The minimum area of convex lattice n-gons

Imre Bárány *
Rényi Institute of Mathematics,
Hungarian Academy of Sciences
H-1364 Budapest Pf. 127 Hungary
baranyi@renyi.hu
and
Department of Mathematics
University College London
Gower Street, London, WC1E 6BT, UK

Norihide Tokushige
College of Education
Ryukyu University
Nishihara, Okinawa, 903-0213 Japan
hide@edu.u-ryukyu.ac.jp

June 28, 2002

Abstract

Let $A(n)$ be the minimum area of convex lattice n-gons. We prove that $\lim A(n)/n^3$ exists. Our computations suggest that the value of the limit is very close to $0.0185067\ldots$.

*Partially supported by Hungarian National Science Foundation Grants T 032452 and t 029255

1 Introduction

What is the minimal area $A(n)$ a convex lattice polygon with n vertices can have? The first to answer this question was G.E. Andrews [An63]. He proved that $A(n) \geq cn^3$ with some universal constant c. V.I. Arnol'd arrived to the same question from another direction [Ar80], and proved the same estimate. Further proofs are due to W. Schmidt [Schm85], Bárány-Pach [BP92]. The best lower bound comes form Rabinowitz [Ra93] via an inequality of Rényi-Sulanke [RS63]

$$\frac{1}{8\pi^2} < \frac{A(n)}{n^3} \leq \frac{1}{54}(1 + o(1)).$$

The upper bound follows from Remark 2 below.

Our main result is:

Theorem 1 $\lim A(n)/n^3$ *exists.*

The value of the limit — as we are going to show — equals the minimum of finitely many explicit extremal problems. But the finitely many is about 10^{10}, too many to solve. Our computations show, however, that most likely

$$\lim \frac{A(n)}{n^3} = 0.0185067\ldots.$$

We will also see that the convex lattice n-gon P with area $A(n)$ has elongated shape: after applying a suitable lattice preserving affine transformation, P has lattice width $c_1 n$ in direction $(0, 1)$ and has width $c_2 n^2$ in direction $(1, 0)$ where c_1, c_2 positive constants. Almost all the paper is devoted to the proof of Theorem 1.

Remark 1 Actually, Andrews [An63] showed much more, namely the following (see also [Schm85], [KS84]). If $P \subset \mathbf{R}^d$ is a convex lattice polytope with n vertices and volume $V > 0$, then

$$cn^{\frac{d+1}{d-1}} \leq V,$$

where c is a constant depending only on dimension.

2 Reduction

Define $\mathcal{P}_n$ as the set of all convex lattice n-gons in $\mathbf{R}^2$, then

$$A(n) = \min\{\operatorname{Area} P : P \in \mathcal{P}_n\}.$$

In the next two claims, whose proof is given at the end of this section, we reduce the search for $A(n)$. As $A(n)$ is increasing it is enough to work with even n.

Claim 1 *For even n, there exists a centrally symmetric $P \in \mathcal{P}_n$ with $A(n) = \operatorname{Area} P$.*

Fix a centrally symmetric $P \in \mathcal{P}_n$ with $A(n) = \operatorname{Area} P$ ($n = 2k$ even). The edges are $z_1, z_2, \dots, z_k, -z_1, \dots, -z_k$ in this order. Clearly, each z_i is a primitive vector, i.e., its components are coprime. Write $\mathbf{P}$ for the set of all primitive vectors in $\mathbf{Z}^2$. Define

$$C = \operatorname{conv}\{z_1, z_2, \dots, z_k, -z_1, \dots, -z_k\}.$$

Then P is the zonotope spanned by $\{z_1, \dots, z_k\}$, i.e., $P = \sum_{i=1}^k [0, z_i]$ As it is well-known and easy to check

$$\operatorname{Area} P(C) = \sum_{1 \le i < j \le k} |\det(z_i, z_j)|.$$

Write $\mathcal{C}$ for the set of 0-symmetric convex bodies in $\mathbf{R}^2$. So $C \in \mathcal{C}$ and define

$$A(C) = \frac{1}{8} \sum_{u \in C \cap \mathbf{P}} \sum_{v \in C \cap \mathbf{P}} |\det(u, v)|.$$

The following claim shows that $A(C) = \operatorname{Area} P(C)$.

Claim 2 *If $z \in C \cap \mathbf{P}$ then $z = z_i$, or $-z_i$ for some i.*

This means that the search for $A(n)$, or for minimal $P \in \mathcal{P}_n$ is reduced to the following minimization problem.

$$\operatorname{Min}(n) = \min\{A(C) : C \in \mathcal{C} \text{ with } |C \cap \mathbf{P}| = n\}.$$

Observe that the solution C to the problem Min(n) is invariant under lattice preserving linear transformation. Thus we may fix C in *standard position.* This means that the lattice width of C is $2b = 2b(C)$ and is taken in direction $(0, 1)$. Recall (from [KL88], say,) that the *width* of $K \subset \mathbf{R}^2$ in direction $z \in \mathbf{Z}^2$, $z \neq 0$ is

$$w(z, K) = \max\{zx - zy : \ x, y \in K\},$$

and the *lattice width* of K is, by definition,

$$w(K) = \min\{w(z, K) : \ z \in \mathbf{Z}^2, z \neq 0\}.$$

Now let $[-a, a]$ be the intersection of C with the x axis. We may further assume that the tangent line to C at $(a, 0)$ has slope ≥ 1. A simple computation, (using the fact that the width of C in direction $(1, 0)$ and $(1, -1)$ is at least $2b$) shows that $2a \geq b$. We fix C in this standard position. We record the following inequalities:

$$2a \geq b, \quad 2ab \leq \operatorname{Area} C \leq 4ab.$$

Remark 2 From now on we may assume $b \geq 2$ since for $b = 1$, according to Claim 2, the minimal C is (with $n = 2k$)

$$\operatorname{conv}\{\pm(0, 1), \pm(1, 1), \pm(2, 1), \ldots, \pm(k-2, 1), \pm(1, 0)\}$$

which gives $\lim \frac{A(C)}{n^3} = \frac{1}{48}$, the example found in [Ra93].

When $C \in \mathcal{C}$ is a circle with $|C \cap \mathbf{P}| = n$, its radius, and then $A(C)$ are estimated easily showing $\lim \frac{A(C)}{n^3} = \frac{1}{54}$. This is the estimate given in the introduction.

Proof of Claim 1. Let $Q \in \mathcal{P}_n$ with vertices $v_1, v_2, \ldots, v_{2k}$ $(n = 2k)$ in this order. The diagonal $[v_i, v_{i+k}]$ cuts Q into two parts. Reflecting the part with smaller (or equal) area to the point $(v_i + v_{i+k})/2$ produces a lattice polygon with area $\leq \operatorname{Area} Q$. So it is enough to show that, for some $i \in \{1, \ldots, k\}$, the reflected n-gon is convex. It is certainly convex if there are parallel tangent lines to Q at v_i and v_{i+k}.

If there are no such tangents then the lines of the edges incident to v_i intersect those of incident to v_{i+k} on the same side of the line $v_i v_{i+k}$, on the left side, say. Then the lines of the edges, incident to v_{i+k} intersect those of

v_{i+k+1} on the left side of the line $v_{i+1}v_{i+k+1}$, again. Starting with $i = 1$ a contradiction is reached at $i = k + 1$. ∎

We prove Claim 2 in stronger form:

Claim 2′ *Assume that $x_1, \ldots, x_k \in \mathbf{R}^2$, and no two of them collinear. If $x \in \text{conv}\{\pm x_1, \ldots, \pm x_k\}$ and $x \neq \pm x_i$ $(\forall i)$, then there is a j such that replacing x_j by x gives a zonotope with smaller area.*

Proof. Assume first that x is on the boundary of $\text{conv}\{\pm x_1, \ldots, \pm x_k\}$. Then $x = (1-u)x_s + ux_t$ for some $0 < u < 1$. We may also assume that $\sum_{i=1}^k |\det(x_i, x_s)| \geq \sum_{i=1}^k |\det(x_i, x_t)|$. Let $y_s = x$ and $y_i = x_i$ if $i \neq s$. Then

$$\begin{aligned}
\sum_{i=1}^k |\det(y_i, y_s)| &= \sum_{i \neq s} |\det(x_i, x)| < \sum_{i=1}^k |\det(x_i, x)| \\
&= \sum_{i=1}^k |\det(x_i, (1-u)x_s + ux_t)| \\
&= \sum_{i=1}^k |(1-u)\det(x_i, x_s) + u\det(x_i, x_t)| \\
&\leq (1-u)\sum_{i=1}^k |\det(x_i, x_s)| + u\sum_{i=1}^k |\det(x_i, x_t)| \\
&\leq (1-u)\sum_{i=1}^k |\det(x_i, x_s)| + u\sum_{i=1}^k |\det(x_i, x_s)| \\
&= \sum_{i=1}^k |\det(x_i, x_s)|.
\end{aligned}$$

Thus, replacing x_s by x makes the area smaller.

If x is in the interior of $\text{conv}\{\pm x_1, \ldots, \pm x_k\}$, then λx is on the boundary of this set with a unique $\lambda > 1$ (apart from the trivial case $x = 0$). The previous argument shows that replacing x_s by λx makes the area smaller, and consequently, replacing x_s by x makes it smaller, too. ∎

In the next two sections we approximate $|C \cap \mathbf{P}|$ and $A(C)$ using that the density of $\mathbf{P}$ in $\mathbf{Z}^2$ is $6/\pi^2$ (cf. [HW79]). We need to measure approximation by a quantity invariant under lattice preserving linear transformations. This is going to be the lattice width $2b = 2b(C)$.

3 Approximating $|C \cap \mathbf{P}|$

Lemma 1

$$\left||C \cap \mathbf{P}| - \frac{6}{\pi^2}\operatorname{Area} C\right| \ll \operatorname{Area} C \cdot \frac{\log b}{b}.$$

Here and in what follows we use Vinogradov's $\ll$ notation. Thus $f(n) \ll g(n)$ means that $f(n) \leq Dg(n)$ with some universal constant D.

Proof. The proof is standard and uses the Möbius function $\mu(d)$ see [HW79]. Set

$$C^+ = C \cap \{(x,y) \in \mathbf{R}^2 : y > 0\}.$$

Clearly, $|C \cap \mathbf{P}| = 2 + 2|C^+ \cap \mathbf{P}|$ and

$$\begin{aligned} |C^+ \cap \mathbf{P}| &= \sum_{(u,v)\in C^+\cap \mathbf{Z}^2} \sum_{\substack{d|u \\ d|v}} \mu(d) = \sum_{d=1}^{\infty} \mu(d) \sum_{\substack{(u,v) \in C^+ \cap \mathbf{Z}^2 \\ d|u,\, d|v}} 1 = \\ &= \sum_{d=1}^{b} \mu(d) |\frac{1}{d} C^+ \cap \mathbf{Z}^2|. \end{aligned}$$

Claim 3

$$\left||\frac{1}{d}C^+ \cap \mathbf{Z}^2| - \frac{1}{d^2}\operatorname{Area} C^+\right| \leq \frac{12a}{d}.$$

Proof. It suffices to show this for $d = 1$. Let $Q(z)$ denote the unit square centered at $z \in \mathbf{Z}^2$. Call $z \in \mathbf{Z}^2$ *inside* if $Q(z) \subset C^+$, *boundary* if $z \in C^+$ but $Q(z) \not\subset C^+$, and *outside* if $z \notin C^+$ and $Q(z) \cap \operatorname{int} C^+ \neq \emptyset$. (Note that we will use the same inside, boundary, and outside squares $Q(z)$ in the next section.) Clearly

$$\begin{aligned} |C^+ \cap \mathbf{Z}^2| &= |\{z \in \mathbf{Z}^2 : \text{inside}\}| + |\{z \in \mathbf{Z}^2 : \text{boundary}\}| = \\ &= \operatorname{Area} C^+ + \sum_{z \text{ boundary}} \operatorname{Area}(Q(z) \setminus C^+) - \sum_{z \text{ outside}} \operatorname{Area}(Q(z) \cap C^+), \end{aligned}$$

and the number of boundary and outside $z \in \mathbf{Z}^2$ is at most the perimeter of the smallest aligned box containing C^+, which is $2b + 2(a+b) \leq 12a$. ∎

With Claim 3 we have

$$\begin{aligned}
&\left|\sum_{d=1}^{b}\mu(d)\left(|\frac{1}{d}C^+\cap \mathbf{Z}^2| - \frac{1}{d^2}\mathrm{Area}\,C^+\right)\right| \\
&\leq \sum_{d=1}^{b}\left||\frac{1}{d}C^+\cap \mathbf{Z}^2| - \frac{1}{d^2}\mathrm{Area}\,C^+\right| \\
&\leq \sum_{d=1}^{b}\frac{1}{d}(2b+2(a+b)) \leq 12a(1+\log b).
\end{aligned}$$

Thus

$$\left||C^+\cap \mathbf{P}| - \sum_{d=1}^{b}\frac{\mu(d)}{d^2}\mathrm{Area}\,C^+\right| \leq 12a(1+\log b).$$

Note that $|\sum_{d=1}^{b}\frac{\mu(d)}{d^2} - \frac{6}{\pi^2}| \leq \sum_{d=b+1}^{\infty}\frac{1}{d^2} < \frac{1}{b}$. Then

$$\begin{aligned}
\left||C\cap \mathbf{P}| - \frac{6}{\pi^2}\mathrm{Area}\,C\right| &\leq 2 + \frac{1}{b}\mathrm{Area}\,C + 24a(1+\log b) \\
&\leq \mathrm{Area}\,C\left(\frac{2}{2ab} + \frac{1}{b} + \frac{12(1+\log b)}{b}\right) \\
&\ll \mathrm{Area}\,C \cdot \frac{\log b}{b}.
\end{aligned}$$

This completes the proof of Lemma 1. ∎

4 Approximating $A(C)$

Lemma 2

$$\left|\sum_{u\in C^+\cap \mathbf{P}}\sum_{v\in C^+\cap \mathbf{P}}|\det(u,v)| - \left(\frac{6}{\pi^2}\right)^2\int_{C^+}\int_{C^+}|\det(x,y)|\,dx\,dy\right| \ll (\mathrm{Area}\,C^+)^3\frac{\log b}{b}.$$

Proof. It is very simple to see that

$$\det(u,v) = \int_{x\in Q(u)}\int_{y\in Q(v)}\det(x,y)\,dx\,dy.$$

Then this holds for $|\det(u,v)|$ with $|\det(x,y)|$ as the integrand if $\det(x,y)$ has constant sign on $Q(u)\times Q(v)$. This happens if $Q(u)$ and $Q(v)$ are separated by a line going through the origin. In case they are not separated, there are $\xi, \eta \in \mathbf{R}^2$ with $\|\xi\|_{\max}, \|\eta\|_{\max} \le \frac{1}{2}$ such that

$$0 = \det(u+\xi, v+\eta) = (u_1+\xi_1)(v_2+\eta_2) - (u_2+\xi_2)(v_1+\eta_1).$$

This shows that

$$\det(u,v) = u_1v_2 - u_2v_1 = -u_1\eta_2 - \xi_1 v_2 - \xi_1\eta_2 + u_2\eta_1 + v_1\xi_2 + \xi_2\eta_1$$

implying $|\det(u,v)| \le \frac{1}{2}(|u_1|+|u_2|+|v_1|+|v_2|+1) \le \frac{1}{2}(2a+2b+1) \le 4a$ if $u, v \in \mathbf{Z}^2 - \{0\}$.

Similarly, if $Q(u)$ and $Q(v)$ are not separated, then for all $(x,y) \in Q(u) \times Q(v)$ (with $u, v \in \mathbf{Z}^2 - \{0\}$ again)

$$|\det(x,y)| \le 8a.$$

We start estimating $\sum\sum |\det(u,v)|$ via

$$\begin{aligned}
\sum_{u\in C^+\cap \mathbf{P}} \sum_{v\in C^+\cap\mathbf{P}} |\det(u,v)| &= \sum_{u\in C^+\cap\mathbf{Z}^2} \sum_{v\in C^+\cap\mathbf{Z}^2} |\det(u,v)| \sum_{s|u_1, s|u_2} \mu(s) \sum_{t|v_1,t|v_2} \mu(t) \\
&= \sum_{s=1}^{\infty}\sum_{t=1}^{\infty} \mu(s)\mu(t) \sum_{\substack{u\in C^+\cap\mathbf{Z}^2\\ s|u_1,\, s|u_2}} \sum_{\substack{v\in C^+\cap\mathbf{Z}^2\\ t|v_1,\, t|v_2}} |\det(u,v)| \\
&= \sum_{s=1}^{b}\sum_{t=1}^{b} s\mu(s)t\mu(t) \sum_{u\in\frac{1}{s}C^+\cap\mathbf{Z}^2} \sum_{v\in\frac{1}{t}C^+\cap\mathbf{Z}^2} |\det(u,v)| \\
&= \sum(1) - \sum(2) + \sum(3)
\end{aligned}$$

where

$$\sum(1) = \sum_s\sum_t s\mu(s)t\mu(t) \sum_{u\in\frac{1}{s}C^+\cap\mathbf{Z}^2} \sum_{v\in\frac{1}{t}C^+\cap\mathbf{Z}^2} \int_{Q(u)}\int_{Q(v)} |\det(x,y)|\, dx\, dy,$$

$\sum(2)$ is the same as $\sum(1)$ but for non-separated $Q(u), Q(v)$, and

$$\sum(3) = \sum_s\sum_t s\mu(s)t\mu(t) \sum_{u\in\frac{1}{s}C^+\cap\mathbf{Z}^2} \sum_{v\in\frac{1}{t}C^+\cap\mathbf{Z}^2} |\det(u,v)|\, dx\, dy,$$

again for non-separated $Q(u), Q(v)$. For fixed s and t, the number, $N(s,t)$, of non-separated pairs $Q(u), Q(v)$ with $u \in \frac{1}{s}C^+ \cap \mathbf{Z}^2$ and $v \in \frac{1}{t}C^+ \cap \mathbf{Z}^2$ can be estimated generously via Claim 3:

$$\begin{aligned} N(s,t) &\le \left(\text{Area}\,\frac{1}{s}C^+ + \frac{10a}{s}\right)\left(\text{Area}\,\frac{1}{t}C^+ + \frac{10a}{t}\right) \\ &\ll (\text{Area}\,C^+)^2 \frac{1}{s^2t^2}. \end{aligned}$$

In $\sum(3)$, $|\det(u,v)| \ll \frac{a}{s} + \frac{a}{t}$, and in $\sum(2)$ the integrand $|\det(x,y)| \ll \frac{a}{s} + \frac{a}{t}$ as well. Consequently

$$\begin{aligned} |\textstyle\sum(2)|, |\sum(3)| &\ll \displaystyle\sum_{s=1}^{b}\sum_{t=1}^{b} st\,(\text{Area}\,C^+)^2\,\frac{1}{s^2t^2}\,(\frac{1}{s} + \frac{1}{t})\,a \\ &\ll (\text{Area}\,C^+)^3\,\frac{\log b}{b}. \end{aligned}$$

Now we turn to $\sum(1)$. Define $R(s) = \bigcup_{u \text{ boundary}}(Q(u) \setminus \frac{1}{s}C^+)$ and $T(s) = \bigcup_{u \text{ outside}}(Q(u) \cap \frac{1}{s}C^+)$. We have to integrate over

$$\left[\frac{1}{s}C^+ \cup R(s) \setminus T(s)\right] \times \left[\frac{1}{t}C^+ \cup R(t) \setminus T(t)\right].$$

The main term comes from integrating over $\frac{1}{s}C^+ \times \frac{1}{t}C^+$. We are going to estimate the remaining 8 integrals.

It is readily seen that for $x, y \in C^+$ $|\det(x,y)| \le \text{Area}\,C^+$. We need a slight strengthening of this (whose simple proof is omitted).

Claim 4 *When $x \in R(s) \cup T(s)$ and $y \in R(t) \cup T(t)$ and $s, t \le b$, then*

$$|\det(x,y)| \ll \frac{\text{Area}\,C^+}{st}.$$

Using Claim 4

$$\begin{aligned} \int_{\frac{1}{s}C^+}\int_{R(t)} |\det(x,y)|\,dx\,dy &\ll \frac{\text{Area}\,C^+}{st}\int_{\frac{1}{s}C^+} dx \int_{R(t)} dy \\ \ll \frac{\text{Area}\,C^+}{st}\,\frac{1}{s^2}\,\text{Area}\,C^+\,\frac{a}{t} &\ll \frac{(\text{Area}\,C^+)^3}{b}\,\frac{1}{s^3t^2}. \end{aligned}$$

So the sum of these terms multiplied by st is

$$\ll \sum_{s=1}^{b}\sum_{t=1}^{b} st\,\frac{(\operatorname{Area} C^+)^3}{b}\,\frac{1}{s^3t^2} \ll (\operatorname{Area} C^+)^3\,\frac{\log b}{b}.$$

The same applies to the integral over $\frac{1}{s}C^+ \times T(t)$ and when t and s are interchanged. Similarly

$$\int_{T(s)}\int_{T(t)} |\det(x,y)|\,dx\,dy \ll \frac{\operatorname{Area} C^+}{st}\,\frac{a}{s}\,\frac{a}{t} \ll \frac{(\operatorname{Area} C^+)^3}{b^2}\,\frac{1}{s^2}\,\frac{1}{t^2},$$

and the same works for the remaining three integrals. Thus we have

$$\left|\sum_{u\in C^+\cap\mathbf{P}}\ \sum_{v\in C^+\cap\mathbf{P}} |\det(u,v)| - \sum_{s=1}^{b}\sum_{t=1}^{b} s\mu(s)t\mu(t)\int_{\frac{1}{s}C^+}\int_{\frac{1}{t}C^+} |\det(x,y)|\,dx\,dy\right| \ll$$

$$\ll (\operatorname{Area} C^+)^3\,(\frac{\log b}{b} + \frac{\log^2 b}{b^2}) \ll (\operatorname{Area} C^+)^3\,\frac{\log b}{b}.$$

Here

$$\begin{aligned}
&\sum_{s=1}^{b}\sum_{t=1}^{b} s\mu(s)t\mu(t)\int_{\frac{1}{s}C^+}\int_{\frac{1}{t}C^+} |\det(x,y)|\,dx\,dy \\
&= \sum_{s=1}^{b}\sum_{t=1}^{b} \frac{\mu(s)}{s^2}\,\frac{\mu(t)}{t^2}\int_{C^+}\int_{C^+} |\det(x,y)|\,dx\,dy \\
&= \left(\frac{6}{\pi^2} - \sum_{s=b+1}^{\infty}\frac{\mu(s)}{s^2}\right)\left(\frac{6}{\pi^2} - \sum_{t=b+1}^{\infty}\frac{\mu(t)}{t^2}\right)\int_{C^+}\int_{C^+} |\det(x,y)|\,dx\,dy.
\end{aligned}$$

By Claim 4 $\int_{C^+}\int_{C^+} |\det(x,y)|\,dx\,dy \ll (\operatorname{Area} C^+)^3$. Thus

$$\left|\sum_{s=1}^{b}\sum_{t=1}^{b} s\mu(s)t\mu(t)\int_{\frac{1}{s}C^+}\int_{\frac{1}{t}C^+} |\det(x,y)|\,dx\,dy - \left(\frac{6}{\pi^2}\right)^2\int_{C^+}\int_{C^+} |\det(x,y)|\,dx\,dy\right| \ll (\operatorname{Area} C^+)^3\frac{1}{b}$$

finishing the proof of Lemma 2. ∎

5 Symmetrization

Theorem 2 *Assume* $K \in \mathcal{C}$*, and* $B \in \mathcal{C}$ *is a disk with* $\operatorname{Area} K = \operatorname{Area} B$*. Then*

$$\int_K \int_K |\det(x,y)|\, dx\, dy \geq \int_B \int_B |\det(x,y)|\, dx\, dy.$$

Equality holds iff K *is an ellipsoid.*

This theorem is known as Busemann's random simplex inequality [Bu53]. The proof goes by standard symmetrization (see e.g. [Bu53] or [Schn93]), so we only give a sketch. Let K^* be the symmetral of K with respect to a line l passing through O. We may assume, without loss of generality, that l is the x axis. To prove the theorem, it suffices to show the following:

Claim 5

$$\int_K \int_K |\det(x,y)|\, dx\, dy \geq \int_{K^*} \int_{K^*} |\det(x,y)|\, dx\, dy.$$

Proof. Write $x = (x_1, x_2)$ and $y = (y_1, y_2)$. Fix $x_1 \in \mathbf{R}$ and define $I(x_1) = \{x_2 \in \mathbf{R} : (x_1, x_2) \in K\}$ which is clearly an interval, $I(x_1) = [A, B]$, say. Then $\{x_2 \in \mathbf{R} : (x_1, x_2) \in K^*\} = [-(B-A)/2, (B-A)/2]$ is an interval again which we denote by $I^*(x_1)$. Similarly, for fixed y_1, $\{y_2 : (y_1, y_2) \in K\} = I(y_1)$ is an interval. Now the part of the integral $\int_K \int_K |\det(x,y)|\, dx\, dy$ with x_1 and y_1 fixed and the same part of $\int_{K^*} \int_{K^*} |\det(x,y)|\, dx\, dy$ can be compared easily. First, $\det(x,y) = x_1 y_2 - x_2 y_1$, and

$$\int_{I(x_1)} \int_{I(y_1)} |x_1 y_2 - x_2 y_1|\, dx_2\, dy_2 \leq \int_{I^*(x_1)} \int_{Y^*(y_1)} |x_1 y_2 - x_2 y_1|\, dx_2\, dy_2.$$

This is true since the integrand is the absolute value of a linear function on the rectangle $(x_2, y_2) \in I(x_1) \times I(y_1)$ which is clearly the smallest when the linear function is 0 at the center of the rectangle. This is the case exactly for the symmetral. ∎

If B is a disk centered at the origin, then

$$\int_B \int_B |\det(x,y)|\, dx\, dy = \frac{8}{9\pi^2} (\operatorname{Area} B)^3.$$

Thus Lemmas 1 and 2, and Theorem 2 give

Corollary 1 *If $C \in \mathcal{C}$ with $|C \cap \mathbf{P}| = n$, then*

$$A(C) \geq \left(\frac{1}{54} - D\,\frac{\log b}{b}\right) n^3$$

where D is a universal constant.

Remark 3 It turns out that one can take $D = 5000$ here when working with explicit constants instead of $\ll$.

6 The value of $\lim A(n)/n^3$

Set

$$f(z_1, \dots, z_b) = \frac{2}{3}\sum_{i=1}^{b}\left(2\frac{\phi^2(i)}{i} + \frac{\phi(i)}{i^2}(\sum_{j=1}^{i-1} j\phi(j))\right) z_i^3 + 2\sum_{i=1}^{b}\phi(i)z_i\left(\sum_{j=1}^{i-1}\frac{\phi(j)}{j}z_j^2\right).$$

We say that $z_1, \dots, z_b$ is *special* if the z_i are decreasing, $z_b \geq 0$ and they are convex, i.e., they satisfy $2z_i \geq z_{i-1} + z_{i+1}$ for $i = 2, \dots, b-1$. Define the following minimization problem:

$$\text{minimize } f(z_1, \dots, z_b) \text{ subject to } 4\sum_{i=1}^{b}\frac{\phi(i)}{i}z_i = 1, z_1, \dots, z_b \text{ is special.}$$

The minimum, which clearly exists, will be denoted by $M(b)$.

Theorem 3 $\lim A(n)/n^3 = \min_{b \leq 10^{10}} M(b)$.

Before the proof we give the following construction. Assume $b > 1$, and $x_1, \dots, x_b$ is special with $x_1 > 0$. Define

$$K = \text{conv}\{(\pm x_i, \pm i) \in \mathbf{R}^2 : i = 1, \dots, b\}.$$

K is a convex set which is symmetric with respect to both axes. Define the 2 by 2 diagonal matrix H_n, with diagonal elements λ_n and 1, that satisfies $|\mathbf{P} \cap H_nK| = n$. (There might be a little ambiguity in this definition since several, but at most $4b$, elements may appear on the boundary of H_nK. Resolve it by considering some of these points as belonging, while some others as not belonging, to H_nK.) Setting $K_n = H_nK$ we see that $|\mathbf{P} \cap K_n| = n$.

Define $I(e,f;i) = \{(x,i) \in \mathbf{Z}^2 : e \le x < f\}$. It is clear that the density of $\mathbf{P}$ on the line $y = i$ is $\phi(i)/i$, and there are exactly $\phi(i)$ primitive points on an interval of the form $I(si,(s+1)i;i)$. Write $I(i,s)$ for this interval. Now

$$|n - 4\sum_{i=1}^{b} \frac{\phi(i)}{i}\lambda_n x_i| \le 4\sum_{1}^{b} i \le 4b^2,$$

since, for each $i > 0$, the error term comes from the two subintervals $I_{left}(i)$ and $I_{right}(i)$ of $I(-\lambda_n x_i, \lambda_n x_i; i)$ that remain after deleting all $I(i,s)$ contained in it. A similar argument gives the following claim.

Claim 6

$$A(K_n) = \lambda_n^3 f(x_1,\dots,x_b) + O(\lambda_n^2),$$

where the implied constant depends only on b.

Proof. We only give a sketch. The basic observation is that $\sum |\det(u,v)|$ over $(u,v) \in (I(i,s)\cap\mathbf{P}) \times (I(j,t)\cap\mathbf{P})$ is the same as $\phi(i)\phi(j)/(ij)$ times the same sum over $(u,v) \in I(i,s) \times I(j,t)$, provided $\det(u,v)$ does not change sign on the box $I(i,s)\times I(j,t)$. The error terms come from two sources: First, from boxes where sign change occurs, but there are few of those, and there $|\det(u,v)| \le ij$. Secondly, when $u \in I_{left}(i)$ or $I_{right}(i)$, and similarly for v. But these intervals are short, with $|\det(u,v)|$ at most $4b\lambda_n$ (cf Claim 4). The statement follows by summing $\phi(i)\phi(j)|\det(u,v)|/(ij)$ over all $(u,v) \in (K_n \cap \mathbf{Z}^2) \times (K_n \cap \mathbf{Z}^2)$. ∎

So this construction satisfies $|\mathbf{P} \cap K_n| = n$ and

$$\lim \frac{A(K_n)}{n^3} = f(x_1,\dots,x_b)\left(4\sum_{i=1}^{b}\frac{\phi(i)}{i}x_i\right)^{-3}.$$

Proof of Theorem 1. Let C_n be the solution of the extremal problem Min(n) from Section 2. Let n_j be a sequence along which Min(n_j) tends to $M = \liminf \mathrm{Min}(n)$. If $b(C_n) \to \infty$ along a subsequence of n_j, then, according to Corollary 1

$$\lim \frac{A(C_n)}{n^3} = \frac{1}{54}$$

along the sequence n_j. But then this is true along the sequence n as well since, for the disk B_n containing n primitive points, $\lim A(B_n)/n^3 = 1/54$.

Assume now that $b(C_n)$ is bounded along n_j. Then we can choose a subsequence of n_j along which $b(C_n) = b$ for some fixed $b > 1$. To save writing, we denote this subsequence by n_j as well.

Let C_n^* denote the symmetral of C_n with respect to the y axis. The discrete analogue of the proof of Theorem 2 shows (we omit the straightforward details) that, along the sequence n_j,

$$\lim \frac{A(C_n)}{n^3} = \lim \frac{A(C_n^*)}{n^3}.$$

For $n = n_j$ and $i = 0, 1, \dots, b$ define $x_i(n) \geq 0$ by

$$[-x_i(n), x_i(n)] = C_n^* \cap \{(x, i) : x \in \mathbf{R}\}.$$

Then with our previous notation $x_0(n) = a(C_n^*) = a(C_n)$, and $a(C_n) \to \infty$ (along n_j) since $\operatorname{Area} C_n \leq 4a(C_n)b$. Choose now a subsequence of n_j along which $x_i(n)/a(C_n)$ is convergent, with limit x_i, for $i = 1, \dots, b$. As the sequence $x_1, \dots, x_b$ is special and $x_1 > 0$, the above construction works and gives the sequence K_n. It is obvious that, along the last subsequence of n_j, $\lim A(K_n)/n^3 = \lim A(C_n^*)/n^3 = M$. Then

$$\lim_{n \to \infty} \frac{A(K_n)}{n^3} = M$$

as well. ∎

Proof of Theorem 3. Given a special $x_1, \dots, x_b$ with $4 \sum_1^b \frac{\phi(i)}{i} x_i = 1$, we constructed a sequence of bodies K_n with $|K_n \cap \mathbf{P}| = n$ and $\lim A(K_n)/n^3 = f(x_1, \dots, x_b)$. So the value of the limit in Theorem 1 is less than $1/54$ if we find a single special sequence on which f is smaller than $1/54$. Here is such a sequence with $b = 15$:

$$\begin{aligned}
x_1 &= 0.03352589244, & x_2 &= 0.03335447314, & x_3 &= 0.03300806459,\\
x_4 &= 0.03251038169, & x_5 &= 0.03186074614, & x_6 &= 0.03104245531,\\
x_7 &= 0.03004944126, & x_8 &= 0.02886386937, & x_9 &= 0.02745127878,\\
x_{10} &= 0.02577867736, & x_{11} &= 0.02380388582, & x_{12} &= 0.02143223895,\\
x_{13} &= 0.01851220227, & x_{14} &= 0.01470243266, & x_{15} &= 0.008861427136,
\end{aligned}$$

giving

$$f(x_1, \dots, x_{15}) = 0.0185067386955\dots$$

which is smaller than $1/54 = 0.0185185185...$ by about 10^{-5}.

To see the bound $b \leq 10^{10}$, we use Corollary 1: if $b > 10^{10}$, and $C \in \mathcal{C}$ with $|C \cap \mathbf{P}| = n$, then

$$\frac{A(C)}{n^3} \geq \frac{1}{54} - 5000\frac{\log 10^{10}}{10^{10}} > 0.018507.$$

∎

Remark 4 The 10^{10} bound can be improved to about 10^7 by proving a stability version of Theorem 2: informally stated, this would say that if the left hand side of the inequality in Theorem 2 is smaller than $1+\varepsilon$ times its right hand side, then K can be sandwiched between two ellipsoids E and $(1+c\sqrt{\varepsilon})E$.

7 Remarks on computation

It seems hard to solve the minimization problems explicitly. We used the following heuristics. Let $x_1, \ldots, x_b$ be a solution to the problem. What can we expect about $x_1, \ldots, x_b$? According to Theorem 2, it is reasonable to assume that $(x_1, 1), \ldots, (x_b, b)$ are almost on the boundary of an ellipsoid. So let E_t be an ellipsoid whose half-axes are of length 1 and t with $b \leq t \leq b+1$, and define

$$w(i,t) = \sqrt{1 - \left(\frac{i}{t}\right)^2} \quad (0 \leq i \leq t).$$

Set $W = 4\sum_{i=1}^{b} \frac{\phi(i)}{i} w(i,t)$ and $z_i = w(i,t)/W$ for $i = 1, \ldots, b$. Then these $z_1, \ldots, z_b$ form a good approximation for the solution of the minimization problem. In fact, this method gives, with $t = 15.56$ (then $b = 15$) the points $z_1, \ldots, z_{15}$, that already satisfy $f(z_1, \ldots, z_{15}) < 1/54$. The even better solution $x_1, \ldots, x_{15}$ giving $f(x_1, \ldots, x_{15}) = 0.0185067\ldots$ was found near the previous $z_1, \ldots, z_{15}$ by solving the set of equations that constitute the necessary conditions for the extremum.

Using this heuristics we have checked the value of f near the ellipsoid E_t for $b = 1, \ldots, 100$ carefully (Figure 1) and for $b = 101, \ldots, 1000$ roughly (Figure 2). The computation suggests that the true limit of $A(n)/n^3$ is very close to the above value $0.0185067\ldots$. If this is the case then the minimizer P_n is very close, but not equal to, the ellipsoid with equation $x^2/A^2 + y^2/B^2 = 1$ where $A = 0.003573n^2$ and $B = 1.656n$. But even if the minimum is different, the shape of the minimizer P_n is oblong: it is $c_1 n$ wide $c_2 n^2$ long in its lattice width direction.

Figure 1

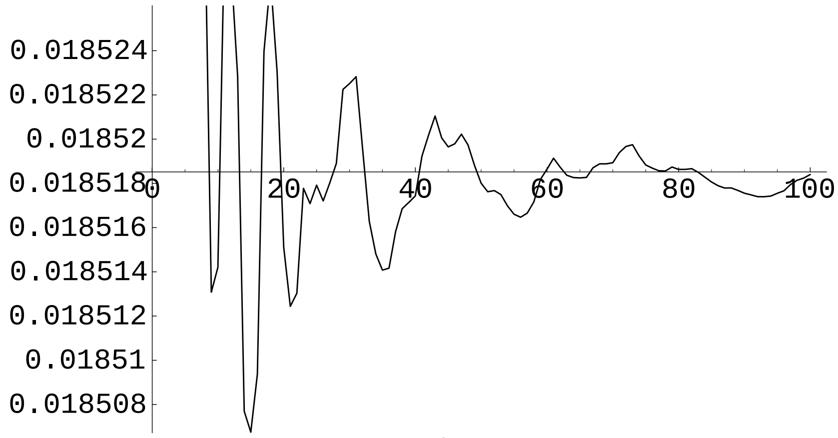
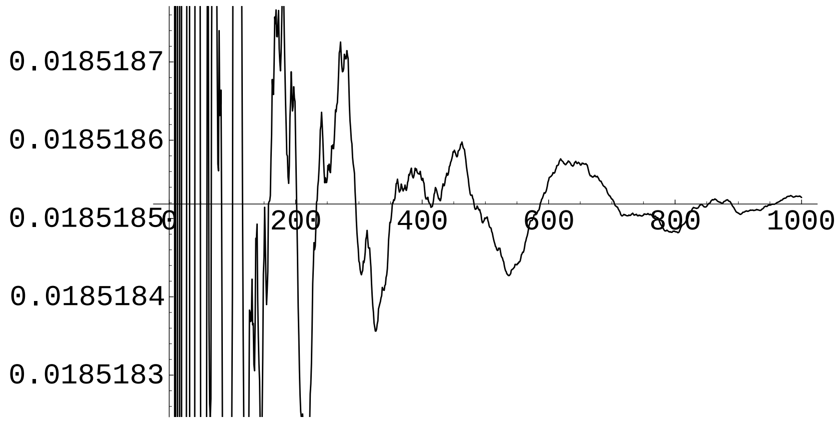


Figure 2



References

[An63] G. E. Andrews: *A lower bound for the volumes of strictly convex bodies with many boundary points.* Trans. AMS, **106** (1963), 270–279.

[Ar80] V. I. Arnol'd: *Statistics of integral convex polytopes.* (in Russian) Funk. Anal. Pril., **14** (1980), 1–3. English translation: Funct. Anal. Appl., **14** (1980), 79–84.

[BP92] I. Bárány, J. Pach: *On the number of convex lattice polygons.* Comb. Prob. Comp., **1** (1992), 295–302.

[Bu53] H. Busemann: *Volume in terms of concurrent cross-sections,* Pacific J. Math., **3** (1953), 1–12.

[HW79] G. H. Hardy, E. M. Wright: *An introduction to the theory of numbers,* Clarendon Press, Oxford (1979).

[KL88] R. Kanann, L. Lovász. *Covering minima and lattice-point-free convex bodies,* Annals of Math., **128** (1988), 577–602.

[KS84] S. B. Konyagin, K. A. Sevastyanov: *Estimation of the number of vertices of a convex integral polyhedron in terms of its volume.* (in Russian) Funk. Anal. Pril., **18** (1984), 13–15. English translation: Funct. Anal. Appl., **18** (1984), 11–13.

[Ra93] S. Rabinowitz: $O(n^3)$ *bounds for the area of a convex lattice* n*-gon.* Geombinatorics, **2** (1993), 85–88.

[RS63] A. Rényi, R. Sulanke: *Über die konvexe Hülle von* n *zufällig gewählten Punkten.* Z. Wahrsch. Verw. Geb., **2** (1963), 75-84.

[Schm85] W. Schmidt: *Integral points on surfaces and curves.* Monatshefte Math., **99** (1985), 45–82.

[Schn93] R. Schneider: *Convex bodies: the Brunn-Minkowski theory.* Encyclopedia Math. Appl. vol 44. Cambridge University Press, Cambridge, 1993.