THE MAHLER CONJECTURE IN TWO DIMENSIONS VIA THE PROBABILISTIC METHOD

MATTHEW C. H. TOINTON

ABSTRACT. The *Mahler volume* is, intuitively speaking, a measure of how "round" a centrally symmetric convex body is. In one direction this intuition is given weight by a result of Santaló, who in the 1940s showed that the Mahler volume is maximized, in a given dimension, by the unit sphere and its linear images, and only these. A counterpart to this result in the opposite direction is proposed by a conjecture, formulated by Kurt Mahler in the 1930s and still open in dimensions 4 and greater, asserting that the Mahler volume should be minimized by a cuboid. In this article we present a seemingly new proof of the 2-dimensional case of this conjecture via the probabilistic method. The central idea is to show that either deleting a random pair of edges from a centrally symmetric convex polygon, or deleting a random pair of vertices, reduces the Mahler volume with positive probability.

1. INTRODUCTION

A *convex body* $A \subset \mathbb{R}^d$ is a compact convex set with non-empty interior; it is said to be *centrally symmetric* if $x \in A$ precisely when $-x \in A$. This article concerns a long-standing and seemingly difficult question in convex geometry—the *Mahler conjecture*—that attempts to give a rigorous version of the intuitively reasonable statement that cubes and octahedra are the "least round" centrally symmetric convex bodies.

Given a centrally symmetric convex body $A \subset \mathbb{R}^d$, the *polar body* $A^\circ \subset \mathbb{R}^d$ is defined by

$$A^\circ = \{x \in \mathbb{R}^d : \langle x, a \rangle \leq 1 \text{ for all } a \in A\}.$$

The *Mahler volume* $M(A)$ of A is then defined to be

$$M(A) := \text{vol}(A)\,\text{vol}(A^\circ).$$

Here, of course, $\text{vol}(A)$ means the volume of A (or, more formally, its Lebesgue measure). Note that the Mahler volume is invariant under invertible linear transformations of $\mathbb{R}^d$, since if T is such a transformation then

$$(1.1) \qquad\qquad T(A)^\circ = (T^*)^{-1}(A^\circ).$$

The Mahler volume can be thought of as measuring how "round" A is; the Mahler conjecture seeks to justify this line of thinking by showing that Euclidean balls and their linear images maximize the Mahler volume, whereas cubes and cross-polytopes and their linear images minimize it.

Write B^d for the unit Euclidean ball in $\mathbb{R}^d$, write $Q^d = [-1,1]^d$ for the standard centred cube, and write $O_d = \{x \in \mathbb{R}^d : \|x\|_1 \leq 1\}$ for the standard cross-polytope. Note that $O_d = (Q^d)^\circ$ and

$Q^d = (O^d)^\circ$, so that $M(Q^d) = M(O_d)$. Precisely, then, the Mahler conjecture states that for any centrally symmetric convex body $A \subset \mathbb{R}^d$ we have $M(Q^d) \leq M(A) \leq M(B^d)$.

The second of these inequalities was proved by Santaló [11] in the 1940s, having been previously proved by Blaschke [2] in the cases $d = 2, 3$. The Mahler conjecture thus reduces to the lower bound, and in fact for the remainder of this article we refer to the lower bound only as the "Mahler conjecture".

Conjecture (Mahler [9])**.** *Let $A \subset \mathbb{R}^d$ be a centrally symmetric convex body. Then*

$$M(A) \geq M(Q^d).$$

The $d = 2$ case of this conjecture was proved by Mahler [8]. Iriyeh and Shibata [4] have very recently released a proof the $d = 3$ case. The conjecture seems to be open for $d \geq 4$ [13], although there are some partial results. For example, the conjecture is known to hold for certain specific classes of convex body (see [7], for example), and the weaker inequality $M(A) \geq (\pi/4)^{d-1} M(Q^d)$ has been shown by Kuperberg [6] to hold in full generality. Moreover, it is known that the cube is a *local* minimizer of the Mahler volume [10], and more generally that so-called *Hanner polytopes* (which have the same Mahler volume as the cube) are local minimizers [5]. For more information on progress on the Mahler conjecture the reader may consult the expository article of Tao contained in [13], or the dissertation [3] of Henze, which also gives a detailed account of Mahler's proof of the $d = 2$ case and a sketch proof of the upper bound $M(A) \leq M(B^d)$.

In proving the $d = 2$ case, Mahler actually proved the following theorem (see [3, Lemma 2.9]).

Theorem 1 (Mahler)**.** *Let $n > 2$, and let A be a centrally symmetric convex polygon with $2n$ edges. Then there exists a centrally symmetric polygon A' with $2(n-1)$ edges such that $M(A') < M(A)$.*

This is sufficient since by limiting arguments we may always assume that the convex body A in the Mahler conjecture is a polytope, and any centrally symmetric quadrilateral is a linear image of Q^2.

The purpose of this article is to present a new and, we hope, entertaining proof of Theorem 1, via a somewhat different method from that used by Mahler. In fact, we give a slight refinement of Theorem 1, which we now describe.

We may view a centrally symmetric convex polytope A in $\mathbb{R}^d$ as the region bounded by a finite set $\mathcal{H}$ of hyperplanes, with central symmetry in particular implying that $H \in \mathcal{H}$ if and only if $-H \in \mathcal{H}$; let us call the minimal such $\mathcal{H}$ the *hyperplane presentation* of A. Note that if the number of pairs of hyperplanes is greater than d then there is at least one way to remove a pair of hyperplanes from the presentation in such a way that the remaining hyperplanes still define a centrally symmetric convex polytope. Alternatively, we may view a centrally symmetric convex polytope A in $\mathbb{R}^d$ as the convex hull of a finite set $\mathcal{P}$ of points, with central symmetry in particular implying that $x \in \mathcal{P}$ if and only if $-x \in \mathcal{P}$; let us call the minimal such $\mathcal{P}$ the *vertex presentation* of A. Note that if the number of pairs of points is greater than d then there is at least one way to remove a pair of points from the presentation in such a way that the remaining points still define a centrally symmetric convex polytope.

Our refinement of Theorem 1 is then as follows.

Theorem 2. *Let $n > 2$, and let A be a centrally symmetric convex polygon with $2n$ edges. Then there is either a way to remove an opposite pair of lines from the hyperplane presentation A, or a way to remove an opposite pair of points from the vertex presentation of A, in such a way that the Mahler volume decreases.*

One of the most interesting aspects of our proof of Theorem 2 is that it uses the so-called *probabilistic method*. The probabilistic method is a powerful tool, most often associated with combinatorics, that can be used to prove the existence of a certain mathematical object with certain given properties. The crucial insight of the probabilistic method is that in order to show that it is possible for a given object to have certain properties, it is sufficient to show that if the object is generated at random according to some probability distribution then it has the required properties with positive probability. The reader may consult the book [1] of Alon and Spencer for a more thorough description of the method and of its rich history and numerous applications.

Acknowledgements. I am grateful to two anonymous referees for careful readings of an earlier version of this article and a number of suggestions that have significantly improved the exposition.

2. Outline of the argument

Given a centrally symmetric convex polytope A in $\mathbb{R}^d$ we write $V(A)$ for the set of vertices of A, and $F(A)$ for the set of *facets* of A, which is to say the set of faces of A of dimension $d - 1$. In the two-dimensional setting we write $E(A)$ instead of $F(A)$ to emphasize that facets are simply edges.

Each $s \in F(A)$ lies within a hyperplane in $\mathbb{R}^d$ appearing in the hyperplane presentation of A. Thus, there exists a unit vector $u(s) \in \mathbb{R}^d$ that is perpendicular to s, and some real number $c(s) > 0$ such that $s \subset \{x \in \mathbb{R}^d : \langle x, u(s) \rangle = c(s)\}$, and we have

$$(2.1) \qquad A = \{x \in \mathbb{R}^d : \langle x, u(s) \rangle \leq c(s) \text{ for all } s \in F(A)\}.$$

It follows that A° is a centrally symmetric convex polytope in $\mathbb{R}^d$, and the set $V(A^\circ)$ of vertices of A° is precisely $\{v^\circ(s) : s \in F(A)\}$, where for each $s \in F(A)$ we define the point $v^\circ(s) \in \mathbb{R}^d$ via

$$(2.2) \qquad v^\circ(s) = u(s)/c(s).$$

For each vertex $v \in V(A^\circ)$ we denote by $e^\circ(v)$ the facet of A that gave rise to v; thus $v = v^\circ(e^\circ(v))$. This notation is illustrated in Figure 1 for the two-dimensional case.

Throughout our proof of Theorem 2 we consider two types of parallelogram formed from the vertices of a given centrally symmetric convex polygon A, which we call *type-1* parallelograms and *type-2* parallelograms. We define the *type-1 parallelogram based at the edge $e \in E(A)$*, written
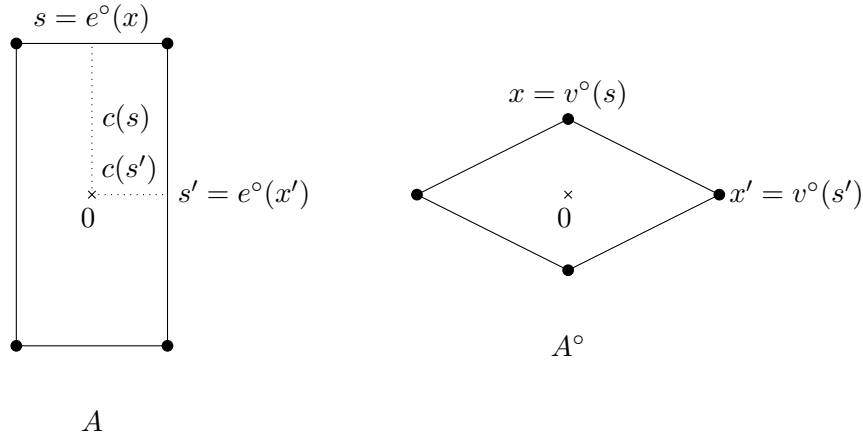

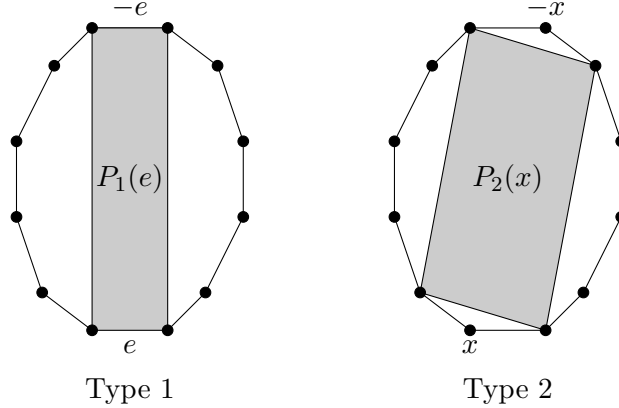FIGURE 1. Illustration of the vertex and facet notation in two dimensions.

FIGURE 2. The two types of parallelogram.

$P_1(e)$, to be the convex hull of the edges $\pm e$. We define the *type-2 parallelogram based at the vertex* $x \in V(A)$, written $P_2(x)$, to be the convex hull of the four vertices adjacent to the vertices $\pm x$. Note that the definition of a type-2 parallelogram is degenerate when the ambient polygon is itself a parallelogram. These definitions are illustrated in Figure 2.

The first stage of our argument is to show that the type-1 parallelograms in a centrally symmetric convex polygon are smaller on average than the type-2 parallelograms, as follows.

Proposition 2.1. *Let A be a centrally symmetric convex polygon in $\mathbb{R}^2$ with at least 6 sides. Then*

$$(2.3) \qquad \sum_{s \in E(A)} \mathrm{area}(P_1(s)) \leq \sum_{x \in V(A)} \mathrm{area}(P_2(x)),$$

with equality if and only A has exactly 6 sides.

The second stage of our argument shows how to pass from a certain comparison of type-1 parallelograms in A with type-2 parallelograms in A° to the conclusion required by Theorem 2, as follows.

Proposition 2.2. *Let A be a centrally symmetric convex polygon in $\mathbb{R}^2$ with at least 6 sides, and let $s \in E(A)$. Let $A_s = \{x \in \mathbb{R}^d : \langle x, u(s')\rangle \leq c(s') \text{ for } s' \in E(A)\backslash\{\pm s\}\}$ be the set obtained from A by removing $\pm s$ from $E(A)$ in the presentation (2.1). Suppose that*

$$(2.4) \qquad \frac{\mathrm{area}(P_1(s))}{\mathrm{area}(A)} \leq \frac{\mathrm{area}(P_2(v^\circ(s)))}{\mathrm{area}(A^\circ)}.$$

Then $M(A_s) < M(A)$.

We prove these results in the next section using elementary geometry. For now, let us see how they combine to imply Theorem 2.

Proof of Theorem 2. It is not difficult to check that if A is a centrally symmetric convex body in $\mathbb{R}^d$ then $(A^\circ)^\circ = A$ (see [12, Proposition 1.1], for example). In the case $d = 2$, note also that A and A° are each centrally symmetric polygons with the same number of edges, and removing a pair of edges from the hyperplane presentation of one corresponds to removing a pair of vertices from the

vertex presentation of the other. In proving Theorem 2, therefore, we may interchange A and A° without loss of generality. In particular, we may assume that

$$\mathbb{E}_{s\in E(A)}\frac{\mathrm{area}(P_1(s))}{\mathrm{area}(A)} \le \mathbb{E}_{t\in E(A^\circ)}\frac{\mathrm{area}(P_1(t))}{\mathrm{area}(A^\circ)}.$$

Combining this with Proposition 2.1 implies that

$$\mathbb{E}_{s\in E(A)}\frac{\mathrm{area}(P_1(s))}{\mathrm{area}(A)} \le \mathbb{E}_{s\in E(A)}\frac{\mathrm{area}(P_2(v^\circ(s)))}{\mathrm{area}(A^\circ)},$$

and so—and this is where we apply the probabilistic method—there must exist some s such that the inequality (2.4) holds. Proposition 2.2 implies that for this s we have $M(A_s) < M(A)$, and so the theorem is proved. $\qquad\square$

Remark. In dimension greater than 2 we would no longer be able to interchange A and A° in quite the same way, since they would not in general have the same number of facets or vertices. In three dimensions, for example, if A were the cube and A° the cross-polytope then interchanging A and A° would increase the number of facets of A. Indeed, deleting a pair of facets from the three-dimensional cross-polytope produces a linear image of the cube, and so the process of deleting pairs of facets would not even necessarily terminate if we allowed interchanges.

3. The details of the argument

In this section we prove Propositions 2.1 and 2.2. Throughout, given points $x_1, \ldots, x_r \in \mathbb{R}^d$ we write $[x_1, \ldots, x_r]$ for their convex hull. Moreover, given a centrally symmetric convex polytope A, for each $x \in V(A)$ we write $x + 1$ for the vertex neighbouring x in a clockwise direction, and $x - 1$ for the vertex neighbouring x in an anticlockwise direction. More generally, for $k \in \mathbb{N}$ we define $x + k$ and $x - k$ recursively via $x + k = (x + (k - 1)) + 1$ and $x - k = (x - (k - 1)) - 1$. Note that if A has $2n$ vertices then $x + n = -x$ for every $x \in V(A)$.

Proof of Proposition 2.1. Equality is trivial when A has 6 sides, so we may assume it has $2n$ sides with $n \ge 4$ and prove that the inequality (2.3) holds and is strict.

Given a vertex x of A write $H(x)$ for the hexagon $[x - 1, x, x + 1, x + n - 1, x + n, x + n + 1]$. Write $e(x) \in E(A)$ for the edge $[x - 1, x]$, and note that each of $P_1(e(x))$ and $P_2(x)$ is a subset of $H(x)$. Carving $P_1(e(x))$ out from $H(x)$ leaves triangles $T_1(x) = [x, x + 1, x + n - 1]$ and $-T_1(x)$, as illustrated in Figure 3. Carving out $P_2(x)$ from $H(x)$ leaves the triangles $T_2(x) = [x - 1, x, x + 1]$ and $-T_2(x)$, as also illustrated in Figure 3. The desired conclusion is therefore equivalent to the statement that

$$\sum_{x\in V(A)} \mathrm{area}(T_1(x)) > \sum_{x\in V(A)} \mathrm{area}(T_2(x)).$$

However, since we are summing over all vertices x we may replace x by $x - 1$ on the left-hand side, meaning that this is equivalent to the statement that

$$\sum_{x\in V(A)} \mathrm{area}(T_1(x - 1)) > \sum_{x\in V(A)} \mathrm{area}(T_2(x)).$$

In fact, we claim that this inequality holds not just in the sum, but term by term, in the sense that

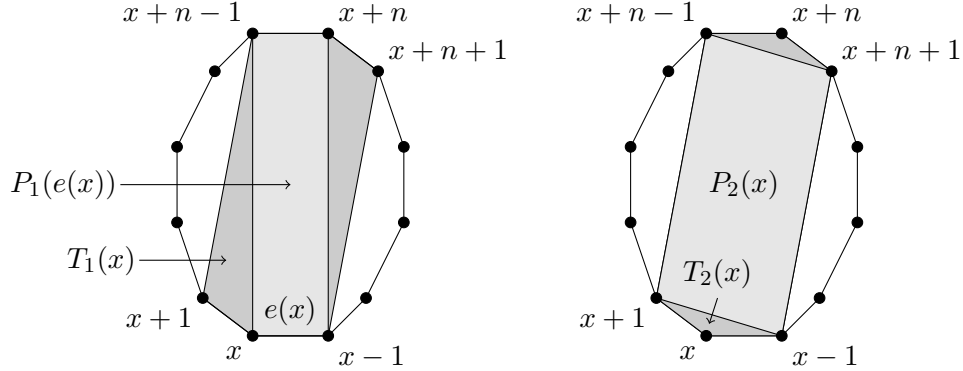$$(3.1) \qquad\qquad \mathrm{area}(T_1(x - 1)) > \mathrm{area}(T_2(x))$$

FIGURE 3. The two parallelograms from the proof of Proposition 2.1 carved out of the same hexagon.

for every $x \in V(A)$. To see that this is true, note that the triangles $T_1(x-1)$ and $T_2(x)$ may both be thought of as having the edge $[x-1, x]$ as a base. Without loss of generality, we may assume that this edge is horizontal and that the body A lies above it, as illustrated in Figure 4.

Now by symmetry the edge $[x+n-1, x+n]$ is horizontal, and so by convexity none of the edges lying on the clockwise path from $[x-1, x]$ to $[x+n-1, x+n]$ can be horizontal or downward sloping from $x+j$ to $x+j+1$. The vertical components of $x, x+1, \ldots, x+n-1$ are therefore strictly increasing. Since $n \geq 4$, we may therefore conclude that $x+n-2$ has a vertical component strictly greater than that of $x+1$. The triangle $T_1(x-1)$ therefore has a greater height than $T_2(x)$ and the same base, as illustrated in Figure 4, and so (3.1) holds and the lemma is proved. $\square$

The proof of Proposition 2.2 is also relatively straightforward, but we make it easier to follow with two simple lemmas.

Lemma 3.1. *Let A be a centrally symmetric convex polygon in $\mathbb{R}^2$. Suppose that $s \in E(A)$ is horizontal, let $s' \in E(A)$ be an edge adjacent to s, and write z for the intersection of the vertical*
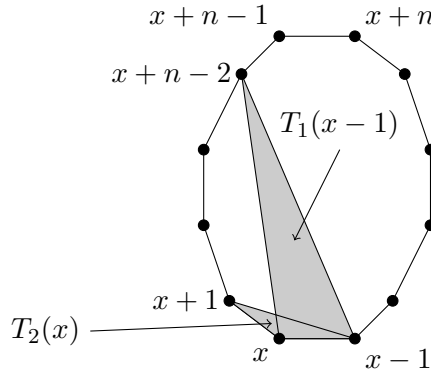

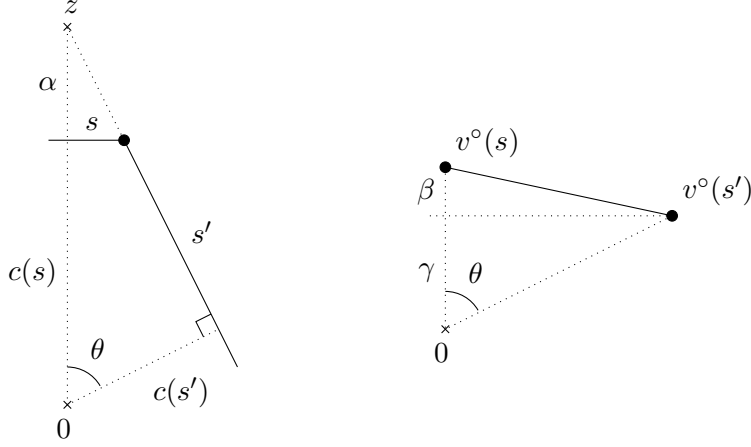FIGURE 4. Illustration of the triangles appearing in the inequality (3.1).

FIGURE 5. Illustration of the hypotheses of Lemma 3.1.

axis and the line in $\mathbb{R}^2$ containing s', as illustrated in Figure 5. Write α for the vertical distance from z to s, write β for the vertical distance from $v^\circ(s)$ to $v^\circ(s')$, and write γ for the vertical distance from $v^\circ(s')$ to the origin. Then

$$\frac{\alpha}{c(s)} = \frac{\beta}{\gamma}.$$

Proof. We have $\|v^\circ(s')\| = 1/c(s')$ by (2.2), and by the equality of the angles labelled θ in Figure 5 we may compare similar triangles to conclude that $\|v^\circ(s')\|/\gamma = (\alpha + c(s))/c(s')$. Combining these gives $\alpha + c(s) = 1/\gamma$. Another instance of (2.2) implies that $\beta + \gamma = 1/c(s)$. These last two equalities give $\alpha\gamma = 1 - c(s)\gamma = \beta c(s)$, which gives the desired result. □

An immediate consequence of Lemma 3.1 is the following fact.

Lemma 3.2. *Let A be a centrally symmetric convex polygon in $\mathbb{R}^2$, and suppose that $s \in E(A)$ is horizontal and that the lines containing the edges adjacent to the edge s intersect on the vertical axis. Then the vertices adjacent to $v^\circ(s)$ lie on a common horizontal line.*

Proof of Proposition 2.2. It follows from (1.1) that applying an invertible linear transormation to A not affect the value of either side of the inequality (2.4). We are therefore free to apply such a transformation to A. In particular, by rotating we may assume that s is horizontal, and by applying a suitable shear we may assume that the intersection z of the lines containing the edges adjacent to s lies on the vertical axis, as illustrated in Figure 6. Note that the edge $[v^\circ(s) - 1, v^\circ(s) + 1]$ of $P_2(v^\circ(s))$ is also horizontal by Lemma 3.2. Having applied these transformations, we write β for the vertical distance from $v^\circ(s)$ to $P_2(v^\circ(s))$, and γ for the distance from $v^\circ(s)$ to the origin minus β, as in Lemma 3.1. This notation is also recorded in Figure 6.

The body A_s is obtained from the body A by removing the edges $\pm s$ and extending their neighbouring edges to the points $\pm z$, which will form a pair of vertices of A_s. The body $(A_s)^\circ$ is obtained from the body A° by removing the vertices $\pm v^\circ(s)$ and taking the convex hull of the remaining vertices. Again, the reader may find it helpful to refer to Figure 6.
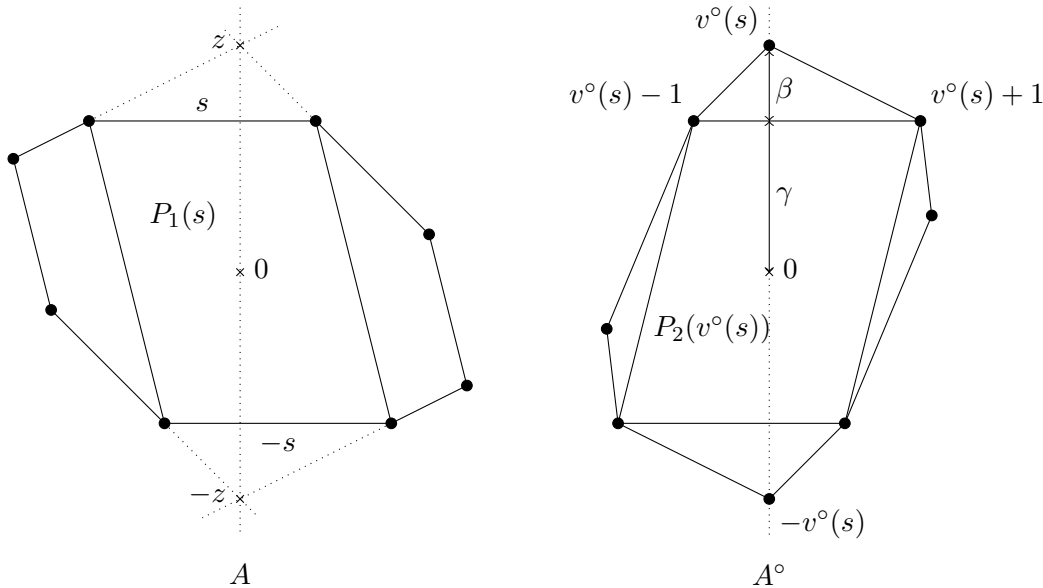
FIGURE 6. Illustration of the proof of Proposition 2.2.

This all implies that

$$\operatorname{area}((A_s)^\circ) = \operatorname{area}(A^\circ) - \frac{\beta}{2\gamma}\operatorname{area}(P_2(v^\circ(s))),$$

and combined with Lemma 3.1 implies that

$$\operatorname{area}(A_s) = \operatorname{area}(A) + \frac{\beta}{2\gamma}\operatorname{area}(P_1(s)).$$

Multiplying these equations together, we see that

$$M(A_s) = M(A) + \frac{\beta}{2\gamma}\Big(\operatorname{area}(A^\circ)\operatorname{area}(P_1(s)) - \operatorname{area}(A)\operatorname{area}(P_2(v^\circ(s)))\Big)$$

$$-\frac{\beta^2}{4\gamma^2}\operatorname{area}(P_1(s))\operatorname{area}(P_2(v^\circ(s))).$$

Since the last term of this expression is always negative, the proposition follows from (2.4).   □

References

[1] N. Alon and J. H. Spencer. *The probabilistic method* (second edition), Wiley–Interscience, New York (2000).

[2] W. Blaschke. *Vorlesungen über Differentialgeometrie II*, Springer, Berlin (1923).

[3] M. Henze. *The Mahler conjecture*, diploma thesis, Otto-von-Guericke-Universität Magdeburg (2008).

[4] H. Iriyeh and M. Shibata. Symmetric Mahler's conjecture for the volume product in the three dimensional case, preprint, arXiv:1706.01749.

[5] J. Kim. Minimal volume product near Hanner polytopes, *J. Funct. Anal.* **226**(4) (2014), 2360–2402. arXiv:1212.2544.

[6] G. Kuperberg. From the Mahler conjecture to Gauss linking integrals, *Geom. Funct. Anal.* **18**(3) (2008), 870–892. arXiv:math/0610904.

[7] M. A. Lopez and S. Reisner. A special case of Mahler's conjecture, *Discrete Comput. Geom.* **20**(2) (1998), 163–177.

[8] K. Mahler. Ein Minimalproblem für konvexe Polygone, *Mathematica (Zutphen)* **B** (1939), 118–127.

[9] K. Mahler. Ein Übertragungsprinzip für konvexe Körper, *Casopis Pest. Mat. Fys.* **68** (1939), 93–102.

[10] F. Nazarov, F. Petrov, D. Ryabogin and A. Zvavitch. A remark on the Mahler conjecture: local minimality of the unit cube, *Duke Math. J.* **154**(3) (2010), 419–430. arXiv:0905.0867.

[11] L. A. Santaló. An affine invariant for convex bodies of n-dimensional space (in Spanish), *Portugaliae Math.* **8**, (1949), 155–161.

[12] T. C. Tao. *Santaló's inequality*, expository note available from the author's website at
http://www.math.ucla.edu/~tao/preprints/Expository/santalo.dvi.

[13] T. C. Tao. *Structure and Randomness: pages from year one of a mathematical blog*, American Mathematical Society, Providence, RI (2008).

Insitut de Mathématiques, Université de Neuchâtel, Rue Emile-Argand 11, CH-2000 Neuchâtel, Switzerland

E-mail address: matthew.tointon@unine.ch