewed as the sum (in the sense of 1-cycles) of two closed polygonal paths, one of which has diameter at most ε_0 . Deleting the latter path and iterating, we conclude that γ_n can be viewed as the sum of a simple closed polygonal path γ_n^0 and a finite number of closed polygonal paths $\gamma_n^1, \dots, \gamma_n^k$ of diameter at most ε_0 ; furthermore, the total lengths of $\gamma_n^0, \gamma_n^1, \dots, \gamma_n^k$ sum up to at most the length of γ , and from (3.2) we see that the curves $\gamma_n^1, \dots, \gamma_n^k$ lie within the $2\varepsilon_0$ -neighbourhood of γ .

If ε_0 is small enough, we can find a point z in Ω that is at a distance at least $10\varepsilon_0$ from γ . The winding number of γ around z is equal to 1. By Rouché's theorem, the winding number of γ_n around z is then also equal to 1, while the winding numbers of $\gamma_n^1, \dots, \gamma_n^k$ around z are equal to 0. Thus the winding number of γ_n^0 around z is equal to 1; thus γ_n^0 has an anticlockwise orientation, and z lies in the region Ω_n enclosed by γ_n^0 . This argument also shows that the symmetric difference $\Omega \Delta \Omega_n$ between Ω and Ω_n lies in the ε_0 -neighbourhood of γ . As γ is

rectifiable, this implies that

$$|\Omega_n| = |\Omega| + O(\varepsilon_0)$$

where the implied constant in the $O(\cdot)$ notation depends on the length of γ . On the other hand, from (3.1) one has

$$\int_{\gamma_n^0} y \, dx = -|\Omega_n|.$$

For each $i = 1, \dots, k$, we see that

$$\int_{\gamma_n^i} y \, dx = \int_{\gamma_n^i} (y - y^i) \, dx = O(\varepsilon_0 \ell(\gamma_n^i))$$

where y_i is the y -coordinate of an arbitrary point in γ_n^i , and $\ell(\gamma_n^i)$ denotes the length of γ_n^i . Summing, we conclude that

$$\sum_{i=1}^k \int_{\gamma_n^i} y \, dx = O(\varepsilon_0).$$

Finally, for n sufficiently large, we have from the rectifiability of γ that

$$\int_{\gamma_n} y \, dx = \int_{\gamma} y \, dx + O(\varepsilon_0).$$

Putting all these bounds together, we conclude that

$$\int_{\gamma} y \, dx = -|\Omega| + O(\varepsilon_0);$$

since $\varepsilon > 0$ can be made arbitrarily small, the claim follows. \square

The relevance of this area concept to the square peg problem lies in the following simple identity.

LEMMA 3.5 (Conserved integral of motion for squares). *Let*

$$\gamma_1, \gamma_2, \gamma_3, \gamma_4 : [t_0, t_1] \rightarrow \mathbb{R}^2$$

be rectifiable curves such that $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ traverses squares (as defined in Definition 3.1). Then we have the identity

$$\int_{\gamma_1} y \, dx - \int_{\gamma_2} y \, dx + \int_{\gamma_3} y \, dx - \int_{\gamma_4} y \, dx = \frac{a(t_1)^2 - b(t_1)^2}{2} - \frac{a(t_0)^2 - b(t_0)^2}{2}, \quad (3.3)$$

where the functions $x, y, a, b : [t_0, t_1] \rightarrow \mathbb{R}$ are as in Definition 3.1.

Proof. From Definitions 3.1 and 3.2 we have

$$\begin{aligned}\int_{\gamma_1} y \, dx &= \int_{t_0}^{t_1} y(t) \, dx(t) \\ \int_{\gamma_2} y \, dx &= \int_{t_0}^{t_1} (y(t) + b(t)) (dx(t) + da(t)) \\ \int_{\gamma_3} y \, dx &= \int_{t_0}^{t_1} (y(t) + a(t) + b(t)) (dx(t) + da(t) - db(t)) \\ \int_{\gamma_4} y \, dx &= \int_{t_0}^{t_1} (y(t) + a(t)) (dx(t) - db(t));\end{aligned}$$

note from the rectifiability of $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ that $x, y, x + a, y + b$ (and hence a, b) are of bounded variation. After some cancelling, we may then write the left-hand side of (3.3) as

$$\int_{t_0}^{t_1} a(t) \, da(t) - \int_{t_0}^{t_1} b(t) \, db(t).$$

Since a, b are Lipschitz continuous, one has $a(s')(a(s') - a(s)) = \frac{1}{2}a(s')^2 - \frac{1}{2}a(s)^2 + O(|s - s'|^2)$ for any $s, s' \in [t_0, t_1]$, which easily implies that

$$\int_{t_0}^{t_1} a(t) \, da(t) = \frac{1}{2}a^2(t_1) - \frac{1}{2}a^2(t_0);$$

similarly we have

$$\int_{t_0}^{t_1} b(t) \, db(t) = \frac{1}{2}b^2(t_1) - \frac{1}{2}b^2(t_0),$$

and the claim follows. \square

REMARK 3.6. Geometrically, this conserved integral reflects the following elementary fact: if a square with vertices p_1, p_2, p_3, p_4 (traversed in anticlockwise order) and sidelength l is deformed to a nearby square with vertices $p_1 + dp_1, p_2 + dp_2, p_3 + dp_3, p_4 + dp_4$ and sidelength $l + dl$, then the difference of the areas of the two quadrilaterals with vertices $p_1, p_1 + dp_1, p_4 + dp_4, p_4$ and $p_2, p_2 + dp_2, p_3 + dp_3, p_3$ respectively add up to exactly half the difference between the areas $l^2, (l + dl)^2$ of the two squares (see Figure 3).

REMARK 3.7. One can interpret Lemma 3.5 in the language of differential forms as follows. For $i = 1, 2, 3, 4$, let $\pi_i : \text{Squares} \rightarrow \mathbb{R}^2$ be the i th coordinate

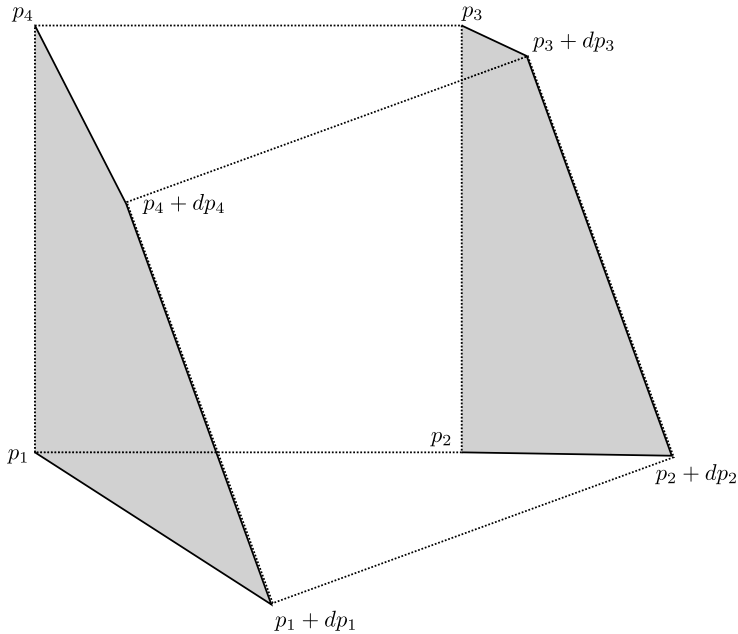


Figure 3. The difference in areas between the shaded regions is half the difference in areas between the squares. Taking ‘Riemann sums’ of this fact will yield Lemma 3.5.

projection, then we can pull back the 1-form $y dx$ on \mathbb{R}^2 by π_i to create a 1-form $\pi_i^*(y dx)$ on the 4-manifold $\overline{\text{Squares}}$. Then the identity (3.3) may be rewritten as

$$\pi_1^*(y dx) - \pi_2^*(y dx) + \pi_3^*(y dx) - \pi_4^*(y dx) = d\phi$$

where $\phi : \overline{\text{Squares}} \rightarrow \mathbb{R}$ is the 0-form that takes a square

$$((x, y), (x + a, y + b), (x + a - b, y + a + b), (x - b, y + a))$$

to the quantity $(a^2 - b^2)/2$, and d denotes the exterior derivative. In particular, the 1-form $\pi_1^*(y dx) - \pi_2^*(y dx) + \pi_3^*(y dx) - \pi_4^*(y dx)$ is exact.

Now we prove Theorem 1.2. Let $[t_0, t_1]$, f, g, ε be as in that theorem. It is convenient to extend the functions $f, g : [t_0, t_1] \rightarrow \mathbb{R}$ by constants to the whole real line \mathbb{R} to form extended functions $\tilde{f}, \tilde{g} : \mathbb{R} \rightarrow \mathbb{R}$. That is to say, we define $\tilde{f}(t) = \tilde{g}(t) := f(t_1) = g(t_1)$ for $t > t_1$, $\tilde{f}(t) = \tilde{g}(t) := f(t_0) = g(t_0)$ for all

$t < t_0$, and $\tilde{f}(t) := f(t)$ and $\tilde{g}(t) := g(t)$ for $t_0 \leq t \leq t_1$. Clearly $\tilde{f}, \tilde{g} : \mathbb{R} \rightarrow \mathbb{R}$ continue to be $(1 - \varepsilon)$ -Lipschitz and of bounded variation.

For any $t \in \mathbb{R}$, the map

$$(a, b) \mapsto (\tilde{g}(t - b) - \tilde{f}(t), \tilde{f}(t + a) - \tilde{f}(t)) \quad (3.4)$$

is a strict contraction on \mathbb{R}^2 (with the usual Euclidean metric) with contraction constant at most $1 - \varepsilon$. Hence, by the contraction mapping theorem (or Banach fixed point theorem) applied to the complete metric space \mathbb{R}^2 , there is a unique solution $(a(t), b(t)) \in \mathbb{R}^2$ to the fixed point equation

$$(a(t), b(t)) = (\tilde{g}(t - b(t)) - \tilde{f}(t), \tilde{f}(t + a(t)) - \tilde{f}(t)); \quad (3.5)$$

furthermore, $a(t)$ and $b(t)$ depend in a Lipschitz fashion on t (the Lipschitz constant can be as large as $O(1/\varepsilon)$, but this will not concern us). If we then define (similar curves also appear in the arguments of Jerrard [11]) the functions $\gamma_1, \gamma_2, \gamma_3, \gamma_4 : [t_0, t_1] \rightarrow \mathbb{R}^2$ by

$$\gamma_1(t) = (t, \tilde{f}(t)) \quad (3.6)$$

$$\gamma_2(t) = (t + a(t), \tilde{f}(t) + b(t)) \quad (3.7)$$

$$\gamma_3(t) = (t + a(t) - b(t), \tilde{f}(t) + a(t) + b(t)) \quad (3.8)$$

$$\gamma_4(t) = (t - b(t), \tilde{f}(t) + a(t)) \quad (3.9)$$

for all $t \in \mathbb{R}$, then $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ is a quadruple of Lipschitz (and thus locally rectifiable) curves that traverse squares. From (3.5)–(3.7), (3.9) we have

$$\gamma_1(t) = \text{Graph}_{\tilde{f}}(t) \quad (3.10)$$

$$\gamma_2(t) = \text{Graph}_{\tilde{f}}(t + a(t)) \quad (3.11)$$

$$\gamma_4(t) = \text{Graph}_{\tilde{g}}(t - b(t)) \quad (3.12)$$

for all $t \in \mathbb{R}$. In particular, $\gamma_1, \gamma_2, \gamma_4$ take values in $\text{Graph}_{\tilde{f}}(\mathbb{R})$, $\text{Graph}_{\tilde{f}}(\mathbb{R})$, and $\text{Graph}_{\tilde{g}}(\mathbb{R})$ respectively; see Figure 4. As for γ_3 , we can use the hypothesis of small Lipschitz constant to establish the following key fact:

PROPOSITION 3.8. *The curve $\gamma_3 : [t_0, t_1] \rightarrow \mathbb{R}^2$ is simple.*

Proof. Suppose for contradiction that $t, t' \in [t_0, t_1]$ are distinct points such that $\gamma_3(t) = \gamma_3(t')$. Then if one rotates the curve $\text{Graph}_{\tilde{f}}$ clockwise by $\pi/2$ around $\gamma_3(t) = \gamma_3(t')$, the rotated curve will intersect $\text{Graph}_{\tilde{g}}$ at the two distinct points $\gamma_2(t)$ and $\gamma_2(t')$ (see Figure 4). As \tilde{g} is $1 - \varepsilon$ -Lipschitz, we conclude that the line segment connecting these two points has slope of magnitude at most $1 - \varepsilon$ with

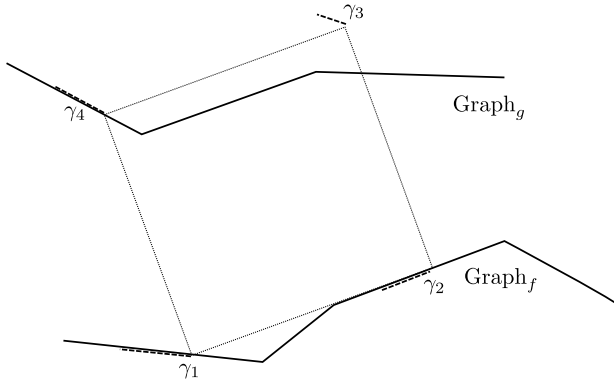


Figure 4. Portions of the curves $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \text{Graph}_f,$ and Graph_g . (In some cases, $\gamma_1, \gamma_2, \gamma_4$ may lie on the enlargements $\text{Graph}_{\tilde{f}}, \text{Graph}_{\tilde{g}}$ of $\text{Graph}_f, \text{Graph}_g$, which are not shown here.)

respect to the x -axis; as \tilde{f} is also $1 - \varepsilon$ -Lipschitz, we similarly conclude that the same line segment has slope of magnitude at most $1 - \varepsilon$ with respect to the y -axis. The two claims are inconsistent, giving the required contradiction. \square

When $t = t_0$ or $t = t_1$, we have $(a(t), b(t)) = (0, 0)$ as the unique fixed point of (3.5). Applying Lemma 3.5, we conclude the identity

$$\int_{\gamma_1} y \, dx - \int_{\gamma_2} y \, dx + \int_{\gamma_3} y \, dx - \int_{\gamma_4} y \, dx = 0. \quad (3.13)$$

By (3.10) and Example 3.3 we have

$$\int_{\gamma_1} y \, dx = \int_{t_0}^{t_1} f(t) \, dt. \quad (3.14)$$

From (3.11) and a change of variables (this change of variables is easy to justify if the map $t \mapsto t + a(t)$ is piecewise linear, and the general case follows by an approximation argument (noting that all functions involved are Lipschitz continuous)) $\tilde{t} := t + a(t)$ we also have

$$\int_{\gamma_2} y \, dx = \int_{t_0}^{t_1} f(t) \, dt; \quad (3.15)$$

note that $t + a(t)$ may be temporarily decreasing instead of increasing as t increases from t_0 to t_1 , but the net contributions of such excursions cancel out

(by the fundamental theorem of calculus, or equivalently because 1-forms on a line are automatically exact). Similarly, from (3.12) and the change of variables $\tilde{t} := t - b(t)$ we have

$$\int_{\gamma_4} y \, dx = \int_{t_0}^{t_1} g(t) \, dt. \quad (3.16)$$

Thus from (3.13) we must have

$$\int_{\gamma_3} y \, dx = \int_{t_0}^{t_1} g(t) \, dt$$

or equivalently

$$\int_{\gamma_3 + (-\text{Graph}_g)} y \, dx = 0$$

where $\gamma_3 + (-\text{Graph}_g)$ denotes the concatenation of γ_3 with the reversal of the graph Graph_g . This is a closed curve, hence by Lemma 3.4 this curve cannot be simple. Since γ_3 and Graph_g are separately simple (the former thanks to Proposition 3.8), we conclude that there exists $t_0 < t, t' < t_1$ such that

$$\gamma_3(t) = \text{Graph}_g(t').$$

In particular, $\gamma_1(t), \gamma_2(t), \gamma_3(t), \gamma_4(t)$ all lie in the set $\text{Graph}_{\tilde{f}} \cup \text{Graph}_{\tilde{g}}$. Since $\tilde{g}(t) > \tilde{f}(t)$, we see from (3.5) that $a(t)$ and $b(t)$ cannot both vanish; thus $(\gamma_1(t), \gamma_2(t), \gamma_3(t), \gamma_4(t))$ lie in Squares. Now we claim that all four vertices of this square in fact lie in the set (1.1). Indeed, suppose for contradiction that one of the vertices, call it v , was outside of (1.1), then it lies on the ray $\{(t, f(t_0)) : t < t_0\}$ or on the ray $\{(t, f(t_1)) : t > t_1\}$. But in either case, the set $\text{Graph}_{\tilde{f}} \cup \text{Graph}_{\tilde{g}} \setminus \{v\}$ is contained in the open double sector $v + \{(x, y) : |y| < |x|\}$, and hence $\text{Graph}_{\tilde{f}} \cup \text{Graph}_{\tilde{g}}$ cannot inscribe any square containing v as a vertex (as one cannot subtend a right angle at v). This implies that the set (1.1) inscribes a square as required, and Theorem 1.2 follows.

REMARK 3.9. It is instructive to compare the above argument with the following homological argument, which requires additional regularity hypotheses on f, g at the boundary points t_0, t_1 . Namely, suppose in addition to the hypotheses of Theorem 1.2 that f, g are differentiable at t_0, t_1 with $g'(t_0) > f'(t_0)$ and $g'(t_1) < f'(t_1)$; this corresponds the curve (1.1) being ‘locally monotone’ in the sense of [25] or [27], even at the endpoints $\text{Graph}_f(t_0) = \text{Graph}_g(t_0)$ and $\text{Graph}_f(t_1) = \text{Graph}_g(t_1)$. A local analysis then reveals that the curve $t \mapsto \gamma_3(t)$ defined above lies in the interior of (1.1) for t close to t_0 , and in the exterior of (1.1) for t close to t_1 , and so it must cross (1.1) at some point; indeed, if all

intersections were transversal, then it must cross this curve an odd number of times. (Actually, it is not difficult to use the Lipschitz hypotheses to show that this curve can only cross Graph_g and not Graph_f .) In contrast, the integral argument based on the conserved integral (3.3) does not give any information on the parity of crossings (indeed, if f, g are not differentiable at the endpoints, one could conceivably have an infinite number of transverse crossings near the endpoints $\text{Graph}_f(t_0) = \text{Graph}_g(t_0)$ and $\text{Graph}_f(t_1) = \text{Graph}_g(t_1)$), but do not require the functions f, g to be so regular at the endpoints t_0, t_1 that a local analysis is possible there.

REMARK 3.10. The following observations are due to Benjamin Matschke (private communication). The above arguments can be generalized to show that for any fixed $s, r > 0$, and with f, g as in Theorem 1.2 but with the Lipschitz constant $1 - \varepsilon$ replaced by $\tan(\alpha/2) - \varepsilon$ with $\alpha := \arctan(r/s) \in (0, \pi]$, the set $\text{Graph}_f \cup \text{Graph}_g$ inscribes a quadruple similar to the equilateral trapezoid

$$(0, 0), (1, 0), (s + 1, r), (-s, r)$$

or equivalently a quadruple of the form

$$(x, y), (x + a, y + b), (x + (s + 1)a - rb, y + (s + 1)b + ra), \\ (x + (-s)a - rb, y + (-s)b + ra).$$

Theorem 1.2 corresponds to the endpoint case $s = 0, r = 1$ of this more general claim. Indeed, by repeating the above arguments one can find Lipschitz curves $\gamma_1, \gamma_2, \gamma_3, \gamma_4 : [t_0, t_1] \rightarrow \mathbb{R}^2$ of the form

$$\begin{aligned} \gamma_1(t) &= (t, \tilde{f}(t)) \\ \gamma_2(t) &= (t + a(t), \tilde{f}(t) + b(t)) \\ \gamma_3(t) &= (t + (s + 1)a(t) - rb(t), y + (s + 1)b(t) + ra(t)) \\ \gamma_4(t) &= (t + (-s)a(t) - rb(t), y + (-s)b(t) + ra(t)) \end{aligned}$$

for some Lipschitz functions $a, b, \tilde{f} : [t_0, t_1] \rightarrow \mathbb{R}$ obeying (3.10)–(3.12), then one can again verify that γ_3 is simple, and a variant of the calculation used to prove Lemma 3.5 establishes the identity

$$\begin{aligned} (2s + 1) \int_{\gamma_1} y \, dx - (2s + 1) \int_{\gamma_2} y \, dx + \int_{\gamma_3} y \, dx - \int_{\gamma_4} y \, dx \\ = \frac{r(2s + 1)}{2} ((a(t_1)^2 - b(t_1)^2) - (a(t_0)^2 - b(t_0)^2)) \end{aligned}$$

and one can then conclude the claim by repeating the remaining arguments of this section; we leave the details to the interested reader. On the other hand, when

the equilateral trapezoid is not a rectangle or square, the known homological arguments do not seem to force the existence of an inscribed copy of the trapezoid even when the curve is smooth, because there are no symmetry reductions available to make the number of inscribed copies odd rather than even.

4. Periodic variants of the square peg problem

We now discuss periodic versions of the square peg problem, in which the plane \mathbb{R}^2 is replaced by the cylinder

$$\text{Cyl}_L := (\mathbb{R}/L\mathbb{Z}) \times \mathbb{R}$$

for some fixed period (one could easily normalize L to be 1 if desired, but we will find it convenient to allow L to be a parameter at our disposal) $L > 0$. There is an obvious projection map $\pi_L : \mathbb{R}^2 \rightarrow \text{Cyl}_L$ from the plane to the cylinder, which induces a projection $\pi_L^{\oplus 4} : (\mathbb{R}^2)^4 \rightarrow \text{Cyl}_L^4$; we let Squares_L and $\overline{\text{Squares}}_L$ be the images of Squares and $\overline{\text{Squares}}$ under this latter projection. More explicitly, we have

$$\begin{aligned} \text{Squares}_L := \{ & ((x, y), (x + a, y + b), (x + a - b, y + a + b), \\ & (x - b, y + a)) : x \in \mathbb{R}/L\mathbb{Z}; y, a, b \in \mathbb{R}; (a, b) \neq (0, 0) \} \end{aligned}$$

and

$$\begin{aligned} \overline{\text{Squares}}_L := \{ & ((x, y), (x + a, y + b), (x + a - b, y + a + b), \\ & (x - b, y + a)) : x \in \mathbb{R}/L\mathbb{Z}; y, a, b \in \mathbb{R} \}, \end{aligned}$$

where we define the sum $x + a$ of an element $x \in \mathbb{R}/L\mathbb{Z}$ and a real number $a \in \mathbb{R}$ in the obvious fashion. Again note that $\overline{\text{Squares}}_L$ is an oriented 4-manifold in Cyl_L^4 . As before, a subset Γ of Cyl_L is said to *inscribe a square* if Γ^4 intersects Squares_L . We give Cyl_L and Cyl_L^4 the usual flat Riemannian metric, which is then inherited by Squares_L and $\overline{\text{Squares}}_L$.

We have a standard closed curve $\text{Graph}_{0,L} : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ in Cyl_L defined by

$$\text{Graph}_{0,L}(t) := (t, 0);$$

one can think of $\text{Graph}_{0,L}$ homologically as a 1-cycle generating the first homology of Cyl_L . Any other closed curve $\gamma : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ will be homologous to $\text{Graph}_{0,L}$ if and only if it takes the form

$$\gamma(t \bmod L) = \pi_L(\tilde{\gamma}(t))$$

for some continuous lift $\tilde{\gamma} : \mathbb{R} \rightarrow \mathbb{R}^2$ of γ that is $L\mathbb{Z}$ -equivariant in the sense that

$$\tilde{\gamma}(t + L) = \tilde{\gamma}(t) + (L, 0) \tag{4.1}$$

for all $t \in \mathbb{R}$.

Amongst all the curves γ in Cyl_L , we isolate the *polygonal curves*, in which γ is piecewise linear (possibly after reparameterization), that is to say γ is the concatenation of finitely many line segments in Cyl_L .

We now introduce the following variant of Conjecture 1.1.

CONJECTURE 4.1 (Periodic square peg problem). *Let $L > 0$, and let $\sigma_1, \sigma_2 : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ be simple curves in Cyl_L homologous to $\text{Graph}_{0,L}$. Suppose also that the sets $\sigma_1(\mathbb{R}/L\mathbb{Z})$ and $\sigma_2(\mathbb{R}/L\mathbb{Z})$ are disjoint. Then $\sigma_1(\mathbb{R}/L\mathbb{Z}) \cup \sigma_2(\mathbb{R}/L\mathbb{Z})$ inscribes a square.*

In contrast to Conjecture 1.1, we do not know the answer to Conjecture 4.1 even when σ_1, σ_2 are smooth or piecewise polygonal (and we in fact suspect that resolving this case would soon resolve Conjecture 4.1 in full generality, in analogy to Corollary 5.5 below). This is because the intersection numbers produced by homological arguments become *even* in the periodic setting, rather than *odd*. Of course, by rescaling we could normalize $L = 1$ without loss of generality in Conjecture 4.1 if desired, but we find it preferable to not enforce this normalization.

We are able to relate Conjecture 1.1 to a special case of Conjecture 4.1. To state this special case, we need a further definition:

DEFINITION 4.2 (Infinitesimally inscribed squares). Let $L > 0$. A closed subset Γ of Cyl_L is said to *inscribe infinitesimal squares* if there exists a sequence of squares

$$S_n = ((x_n, y_n), (x_n + a_n, y_n + b_n), (x_n + a_n - b_n, y_n + a_n + b_n), (x_n - b_n, y_n + a_n)) \tag{4.2}$$

in Squares_L converging to a degenerate square $((x, y), (x, y), (x, y), (x, y))$ for some $(x, y) \in \Gamma$, such that

$$\text{dist}(S_n, \Gamma^4) = o(|a_n| + |b_n|)$$

as $n \rightarrow \infty$.

Note that the property of infinitesimally inscribing squares is a purely local property: to show that a set Γ does not infinitesimally inscribe squares, it suffices

to show that for every $p \in \Gamma$, there is a set Γ_p that agrees with Γ in a neighbourhood of p that does not infinitesimally inscribe squares.

We now give two examples of sets with the property of not inscribing infinitesimal squares.

LEMMA 4.3. *Let $f_1, \dots, f_k : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{R}$ be C -Lipschitz functions for some $C < \tan(3\pi/8) = 1 + \sqrt{2}$, such that the graphs $\text{Graph}_{f_i}(\mathbb{R}/L\mathbb{Z})$ for $i = 1, \dots, k$ are all disjoint. Then the set $\bigcup_{i=1}^k \text{Graph}_{f_i}(\mathbb{R}/L\mathbb{Z})$ does not inscribe infinitesimal squares.*

Proof. By the local nature of infinitesimally inscribing squares, it suffices to show that each $\text{Graph}_{f_i}(\mathbb{R}/L\mathbb{Z})$ does not infinitesimally inscribe squares. Suppose for contradiction that there was some $i = 1, \dots, k$ and a sequence of squares (4.2) with $(x_n, y_n) \rightarrow (x, y) \in \text{Graph}_{f_i}(\mathbb{R}/L\mathbb{Z})$, $(a_n, b_n) \rightarrow (0, 0)$, and all vertices staying within $o(|a_n| + |b_n|)$ of $\text{Graph}_{f_i}(\mathbb{R}/L\mathbb{Z})$. As f_i is C -Lipschitz continuous, this implies that the eight points $\pm(a_n, b_n)$, $\pm(-b_n, a_n)$, $\pm(a_n - b_n, a_n + b_n)$, $\pm(a_n + b_n, a_n - b_n)$ all lie in the double sector $\{(t, u) : |u| \leq (C + o(1))|t|\}$. However, the arguments of these eight points (viewed as complex numbers) form a coset of the eighth roots of unity, while the double sector omits all the complex numbers of argument in $[3\pi/8, 5\pi/8]$ if n is large enough; but these two facts are in contradiction. \square

REMARK 4.4. If one rotates the standard unit square $[0, 1]^2$ by $\pi/8$, one obtains a square with the property that all its sides and diagonals have slope between $-\tan(3\pi/8)$ and $\tan(3\pi/8)$; in particular, the vertices of this square can be traversed by the graph of a $\tan(3\pi/8)$ -Lipschitz function. Gluing together infinitely many rescaled copies of this function, it is not difficult to show that the condition $C < \tan(3\pi/8)$ in Lemma 4.3 cannot be improved.

LEMMA 4.5. *Let $\Gamma_1, \dots, \Gamma_k$ be disjoint simple polygonal paths in \mathbb{R}^2 (either open or closed). Then $\Gamma_1 \cup \dots \cup \Gamma_k$ does not infinitesimally inscribe squares.*

Proof. Again it suffices to verify that a single Γ_i does not inscribe squares. Suppose for contradiction that there was a sequence of squares with vertices (4.2) with $(x_n, y_n) \rightarrow (x, y) \in \Gamma_i$, $(a_n, b_n) \rightarrow (0, 0)$, and all vertices staying within $o(|a_n| + |b_n|)$ of Γ_i .

We can translate so that $(x, y) = (0, 0)$. The origin $(0, 0)$ is either a vertex on the path Γ_i or an interior point of an edge. Suppose first that $(0, 0)$ is an interior point of an edge, which then lies on some line ℓ . Then for n large enough, all four vertices (4.2) stay within $o(|a_n| + |b_n|)$ of this line. Applying a suitable translation

and rescaling, we can then obtain another family of squares of unit length, whose vertices (4.2) are bounded and stay within $o(1)$ of ℓ . Using compactness to extract a limit, we obtain a unit square with all four vertices on ℓ , which is absurd.

Now suppose that $(0, 0)$ is a vertex of Γ_i , which we may take to be the origin. This origin is the meeting point of two edges of Γ_i that lie on two distinct lines ℓ_1, ℓ_2 passing through the origin. If (after passing through a subsequence) all four vertices (4.2) lie within $o(|a_n| + |b_n|)$ of ℓ_1 , then by rescaling and taking limits as before we obtain a unit square with all four vertices on ℓ_1 , which is absurd; similarly if all four vertices lie within $o(|a_n| + |b_n|)$ of ℓ_2 . Thus we must have at least one vertex within $o(|a_n| + |b_n|)$ of ℓ_1 and another within $o(|a_n| + |b_n|)$ of ℓ_2 , which forces the entire square to lie within $O(|a_n| + |b_n|)$ of the origin. Rescaling and taking limits again, we now obtain a unit square with all four vertices on the union $\ell_1 \cup \ell_2$ of the two intersecting lines ℓ_1, ℓ_2 , which is again absurd regardless of what angle ℓ_1 and ℓ_2 make with each other. \square

For an example of a curve that does infinitesimally inscribe squares, one can consider any curve that has the local behaviour of a cusp such as $\{(t^2, t^3) : t \in \mathbb{R}\}$.

We now isolate a special case of Conjecture 4.1:

CONJECTURE 4.6 (Periodic square peg problem, special case). *Conjecture 4.1 is true under the additional hypothesis that $\sigma_1(\mathbb{R}/L\mathbb{Z}) \cup \sigma_2(\mathbb{R}/L\mathbb{Z})$ does not inscribe infinitesimal squares.*

The main result of this section is then

PROPOSITION 4.7. *Conjecture 1.1 implies Conjecture 4.6. In particular (by Lemma 4.5), Conjecture 1.1 implies the special case of Conjecture 4.1 when the curves σ_1, σ_2 are polygonal paths.*

To put it another way, if one wished to disprove Conjecture 1.1, it would suffice to produce a union $\sigma_1(\mathbb{R}) \cup \sigma_2(\mathbb{R})$ of two disjoint periodic curves which did not inscribe any squares or infinitesimal squares.

Proposition 4.7 is an immediate consequences of the following proposition:

PROPOSITION 4.8 (Transforming periodic sets to bounded sets). *Let Γ be a compact subset of $\mathbb{C} \setminus \mathbb{R}$ for some $L > 0$, and let $\pi_L^{-1}(\Gamma)$ be its lift to \mathbb{R}^2 . For every large natural number n , let $\phi_n : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ denote the map (one can view this map as an approximation to the conformal map $z \mapsto n \tanh(z/n)$ in the complex plane, which ‘gently pinches’ the periodic set $\pi_L^{-1}(\Gamma)$ to a bounded set in a manner that almost preserves squares. The use of trigonometric functions here is*

primarily for notational convenience; one could also use other maps than ϕ_n here as long as they obeyed the above-mentioned qualitative features of approximate conformality and mapping periodic sets to bounded sets)

$$\phi_n(x, y) := \left(n \tanh \frac{x}{n}, y \operatorname{sech}^2 \frac{x}{n} \right).$$

Then at least one of the following three statements hold:

- (i) Γ inscribes a square;
- (ii) Γ inscribes infinitesimal squares;
- (iii) for sufficiently large n , $\phi_n(\pi_L^{-1}(\Gamma)) \cup \{(-n, 0), (n, 0)\}$ does not inscribe a square.

Indeed, to establish Conjecture 4.6 assuming Conjecture 1.1, one simply applies Proposition 4.8 to the set $\Gamma := \sigma_1(\mathbb{R}/L\mathbb{Z}) \cup \sigma_2(\mathbb{R}/L\mathbb{Z})$; the conclusion (ii) is ruled out by hypothesis and the conclusion (iii) is ruled out by Conjecture 1.1, leaving (i) as the only possible option.

Proof. We will assume that (iii) fails and conclude that either (i) or (ii) holds.

By hypothesis, we can find a sequence of n going to infinity, and a sequence of squares with vertices (4.2), such that each square (4.2) is inscribed in $\phi_n(\pi_L^{-1}(\Gamma))$. The plan is to transform these squares into squares that either converge to a square inscribed in Γ , or become an infinitesimal inscribed square in Γ .

We first rule out a degenerate case when one of the points $(-n, 0)$, $(n, 0)$ is one of the vertices (4.2). Suppose for instance that (x_n, y_n) was equal to $(n, 0)$. Since Γ is compact, we see that $\pi_L^{-1}(\Gamma)$ is contained in a strip of the form

$$\{(x, y) : y = O(1)\}.$$

Using the identity

$$\operatorname{sech}^2(x) = 1 - \tanh^2(x) = O(1 - |\tanh(x)|),$$

we conclude that $\phi_n(\pi_L^{-1}(\Gamma))$ is contained in the region

$$\left\{ (x, y) : -n < x < n; y = O\left(1 - \frac{|x|}{n}\right) \right\}. \quad (4.3)$$

If $(x_n, y_n) = (n, 0)$ and the remaining three vertices of (4.2) lie in $\phi_n(\pi_L^{-1}(\Gamma))$, this forces $(-a_n, -b_n)$, $(-a_n + b_n, -a_n - b_n)$, $(b_n, -a_n)$ to all have argument $O(1/n)$ when viewed as complex numbers, which is absurd for n large enough,

since these arguments differ by $\pi/4$ or $\pi/2$. Thus we have $(x_n, y_n) \neq (n, 0)$ (after passing to a subsequence of n if necessary); a similar argument precludes any of the vertices in (4.2) being equal to $(-n, 0)$ or $(n, 0)$.

At least one of the four vertices in (4.2) must have a y -coordinate of magnitude at least $(|a_n| + |b_n|)/2$, since two of these y -coordinates differ by $a_n + b_n$ and the other two differ by $a_n - b_n$. Applying (4.3), we conclude that the x -coordinate of that vertex lies at a distance at least $cn|a_n| + |b_n|$ from $(-n, 0)$, $(n, 0)$ for some $c > 0$ independent of n ; by the triangle inequality, we conclude (for n large enough) that all four vertices have this property. In particular, we have

$$|a_n| + |b_n| = O\left(1 - \frac{|x_n|}{n}\right).$$

Observe that in the region (4.3), we may invert ϕ_n by the formula

$$\phi_n^{-1}(x, y) := \left(\frac{n}{2} \log \frac{n+x}{n-x}, \frac{y}{1-x^2/n^2}\right).$$

On (4.3), we can compute the partial derivatives

$$\begin{aligned} \frac{\partial}{\partial x} \phi_n^{-1}(x, y) &= \left(\frac{1}{1-x^2/n^2}, \frac{2x}{n^2} \frac{y}{(1-x^2/n^2)^2}\right) \\ &= \frac{1}{1-x^2/n^2} \left(1, O\left(\frac{1}{n}\right)\right) \end{aligned}$$

and

$$\frac{\partial}{\partial y} \phi_n^{-1}(x, y) = \left(0, \frac{1}{1-x^2/n^2}\right)$$

and so by Taylor expansion we see that

$$\begin{aligned} \phi^{-1}(x_n + a_n, y_n + b_n) &= \phi^{-1}(x_n, y_n) + \frac{(a_n, b_n)}{1-x_n^2/n^2} \\ &\quad + O\left(\frac{|a_n| + |b_n|}{n(1-x_n^2/n^2)}\right) \\ \phi^{-1}(x_n + a_n - b_n, y_n + a_n + b_n) &= \phi^{-1}(x_n, y_n) + \frac{(a_n - b_n, a_n + b_n)}{1-x_n^2/n^2} \\ &\quad + O\left(\frac{|a_n| + |b_n|}{n(1-x_n^2/n^2)}\right) \end{aligned}$$

$$\begin{aligned} \phi^{-1}(x_n - b_n, y_n + a_n) &= \phi^{-1}(x_n, y_n) + \frac{(-b_n, a_n)}{1 - x_n^2/n^2} \\ &\quad + O\left(\frac{|a_n| + |b_n|}{n(1 - x_n^2/n^2)}\right). \end{aligned}$$

Thus, if we set $(\tilde{x}_n, \tilde{y}_n) := \pi_L(\phi^{-1}(x_n, y_n))$ and $(\tilde{a}_n, \tilde{b}_n) := (a_n, b_n)/(1 - x_n^2/n^2)$, we see that $\tilde{a}_n, \tilde{b}_n = O(1)$, and the four vertices of the square

$$\begin{aligned} &((\tilde{x}_n, \tilde{y}_n), (\tilde{x}_n + \tilde{a}_n, \tilde{y}_n + \tilde{b}_n), (\tilde{x}_n + \tilde{a}_n - \tilde{b}_n, \tilde{y}_n + \tilde{a}_n + \tilde{b}_n), (\tilde{x}_n - \tilde{b}_n, \tilde{y}_n + \tilde{a}_n)) \\ &\in \text{Squares}_L \end{aligned}$$

all lie within $O(|\tilde{a}_n| + |\tilde{b}_n|/n)$ of Γ .

By passing to a subsequence, we may assume that $(\tilde{a}_n, \tilde{b}_n)$ converges to some pair (a, b) , which may possibly be equal to $(0, 0)$. By compactness of Γ , we may similarly assume that $(\tilde{x}_n, \tilde{y}_n)$ converges to a limit $(x, y) \in \Gamma$. If $(a, b) \neq (0, 0)$, then on taking limits using the closed nature of Γ we conclude that the nondegenerate square

$$((x, y), (x + a, y + b), (x + a - b, y + a + b), (x - b, y + a)) \in \text{Squares}_L$$

is inscribed in Γ , giving (i); if instead $(a, b) = (0, 0)$ then we obtain (ii). \square

REMARK 4.9. In contrast to the smooth cases of Conjecture 1.1, there are no homological obstructions to establishing a counterexample to Conjecture 4.1. For instance, when the Lipschitz constants of f, g are strictly less than one, the arguments of the previous section can be used to produce a quadruplet $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ of rectifiable curves $\gamma_1, \gamma_2, \gamma_3, \gamma_4 : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ traversing squares, with $\gamma_1, \gamma_2, \gamma_4$ taking values in $\text{Graph}_f(\mathbb{R}/L\mathbb{Z}), \text{Graph}_f(\mathbb{R}/L\mathbb{Z}), \text{Graph}_g(\mathbb{R}/L\mathbb{Z})$ respectively, and all four curves homologous to the standard 1-cycle $\text{Graph}_{0,L}$. In particular, γ_3 would (assuming sufficient transversality and regularity) intersect the graphs $\text{Graph}_f(\mathbb{R}/L\mathbb{Z})$ and $\text{Graph}_g(\mathbb{R}/L\mathbb{Z})$ an even number of times per unit period, rather than an odd number of times. This is of course consistent with the curve not intersecting these graphs at all. The use of an infinite oscillation to switch the parity of intersection from odd to even is reminiscent of the ‘Eilenberg–Mazur swindle’ (see for example [20]).

5. A quadripartite variant

In Conjecture 4.1, the four vertices of the square could be distributed arbitrarily among the two graphs $\text{Graph}_f(\mathbb{R}/L\mathbb{Z})$ and $\text{Graph}_g(\mathbb{R}/L\mathbb{Z})$. It seems to be more natural to force each vertex to lie in just one of the two graphs. To formulate this more precisely, we introduce a further definition:

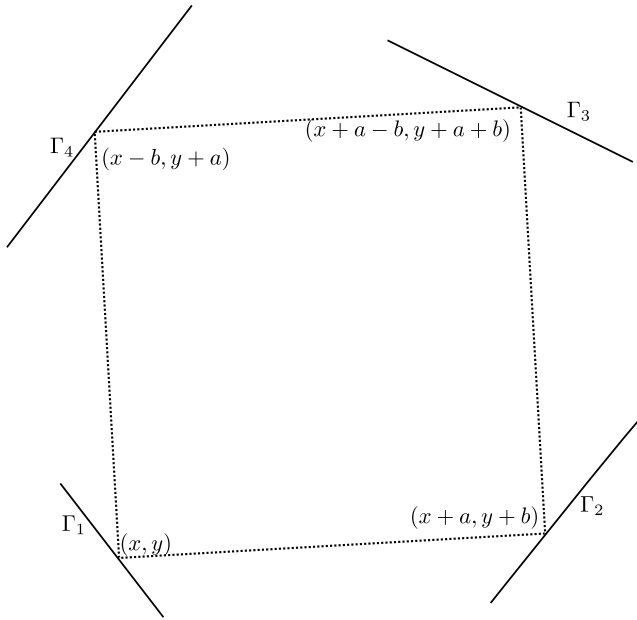


Figure 5. Four line segments $(\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4)$ jointly inscribing a square. The order (up to cyclic permutation) is important; for instance, in the given picture, the quadruple $(\Gamma_1, \Gamma_4, \Gamma_3, \Gamma_2)$ does *not* jointly inscribe a square.

DEFINITION 5.1 (Jointly inscribing squares). Let $L > 0$. Let $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ be four sets in Cyl_L (possibly overlapping). We say that the quadruplet $(\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4)$ *jointly inscribes a square* if $\Gamma_1 \times \Gamma_2 \times \Gamma_3 \times \Gamma_4$ intersects $\overline{\text{Squares}_L}$, or equivalently if there exist $x \in \mathbb{R}/L\mathbb{Z}$ and $y, a, b \in \mathbb{R}$ such that

$$\begin{aligned} (x, y) &\in \Gamma_1 \\ (x + a, y + b) &\in \Gamma_2 \\ (x + a - b, y + a + b) &\in \Gamma_3 \\ (x - b, y + a) &\in \Gamma_4. \end{aligned}$$

See Figure 5.

Note in Definition 5.1 that we now permit the inscribed square to be degenerate. Conjecture 4.1 would then follow from

CONJECTURE 5.2 (Quadripartite periodic square peg problem). *Let $L > 0$, and let $\sigma_1, \sigma_2 : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ be simple closed curves homologous to $\text{Graph}_{0,L}$.*

Then the quadruplet

$$(\sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}))$$

jointly inscribes a square.

Indeed, if σ_1, σ_2 are as in Conjecture 4.1, we can apply Conjecture 5.2 to obtain $x \in \mathbb{R}/L\mathbb{Z}$ and $y, a, b \in \mathbb{R}$ with

$$(x, y), (x + a, y + b) \in \sigma_1(\mathbb{R}/L\mathbb{Z})$$

and $(x + a - b, y + a + b), (x - b, y + a) \in \sigma_2(\mathbb{R}/L\mathbb{Z})$. As $\sigma_1(\mathbb{R}/L\mathbb{Z})$ and $\sigma_2(\mathbb{R}/L\mathbb{Z})$ are assumed disjoint in Conjecture 4.1, we have $(a, b) \neq (0, 0)$, and Conjecture 4.1 follows.

We can reverse this implication in some cases:

PROPOSITION 5.3. *Conjecture 4.1 implies Conjecture 5.2 in the special case that $\sigma_i(\mathbb{R}/L\mathbb{Z})$ does not inscribe squares for $i = 1, 2$.*

Note that the hypothesis that $\sigma_i(\mathbb{R}/L\mathbb{Z})$ does not inscribe squares is satisfied in many cases; for instance, by modifying the proof of Lemma 4.3 we see that this is the case if σ_i is the graph of a C -Lipschitz function for some $C < \tan(3\pi/8)$.

Proof. Let σ_1, σ_2 be as in Conjecture 5.2, and assume that $\sigma_1(\mathbb{R}/L\mathbb{Z})$ and $\sigma_2(\mathbb{R}/L\mathbb{Z})$ do not separately inscribe squares. Let m be a sufficiently large natural number, then $\sigma_1(\mathbb{R}/L\mathbb{Z})$ and $\sigma_2(\mathbb{R}/L\mathbb{Z}) + (0, mL)$ will be disjoint. Applying Conjecture 4.1, we may find a square

$$((x, y), (x + a, y + b), (x + a - b, y + a + b), (x - b, y + a)) \in \text{Squares}_L$$

inscribed in $\sigma_1(\mathbb{R}/L\mathbb{Z}) \cup (\sigma_2(\mathbb{R}/L\mathbb{Z}) + (0, mL))$. In particular, we have

$$y, y + b, y + a + b, y + a \in [-C, C] \cup [mL - C, mL + C]$$

for some $C > 0$ independent of m . If m is large enough, this forces the quadruple $(y, y + b, y + a + b, y + a)$ to be of the form $(O(C), O(C), O(C), O(C))$, the form $(mL + O(C), mL + O(C), mL + O(C), mL + O(C))$, or some cyclic permutation of $(O(C), O(C), mL + O(C), mL + O(C))$. In the first case, we have a square inscribed in $\sigma_1(\mathbb{R}/L\mathbb{Z})$, and in the second case we have (after translation by $(0, mL)$) a square inscribed by $\sigma_2(\mathbb{R}/L\mathbb{Z})$; both these cases are ruled out by hypothesis. Thus, after cyclic permutation, we may assume that

$$(y, y + b, y + a + b, y + a) = (O(C), O(C), mL + O(C), mL + O(C))$$

which implies that the (possibly degenerate) square

$$\begin{aligned} &((x, y), (x + a - mL, y + b), (x + a - b - mL, y + a + b), \\ &(x - b, y + a - mL)) \in \overline{\text{Squares}_L} \end{aligned}$$

is jointly inscribed by $(\sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}))$, giving Conjecture 5.2 in this case. \square

Because the squares in Definition 5.1 are now permitted to be degenerate, Conjecture 5.2 enjoys good convergence properties with respect to limits:

PROPOSITION 5.4 (Stability of not jointly inscribing squares). *Let $L > 0$. Let $\sigma_{1,n}, \sigma_{2,n} : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ be sequences of simple closed curves which converge uniformly to simple closed curves $\sigma_1, \sigma_2 : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ respectively as $n \rightarrow \infty$. If each of the quadruples*

$$(\sigma_{1,n}(\mathbb{R}/L\mathbb{Z}), \sigma_{1,n}(\mathbb{R}/L\mathbb{Z}), \sigma_{2,n}(\mathbb{R}/L\mathbb{Z}), \sigma_{2,n}(\mathbb{R}/L\mathbb{Z}))$$

jointly inscribe a square, then so does

$$(\sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z})).$$

It is possible to weaken the requirement of uniform convergence (for instance, one can just assume pointwise convergence if the curves $\sigma_{1,n}, \sigma_{2,n}$ are uniformly bounded), but we will not need to do so here.

Proof. By hypothesis, one can find a sequence $p_n \in \overline{\text{Squares}_L}$ such that $p_n \in \sigma_{1,n}(\mathbb{R}/L\mathbb{Z}) \times \sigma_{1,n}(\mathbb{R}/L\mathbb{Z}) \times \sigma_{2,n}(\mathbb{R}/L\mathbb{Z}) \times \sigma_{2,n}(\mathbb{R}/L\mathbb{Z})$ for all n . As $\sigma_{1,n}, \sigma_{2,n}$ converge uniformly to σ_1, σ_2 , the p_n are bounded and thus have at least one limit point p , which must lie in both $\overline{\text{Squares}_L}$ and $\sigma_1(\mathbb{R}/L\mathbb{Z}) \times \sigma_1(\mathbb{R}/L\mathbb{Z}) \times \sigma_2(\mathbb{R}/L\mathbb{Z}) \times \sigma_2(\mathbb{R}/L\mathbb{Z})$, giving the claim. \square

As one application of this proposition, we have

COROLLARY 5.5. *In order to prove Conjecture 5.2, it suffices to do so in the case that the simple closed curves $\sigma_1, \sigma_2 : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ are polygonal paths.*

This can be compared with the situation with Conjecture 1.1, which is known to be true for polygonal paths, but for which one cannot take limits to conclude the general case, due to the possibility of the inscribed squares degenerating to zero.

Proof. By Proposition 5.4, it suffices to show that any simple closed curve $\gamma : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ can be uniformly approximated to any desired accuracy $O(\varepsilon)$ for $\varepsilon > 0$ by a simple polygonal closed curve $\tilde{\gamma} : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ (note from a winding number argument that if γ is homologous to $\text{Graph}_{0,L}$, then $\tilde{\gamma}$ will be also if ε is small enough). By uniform continuity, there exists a natural number N such that $d_{\text{Cyl}_L}(\gamma(t), \gamma(t')) \leq \varepsilon$ whenever $d_{\mathbb{R}/L\mathbb{Z}}(t, t') \leq L/N$, where d_{Cyl_L} , $d_{\mathbb{R}/L\mathbb{Z}}$ denote the Riemannian distance functions on Cyl_L , $\mathbb{R}/L\mathbb{Z}$ respectively; as γ is simple and continuous, a compactness argument shows that there also exists $0 < \delta < \varepsilon$ such that $d_{\text{Cyl}_L}(\gamma(t), \gamma(t')) > 4\delta$ whenever $d_{\mathbb{R}/L\mathbb{Z}}(t, t') \geq L/N$. Finally, by uniform continuity again, there exists a natural number $M \geq N$ such that $d_{\text{Cyl}_L}(\gamma(t), \gamma(t')) \leq \delta$ whenever $d_{\mathbb{R}/L\mathbb{Z}}(t, t') \leq L/M$.

Let $\gamma_1 : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{R}^2$ be the polygonal path such that $\gamma_1(jL/M) := \gamma(jL/M)$ for every $j \in \mathbb{Z}/M\mathbb{Z}$, with γ_1 linear on each interval $jL/M + [0, 1/M]$. From the triangle inequality we see that $d_{\text{Cyl}_L}(\gamma_1(t), \gamma(t)) \leq 2\delta$ for all $t \in \mathbb{R}$. Unfortunately, γ_1 need not be simple. However, if t, t' are such that $\gamma_1(t) = \gamma_1(t')$, then by the triangle inequality we have $d_{\text{Cyl}_L}(\gamma(t), \gamma(t')) \leq 4\delta$, and hence $d_{\mathbb{R}/L\mathbb{Z}}(t, t') \leq L/N$. Using a greedy algorithm to iteratively remove loops from the polygonal path γ_1 (with each removal decreasing the number of edges remaining in γ_1 in a unit period), we may thus find a finite family I_1, \dots, I_k of disjoint closed intervals in $\mathbb{R}/L\mathbb{Z}$, each of length at most L/N , such that paths $\gamma_1|_{I_j} : I_j \rightarrow \text{Cyl}_L$ is closed for each $1 \leq j \leq k$ (that is, γ_1 evaluates to the same point at the left and right endpoints of I_j), and such that γ_1 becomes simple once each of the intervals I_j is contracted to a point. Note also that all of the loops of γ_1 removed by this process have diameter $O(\varepsilon)$. If one then chooses a sufficiently small neighbourhood interval \tilde{I}_j for each I_j , and defines $\tilde{\gamma}$ to equal γ_1 outside $\bigcup_{j=1}^k \tilde{I}_j$, and linear on each of the \tilde{I}_j , then we see that $\tilde{\gamma}$ is a simple $L\mathbb{Z}$ -equivariant polygonal path that lies within $O(\varepsilon)$ of γ , as required. \square

We in fact believe the following stronger claim than Conjecture 5.2 to hold:

CONJECTURE 5.6 (Area inequality). *Let $L > 0$, and let $\sigma_1, \sigma_2, \sigma_3, \sigma_4 : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ be simple closed polygonal paths homologous to $\text{Graph}_{0,L}$. If $(\sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}), \sigma_3(\mathbb{R}/L\mathbb{Z}), \sigma_4(\mathbb{R}/L\mathbb{Z}))$ does not jointly inscribe a square, then*

$$\int_{\sigma_1} y \, dx - \int_{\sigma_2} y \, dx + \int_{\sigma_3} y \, dx - \int_{\sigma_4} y \, dx \neq 0. \quad (5.1)$$

Note that the 1-form $y \, dx$ is well defined on Cyl_L and $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ can be viewed as 1-cycles, so the integrals in (5.1) make sense as integration of differential forms (but one could also use Definition 3.2 with the obvious

modifications if desired). One can also simplify the left-hand side of (5.1) as

$$\int_{\sigma_1 - \sigma_2 + \sigma_3 - \sigma_4} y \, dx \tag{5.2}$$

where $\sigma_1 - \sigma_2 + \sigma_3 - \sigma_4$ is interpreted as a 1-cycle. Viewed contrapositively, Conjecture 5.6 then asserts that $(\sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}), \sigma_3(\mathbb{R}/L\mathbb{Z}), \sigma_4(\mathbb{R}/L\mathbb{Z}))$ must inscribe a square whenever the integral (5.2) vanishes.

Clearly, Conjecture 5.2 follows from Conjecture 5.6 by first using Corollary 5.5 to reduce to the case where σ_1, σ_2 are polygonal paths, and then applying Conjecture 5.6 with $(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$ replaced by $(\sigma_1, \sigma_1, \sigma_2, \sigma_2)$.

Conjecture 5.6 is somewhat strong compared with the other conjectures in this paper. Nevertheless, we can repeat the arguments in Section 3 to obtain some evidence for it:

THEOREM 5.7. *Conjecture 5.6 holds when $\sigma_2 = \text{Graph}_{f_2}$ and $\sigma_4 = \text{Graph}_{f_4}$ for some $(1 - \varepsilon)$ -Lipschitz functions $f_2, f_4 : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ and some $\varepsilon > 0$.*

Of course, by cyclic permutation one can replace the role of σ_2, σ_4 here by σ_1, σ_3 .

Proof. Write $\sigma_1(t) := (x_1(t), y_1(t))$. For any $t \in \mathbb{R}/L\mathbb{Z}$, the map

$$(a, b) \mapsto (f_4(x_1(t) - b) - y_1(t), f_2(x_1(t) + a) - y_1(t))$$

is a contraction on \mathbb{R}^2 with constant at most $1 - \varepsilon$, hence by the contraction mapping theorem there exist unique continuous functions $a, b : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{R}$ such that

$$(a(t), b(t)) = (f_4(x_1(t) - b(t)) - y_1(t), f_2(x_1(t) + a(t)) - y_1(t))$$

for all $t \in \mathbb{R}/L\mathbb{Z}$. As $\sigma_1, \sigma_2, \sigma_4$ are polygonal paths, f_2, f_4 are piecewise linear functions, which implies that a, b are also piecewise linear. We then set $\gamma_1, \gamma_2, \gamma_3, \gamma_4 : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ to be the polygonal paths

$$\begin{aligned} \gamma_1(t) &:= \sigma_1(t) \\ \gamma_2(t) &:= \sigma_1(t) + (a(t), b(t)) \\ &= \text{Graph}_{f_2}(x_1(t) + a(t)) \\ \gamma_3(t) &:= \sigma_1(t) + (a(t) - b(t), a(t) + b(t)) \\ \gamma_4(t) &:= \sigma_1(t) + (-b(t), a(t)) \\ &= \text{Graph}_{f_4}(x_1(t) - b(t)). \end{aligned}$$

Clearly $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ traverses squares. Applying Lemma 3.5, we conclude that

$$\int_{\gamma_1} y \, dx - \int_{\gamma_2} y \, dx + \int_{\gamma_3} y \, dx - \int_{\gamma_4} y \, dx = 0.$$

Since $t \mapsto \text{Graph}_{f_2}(x_1(t)+a(t))$ and $t \mapsto \text{Graph}_{f_4}(x_1(t)-b(t))$ are homologous to σ_2 and σ_4 respectively, and all 1-forms on curves such as $\text{Graph}_{f_2}(\mathbb{R}/L\mathbb{Z})$ or $\text{Graph}_{f_4}(\mathbb{R}/L\mathbb{Z})$ are automatically closed, we have $\int_{\gamma_i} y \, dx = \int_{\sigma_i} y \, dx$ for $i = 1, 2, 4$. Thus it suffices to show that

$$\int_{\gamma_3 - \sigma_3} y \, dx \neq 0. \quad (5.3)$$

The argument in Proposition 3.8 (unwrapping the curves from Cyl_L to \mathbb{R}^2) shows that γ_3 is simple, and the graph σ_3 is of course also simple. As we are assuming $(\sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}), \sigma_3(\mathbb{R}/L\mathbb{Z}), \sigma_4(\mathbb{R}/L\mathbb{Z}))$ to not jointly inscribe squares, $\gamma_3(\mathbb{R}/L\mathbb{Z})$ and $\sigma_3(\mathbb{R}/L\mathbb{Z})$ must stay disjoint. The curve γ_3 is homotopic to σ_1 and thus also homologous to $\text{Graph}_{0,L}$. Applying the Jordan curve theorem, we see that the closed polygonal paths γ_3, σ_3 enclose some nonempty polygonal region Ω in Cyl_L . By Stokes' theorem, we conclude that the left-hand side of (5.3) is equal to some sign times the Lebesgue measure of Ω , giving (5.3) as required. \square

We also have an analogue of Corollary 5.5:

PROPOSITION 5.8 (Stability of area inequality). *Let $L > 0$. Suppose that $\sigma_{1,n}, \sigma_{2,n}, \sigma_{3,n}, \sigma_{4,n} : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ are simple closed polygonal paths which converge uniformly to simple closed polygonal paths $\sigma_1, \sigma_2, \sigma_3, \sigma_4 : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ respectively. If, for all sufficiently small $h \in \mathbb{R}$ Conjecture 5.6 holds for each of the quadruples $(\sigma_{1,n}, \sigma_{2,n}, \sigma_{3,n}, \sigma_{4,n} + (0, h))$ for all $n \geq 1$, then it also holds for $(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$.*

This proposition will be useful for placing the polygonal paths $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ in 'general position'.

Proof. We can of course assume that $(\sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}), \sigma_3(\mathbb{R}/L\mathbb{Z}), \sigma_4(\mathbb{R}/L\mathbb{Z}))$ does not jointly inscribe a square, as the claim is trivial otherwise. By Proposition 5.4 applied in the contrapositive, we see that there exists an $\varepsilon > 0$ and $N \geq 1$ such that $(\sigma_{1,n}(\mathbb{R}/L\mathbb{Z}), \sigma_{1,n}(\mathbb{R}/L\mathbb{Z}), \sigma_{1,n}(\mathbb{R}/L\mathbb{Z}), \sigma_{1,n}(\mathbb{R}) + (0, h))$ does not jointly inscribe squares for any $|h| \leq \varepsilon$ and $n \geq N$. Applying the hypothesis (and shrinking ε if necessary), we conclude that

$$\int_{\sigma_{1,n} - \sigma_{2,n} + \sigma_{3,n} - \sigma_{4,n}} y \, dx - Lh \neq 0$$

for $n \geq N$ and $|y| \leq \varepsilon$, and hence

$$\left| \int_{\sigma_{1,n} - \sigma_{2,n} + \sigma_{3,n} - \sigma_{4,n}} y \, dx \right| \geq L\varepsilon;$$

taking limits as $n \rightarrow \infty$ we conclude that

$$\left| \int_{\sigma_1 - \sigma_2 + \sigma_3 - \sigma_4} y \, dx \right| \geq L\varepsilon$$

giving the claim. □

In the remainder of this section, we discuss how to interpret (this discussion is not used elsewhere in the paper and may be safely skipped by the reader if desired) the area inequality conjecture (Conjecture 5.6) using the language of homology. Let $L > 0$, and let $\sigma_1, \sigma_2, \sigma_3, \sigma_4 : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ be simple closed polygonal paths. We say that $\sigma_1, \dots, \sigma_4$ are in *general position* if the following hold for any distinct $i, j, k \in \{1, 2, 3, 4\}$:

- (i) for any edge e of σ_i and any edge f of σ_j , the angle between the direction of e and the direction of f is not an integer multiple of $\pi/4$;
- (ii) for any vertices u, v, w of $\sigma_i, \sigma_j, \sigma_k$ respectively, there does not exist a square with u, v, w as three of its four vertices.

It is easy to see that one can perturb $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ by an arbitrarily small amount to be in general position (for example, a random perturbation will almost surely work); furthermore one can ensure that this general position will persist even after shifting σ_4 vertically by an arbitrary amount. Hence by Proposition 5.8, to prove Conjecture 5.6 it suffices to do so under the hypothesis of general position.

The Cartesian product $\sigma_1 \times \sigma_2 \times \text{Cyl}_L \times \sigma_4$ can be viewed as a (polyhedral) 5-cycle in Cyl_L^4 ; by the hypotheses of general position, this cycle intersects the oriented 4-manifold $\overline{\text{Squares}_L}$ transversely (and in a compact set), giving rise to a 1-cycle σ_{124} in $\overline{\text{Squares}_L}$. For $j = 1, 2, 3, 4$, let $\pi_j : \overline{\text{Squares}_L} \rightarrow \text{Cyl}_L$ be the projection map to the j th coordinate. Because the 5-cycle $\sigma_1 \times \sigma_2 \times \text{Cyl}_L \times \sigma_4$ is homologous to the 5-cycle $\text{Graph}_{0,L} \times \text{Graph}_{0,L} \times \text{Cyl}_L \times \text{Graph}_{0,L}$, we see on restricting to $\overline{\text{Squares}_L}$ that the 1-cycle σ_{124} is homologous to the 1-cycle

$$\text{Graph}_{0,L}^A := \{(p, p, p, p) : p \in \text{Graph}_{0,L}(\mathbb{R}/L\mathbb{Z})\} \tag{5.4}$$

(with the usual orientation), and hence the pushforwards $\gamma_j := (\pi_j)_* \sigma_{124}$ (which are polygonal 1-cycles on Cyl_L) are homologous to $\text{Graph}_{0,L}$ and thus to σ_j for

$j = 1, 2, 3, 4$. For $j = 1, 2, 4$, γ_j takes values in the curve $\sigma_j(\mathbb{R}/L\mathbb{Z})$; as $y dx$ is closed on that curve, we thus have

$$\int_{\sigma_{124}} \pi_j^*(y dx) = \int_{(\pi_j)_*\sigma_{124}} y dx = \int_{\sigma_j} y dx$$

for $j = 1, 2, 4$. On the other hand, from Remark 3.7 (adapted to the cylinder Cyl_L in the obvious fashion), the 1-form

$$\pi_1^*(y dx) - \pi_2^*(y dx) + \pi_3^*(y dx) - \pi_4^*(y dx) \quad (5.5)$$

is exact on $\overline{\text{Squares}_L}$, and hence

$$\int_{\sigma_{124}} \pi_1^*(y dx) - \pi_2^*(y dx) + \pi_3^*(y dx) - \pi_4^*(y dx) = 0.$$

Putting all this together, we see that the claim (5.1) can be rewritten as

$$\int_{\gamma_3 - \sigma_3} y dx \neq 0. \quad (5.6)$$

Meanwhile, the hypothesis that $(\sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}), \sigma_3(\mathbb{R}/L\mathbb{Z}), \sigma_4(\mathbb{R}/L\mathbb{Z}))$ do not jointly inscribe squares is equivalent to the assertion that the 1-cycles γ_3 and σ_3 are disjoint.

The 4-manifold $\overline{\text{Squares}_L}$ is homeomorphic to $\mathbb{R}/L\mathbb{Z} \times \mathbb{R}^3$, and so its first homology is generated by $\text{Graph}_{0,L}^\Delta$. One can decompose the 1-cycle σ_{124} as an integer linear combination of finitely many closed polygonal curves in $\overline{\text{Squares}_L}$ (which are allowed to intersect each other); as σ_{124} is homologous to $\text{Graph}_{0,L}^\Delta$, one of these curves, call it $\sigma_{124}^0 : \mathbb{R}/L\mathbb{Z} \rightarrow \overline{\text{Squares}_L}$, must be homologous to $m\text{Graph}_{0,L}^\Delta$ for some nonzero integer m , thus it obeys the equivariance $\sigma_{124}^0(t + L) = \gamma_{124}^0(t) + (mL, 0)$. By reversing orientation we may assume m is positive.

We now lift the cylinder Cyl_L up to the larger cylinder Cyl_{mL} , which is an m -fold cover of the original cylinder; one can similarly lift $\overline{\text{Squares}_L}$ to the m -fold cover $\overline{\text{Squares}_{mL}}$. The curves $\sigma_j : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ lift to curves $\tilde{\sigma}_j : \mathbb{R}/mL\mathbb{Z} \rightarrow \text{Cyl}_{mL}$. The 1-cycle σ_{124} lifts to a 1-cycle $\tilde{\sigma}_{124}$ homologous to $\text{Graph}_{0,mL}^\Delta$; meanwhile, the curve σ_{124}^0 lifts to m copies of a curve $\tilde{\sigma}_{124}^0$ homologous to $\text{Graph}_{0,mL}^\Delta$ (and parameterized by $\mathbb{R}/mL\mathbb{Z}$), and contained (as a set) in $\tilde{\sigma}_{124}$. We conclude that $\tilde{\sigma}_{124} - \tilde{\sigma}_{124}^0$ is a 1-boundary, thus

$$\tilde{\sigma}_{124} = \tilde{\sigma}_{124}^0 + \partial U$$

for some 2-cycle U in $\overline{\text{Squares}_{mL}}$. We can then define curves $\tilde{\gamma}_j^0 : \mathbb{R}/mL\mathbb{Z} \rightarrow \text{Cyl}_{mL}$ for $j = 1, 2, 3, 4$ by $\tilde{\gamma}_j^0 := \pi_j \circ \tilde{\sigma}_{124}^0$; these curves are the analogues of the

curves $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ from Section 3. As the 1-form $y dx$ is closed on the curves $\sigma_j(\mathbb{R}/L\mathbb{Z})$, as well as their lifts $\tilde{\sigma}_j(\mathbb{R}/mL\mathbb{Z})$ to Cyl_{mL} , we have

$$\int_{\partial U} \pi_j^*(y dx) = \int_{\partial(\pi_j)_*U} y dx = 0$$

for $j = 1, 2, 4$, while from the exact nature of (5.5) gives

$$\int_{\partial U} \pi_1^*(y dx) - \pi_2^*(y dx) + \pi_3^*(y dx) - \pi_4^*(y dx) = 0$$

and hence

$$\int_{\partial(\pi_3)_*U} y dx = 0.$$

Hence one can also express (5.6) as

$$\int_{\tilde{\gamma}_3^0 - \tilde{\sigma}_3} y dx \neq 0. \quad (5.7)$$

The 1-cycle $\tilde{\gamma}_3^0 - \tilde{\sigma}_3$ is homologous to $\text{Graph}_{0,mL} - \text{Graph}_{0,mL} = 0$ and thus can be expressed as a 1-boundary $\tilde{\gamma}_3^0 - \tilde{\sigma}_3 = \partial\Omega$ for some 2-cycle Ω in Cyl_{mL} . By Stokes' theorem, (5.7) can now be expressed as

$$-\int_{\Omega} dx \wedge dy \neq 0. \quad (5.8)$$

In the case when the σ_2, σ_4 were graphs of Lipschitz functions of constant less than 1, the closed path $\tilde{\gamma}_3^0$ was necessarily simple (and one could take $m = 1$); if $(\sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}), \sigma_3(\mathbb{R}/L\mathbb{Z}), \sigma_4(\mathbb{R}/L\mathbb{Z}))$ did not jointly inscribe squares, then $\tilde{\gamma}_3^0$ avoided σ_3 , and so by the Jordan curve theorem the 2-cycle Ω had a definite sign which yielded (5.8) and thus (5.7), (5.6), (5.1). Unfortunately, in the general case it is possible for the 2-cycle Ω to contain both positive and negative components, even after stripping out the 1-boundaries ∂U from $\tilde{\sigma}_{124}$ and working just with $\tilde{\sigma}_{124}^0$. However, from working with numerous examples, it appears to the author that there is always an imbalance between the positive and negative components of Ω that leads to the inequality (5.8) and hence to Conjecture 5.6. Unfortunately, the author was unable to locate an argument to establish this claim rigorously.

REMARK 5.9. The Jordan curve theorem does imply that the simple closed curve $\tilde{\sigma}_3$ partitions the cylinder Cyl_{mL} into two connected components, the region 'above' $\tilde{\sigma}_3$ (which contains all the points in Cyl_{mL} with sufficiently large

positive y coordinate) and the region ‘below’ $\tilde{\sigma}_3$ (which contains all the points in $\text{CY}_{\perp mL}$ with sufficiently large negative y coordinate). Suppose that $(\sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}), \sigma_3(\mathbb{R}/L\mathbb{Z}), \sigma_4(\mathbb{R}/L\mathbb{Z}))$ does not jointly inscribe a square, so that γ_3 avoids σ_3 . If we let K denote the connected component of $\tilde{\gamma}_3$ (viewed as a subset of $\text{CY}_{\perp L}$) that contains (the image of) $\tilde{\gamma}_3^0$, then K must then either lie entirely in the region above $\tilde{\sigma}_3$ or the region below $\tilde{\sigma}_3$. We conjecture that this determines the sign in (5.1) (or (5.6), (5.7), (5.8)). Namely, if K is in the region above $\tilde{\sigma}_3$, we conjecture that left-hand side of (5.1) (or (5.6), (5.7), (5.8)) must be strictly positive, and if K is in the region below $\tilde{\sigma}_3$, then these left-hand sides must be strictly negative. An equivalent statement of this is that if $\tilde{\gamma}_3^+$, $\tilde{\gamma}_3^- : \mathbb{R}/mL\mathbb{Z} \rightarrow \text{CY}_{\perp mL}$ are simple closed polygonal paths that traverse the upper and lower boundary of K respectively (by which we mean the portions of the boundary of K that can be connected by a path to points in $\text{CY}_{\perp mL}$ with arbitrarily large positive or negative y coordinate respectively), then we have the inequalities

$$\int_{\tilde{\gamma}_3^-} y \, dx \leq \int_{\tilde{\gamma}_3^0} y \, dx \leq \int_{\tilde{\gamma}_3^+} y \, dx. \quad (5.9)$$

This claim, which implies Conjecture 5.6, is an assertion about the relative sizes of the ‘holes’ in K ; see Figure 6.

6. A combinatorial variant

Conjecture 5.6 appears difficult to resolve in general. However, there is a more tractable-seeming special case of Conjecture 5.6 which captures many of the key features of the full conjecture:

CONJECTURE 6.1 (Special case of area inequality). *Conjecture 5.6 holds when $\sigma_1 = \text{Graph}_{0,L}$.*

Once one makes the restriction $\sigma_1 = \text{Graph}_{0,L}$, Conjecture 5.6 turns out to collapse from a two-dimensional problem to a more tractable one-dimensional one. Indeed, suppose that the tuple $(\sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}), \sigma_3(\mathbb{R}/L\mathbb{Z}), \sigma_4(\mathbb{R}/L\mathbb{Z}))$ does not jointly inscribing a square, with $\sigma_1 = \text{Graph}_{0,L}$. That is to say, there does not exist $x \in \mathbb{R}/L\mathbb{Z}$ and $y, a, b \in \mathbb{R}$ with

$$\begin{aligned} (x, y) &\in \sigma_1(\mathbb{R}/L\mathbb{Z}) \\ (x + a, y + b) &\in \sigma_2(\mathbb{R}/L\mathbb{Z}) \\ (x + a - b, y + a + b) &\in \sigma_3(\mathbb{R}/L\mathbb{Z}) \\ (x - b, y + a) &\in \sigma_4(\mathbb{R}/L\mathbb{Z}). \end{aligned}$$

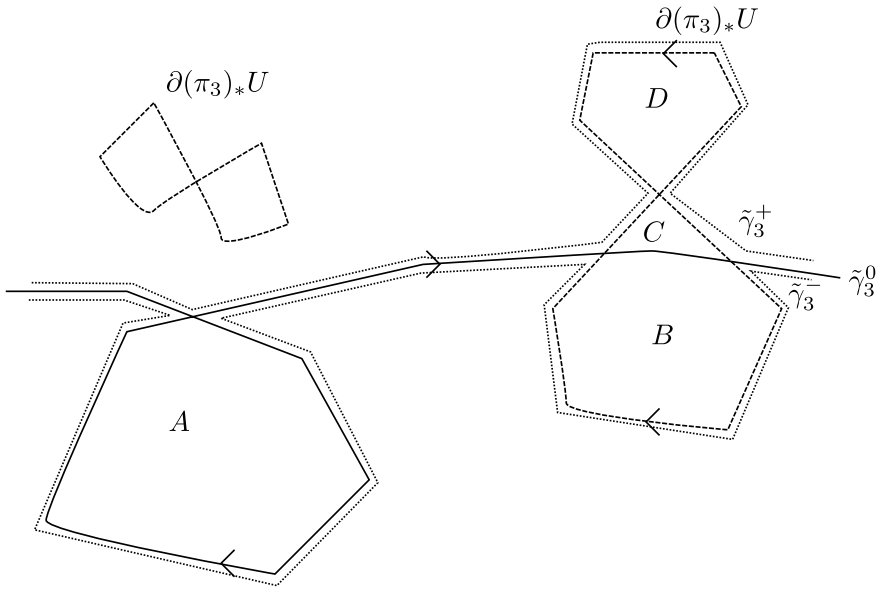


Figure 6. A polygonal path $\tilde{\gamma}_3^0$ (drawn as a solid line), together with some additional 1-boundaries $\partial(\pi_3)_*U$ (the two dashed lines). Here, K is $\tilde{\gamma}_3^0$ together with the component of $\partial(\pi_3)_*U$ that intersects $\tilde{\gamma}_3^0$. The paths $\tilde{\gamma}_3^-$ and $\tilde{\gamma}_3^+$ are drawn as dotted lines; they have been moved slightly away from $\tilde{\gamma}_3^0$ for visibility. The area inequalities (5.9) can then be written as $0 \leq 2|A| + |B| \leq |A| + |B| + |C| + |D|$, where $|A|$ denotes the Lebesgue measure of the region A in the figure, and similarly for $|B|, |C|, |D|$. Each 1-boundary gives a zero contribution to the area under $\tilde{\gamma}_3$, so one also has $|B| + |C| = |D|$. In the depicted scenario, the first area inequality is automatically true, but the second one is not necessarily so.

The first condition $(x, y) \in \sigma_1(\mathbb{R}/L\mathbb{Z})$ simply asserts that $y = 0$. If one now defines the linearly transformed closed polygonal curves $\tilde{\sigma}_2, \tilde{\sigma}_3, \tilde{\sigma}_4 : \mathbb{R}/L\mathbb{Z} \rightarrow \text{Cyl}_L$ by the formulae

$$\begin{aligned} \tilde{\sigma}_2(t) &= (x_2(t) - y_2(t), y_2(t)) \\ \tilde{\sigma}_3(t) &= (x_3(t), -y_3(t)) \\ \tilde{\sigma}_4(t) &= (x_4(t) + y_4(t), y_4(t)) \end{aligned}$$

where $x_i, y_i : \mathbb{R} \rightarrow \mathbb{R}$ are the components of σ_i for $i = 1, 2, 3, 4$, then $\tilde{\sigma}_2, \tilde{\sigma}_3, \tilde{\sigma}_4$ remain simple, and (on setting $\tilde{x} := x + a - b$) we see that the property of the

quadruple

$$(\sigma_1(\mathbb{R}/L\mathbb{Z}), \sigma_2(\mathbb{R}/L\mathbb{Z}), \sigma_3(\mathbb{R}/L\mathbb{Z}), \sigma_4(\mathbb{R}/L\mathbb{Z}))$$

not jointly inscribing squares is equivalent to the nonexistence of real numbers \tilde{x} , a , b such that

$$\begin{aligned}(\tilde{x}, b) &\in \tilde{\sigma}_2(\mathbb{R}/L\mathbb{Z}) \\(\tilde{x}, -a - b) &\in \tilde{\sigma}_3(\mathbb{R}/L\mathbb{Z}) \\(\tilde{x}, a) &\in \tilde{\sigma}_4(\mathbb{R}/L\mathbb{Z}).\end{aligned}$$

Also, from the change of variables we see that

$$\int_{\tilde{\sigma}_j} y \, dx = (-1)^j \int_{\sigma_j} y \, dx$$

for $j = 2, 3, 4$. Relabelling $\tilde{\sigma}_2, \tilde{\sigma}_3, \tilde{\sigma}_4$ as $\gamma_1, \gamma_2, \gamma_3$, and writing $b, -a - b, a$ as y_1, y_2, y_3 respectively, we thus see that Conjecture 6.1 is equivalent to the following more symmetric version:

CONJECTURE 6.2 (Special case of area inequality, symmetric form). *Let $L > 0$, and let $\gamma_1, \gamma_2, \gamma_3 : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{C}\mathbb{Y}1_L$ be simple closed polygonal paths homologous to $\text{Graph}_{0,L}$, such that there does not exist points $(x, y_i) \in \gamma_i(\mathbb{R}/L\mathbb{Z})$ with $x, y_1, y_2, y_3 \in \mathbb{R}$ and $y_1 + y_2 + y_3 = 0$. Then one has*

$$\int_{\gamma_1 + \gamma_2 + \gamma_3} y \, dx \neq 0.$$

In this section we show that Conjecture 6.2 is equivalent to an almost purely combinatorial statement. To formulate it, we need some definitions. Recall that the *signum function* $\text{sgn} : [-\infty, +\infty] \rightarrow \{-1, +1\}$ on the extended real line $[-\infty, +\infty]$ is defined to equal -1 of $[-\infty, 0)$, 0 on 0 , and $+1$ on $(0, +\infty]$.

DEFINITION 6.3 (Noncrossing sums). Let $m = 2, 3$, and for each $1 \leq i \leq m$, let $y_{i,1}, y_{i,2} \in [-\infty, +\infty]$ be extended reals. We say that the pairs $\{y_{i,1}, y_{i,2}\}$ for $i = 1, \dots, m$ have *noncrossing sums* if the following axioms are obeyed:

- (i) either all of the $y_{i,1}, y_{i,2}$ avoid $+\infty$, or they all avoid $-\infty$;
- (ii) for any $j_1, \dots, j_m \in \{1, 2\}$, the sum $y_{1,j_1} + \dots + y_{m,j_m}$ (which is well defined by (i)) is nonzero;
- (iii) one has the cancellation

$$\sum_{j_1, \dots, j_m \in \{1, 2\}} (-1)^{j_1 + \dots + j_m} \text{sgn}(y_{1,j_1} + \dots + y_{m,j_m}) = 0. \quad (6.1)$$

or equivalently (by (ii))

$$\sum_{j_1, \dots, j_m \in \{1, 2\}: y_{1, j_1} + \dots + y_{m, j_m} > 0} (-1)^{j_1 + \dots + j_m} = 0.$$

That is to say, there are as many positive sums $y_{1, j_1} + \dots + y_{m, j_m} > 0$ with the index sum $j_1 + \dots + j_m$ even as there are positive sums $y_{1, j_1} + \dots + y_{m, j_m} > 0$ with $j_1 + \dots + j_m$ odd (and similarly with ‘positive’ replaced by ‘negative’).

Otherwise, we say that the pairs $\{y_{i,1}, y_{i,2}\}$ for $i = 1, \dots, m$ have *crossing sums*.

REMARK 6.4. The notion of noncrossing sums is invariant with respect to interchanges between $y_{i,1}$ and $y_{i,2}$ for $i = 1, \dots, m$, or between the pairs $\{y_{i,1}, y_{i,2}\}$ for $i = 1, \dots, m$, or by replacing each of the $y_{i,j}$ with their negations $-y_{i,j}$. One could define this concept for other values of m than $m = 2, 3$, but these are the only two values of m we will need here.

EXAMPLE 6.5. Let a_1, a_2, b_1, b_2 be distinct elements of $\mathbb{R} \cup \{-\infty\}$. Then the pairs $\{a_1, a_2\}$ and $\{-b_1, -b_2\}$ have noncrossing sums if and only if the number of pairs $(i, j) \in \{1, 2\} \times \{1, 2\}$ with $a_i < b_j$ is even, thus for instance $\{a_1, a_2\}$ and $\{-b_1, -b_2\}$ will have noncrossing sums if

$$a_1 < a_2 < b_1 < b_2$$

or

$$a_1 < b_1 < b_2 < a_2$$

but not if

$$a_1 < b_1 < a_2 < b_2.$$

In particular, if $-\infty < b_1 < b_2 < +\infty$, $\{-\infty, a\}$ and $\{-b_1, -b_2\}$ have noncrossing sums if and only if a lies outside of $[b_1, b_2]$.

In a more topological form: the pairs $\{a_1, a_2\}$ and $\{-b_1, -b_2\}$ have noncrossing sums if it is possible to connect $(0, a_1)$ and $(0, a_2)$ (respectively $(0, b_1)$ and $(0, b_2)$) by a curve γ_a (respectively γ_b) in the (one-point compactification of the) right half-plane $[0, +\infty) \times \mathbb{R}$, in such a manner that γ_a and γ_b do not cross. See Figures 7, 8.

EXAMPLE 6.6. The pairs $\{0, 1\}, \{0, 6\}, \{-5, 1\}$ have noncrossing sums because there are as many positive sums $0 + 0 + 1, 0 + 6 - 5, 1 + 6 + 1$ with an even index sum as there are positive sums $1 + 0 + 1, 1 + 6 - 5, 0 + 6 + 1$ with an odd index sum. On the other hand, the pairs $\{-3, 5\}, \{-3, 5\}, \{-3, 5\}$ have crossing

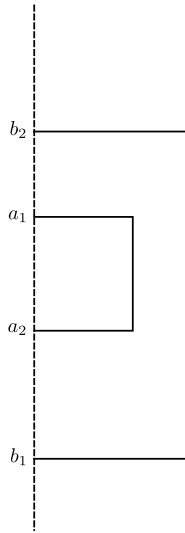


Figure 7. The pairs $\{a_1, a_2\}$ and $\{-b_1, -b_2\}$ have noncrossing sums: the sums $a_1 + (-b_1), a_2 + (-b_1)$ are positive, while the sums $a_1 + (-b_2), a_2 + (-b_2)$ are negative. Note that the path connecting $(0, a_1)$ to $(0, a_2)$ does not cross the path connecting $(0, b_1)$ to $(0, b_2)$.

sums because there are fewer positive sums $5 + 5 + 5$ with an even index sum than positive sums $-3 + 5 + 5, 5 - 3 + 5, 5 + 5 - 3$ with an odd index sum.

As we shall see later, the notion of $\{y_{i,1}, y_{i,2}\}$ for $i = 1, 2, 3$ having noncrossing sums also has a topological interpretation (assuming axiom (i)), namely that there are curves γ_i in the one-point compactification of $[0, +\infty) \times \mathbb{R}$ connecting $(0, y_{i,1})$ to $(0, y_{i,2})$ for $i = 1, 2, 3$ such that there do not exist $x, y_1, y_2, y_3 \in \mathbb{R}$ with $y_1 + y_2 + y_3 = 0$ and (x, y_i) in γ_i for $i = 1, 2, 3$. This may help explain the terminology ‘noncrossing sums’.

EXAMPLE 6.7. Suppose that $0 < a < b < c$ and x, y, z are real numbers such that $\{x, x + a\}, \{y, y + b\}, \{z, z + c\}$ have noncrossing sums. The 2^3 sums formed from this triplet may be almost completely ordered as

$$\begin{aligned}
 &x + y + z < x + a + y + z < x + y + b + z \\
 &< x + a + y + b + z, x + y + z + c \\
 &< x + a + y + z + c < x + y + b + z + c < x + a + y + b + z + c
 \end{aligned}$$

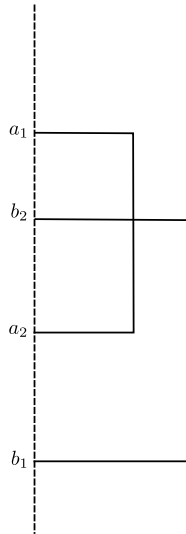


Figure 8. The pairs $\{a_1, a_2\}$ and $\{-b_1, -b_2\}$ have crossing sums: the sums $a_1 + (-b_1)$, $a_2 + (-b_1)$, $a_1 + (-b_2)$ are positive, while the sum $a_2 + (-b_2)$ is negative. Note that the path connecting $(0, a_1)$ to $(0, a_2)$ crosses the path connecting $(0, b_1)$ to $(0, b_2)$.

(the reader may wish to first see this in the case $x = y = z = 0$). The sums $x + a + y + z$, $x + y + b + z$, $x + y + z + c$, $x + a + y + b + z + c$ have even index sum, and the other four sums have odd index sum. Therefore, $\{x, x + a\}$, $\{y, y + b\}$, $\{z, z + c\}$ has noncrossing sums precisely when the origin 0 falls in one of the five intervals

$$\begin{aligned}
 &(-\infty, x + y + z), \\
 &(x + a + y + z, x + y + b + z), \\
 &(x + a + y + b + z, x + y + z + c), \\
 &(x + a + y + z + c, x + y + b + z + c), \\
 &(x + a + y + b + z + c, +\infty)
 \end{aligned}$$

(with the third interval deleted if $x + y + z + c < x + a + y + b + z$). In particular, we see that pair $\{x, x + a\}$ with the smallest difference has no influence on the sign of the triple sums, that is to say

$$\text{sgn}(x + y_2 + y_3) = \text{sgn}(x + a + y_2 + y_3)$$

for $y_2 \in \{y, y + b\}$ and $y_3 \in \{z, z + c\}$. Conversely, if this pair has no influence on the sign of triple sums then the pairs $\{x, x + a\}$, $\{y, y + b\}$, $\{z, z + c\}$ have noncrossing sums. This lack of influence by the pair with the smallest difference can thus be used as an alternate definition of noncrossing sums in the $m = 3$ case (and it also works in the $m = 2$ case).

One corollary of this analysis is that if $\{y, y + b\}$ has no influence on the sign of the triple sums, then neither does $\{x, x + a\}$; similarly, if $\{z, z + c\}$ has no influence on the noncrossing sums, then neither does $\{y, y + b\}$.

We are now ready to give the combinatorial formulation of Conjecture 6.2.

CONJECTURE 6.8 (Combinatorial formulation). *Let k_1, k_2, k_3 be odd natural numbers, and for each $i = 1, 2, 3$, let $y_{i,1}, \dots, y_{i,k_i}$ be distinct real numbers. Adopt the convention that $y_{i,0} = y_{i,k_i+1} = -\infty$. Assume the following axioms:*

- (i) *(noncrossing) for any $1 \leq i \leq 3$ and $0 \leq p < q \leq k_i$ with p, q the same parity, the pairs $\{y_{i,p}, y_{i,p+1}\}$ and $\{-y_{i,q}, -y_{i,q+1}\}$ have noncrossing sums;*
- (ii) *(noncrossing sums) for any $0 \leq p_1 \leq k_1, 0 \leq p_2 \leq k_2, 0 \leq p_3 \leq k_3$ with p_1, p_2, p_3 the same parity, the pairs $\{y_{1,p_1}, y_{1,p_1+1}\}$, $\{y_{2,p_2}, y_{2,p_2+1}\}$, $\{y_{3,p_3}, y_{3,p_3+1}\}$ have noncrossing sums.*

Then one has the inequality

$$\sum_{i=1}^3 \sum_{j=1}^{k_i} (-1)^{j-1} y_{i,j} < 0. \tag{6.2}$$

REMARK 6.9. In the language of Arnóld, the hypothesis (i) shows that the ordering of the extended real numbers $-\infty, y_{i,1}, \dots, y_{i,k_i}$ is given by the permutation of a *meander* (formed by gluing together two noncrossing matchings); see [13].

The main result of this section is then

THEOREM 6.10. *Conjecture 6.2 (and hence Conjecture 6.1) is equivalent to Conjecture 6.8.*

6.1. Forward direction. Let us first assume Conjecture 6.2 and see how it implies Conjecture 6.8. Let k_1, k_2, k_3 and $y_{i,j}$ obey the assumptions of Conjecture 6.8, but suppose for contradiction that (6.2) failed. The plan is

then to use the quantities $y_{i,j}$ to build simple closed polygonal paths $\gamma_1, \gamma_2, \gamma_3 : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{C}Y1_L$ for some $L > 0$ to which Conjecture 6.2 may be applied.

By perturbing one of the $y_{i,j}$ slightly (noting that all the hypotheses on the $y_{i,j}$ are open conditions) we may assume that the quantity

$$Q := \sum_{i=1}^3 \sum_{j=1}^{k_i} (-1)^{j-1} y_{i,j} \tag{6.3}$$

is strictly positive. Similarly, we may assume that the differences $|y_{i,p+1} - y_{i,p}|$ with $i = 1, 2, 3$ and $1 \leq p < k_i$ are all distinct.

We will need a strictly monotone decreasing function $\phi : [0, +\infty) \rightarrow [1, 2]$; the exact choice of ϕ is unimportant, but for concreteness one can take for instance $\phi(t) := 1 + (1/1 + t)$.

Let $L > 0$ be a sufficiently large quantity to be chosen later. We will also need a certain large and negative quantity $-R$ (depending on Q, L , and the $y_{i,j}$) whose precise value will be specified later. By applying Conjecture 6.8(ii) with $p_1 = p_2 = p_3 = 0$, we see that $\{-\infty, y_{1,1}\}, \{-\infty, y_{2,1}\}, \{-\infty, y_{3,1}\}$ have noncrossing sums, which implies that

$$y_{1,1} + y_{2,1} + y_{3,1} < 0.$$

Similarly if we apply Conjecture 6.8(ii) with $p_1 = k_1, p_2 = k_2, p_3 = k_3$, we see that

$$y_{1,k_1} + y_{2,k_2} + y_{3,k_3} < 0.$$

As a consequence, we may find piecewise linear continuous functions $f_1, f_2, f_3 : [-1, 1] \rightarrow \mathbb{R}$ such that

$$f_i(-1) = y_{i,k_i}; \quad f_i(1) = y_{i,1} \tag{6.4}$$

for $i = 1, 2, 3$, and such that

$$f_1(t) + f_2(t) + f_3(t) < 0 \tag{6.5}$$

for all $-1 \leq t \leq 1$. For instance, we can set f_1, f_2, f_3 to be the linear functions given by the boundary conditions (6.4). But we can also subtract an arbitrary positive multiple of $1 - |t|$ from any of f_1, f_2, f_3 and obey the above requirements. In particular, there is a quantity $-C_0$ (independent of L) such that if the quantity $-R$ mentioned previously is less than or equal to $-C_0$, one can find f_1, f_2, f_3 solving the above conditions such that

$$\sum_{i=1}^3 \int_{-1}^1 f_i(t) dt = -R. \tag{6.6}$$

Henceforth we fix f_1, f_2, f_3 with these properties.

Let $i = 1, 2, 3$. We construct some polygonal paths $\gamma_{i,1}, \gamma_{i,1 \rightarrow 2}, \dots, \gamma_{i,k_i-1 \rightarrow k_i}, \gamma_{i,k_i}, \gamma_{i,k_i \rightarrow 1}$ in $\mathbb{C}Y \perp_L$ by the following recipes:

DEFINITION 6.11. (i) $\gamma_{i,1}$ is the rightward horizontal line segment from $(-L/2 + 1, y_{i,1})$ to $(0, y_{i,1})$, projected to $\mathbb{C}Y \perp_L$.

(ii) For any odd $1 \leq p < k_i$, $\gamma_{i,p \rightarrow p+1}$ is the piecewise linear path traversing the vertices

$$(0, y_{i,p}), \left(\frac{L}{2} - \phi(|y_{i,p} - y_{i,p+1}|), y_{i,p} \right),$$

$$\left(\frac{L}{2} - \phi(|y_{i,p} - y_{i,p+1}|), y_{i,p+1} \right), \quad (0, y_{i,p+1})$$

in that order (that is to say, the concatenation of a rightward horizontal line segment, a vertical line segment, and a leftward horizontal line segment, if L is large enough), and then projected to $\mathbb{C}Y \perp_L$.

(iii) For any even $1 < p < k_i$, $\gamma_{i,p \rightarrow p_1}$ is the piecewise linear path traversing the vertices

$$(0, y_{i,p}), \left(-\frac{L}{2} + \phi(|y_{i,p} - y_{i,p+1}|), y_{i,p} \right),$$

$$\left(-\frac{L}{2} + \phi(|y_{i,p} - y_{i,p+1}|), y_{i,p+1} \right), \quad (0, y_{i,p+1})$$

in that order (that is to say, the concatenation of a leftward horizontal line segment, a vertical line segment, and a rightward horizontal line segment, if L is large enough), and then projected to $\mathbb{C}Y \perp_L$.

(iv) γ_{i,k_i} is the rightward horizontal line segment from $(0, y_{i,k_i})$ to $(L/2 - 1, y_{i,k_i})$, projected to $\mathbb{C}Y \perp_L$.

(v) $\gamma_{i,k_i \rightarrow 1}$ is the graph

$$\left\{ \left(\frac{L}{2} + t, f_i(t) \right) : -1 \leq t \leq 1 \right\}$$

traversed from left to right and then projected to $\mathbb{C}Y \perp_L$, thus it begins at $\pi_L(L/2 - 1, y_{i,k_i})$ and ends at $\pi_L(L/2 + 1, y_{i,1})$.

See Figure 9.

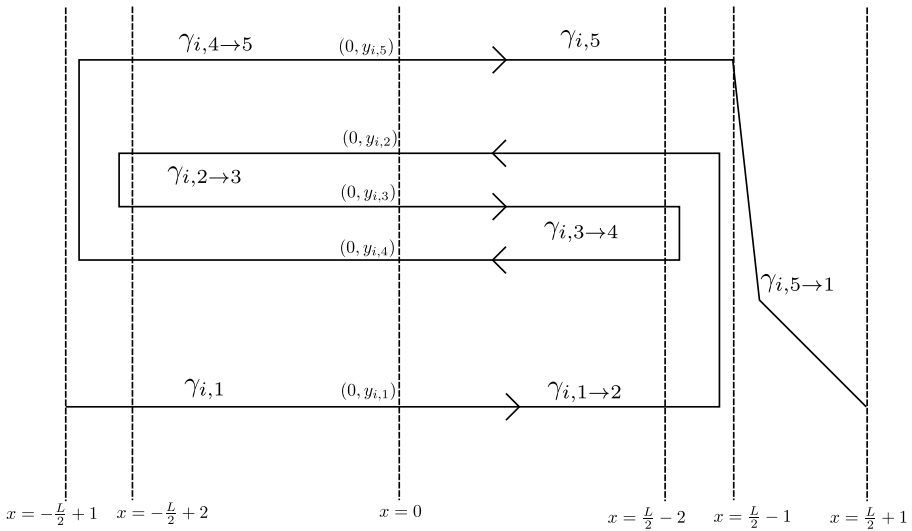


Figure 9. Construction of the components of γ_i , in the case $k_i = 5$. Notice the interlacing between the $(0, y_{i,p})$ with p odd and the $(0, y_{i,p})$ with p even, and the alternating orientations of γ_i at these locations.

Clearly, one can concatenate the paths $\gamma_{i,1}, \gamma_{i,1 \rightarrow 2}, \dots, \gamma_{i,k_i-1 \rightarrow k_i}, \gamma_{i,k_i}, \gamma_{i,k_i \rightarrow 1}$ to form a closed polygonal path γ_i in Cyl_L (which one can parameterize by $\mathbb{R}/L\mathbb{Z}$ after a suitable rescaling). Using the convex fundamental domain

$$\left[-\frac{L}{2} + 1, \frac{L}{2} + 1 \right] \times \mathbb{R}$$

of Cyl_L , we see that γ_i is homotopic in this domain to the horizontal line segment from $(-L/2 + 1, y_{i,1})$ to $(L/2 + 1, y_{i,1})$, and hence γ_i is homologous in Cyl_L to $\text{Graph}_{0,L}$.

Using Conjecture 6.8(i), we can show

LEMMA 6.12. *Suppose L is sufficiently large. Then for any $i = 1, \dots, 3$, the path γ_i is simple.*

Proof. Each of the components $\gamma_{i,1}, \gamma_{i,1 \rightarrow 2}, \dots, \gamma_{i,k_i-1 \rightarrow k_i}, \gamma_{i,k_i}, \gamma_{i,k_i \rightarrow 1}$ of γ_i are separately simple, and the endpoints are all distinct except for the endpoints of adjacent paths, so it suffices to show that no two of these components meet in the interior.

The interior of the path $\gamma_{i,k_i \rightarrow 1}$ lies in the strip $(L/2 - 1, L/2 + 1) \times \mathbb{R}$ (viewed as a subset of $C_{\mathbb{Y} \perp L}$), while the interior of the other paths lie in either $(-L/2 + 1, 0) \times \mathbb{R}$ or $(0, L/2 - 1) \times \mathbb{R}$, again only touching the boundary at endpoints. So it suffices to show that there are no crossings in $(-L/2 + 1, 0) \times \mathbb{R}$ or $(0, L/2 - 1) \times \mathbb{R}$.

We just verify the claim for $(-L/2 + 1, 0) \times \mathbb{R}$, as the case of $(0, L/2 - 1) \times \mathbb{R}$ is completely analogous. The only path components here are $\gamma_{i,1}$ and $\gamma_{i,p \rightarrow p+1}$ for $1 < p < k_i$ even. To check that $\gamma_{i,1}$ and $\gamma_{i,p \rightarrow p+1}$ do not cross, it suffices from Definitions 6.11(i), (iii) to show that $y_{i,1}$ does not lie between $y_{i,p}$ and $y_{i,p+1}$. But from Conjecture 6.8(i) we see that $\{-\infty, y_{i,1}\}$ and $\{-y_{i,p}, -y_{i,p+1}\}$ have noncrossing sums, which gives the claim by Example 6.5.

Now we need to check that $\gamma_{i,p \rightarrow p+1}$ and $\gamma_{i,q \rightarrow q+1}$ do not cross when $1 < p, q < k_1$ are even and distinct. By Conjecture 6.8(i), the pairs $\{y_{i,p}, y_{i,p+1}\}$ and $\{-y_{i,q}, -y_{i,q+1}\}$ have noncrossing sums; thus the interval spanned by $\{y_{i,q}, y_{i,q+1}\}$ either is disjoint from, contains, or is contained in the interval spanned by $\{y_{i,p}, y_{i,p+1}\}$. In the former case, the paths $\gamma_{i,p \rightarrow p+1}$ and $\gamma_{i,q \rightarrow q+1}$ are clearly disjoint because from Definition 6.11(ii), the y coordinate of any point on the first path lies in the interval spanned by $\{y_{i,p}, y_{i,p+1}\}$, and the y coordinate on any point on the second path lies in the interval spanned by $\{y_{i,q}, y_{i,q+1}\}$. By symmetry, the only remaining case to check is when the interval spanned by $\{y_{i,p}, y_{i,p+1}\}$ is contained in the interval spanned by $\{y_{i,q}, y_{i,q+1}\}$. But in this case, we have $|y_{i,p} - y_{i,p+1}| < |y_{i,q} - y_{i,q+1}|$, so by Definition 6.11(iii) and the monotone decreasing nature of ϕ , the vertical segment of the curve $\gamma_{i,p \rightarrow p+1}$ lies to the right of that of $\gamma_{i,q \rightarrow q+1}$. From this we see that the two curves are disjoint. This concludes the demonstration of simplicity in $(-L/2 + 1, 0) \times \mathbb{R}$; the case of $(0, L/2 - 1) \times \mathbb{R}$ is similar. \square

In a similar fashion, we can use Conjecture 6.8(ii) to show

LEMMA 6.13. *There does not exist $x \in \mathbb{R}/L\mathbb{Z}$ and $y_1, y_2, y_3 \in \mathbb{R}$ with $y_1 + y_2 + y_3 = 0$ such that $(x, y_i) \in \gamma_i(\mathbb{R})$ for all $i = 1, 2, 3$.*

Proof. Suppose for contradiction that x, y_1, y_2, y_3 exist with the stated properties.

First suppose that x lies in $[L/2 - 1, L/2 + 1]$ (projected to $\mathbb{R}/L\mathbb{Z}$). Then from Definition 6.11(v) we have $y_i = f_i(x - L/2)$ for $i = 1, 2, 3$, but then from (6.5) we cannot have $y_1 + y_2 + y_3 = 0$, a contradiction.

We now treat the case when x lies in $[0, L/2 - 1]$ (projected to $\mathbb{R}/L\mathbb{Z}$); the remaining case when x lies in $[-L/2 + 1, 0]$ is similar and will be omitted. By Definition 6.11, we see that each of the (x, y_i) lies either on γ_{i,k_i} or on $\gamma_{i,p_i \rightarrow p_i+1}$ for some odd $1 \leq p_i < k_i$.

Suppose first that each of the (x, y_i) lie on $\gamma_{i,p_i \rightarrow p_i+1}$. By hypothesis, the quantities $L/2 - \phi(|y_{i,p_i} - y_{i,p_i+1}|)$ for $i = 1, 2, 3$ are distinct; by cyclic

permutation we may assume that $L/2 - \phi(|y_{1,p_1} - y_{1,p_1+1}|)$ is the smallest of these quantities, or equivalently that $|y_{i,p_i} - y_{i,p_i+1}|$ is minimized at $i = 1$. By Definition 6.11(ii), the x coordinate of $\gamma_{1,p_1 \rightarrow p_1+1}$ does not exceed $L/2 - \phi(|y_{1,p_1} - y_{1,p_1+1}|)$, which implies that

$$x \leq \frac{L}{2} - \phi(|y_{1,p_1} - y_{1,p_1+1}|),$$

and hence by further application of Definition 6.11(ii) we have $y_2 = y_{2,q_2}$ and $y_3 = y_{3,q_3}$ for some $q_2 \in \{p_2, p_2 + 1\}$ and $q_3 \in \{p_3, p_3 + 1\}$; furthermore y_1 lies between y_{1,p_1} and y_{1,p_1+1} inclusive. Since $y_1 + y_2 + y_3 = 0$, this implies that the sums $y_{1,p_1} + y_{2,q_2} + y_{3,q_3}$ and $y_{1,p_1+1} + y_{2,q_2} + y_{3,q_3}$ do not have the same sign. Because $i = 1$ minimizes $|y_{i,p_i} - y_{i,p_i+1}|$, we conclude (from Example 6.7) that the pairs $\{y_{1,p_1}, y_{1,p_1+1}\}$, $\{y_{2,p_2}, y_{2,p_2+1}\}$, $\{y_{3,p_3}, y_{3,p_3+1}\}$ do not have noncrossing sums, contradicting Conjecture 6.8(ii).

The case when one or more of the (x, y_i) lies on γ_{i,k_i} is treated similarly, with k_i now playing the role of p_i (and recalling the convention $y_{i,k_i+1} = -\infty$). This concludes the treatment of the case $x \in [0, L/2 - 1]$, and the case $x \in [-L/2 + 1, 0]$ is similar. \square

From the previous two lemmas and Conjecture 6.2, we conclude that

$$\int_{\gamma_{1+y_2+y_3}} y \, dx \neq 0. \quad (6.7)$$

We work in the fundamental domain

$$\left[-\frac{L}{2} + 1, \frac{L}{2} + 1\right] \times \mathbb{R}$$

of Cyl_L . On the strip $[L/2 - 1, L/2 + 1] \times \mathbb{R}$, the contribution to $\int_{\gamma_{1+y_2+y_3}} y \, dx$ is $\sum_{i=1}^3 \int_{-1}^1 f_i(t) \, dt = -R$ thanks to Definition 6.11(v) and (6.6). On the strip $[-L/2 + 2, L/2 - 2] \times \mathbb{R}$, the curve γ_i for $i = 1, 2, 3$ is simply the union of the line segments $[-L/2 + 2, L/2 - 2] \times \{y_{i,p}\}$ for $p = 1, \dots, k_i$ (traversed from left to right for odd p , and right to left for even p), so the contribution to $\int_{\gamma_{1+y_2+y_3}} y \, dx$ here is

$$\sum_{i=1}^3 (L - 4) \sum_{j=1}^{k_i} (-1)^{j-1} y_{i,j} = (L - 4)Q$$

thanks to (6.3). Finally, the contribution of the remaining strips $[-L/2 + 1, -L/2 + 2] \times \mathbb{R} \cup [L/2 - 2, -L/2 - 1] \times \mathbb{R}$ is some quantity $-C_1$ independent of L and R , as can be seen by translating these strips by $\pm L/2$. The inequality (6.7) thus becomes

$$(L - 4)Q - C_1 - R \neq 0.$$

But as Q is positive, we can make this quantity vanish by choosing L large enough and then setting $-R := C_1 - (L - 4)Q$; note for L large enough that this value of $-R$ will be less than the threshold $-C_0$ needed so that one can arrange the function f_1, f_2, f_3 to obey (6.6). This yields the desired contradiction.

6.2. Backward direction. Now we assume Conjecture 6.8 and see how it implies Conjecture 6.2.

By applying (a slight variant of) Proposition 5.8, we see that to prove Conjecture 6.2, it suffices to do so under the additional nondegeneracy hypothesis that all the vertices of $\gamma_1, \gamma_2, \gamma_3$ have distinct x -coordinates (in particular, these curves do not contain any vertical edges). Write $\gamma_i(t) = (x_i(t), y_i(t))$ for some piecewise linear $x_i : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{R}/L\mathbb{Z}$ and $y_i : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{R}$. As γ_i is homologous to $\text{Graph}_{0,L}$, we can find a continuous lift $\tilde{x}_i : \mathbb{R} \rightarrow \mathbb{R}$ of x_i with $\tilde{x}_i(t + L) = \tilde{x}_i(t) + L$ for all $t \in \mathbb{R}$; we also let $\tilde{y}_i : \mathbb{R} \rightarrow \mathbb{R}$ be the periodic lift of y_i . As $\tilde{x}_i(t) - t$ is periodic and continuous, it is bounded; by multiplying the period L by a large integer if necessary, we may assume that

$$|\tilde{x}_i(t) - t| \leq \frac{L}{10} \tag{6.8}$$

for all $t \in \mathbb{R}$ and $i = 1, 2, 3$.

Using the nondegeneracy hypothesis, we see that for any $i = 1, 2, 3$ and any $x \in \mathbb{R}$, the fibre $\{y \in \mathbb{R} : (x, y) \in \gamma_i(\mathbb{R})\}$ consists of a finite number $k_i(x)$ of real numbers, where the function k_i takes values in the odd natural numbers, is periodic in L , and is locally constant for all x outside of finitely many residue classes mod L . We can enumerate these real numbers as

$$\tilde{y}_i(t_{i,1}(x)), \dots, \tilde{y}_i(t_{i,k_i(x)}(x))$$

where $t_{i,1}(x) < \dots < t_{i,k_i(x)}(x)$ are those real numbers t with $\tilde{x}_i(t) = x$, arranged in increasing order, thus

$$\tilde{x}_i(t_{i,j}(x)) = x \tag{6.9}$$

for all $x \in \mathbb{R}$, $i = 1, 2, 3$, and $1 \leq j \leq k_i(x)$. For x outside of finitely many residue classes mod L , the functions $t_{i,j}$ are locally linear, and have the $L\mathbb{Z}$ -equivariance property

$$t_{i,j}(x + L) = t_{i,j}(x) + L$$

for all $x \in \mathbb{R}$, $i = 1, 2, 3$, and $1 \leq j \leq k_i(x)$. Also, from the nondegeneracy hypothesis we see from the intermediate value theorem that outside of finitely many residue classes mod L , $t_{i,j}$ is increasing for odd j and decreasing for even j .

The analogue of the set $\overline{\text{Squares}_L} \subset \text{Cyl}_L^4$ in this context is the oriented 3-dimensional submanifold Sums of Cyl_L^3 defined by

$$\text{Sums} := \{((x_1, y_1), (x_2, y_2), (x_3, y_3)) \in \text{Cyl}_L^3 : x_1 = x_2 = x_3; y_1 + y_2 + y_3 = 0\}.$$

The hypotheses of Conjecture 6.2 then assert that the 3-cycle $\gamma_1 \times \gamma_2 \times \gamma_3$ in Cyl_L^3 does not intersect Sums . We view Sums as an oriented submanifold of the 4-manifold $V \subset \text{Cyl}_L^3$ defined by

$$V := \{((x_1, y_1), (x_2, y_2), (x_3, y_3)) \in \text{Cyl}_L^3 : x_1 = x_2 = x_3\}.$$

As we assumed the vertices of $\gamma_1, \gamma_2, \gamma_3$ to have distinct x coordinates, the 3-cycle $\gamma_1 \times \gamma_2 \times \gamma_3$ intersects V transversely in some 1-cycle γ_{123} . As $\gamma_1, \gamma_2, \gamma_3$ are homologous to $\text{Graph}_{0,L}$, γ_{123} is homologous to the 1-cycle

$$\text{Graph}_{0,L}^\Delta := \{(p, p, p) : p \in \text{Graph}_{0,L}\}$$

with the standard orientation.

Now we argue as in the previous section. The 4-manifold V is homeomorphic to $\mathbb{R}/L\mathbb{Z} \times \mathbb{R}^3$ and thus has first homology generated by $\text{Graph}_{0,L}^\Delta$. By the greedy algorithm, one can express the 1-cycle γ_{123} as a finite integer linear combination of closed paths contained (as a set) in γ_{123} , each of which is either simple or a 1-boundary; one of these, say $\gamma_{123}^0 : \mathbb{R}/L\mathbb{Z} \rightarrow V$, is homologous to $m\text{Graph}_{0,L}^\Delta$ for some nonzero integer m , and is thus simple. From the Jordan curve theorem, m must be $+1$ or -1 ; by reversing the orientation of γ_{123}^0 we can then assume that $m = 1$, thus γ_{123}^0 is homologous to $\text{Graph}_{0,L}^\Delta$ and is contained (as a set) in γ_{123} . In particular, it avoids Sums . If we write

$$\gamma_{123}^0(t) = ((X(t), Y_1(t)), (X(t), Y_2(t)), (X(t), Y_3(t)))$$

then $X : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{R}/L\mathbb{Z}, Y_1, Y_2, Y_3 : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{R}$ are piecewise linear continuous functions with X homologous to the identity function (in the sense that X lifts to a function $\tilde{X} : \mathbb{R} \rightarrow \mathbb{R}$ with $\tilde{X}(t + L) = \tilde{X}(t) + L$ for all $t \in \mathbb{R}$), with the properties that

$$(X(t), Y_i(t)) \in \gamma_i(\mathbb{R}/L\mathbb{Z}) \tag{6.10}$$

and

$$Y_1(t) + Y_2(t) + Y_3(t) \neq 0 \tag{6.11}$$

for all $t \in \mathbb{R}/L\mathbb{Z}$ and $i = 1, 2, 3$.

REMARK 6.14. One can view the functions $X(t), Y_1(t), Y_2(t), Y_3(t)$ from a dynamical perspective by thinking of $(X(t), Y_i(t))$ as the trajectory of a particle

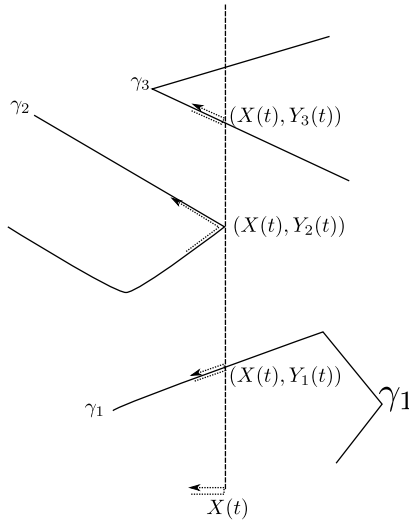


Figure 10. The dynamics of $X(t), Y_1(t), Y_2(t), Y_3(t)$.

P_i that is constrained to lie in $\gamma_i(\mathbb{R}/L\mathbb{Z})$, and with all three particles P_1, P_2, P_3 constrained to lie on a vertical line. We can also constrain the particles P_1, P_2, P_3 to have a constant horizontal speed; the particles move in one horizontal direction until one of the particles P_i cannot move any further due to it hitting a vertex v of γ_i with both edges adjacent to v lying on the same side of the vertical line containing v . Whenever such a collision occurs, the horizontal velocity reverses sign, P_i moves from one edge of γ_i to the next, and the other two particles reverse themselves and retrace their steps; see Figure 10. Note from our hypotheses that only one collision occurs at a time. Because the paths $\gamma_1, \gamma_2, \gamma_3$ have only finitely many edges, these trajectories must be periodic; the above homological considerations ensure that at least one of these trajectories is homologous to $\text{Graph}_{0,L}^A$ (possibly after enlarging the period L).

Recall that $X : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{R}/L\mathbb{Z}$ lifts to a function $\tilde{X} : \mathbb{R} \rightarrow \mathbb{R}$ such that $\tilde{X}(t + L) = \tilde{X}(t) + L$ for all $t \in \mathbb{R}$, thus $\tilde{X}(t) - t$ is periodic and therefore bounded. We also lift $Y_i : \mathbb{R}/L\mathbb{Z} \rightarrow \mathbb{R}$ periodically to $\tilde{Y}_i : \mathbb{R} \rightarrow \mathbb{R}$ for $i = 1, 2, 3$. By lifting and (6.10), we can then find unique continuous functions $T_i : \mathbb{R} \rightarrow \mathbb{R}$ with $T_i(t + L) = T_i(t) + L$ for all $t \in \mathbb{R}$, such that

$$(\tilde{X}(t), \tilde{Y}_i(t)) = (\tilde{x}_i(T_i(t)), \tilde{y}_i(T_i(t))) \tag{6.12}$$

for all $t \in \mathbb{R}$. By replacing L with a large multiple if necessary, we may assume that

$$|\tilde{X}(t) - t|, |T_i(t) - t| \leq \frac{L}{10} \tag{6.13}$$

for all $t \in \mathbb{R}$ and $i = 1, 2, 3$.

By continuity and (6.11), we see that the expression $\tilde{Y}_1(t) + \tilde{Y}_2(t) + \tilde{Y}_3(t)$ is either positive for all t , or negative for all t . By applying the reflection $(x, y) \mapsto (x, -y)$ on $\mathbb{C}Y \perp_L$ we may assume the latter case occurs, thus

$$\tilde{Y}_1(t) + \tilde{Y}_2(t) + \tilde{Y}_3(t) < 0 \tag{6.14}$$

for all $t \in \mathbb{R}/L\mathbb{Z}$. To establish Conjecture 6.2, it suffices to show that

$$\int_{\gamma_1 + \gamma_2 + \gamma_3} y \, dx < 0;$$

integrating fibre by fibre, it will suffice to show that

$$\sum_{i=1}^3 \sum_{j=1}^{k_i(x)} (-1)^{j-1} \tilde{y}_i(t_{i,j}(x)) < 0 \tag{6.15}$$

for almost every $x \in \mathbb{R}$.

Fix $x \in \mathbb{R}$; we may assume that x avoids all the x coordinates of vertices of $\gamma_1, \gamma_2, \gamma_3$. We abbreviate $k_i = k_i(x)$ and

$$y_{i,j} := \tilde{y}_i(t_{i,j})$$

for $i = 1, 2, 3$ and $1 \leq j \leq k_i$, adopting the conventions $y_{i,0} = y_{i,k_i+1} = -\infty$. Applying Conjecture 6.8, it will suffice to verify the hypotheses (i), (ii) of that conjecture.

Let $t_+ = t_+(x)$ denote the largest time $t_+ \in \mathbb{R}$ for which $\tilde{X}(t_+) = x$ (such a time exists thanks to (6.13) and continuity). We claim that $T_i(t_+) = t_{i,k_i}(x)$ for all $i = 1, 2, 3$. Suppose for contradiction that this failed for some $i = 1, 2, 3$, then from (6.10) one has $T_i(t_+) = t_{i,p}(x)$ for some $1 \leq p < k_i$. From (6.13) and the intermediate value theorem we must then have $T_i(t) = t_{i,k_i(x)}(x)$ for some $t > t_+$, which by (6.12), (6.9) gives $\tilde{X}(t) = x$, contradicting the maximality of t_+ . Similarly, if $t_- = t_-(x)$ is the smallest time $t_- \in \mathbb{R}$ for which $\tilde{X}(t_-) = x$, then $T_i(t_-) = t_{i,1}(x)$ for $i = 1, 2, 3$. From (6.14) applied at the times t_+, t_- we have the inequalities

$$y_{1,1} + y_{2,1} + y_{3,1} < 0 \tag{6.16}$$

and

$$y_{1,k_1} + y_{2,k_2} + y_{3,k_3} < 0. \tag{6.17}$$

Having obtained these inequalities, we will have no further need of the functions $X, \tilde{X}, Y_i, \tilde{Y}_i, T_i$ or the curves γ_{123}^0 , although we will introduce a variant of these functions shortly.

We now verify the noncrossing property (i) for any given $i = 1, 2, 3$. We just verify the claim when p, q are odd, as the claim when p, q are even is completely analogous. First suppose that $q < k_i$. Let $\gamma_{x,i,p}$ denote the restriction of path $t \mapsto (\tilde{x}_i(t), \tilde{y}_i(t))$ to the interval $t \in [t_{i,p}(x), t_{i,p+1}(x)]$; define $\gamma_{x,i,q}$ similarly. The path $\gamma_{x,i,p}$ traces out a piecewise linear curve in \mathbb{R}^2 that starts at $(x, y_{i,p})$, ends at $(x, y_{i,p+1})$, and does not encounter the vertical line $x \times \mathbb{R}$ at any point in between; also, it moves to the right for t near $t_{i,p}(x)$ (and to the left for t near $t_{i,p+1}(x)$). Thus, this curve $\gamma_{x,i,p}$ must stay in the right half-plane $[x, +\infty) \times \mathbb{R}$; actually, by (6.8) we see that it stays in the strip $[x, x + L/2] \times \mathbb{R}$ (say). Similarly for $\gamma_{x,i,q}$. On the other hand, as γ_i is simple and $p < q$, the two paths $\gamma_{x,i,p}$ and $\gamma_{x,i,q}$ cannot meet. From the Jordan curve theorem and Example 6.5, this forces the endpoints $\{y_{i,p}, y_{i,p+1}\}$ and $\{y_{i,q}, y_{i,q+1}\}$ of these paths to be noncrossing, giving (i) in this case.

Now suppose that $q = k_i$. In this case we define $\gamma_{x,i,q}$ to be the restriction of $t \mapsto (\tilde{x}_i(t), \tilde{y}_i(t))$ to the interval $[t_{i,k_i(x)}(x), t_{i,1}(x) + L]$ (this interval is well defined by (6.8)). This is a path from (x, y_{i,k_i}) to $(x + L, y_{i,1})$ that does not cross $\{x, x + L\} \times \mathbb{R}$ except at endpoints, and hence lies in the strip $[x, x + L] \times \mathbb{R}$. It cannot cross $\gamma_{x,i,p}$, which also lies in this strip, avoids the right edge, and starts and ends at the points $(x, y_{i,p}), (x, y_{i,p+1})$ on the left edge. By the Jordan curve theorem, this implies that y_{i,k_i} cannot lie between $y_{i,p}$ and $y_{i,p+1}$, which by Example 6.5 implies (since $y_{i,k_i+1} = -\infty$) that $\{y_{i,p}, y_{i,p+1}\}$ and $\{y_{i,k_i}, y_{i,k_i+1}\}$ are noncrossing. This concludes the establishment of (i) when p, q are odd; the case when p, q are even is analogous (working to the left of $\{x\} \times \mathbb{R}$ rather than to the right, and using the convention $y_{i,0} = -\infty$ rather than $y_{i,k_i+1} = -\infty$) and is omitted.

Now we verify (ii) for $0 \leq p_1 \leq k_1, 0 \leq p_2 \leq k_2, 0 \leq p_3 \leq k_3$ with p_1, p_2, p_3 the same parity. We just establish the claim when p_1, p_2, p_3 are all odd, as the case when p_1, p_2, p_3 are all even is completely analogous.

In the case $p_1 = k_1, p_2 = k_2, p_3 = k_3$, we see from (6.17) and the conventions $y_{1,k_1+1} = y_{2,k_2+2} = y_{3,k_3+3} = -\infty$ that the pairs $\{y_{1,k_1}, y_{1,k_1+1}\}, \{y_{2,k_2}, y_{2,k_2+1}\}, \{y_{3,k_3}, y_{3,k_3+1}\}$. Thus we may assume that $p_i < k_i$ for at least one i ; say $p_1 < k_1$.

As in the proof of (i), we can form the curves γ_{x,i,p_i} for $i = 1, 2, 3$, which lie in the strip $[x, x + L] \times \mathbb{R}$, with initial point (x, y_{i,p_i}) and final point (x, y_{i,p_i+1}) . For $i = 1, 2, 3$, let x_i denote the maximum x coordinate attained by γ_{x,i,p_i} , thus $x < x_i \leq x + L$; furthermore $x_i = x + L$ if $p_i = k_i$ and $x_i \leq x + L/2$ otherwise, in particular $x_1 \leq x + L/2$. As the vertices of γ_i have distinct x coordinates, the x_i are distinct in the interval $[x, x + L/2]$; without loss of generality we may then

take $x_1 < x_2, x_3$. For $i = 1, 2, 3$, define γ'_{x,i,p_i} to be the connected component of $\gamma_{x,i,p_i} \cap [x, x_1] \times \mathbb{R}$ that contains (x, y_{i,p_i}) ; thus $\gamma'_{x,1,p_1} = \gamma_{x,1,p_1}$, and for $i = 2, 3$, γ'_{x,i,p_i} is a piecewise path connecting (x, y_{i,p_i}) to some point on the vertical line $\{x_1\} \times \mathbb{R}$.

Consider the set

$$S := \{(x', y'_1, y'_2, y'_3) \in \mathbb{R}^4 : (x', y'_i) \text{ lies in } \gamma'_{x,i,p_i} \text{ for } i = 1, 2, 3\}.$$

The set S is a union of line segments in \mathbb{R}^4 . It contains the point $(x, y_{1,p_1}, y_{2,p_2}, y_{3,p_3})$ with exactly one line segment emanating from it; S similarly contains the point $(x, y_{1,p_1+1}, y_{2,p_2}, y_{3,p_3})$ with exactly one line segment emanating from it. A local analysis (using the nondegeneracy hypothesis that the vertices of $\gamma_1, \gamma_2, \gamma_3$ all have distinct x coordinates) then reveals that every other point (x', y'_1, y'_2, y'_3) in S is either an interior point of a line segment in S (and avoids all other line segments comprising S), or else is a vertex that is the endpoint of exactly two edges in S ; this claim is most delicate in the case where $x' = x_1$, in which the curves $\gamma'_{x,2,p_2}$ and $\gamma'_{x,3,p_3}$ have terminated, but $\gamma_{x,1,p_1}$ leaves (x', y_1) in two leftward directions, thus again forming two edges in S (see Figure 11). Because of this, there must be a path $t \mapsto (X'(t), Y'_1(t), Y'_2(t), Y'_3(t))$ in S from $(x, y_{1,p_1}, y_{2,p_2}, y_{3,p_3})$ to $(x, y_{1,p_1+1}, y_{2,p_2}, y_{3,p_3})$ (see Remark 6.14). By the hypothesis of Conjecture 6.2, we must have

$$Y'_1(t) + Y'_2(t) + Y'_3(t) \neq 0$$

for all t . In particular, we conclude that the sums

$$y_{1,p_1} + y_{2,p_2} + y_{3,p_3}, y_{1,p_1+1} + y_{2,p_2} + y_{3,p_3}$$

have the same sign. A similar argument (using the connected component of γ'_{x,i,p_i} containing (x, y_{i,p_i+1}) rather than (x, y_{i,p_i}) as appropriate) shows more generally that the sums

$$y_{1,p_1} + y_{2,q_2} + y_{3,q_3}, y_{1,p_1+1} + y_{2,q_2} + y_{3,q_3}$$

have the same sign for $q_2 \in \{p_2, p_2 + 1\}$ and $q_3 \in \{p_3, p_3 + 1\}$ (this claim is trivially true when $q_2 = k_2 + 1$ or $q_3 = k_3 + 1$). By Definition 6.3, we conclude that the pairs $\{y_{1,p_1}, y_{1,p_1+1}\}, \{y_{2,p_2}, y_{2,p_2+1}\}, \{y_{3,p_3}, y_{3,p_3+1}\}$ have noncrossing sums, giving (ii) in the case that p_1, p_2, p_3 are all odd; the claim when p_1, p_2, p_3 are all even is proven similarly (using the convention $y_{i,0} = -\infty$ instead of $y_{i,k_i} = -\infty$, working to the left of $\{x\} \times \mathbb{R}$ rather than to the right, and using (6.16) in place of (6.17)) and is omitted. This completes the derivation of Conjecture 6.2 from Conjecture 6.8, and establishes Theorem 6.10.

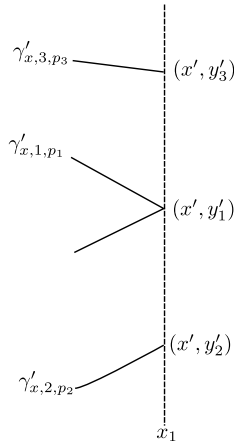


Figure 11. The local behaviour of a point (x', y'_i, y'_2, y'_3) in S when x' equals x_1 .

7. Some special cases of Conjecture 6.8

From Theorem 6.10, we see that any counterexample to Conjecture 6.8 can be converted to counterexamples for Conjectures 6.2, 6.1, and 5.6, and we believe it likely that such counterexamples, should they exist, could then be modified to give counterexamples to Conjectures 5.2, 4.1, 4.6, and hence the original square peg conjecture (Conjecture 1.1). On the other hand, after extensive testing of examples, the author is now inclined to believe that Conjecture 6.8 is true, and a proof of this conjecture is likely to lead to an approach to establish Conjecture 5.6 (and hence Conjectures 5.2, 4.1, 4.6) and perhaps even Conjecture 1.1.

We do not have a proof of Conjecture 6.8 in full generality, however we can verify some special cases. First, we observe the following analogue of Theorem 5.7 for Conjecture 6.2:

THEOREM 7.1. *Conjecture 6.2 is true when one of the curves γ_i is the graph $\gamma_i = \text{Graph}_f$ of a piecewise linear function $f : \mathbb{Z}/L\mathbb{Z} \rightarrow \mathbb{R}$.*

Proof. Suppose that $\gamma_3 = \text{Graph}_f$. By replacing γ_3 with $\text{Graph}_{0,L}$ and γ_2 with the transformed polygonal path $\{(x, y - f(x)) : (x, y) \in \gamma_2(\mathbb{Z}/L\mathbb{Z})\}$, we may assume without loss of generality that $f = 0$. The hypothesis of Conjecture 6.2 then ensures that the reflection $\tilde{\gamma}_2$ of γ_2 across the x axis is disjoint from γ_1 , hence

by the Jordan curve theorem and Stokes' theorem as before we have

$$\int_{\tilde{\gamma}_2} y \, dx \neq \int_{\gamma_1} y \, dx$$

giving Conjecture 6.2 in this case since $\int_{\tilde{\gamma}_2} y \, dx = -\int_{\gamma_2} y \, dx$. □

From this theorem and the construction used in the proof of Theorem 6.10, we see that Conjecture 6.8 holds when one of the k_i , say k_3 , is equal to 1. Of course, as Conjecture 6.8 is largely a combinatorial conjecture, one expects to also be able to verify the $k_3 = 1$ case of Conjecture 6.8 by a direct combinatorial argument, without explicit invocation of the Jordan curve theorem. We can do this by developing some combinatorial analogues of the Jordan curve theorem that are valid even when $k_1, k_2, k_3 > 1$. For any $i \in \{1, 2, 3\}$ and $y \in [-\infty, +\infty]$, define the winding number $W_i(y)$ by the Alexander numbering rule [1]

$$\begin{aligned} W_i(y) &:= \frac{1}{2} + \frac{1}{2} \sum_{j=1}^{k_i} (-1)^{j-1} \operatorname{sgn}(y_{i,j} - y) \\ &= \frac{1}{2} \sum_{j=1}^{k_i} (-1)^{j-1} (1 + \operatorname{sgn}(y_{i,j} - y)) \end{aligned} \tag{7.1}$$

(where we use the hypothesis that k_i is odd), thus W_i is a half-integer on $y_{i,1}, \dots, y_{i,k_i}$, a locally constant integer outside of these points, and jumps by $\pm 1/2$ when one perturbs off of one of the $y_{i,j}$ in either direction. Also observe that $W_i(y)$ equals 0 near $+\infty$, and 1 near $-\infty$. From Fubini's theorem we can relate the winding number to the alternating sum $\sum_{j=1}^{k_i} (-1)^{j-1} y_{i,j}$ by the identity

$$\int_{-T}^{\infty} W_i(y) \, dy = \sum_{j=1}^{k_i} (-1)^{j-1} y_{i,j} + T \tag{7.2}$$

which holds for all sufficiently large T . A similar argument gives

$$\int_{-T}^{\infty} (1 - W_i(-y)) \, dy = -\sum_{j=1}^{k_i} (-1)^{j-1} y_{i,j} + T \tag{7.3}$$

again for sufficiently large T .

We can then use the hypothesis (i) of Conjecture 6.8 to give

LEMMA 7.2 (Combinatorial Jordan curve theorem). *Suppose that the hypotheses of Conjecture 6.8 hold. Let $i = 1, 2, 3$. Then one has $W_i(y_{i,j}) = 1/2$ for all $j = 1, \dots, k_i$, and $W_i(y) \in \{0, 1\}$ for all y in $[-\infty, +\infty] \setminus \{y_{i,1}, \dots, y_{i,k_i+1}\}$.*

Proof. Because W_i is locally constant away from $\{y_{i,1}, \dots, y_{i,k+1}\}$ and jumps by $\pm 1/2$ when it reaches any of the $y_{i,j}$, it suffices to establish the first claim. Let $1 \leq p < k_i$. From Conjecture 6.8(i) we have

$$\sum_{j=q, q+1} (-1)^{j-1} \operatorname{sgn}(y_{i,p} - y_{i,j}) = \sum_{j=q, q+1} (-1)^{j-1} \operatorname{sgn}(y_{i,p+1} - y_{i,j})$$

for all $0 \leq q \leq k_1$ distinct from p with the same parity as p (using the conventions $y_{i,0} = y_{i,k+1} = -\infty$). Direct inspection shows that the claim also holds for $q = p$. Summing over q , and noting that the contributions of $j = 0$ or $j = k_1 + 1$ are the same on both sides, we conclude that

$$\sum_{j=1}^{k_1} (-1)^{j-1} \operatorname{sgn}(y_{i,p} - y_{i,j}) = \sum_{j=1}^{k_1} (-1)^{j-1} \operatorname{sgn}(y_{i,p+1} - y_{i,j})$$

and hence

$$W_i(y_{i,p}) = W_i(y_{i,p+1})$$

for all $1 \leq p < k_i$. Direct computation also shows that $W_i(y_{i,p}) = 1/2$ when $1 \leq p \leq k_1$ maximizes $y_{i,p}$, and the claim follows. \square

Next, for distinct $i, i' \in \{1, 2, 3\}$ and $y \in [-\infty, +\infty]$, we define the further winding number $W_{ii'}(y)$ by the similar formula

$$\begin{aligned} W_{ii'}(y) &:= \frac{1}{2} + \frac{1}{2} \sum_{j=1}^{k_i} \sum_{j'=1}^{k_{i'}} (-1)^{j+j'} \operatorname{sgn}(y_{i,j} + y_{i',j'} - y) \\ &= \frac{1}{2} \sum_{j=1}^{k_i} \sum_{j'=1}^{k_{i'}} (-1)^{j+j'} (1 + \operatorname{sgn}(y_{i,j} + y_{i',j'} - y)). \end{aligned} \quad (7.4)$$

As before, $W_{ii'}(y)$ will be a locally constant integer away from the sums $y_{i,j} + y_{i',j'}$, that equals 0 for sufficiently large positive y and 1 for sufficiently large negative y . From Fubini's theorem we have the analogue

$$\int_{-T}^{\infty} W_{ii'}(y) dy = \sum_{j=1}^{k_i} (-1)^{j-1} y_{i,j} + \sum_{j=1}^{k_{i'}} (-1)^{j-1} y_{i',j'} + T \quad (7.5)$$

for sufficiently large T . Curiously, one has the convolution identity

$$W'_{ii'} = W'_i * W'_{i'}$$

where the primes denote distributional derivatives, although the author was not able to make much use of this identity. The winding numbers $W_{ii'}$ also have

some resemblance to the Steinberg formula [24] for the multiplicity of irreducible representations in a tensor product, although this is likely to be just a coincidence.

The hypothesis (ii) of Conjecture 6.8 allows us to make the winding number $W_{i'}$ vanish at some points, and also give some control on the complementary winding number $W_{i''}$:

PROPOSITION 7.3. *Suppose that the hypotheses of Conjecture 6.8 hold. Let i, i', i'' be distinct elements of $\{1, 2, 3\}$.*

(i) *One has $W_{i'}(-y_{i'',j}) = 0$ for all $j = 0, \dots, k_3 + 1$.*

(ii) *If $0 \leq p \leq k_i$ and $0 \leq q \leq k_{i'}$ have the same parity, then one has*

$$W_{i''}(-y_{i,p} - y_{i',b}) = W_{i''}(-y_{i,p+1} - y_{i',b}) \tag{7.6}$$

for $b = q, q + 1$ if $|y_{i,p} - y_{i,p+1}| \leq |y_{i',q} - y_{i',q+1}|$, and

$$W_{i''}(-y_{i,a} - y_{i',q}) = W_{i''}(-y_{i,a} - y_{i',q+1}) \tag{7.7}$$

for $a = p, p + 1$ if $|y_{i,p} - y_{i,p+1}| \geq |y_{i',q} - y_{i',q+1}|$.

Proof. By permutation we may set $i = 1, i' = 2, i'' = 3$. Suppose that $0 \leq p_1 \leq k_1, 0 \leq p_2 \leq k_2, 0 \leq p_3 \leq k_3$ have the same parity. By Conjecture 6.8(ii) we have

$$\begin{aligned} & \sum_{j_1=p_1, p_1+1} \sum_{j_2=p_2, p_2+1} (-1)^{j_1+j_2} \operatorname{sgn}(y_{1,j_1} + y_{2,j_2} + y_{3,p_3}) \\ &= \sum_{j_1=p_1, p_1+1} \sum_{j_2=p_2, p_2+1} (-1)^{j_1+j_2} \operatorname{sgn}(y_{1,j_1} + y_{2,j_2} + y_{3,p_3+1}); \end{aligned}$$

summing over p_1, p_2 and noting that the contributions of $j_1 = 0, j_1 = k_1 + 1, j_2 = 0, j_2 = k_2 + 1$ are the same on both sides we see that

$$W_{12}(-y_{3,p_3}) = W_{12}(-y_{3,p_3+1})$$

for all $0 \leq p_3 \leq k_3$; since $W_{12}(-y_{3,0}) = W_{12}(+\infty) = +1$, we conclude (i).

Now suppose $0 \leq p \leq k_1$ and $0 \leq q \leq k_2$ have the same parity and $|y_{1,p} - y_{1,p+1}| \leq |y_{2,q} - y_{2,q+1}|$, and let $0 \leq r \leq k_3$ have the same parity as p and q . From Conjecture 6.8(ii) we see that the pairs $\{y_{1,p}, y_{1,p+1}\}, \{y_{2,q}, y_{2,q+1}\}, \{y_{3,r}, y_{3,r+1}\}$ have noncrossing sums, which by Example 6.7 implies that at least one of the pairs $\{y_{1,p}, y_{1,p+1}\}, \{y_{3,r}, y_{3,r+1}\}$ have no influence on the triple sums. This implies that

$$\sum_{j=r, r+1} (-1)^{j-1} \operatorname{sgn}(y_{1,p} + y_{2,b} + y_{3,j}) = \sum_{j=r, r+1} (-1)^{j-1} \operatorname{sgn}(y_{1,p+1} + y_{2,b} + y_{3,j})$$

for $b = q, q + 1$; summing in r we obtain (7.6). The claim (7.7) is proven similarly. □

This proposition is already enough to reprove the $k_3 = 1$ case of Conjecture 6.8 as follows. By adding $y_{3,1}$ to all of the $y_{2,j}$ and then sending $y_{3,1}$ to zero, we may assume that $y_{3,1} = 0$. Then we have $W_{13}(y) = W_1(y)$ and $W_{23}(y) = W_2(y)$ for all y , and hence by Proposition 7.3 we have $W_1(-y_{2,j}) = 0$ for $j = 1, \dots, k_2$ and $W_2(-y_{1,j}) = 0$ for $j = 1, \dots, k_1$. We conclude that on the set $\{y \in \mathbb{R} : W_1(y) = +1\}$, the function $W_2(-y)$ is locally constant and vanishes at the endpoints, thus we have the inclusion

$$\{y \in \mathbb{R} : W_1(y) = +1\} \subset \{y \in \mathbb{R} : W_2(-y) = 0\}. \tag{7.8}$$

This inclusion is strict because the endpoints y_{1,j_1} of the former set cannot match any of the endpoints $-y_{2,j_2}$ of the latter set due to the nonvanishing of the sums $y_{1,j_1} + y_{2,j_2} + y_{3,1} = y_{1,j_1} + y_{2,j_2}$. We conclude (using Lemma 7.2) that for sufficiently large T , we have

$$\int_{-T}^{\infty} W_1(y) dy < \int_{-T}^{\infty} (1 - W_2(-y)) dy$$

and the desired claim (6.2) then follows from (7.2), (7.3).

REMARK 7.4. The above arguments crucially used the hypothesis in Conjecture 6.8(i). Indeed, the conjecture is false without this hypothesis; a simple counterexample is when $k_1 = 3, k_2 = k_3 = 1, y_{1,1} = -1, y_{1,2} = -4, y_{1,3} = -2$, and $y_{2,1} = y_{3,1} = 0$.

We can partially extend these arguments to cover the cases $k_1, k_2, k_3 > 1$ as follows. We use Lemma 7.2 to partition

$$\{0, \dots, k_1 + 1\} \times \{0, \dots, k_2 + 1\} = V_{12}^0 \cup V_{12}^{+1} \tag{7.9}$$

where V_{12}^{+1} (respectively V_{12}^0) consists of those pairs (p, q) for which $W_3(-y_{1,p} - y_{2,q}) = +1$ (respectively $W_3(-y_{1,p} - y_{2,q}) = 0$). We will work primarily on V_{12}^{+1} , although much of the analysis below also applies to V_{12}^0 . The set V_{12}^0 is a combinatorial analogue of the compact set K in the end of Section 5, while V_{12}^{+1} plays the role of the 1-boundaries $\partial(\pi_3)_*U$ that avoid this compact set.

The set V_{12}^{+1} avoids the boundary of $\{0, \dots, k_1 + 1\} \times \{0, \dots, k_2 + 1\}$ and is thus actually a subset of $\{1, \dots, k_1\} \times \{1, \dots, k_2\}$. We place a directed graph $G_{12}^{+1} = (V_{12}^{+1}, E_{12}^{+1})$ on the vertex set V_{12}^{+1} as follows. If $0 \leq p \leq k_1$ and $0 \leq q \leq k_2$ have the same parity and $|y_{1,p} - y_{1,p+1}| \leq |y_{2,q} - y_{2,q+1}|$, we connect (p, b) to $(p + 1, b)$

whenever $b \in \{q, q + 1\}$ is odd with $(p, b) \in V_{12}^{+1}$, and connect $(p + 1, b)$ to (p, b) whenever $b \in \{q, q + 1\}$ is even with $(p, b) \in V_{12}^{+1}$. If instead p, q have the same parity and $|y_{1,p} - y_{1,p+1}| > |y_{2,q} - y_{2,q+1}|$, we connect (a, q) to $(a, q + 1)$ whenever $a \in \{p, p + 1\}$ is odd with $(a, q) \in V_{12}^{+1}$ and connect $(a, q + 1)$ to (a, q) whenever $a \in \{p, p + 1\}$ is even with $(a, q) \in V_{12}^{+1}$. By Lemma 7.3(ii), this construction only produces edges that start and end in V_{12}^{+1} ; indeed, every point (a, b) in $\{p, p + 1\} \times \{q, q + 1\} \cap V_{12}^{+1}$ will be connected to another point in this set, either by an outgoing edge (if $a + b$ has the opposite parity to p or q) or an incoming edge (if $a + b$ has the same parity as p or q). Applying this procedure to each square $\{p, p + 1\} \times \{q, q + 1\}$ with $0 \leq p \leq k_1$ and $0 \leq q \leq k_2$ the same parity, one obtains a directed graph $G_{12}^{+1} = (V_{12}^{+1}, E_{12}^{+1})$ in which each vertex has exactly one outgoing edge and one incoming edge; thus G_{12}^{+1} decomposes into disjoint simple directed cycles. Any one of these cycles γ can enter a vertical line $\{a\} \times \{0, \dots, k_2 + 1\}$ from the left only when the second coordinate is odd, and from the right only when the second coordinate is even; thus γ will intersect such a vertical line at odd second coordinates the same number of times as at even second coordinates; that is to say

$$\sum_{b:(a,b) \in \gamma} (-1)^b = 0.$$

Similarly for every horizontal line, thus

$$\sum_{a:(a,b) \in \gamma} (-1)^a = 0$$

for all $0 \leq b \leq k_2 + 1$. As a consequence, we have

$$\sum_{(a,b) \in \gamma} (-1)^{a+b} (y_{1,a} + y_{2,b}) = 0$$

for each cycle γ , and hence on summing in γ

$$\sum_{(a,b) \in V_{12}^{+1}} (-1)^{a+b} (y_{1,a} + y_{2,b}) = 0$$

and hence by (7.9)

$$\sum_{(a,b) \in V_{12}^0} (-1)^{a+b} (y_{1,a} + y_{2,b}) = \sum_{j=1}^{k_1} (-1)^{j-1} y_{1,j} + \sum_{j=1}^{k_2} (-1)^{j-1} y_{2,j}. \tag{7.10}$$

Next, we claim the identity

$$\sum_{(a,b) \in V_{12}^{+1}} (-1)^{a+b} \operatorname{sgn}(y_{1,a} + y_{2,b} + y_{3,r}) = 0 \tag{7.11}$$

for all $0 \leq r \leq k_3 + 1$. This is certainly the case for $r = 0$, so it suffices to show that

$$\sum_{(a,b) \in V_{12}^+} (-1)^{a+b} \operatorname{sgn}(y_{1,a} + y_{2,b} + y_{3,r}) = \sum_{(a,b) \in V_{12}^+} (-1)^{a+b} \operatorname{sgn}(y_{1,a} + y_{2,b} + y_{3,r+1})$$

for all $0 \leq r \leq k_3$. Fix such a r . By breaking up V_{12}^+ into squares, it suffices to show that

$$\begin{aligned} & \sum_{(a,b) \in \{p,p+1\} \times \{q,q+1\} \cap V_{12}^+} (-1)^{a+b} \operatorname{sgn}(y_{1,a} + y_{2,b} + y_{3,r}) \\ &= \sum_{(a,b) \in \{p,p+1\} \times \{q,q+1\} \cap V_{12}^+} (-1)^{a+b} \operatorname{sgn}(y_{1,a} + y_{2,b} + y_{3,r+1}) \end{aligned} \quad (7.12)$$

whenever $0 \leq p \leq k_1$ and $0 \leq q \leq k_2$ have the same parity as r . Suppose first that $|y_{1,p+1} - y_{1,p}| \leq |y_{2,q+1} - y_{2,q}|$, then by Lemma 7.3(ii), the set $\{p, p+1\} \times \{q, q+1\} \cap V_{12}^+$ is the union of horizontal lines $\{(p, b), (p+1, b)\}$. It then suffices to show that for each such line, we have

$$\sum_{a \in \{p,p+1\}; c \in \{r,r'+1\}} (-1)^{a+c} \operatorname{sgn}(y_{1,a} + y_{2,b} + y_{3,c}) = 0 \quad (7.13)$$

for all $0 \leq r' \leq k_3$ with the same parity as p, q ; but from Conjecture 6.8(ii) and Example 6.7, the sign of the triple sums of $\{y_{1,p}, y_{1,p+1}\}$, $\{y_{2,q}, y_{2,q+1}\}$, $\{y_{3,r'}, y_{3,r'+1}\}$ is not influenced by one of $\{y_{1,p}, y_{1,p+1}\}$ or $\{y_{3,r'}, y_{3,r'+1}\}$, giving (7.13). The case when $|y_{1,p+1} - y_{1,p}| > |y_{2,q+1} - y_{2,q}|$ is treated similarly (using vertical lines in place of horizontal lines).

If we now define the modified winding number

$$W_{12}^0(y) := \sum_{(a,b) \in V_{12}^0} (-1)^{a+b} \operatorname{sgn}(y_{1,a} + y_{2,b} - y)$$

then we see from (7.11) and Proposition 7.3 that

$$W_{12}^0(-y_{3,r}) = 0 \quad (7.14)$$

for all $0 \leq r \leq k_3 + 1$. From (7.10) and Fubini's theorem we see that

$$\int_{-T}^{\infty} W_{12}^0(y) dy = \sum_{j=1}^{k_1} (-1)^{j-1} y_{1,j} + \sum_{j=1}^{k_2} (-1)^{j-1} y_{2,j} + T$$

and

$$\int_{-\infty}^T (1 - W_{12}^0(y)) dy = - \sum_{j=1}^{k_1} (-1)^{j-1} y_{1,j} + \sum_{j=1}^{k_2} (-1)^{j-1} y_{2,j} + T \quad (7.15)$$

for sufficiently large T .

On the set $\{y : W_3(-y) = +1\}$, we now see that the function W_{12}^0 is locally constant (since, by definition of W_{12}^0 , all the discontinuities $y_{1,p} + y_{2,q}$ of W_{12}^0 lie in the set $\{y : W_3(-y) = 0\}$) and equal to 0 on the boundary (thanks to (7.14)). This gives the inclusion

$$\{y : W_3(-y) = +1\} \subset \{y : W_{12}^0(y) = 0\} \quad (7.16)$$

which generalizes (a permutation of) (7.8). This gives some (but not all) cases of Conjecture 6.8:

PROPOSITION 7.5. *Conjecture 6.8 holds under the additional assumption that the function $W_{12}^0(y) \leq 1$ for all y .*

This case of Conjecture 6.8 is analogous to the case of (5.8) when the 2-cycle Ω appearing in that inequality has a definite sign.

Proof. The inclusion (7.16) is strict, because the endpoints of the set $\{y : W_3(-y) = +1\}$ cannot agree with any of the endpoints of $\{y : W_{12}^0(y) = 0\}$. We conclude (using Lemma 7.2 and the hypothesis $W_{12}^0 \leq 1$) that for sufficiently large T that

$$\int_{-\infty}^T W_3(-y) dy < \int_{-\infty}^T (1 - W_{12}^0(y)) dy$$

and the desired inequality (6.2) then follows from (7.2), (7.15). \square

This observation can handle several further cases of Conjecture 6.8 (for example, the perturbative regime in which the $y_{2,1}, \dots, y_{2,k_2}$ are very small compared to the differences between the $y_{1,1}, \dots, y_{1,k_1}$). Unfortunately it is possible for W_{12}^0 to exceed 1, which means that one cannot resolve Conjecture 6.8 purely on the strength of the inclusion (7.16). However, it appears from numerous examples that whenever this occurs, a significant portion of the set $\{y : W_{12}(y)^0 = 0\}$ is ‘closed off’ from W_3 , in that the set $\{y : W_3(-y) = +1\}$ is prohibited from entering that portion, which restores the truth of Conjecture 6.8; the author was able to make this statement rigorous in the case $k_3 = 3$ by a rather lengthy and *ad hoc* argument, which unfortunately does not seem to extend to the general case. Rather than present this (somewhat unenlightening) argument here,

we give an example to illustrate this ‘closing off’ phenomenon. We will take $k_1 = k_2 = 3$ and $y_{1,1} < y_{1,2} < y_{1,3}$ and $y_{2,1} < y_{2,2} < y_{2,3}$ (this ordering is consistent with the noncrossing hypothesis (i)). We will assume that we are in the ‘almost perturbative setting’ in which the nine sums $s_{j_1, j_2} := y_{1, j_1} + y_{2, j_2}$ for $j_1, j_2 = 1, 2, 3$ are ordered by the relations

$$s_{1,1} < s_{1,2} < s_{1,3}, s_{2,1} < s_{2,2} < s_{2,3} < s_{3,1} < s_{3,2} < s_{3,3},$$

thus the only uncertainty in the ordering of these nine sums arises from the relative positions of $s_{1,3}$ and $s_{2,1}$; clearly both orderings are possible. These relations imply the further inequalities

$$y_{2,2} - y_{2,1}, y_{2,3} - y_{2,2} < y_{1,2} - y_{1,1}, y_{1,3} - y_{1,2}.$$

By this and many applications of Proposition 7.3(ii) we can see that $W_3(-y_{1,p} - y_{2,q}) = 0$ for all $p, q \in \{0, 1, 2, 3, 4\}$, hence V_{12}^{+1} is empty and $W_{12}^0 = W_{12}$ in this case.

First suppose one is in the ‘fully perturbative’ setting where $s_{1,3} < s_{2,1}$ (this for instance occurs when all the $y_{2,1}, y_{2,2}, y_{2,3}$ are small compared to the differences $y_{1,2} - y_{1,1}$ and $y_{1,3} - y_{1,2}$). In this case the winding number $W_{12} = W_{12}^0$ only takes the values 0 and 1, with the former occurring on the intervals

$$(s_{1,1}, s_{1,2}) \cup (s_{1,3}, s_{2,1}) \cup (s_{2,2}, s_{2,3}) \cup (s_{3,1}, s_{3,2}) \cup (s_{3,3}, +\infty), \quad (7.17)$$

and Proposition 7.5 gives (6.2) in this case. In this case one can make the error in (6.2) arbitrarily small; for instance if one takes $k_3 = 9$ and

$$\begin{aligned} y_{3,1} &= -s_{1,1} - \varepsilon \\ y_{3,2} &= -s_{1,2} + \varepsilon \\ y_{3,3} &= -s_{1,3} - \varepsilon \\ y_{3,4} &= -s_{2,3} + \varepsilon \\ y_{3,5} &= -s_{2,2} - \varepsilon \\ y_{3,6} &= -s_{2,1} + \varepsilon \\ y_{3,7} &= -s_{3,1} - \varepsilon \\ y_{3,8} &= -s_{3,2} + \varepsilon \\ y_{3,9} &= -s_{3,3} - \varepsilon \end{aligned}$$

for some sufficiently small $\varepsilon > 0$, one can check that the hypotheses of Conjecture 6.8 hold, and the left and right-hand sides of (6.2) differ by 9ε ; see Figure 12.

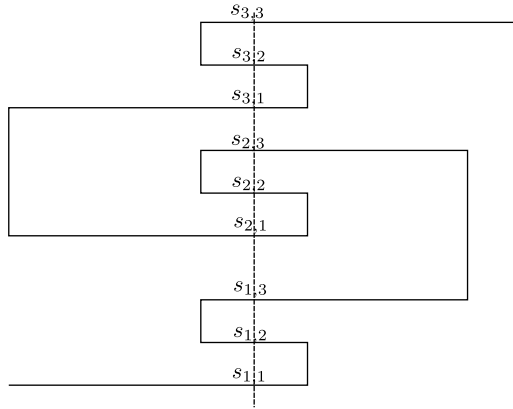


Figure 12. The perturbative case. The solid line represents those sums $(x, y_1 + y_2)$, where (x, y_1) lies on a simple curve passing through $(0, y_{1,1}), (0, y_{1,2}), (0, y_{1,3})$, and (x, y_2) lies on a simple curve passing through $(0, y_{2,1}), (0, y_{2,2}), (0, y_{2,3})$. Note how the entire region (7.17) (viewed as a subset of the y -axis, drawn here as a dashed line) lies above the solid line, in the sense that it is connected to $(0, T)$ for arbitrarily large and negative $-T$. The points $(0, -y_{3,i})$ for $i = 1, \dots, 9$ (not pictured) lie just above the curve in this region.

Now suppose that $s_{1,3} \geq s_{2,1}$. In this case, $W_{12} = W_{12}^0$ now takes the value of 0 on the intervals

$$(s_{1,1}, s_{1,2}) \cup (s_{2,2}, s_{2,3}) \cup (s_{3,1}, s_{3,2}) \cup (s_{3,3}, +\infty)$$

but is additionally equal to +2 on the interval $(s_{3,1}, s_{2,3})$. The argument used to prove Proposition 7.5 then fails to establish (6.2), incurring instead an additional additive error of $s_{1,3} - s_{2,1}$. However, in this case the portion $(s_{2,2}, s_{2,3})$ of $\{y : W_{12}(y) = 1\}$ now becomes ‘closed off’ from the points $-y_{3,1}, \dots, -y_{3,k_3}$, in the sense that none of the $-y_{3,j}$ can lie in this interval, which also implies that $W_3(-y)$ cannot equal 1 in this interval either; see Figure 13. This lets one improve the bound arising from (7.16) by a factor of $s_{2,3} - s_{2,2}$, which exceeds the loss of $s_{1,3} - s_{2,1}$ incurred previously because $s_{2,1} > s_{1,2}$. This restores Conjecture 6.8 in this case. To see why none of the $-y_{3,j}$ lie in $(s_{2,2}, s_{2,3})$, suppose for contradiction that this were not the case, and let $1 \leq p \leq k_3$ be the largest index such that $-y_{3,p} \in (s_{2,2}, s_{2,3})$. This index p cannot equal k_3 , because this would imply that the pairs $\{y_{1,1}, y_{1,2}\}, \{y_{2,3}, y_{2,4}\}$ and $\{y_{3,p}, y_{3,p+1}\}$ have crossing sums (only one of the eight sums from these pairs is positive), contradicting Conjecture 6.8(ii).

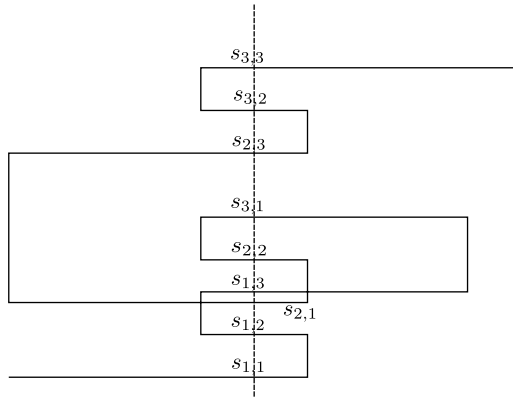


Figure 13. The nonperturbative case. There is now a region of winding number $+2$ between $s_{2,1}$ and $s_{1,3}$. But to compensate for this, the region between $s_{2,2}$ and $s_{3,1}$, which still has a winding number of 0 , has been cut off from $(0, T)$ for large T . This cut-off region is necessarily larger (as measured as a portion of the y -axis) than the region of winding number $+2$.

The same argument excludes the case when p is odd and less than k_3 , since in this case $y_{3,p+1}$ avoids $(s_{2,2}, s_{2,3})$ by hypothesis, and also avoids $(s_{1,3}, s_{2,2})$ since W_{12}^0 equals $+1$ there (here is where we use $s_{1,3} \geq s_{2,1}$), so one has an odd number of positive sums in this case. Finally, the index p cannot be even, because $-y_{3,p+1}$ lies outside $(s_{2,2}, s_{2,3})$ and also outside $(s_{3,2}, s_{3,3})$ (as W_{12} equals $+1$ there) and hence the triple $\{y_{1,2}, y_{1,3}\}, \{y_{2,2}, y_{2,3}\}, \{y_{3,p}, y_{3,p+1}\}$ would be crossing (this has an odd number of positive sums), again contradicting Conjecture 6.8(ii). Thus Conjecture 6.8 holds in all of these cases. More generally, the author has observed numerically that every creation of a region where W_{12}^0 exceeds 1 will invariably be accompanied by a larger region of $\{W_{12}^0(y) = 0\}$ which is now ‘closed off’ from W_3 , and was able to verify this claim rigorously when $k_1 = 3$ or $k_2 = 3$ by *ad hoc* methods, but was unable to see how to establish such a claim in general.

Acknowledgements

The author is supported by NSF grant DMS-1266164, the James and Carol Collins Chair, and by a Simons Investigator Award. He also thanks Mark Meyerson, Albert Hasse, the anonymous referees, and anonymous commenters on his blog for helpful comments, suggestions, and corrections.

References

- [1] J. W. Alexander, ‘Topological invariants of knots and links’, *Trans. Amer. Math. Soc.* **30** (1928), 275–306.
- [2] J. Cantarella, E. Denne and J. McCleary, ‘Transversality for configuration spaces and the ‘square-peg’ theorem’, Preprint (2014), [arXiv:1402.6174](https://arxiv.org/abs/1402.6174).
- [3] C. M. Christensen, ‘A square inscribed in a convex figure’, *Matematisk Tidsskrift B* **1950** (1950), 22–26.
- [4] A. Emch, ‘Some properties of closed convex curves in a plane’, *Amer. J. Math.* **35** (1913), 407–412.
- [5] A. Emch, ‘On the medians of a closed convex polygon’, *Amer. J. Math.* **37** (1915), 19–28.
- [6] A. Emch, ‘On some properties of the medians of closed continuous curves formed by analytic arcs’, *Amer. J. Math.* **38**(1) (1916), 6–18.
- [7] R. Fenn, ‘The table theorem’, *Bull. Lond. Math. Soc.* **2** (1970), 73–76.
- [8] H. W. Guggenheimer, ‘Finite sets on curves and surfaces’, *Israel J. Math.* **3** (1965), 104–112.
- [9] A. Hatcher, *Algebraic Topology*, (Cambridge University Press, Cambridge, 2002).
- [10] C. M. Hebbert, ‘The inscribed and circumscribed squares of a quadrilateral and their significance in kinematic geometry’, *Ann. of Math.* (2) **16**(1–4) (1914/15), 38–42.
- [11] R. P. Jerrard, ‘Inscribed squares in plane curves’, *Trans. Amer. Math. Soc.* **98** (1961), 234–241.
- [12] R. N. Karasëv, ‘On two conjectures of Makeev’, translated from *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **415** (2013); *Geometriya i Topologiya.* **12**, 5–14; *J. Math. Sci. (N.Y.)* **212**(5) (2016), 521–526.
- [13] S. K. Lando and A. K. Zvonkin, ‘Plane and projective meanders, Conference on Formal Power Series and Algebraic Combinatorics (Bordeaux, 1991)’, *Theoret. Comput. Sci.* **117**(1–2) (1993), 227–241.
- [14] V. V. Makeev, ‘On quadrangles inscribed in a closed curve’, *Math. Notes* **57**(1–2) (1995), 91–93.
- [15] B. Matschke, ‘Equivariant topology methods in discrete geometry’, PhD Thesis, Freie Universität Berlin, 2011.
- [16] B. Matschke, ‘A survey on the square peg problem’, *Notices Amer. Math. Soc.* **61**(4) (2014), 346–352.
- [17] M. J. Nielsen and S. E. Wright, ‘Rectangles inscribed in symmetric continua’, *Geom. Dedicata* **56**(3) (1995), 285–297.
- [18] I. Pak, ‘Lectures on discrete and polyhedral geometry’, <http://math.ucla.edu/~pak/book.htm>, 2010.
- [19] V. Pettersson, H. Tverberg and P. Östergård, ‘A note on Toeplitz’ conjecture’, *Discrete Comput. Geom.* **51** (2014), 722–728.
- [20] V. Poénaru, ‘What is . . . an infinite swindle?’, *Notices Amer. Math. Soc.* **54** (2007), 619–622.
- [21] F. Sagols and R. Marín, ‘The inscribed square conjecture in the digital plane’, in *Combinatorial Image Analysis*, Lecture Notes in Computer Science, 5852 (Springer, Berlin, 2009), 411–424.
- [22] F. Sagols and R. Marín, ‘Two discrete versions of the inscribed square conjecture and some related problems’, *Theoret. Comput. Sci.* **412**(15) (2011), 1301–1312.
- [23] L. G. Schnirelman, ‘On some geometric properties of closed curves’, *Uspekhi Mat. Nauk* **10** (1944), 34–44.
- [24] R. Steinberg, ‘A general Clebsch–Gordan theorem’, *Bull. Amer. Math. Soc.* **67** (1961), 406–407.

- [25] W. R. Stromquist, 'Inscribed squares and squarelike quadrilaterals in closed curves', *Mathematika* **36** (1989), 187–197.
- [26] O. Toeplitz, 'Ueber einige Aufgaben der Analysis situs', *Verhandlungen der Schweizerischen Naturforschenden Gesellschaft in Solothurn* **4** (1911), 197.
- [27] S. Vrećica and R. T. Živaljević, 'Fulton–MacPherson compactification, cyclohedra, and the polygonal pegs problem', *Israel J. Math.* **184**(1) (2011), 221–249.
- [28] K. Zindler, 'Über konvexe Gebilde', *Monatshefte für Mathematik und Physik* **31** (1921), 25–56.