Isoperimetric Inequalities and the Alexandrov Theorem

Emmanuel Tsukerman

May 13, 2013

Undergraduate Honors Thesis
Adviser: Simon Brendle
Stanford University

Contents

1 Introduction **2**

2 The Classical Isoperimetric Inequality **3**

2.1 Proof of the Classical Isoperimetric Inequality via Induction . 3

2.2 Proof of the Classical Isoperimetric Inequality via the Brunn-Minkowski Inequality 5

3 The Alexandrov Theorem **10**

3.1 Preliminaries 10

3.2 Proof of Alexandrov's Theorem 15

4 The Isoperimetric Inequality on the Sphere **23**

5 Discrete Isoperimetric Inequalities[1] **25**

5.1 Harper's Vertex Isoperimetric Theorem 26

5.2 Monotonicity of the optimal perimeter in isoperimetric problems on $\mathbb{Z}^k \times \mathbb{N}^d$ 30

5.3 On the Edge-Isoperimetric Problem on $(\mathbb{N}^2, \infty)$ 34

6 Acknowledgements **42**

[1]The contents of sections 5.2 and 5.3 have been submitted for publication [30].

Abstract

In this work, we introduce the reader to the broad field of isoperimetric inequalities. We begin by discussing several methods of solving the classical Isoperimetric Problem in $\mathbb{R}^n$. The first is an elegant proof by induction on the dimension n. The second proof proceeds by developing basics of Brunn-Minkowski theory and then applying the Brunn-Minkowski Inequality to obtain a proof of the Isoperimetric Inequality. In the third proof, we introduce Alexandrov's cutting-planes method and Alexandrov's Theorem classifying compact surfaces of constant mean curvature in $\mathbb{R}^n$. Combined with a short Calculus of Variations argument, we obtain one more proof of the Isoperimetric Inequality.

Next, we consider the Isoperimetric Problem in spaces different from $\mathbb{R}^n$. We prove that spherical caps are optimal sets in $\mathbb{S}^n$, thereby solving the Isoperimetric Problem on the sphere. We then consider discrete isoperimetric problems, which is a relatively young field dealing with isoperimetric problems on graphs. We introduce the technique of "compressions" and use it to prove Harper's Theorem, which solves the Isoperimetric problem on the graph of the hypercube.

Finally, we present original findings in the field of discrete isoperimetric inequalities. We prove general theorems for isoperimetric problems on lattices of the form $\mathbb{Z}^k \times \mathbb{N}^d$ which state that the perimeter of the optimal set is a monotonically increasing function of the volume under certain natural assumptions, such as local symmetry or being induced by an ℓ_p-norm. The proved monotonicity property is surprising considering that solutions are not always nested (and consequently standard techniques such as compressions do not apply). The monotonicity results of this note apply in particular to vertex- and edge-isoperimetric problems in the ℓ_p distances and can be used as a tool to elucidate properties of optimal sets. As an application, we consider the edge-isoperimetric inequality on the graph $\mathbb{N}^2$ in the ℓ_∞-distance. We show that there exist arbitrarily long consecutive values of the volume for which the minimum boundary is the same.

1 Introduction

Isoperimetric problems are classical objects of study in mathematics. In general, such problems ask for sets whose boundary is smallest for a given volume. A classic example dating back to ancient Greece is to determine the shape in the plane for which the perimeter is minimized subject to a volume constraint. The answer, as one would guess, is a circle, and was known already in Ancient Greece. However, the first mathematically rigorous proof of this fact was obtained only in the 19th century.

Though they are classical objects of study. Isoperimetric Problems are an active field of research in a number of areas, such as differential geometry, PDE, discrete geometry, probability and graph theory.

2 The Classical Isoperimetric Inequality

The isoperimetric inequality in $\mathbb{R}^2$ is an inequality involving the square of the circumference of a closed curve in the plane and the area of a plane region it encloses. Specifically, the isoperimetric inequality states that for the length L of a closed curve and the area A of the planar region that it encloses,

$$4\pi A^2 \leq L^2$$

and that equality holds if and only if the curve is a circle. In $\mathbb{R}^n$, the isoperimetric inequality is as follows: For any body K with n-dimensional volume $|K|$ and surface area $|\partial K|$,

$$\frac{|K|^{n-1}}{|\partial K|^n} \leq \frac{|B^n|^{n-1}}{|\partial B^n|^n}.$$

Remark 2.1 *It is interesting to note that for sufficiently smooth domains, the n-dimensional isoperimetric inequality is equivalent to the Sobolev inequality on R^n with optimal constant [15]:*

$$\left(\int_{\mathbb{R}^n} |u|^{\frac{n}{n-1}}\right)^{\frac{n-1}{n}} \leq n^{-1}|B^n|^{-1/n} \int_{\mathbb{R}^n} |\nabla u|$$

for all $u \in W^{1,1}(\mathbb{R}^n)$. Here $|B^n|$ denotes the volume of the unit ball.

2.1 Proof of the Classical Isoperimetric Inequality via Induction

We present an adaptation of the proof given in [2]. The proof proceeds by induction on the dimension n of $\mathbb{R}^n$.

Theorem 1 *(Classical Isoperimetric Inequality). For any body K with n-dimensional volume $|K|$ and surface area $|\partial K|$,*

$$\frac{|K|^{n-1}}{|\partial K|^{n}} \leq \frac{|B^n|^{n-1}}{|\partial B^n|^{n}}. \tag{1}$$

Proof. The base case of $n = 1$ is straightforward, since the boundary ∂K of a closed interval has volume $|\partial K| = 2$ (the counting measure for two boundary points of a line segment) and so does the boundary of the ball. Assume by induction that the inequality holds for dimension $n - 1$. Let $K \subset \mathbb{R}^n$ and ∂K its boundary. Without loss of generality, we may assume that $|K| = |B^n|$. Define $K_t = K \cap \{x_n = t\}$ and $\partial K_t = \partial K \cap \{x_n \cap t\}$. Finally, let $V(t) = |K_t|$ and $A(t) = |\partial K_t|$. Note that because each K_t is a parallel slice, $\int V(t)dt = |K|$. Then

$$V'(t) = \int_{\partial K_t} \frac{1}{\tan\theta},$$

where θ denotes the angle formed by the x_n-axis and the unit normal vector to ∂K. Using Jensen's inequality, we have

$$\int_{\partial K_t} \frac{1}{\sin\theta} = \int_{\partial K_t} \sqrt{1 + \frac{1}{\tan^2\theta}} \geq \sqrt{|\partial K_t|^2 + (\int_{\partial K_t} \frac{1}{\tan\theta})^2} \geq \sqrt{A(t)^2 + V'(t)^2}.$$

Therefore

$$|\partial K| = \int (\int_{\partial K_t} \frac{1}{\sin\theta})dt \geq \int \sqrt{A(t)^2 + V'(t)^2}dt.$$

We now define an auxiliary function $h(\tau)$. For each τ, the number $h(\tau) \in [-1, 1]$ is defined by the requirement that

$$\int_{-\infty}^{\tau} V(t)dt = \int_{-1}^{h(\tau)} |B^{n-1}|(1 - s^2)^{\frac{n-1}{2}} ds.$$

Informally, $h(\tau)$ matches the volume up to height $h(\tau)$ of the ball with the volume up to height τ of K. Since $\int V(t)dt = |B^n| = \int_{-1}^{1} |B^{n-1}|(1-s^2)^{\frac{n-1}{2}} ds$, the function $h(\tau)$ is well-defined. Differentiating, we get

$$V(t) = |B^{n-1}|(1 - h(t)^2)^{\frac{n-1}{2}} h'(t).$$

If we let $f(t) = |B^{n-1}|^{\frac{1}{n-1}} V(t)^{-\frac{1}{n-1}} \sqrt{1-h(t)^2}$, then $f(t)^{n-1} h'(t) = 1$. Using the arithmetic-geometric mean inequality, we conclude that

$$(n-1) f(t) + h'(t) \geq n.$$

On the other hand, it follows from the Cauchy-Schwarz inequality and the induction hypothesis that

$$\begin{aligned}
\sqrt{A(t)^2 + V'(t)^2} &\geq \sqrt{1-h(t)^2} A(t) - h(t) V'(t) \\
&= |B^{n-1}|^{-\frac{1}{n-1}} f(t) V(t)^{\frac{1}{n-1}} A(t) - h(t) V'(t) \\
&\geq (n-1) f(t) V(t) - h(t) V'(t) \\
&\geq n V(t) - h'(t) V(t) - h(t) V'(t).
\end{aligned}$$

This implies

$$\begin{aligned}
|\partial K| &\geq \int \sqrt{A(t)^2 + V'(t)^2} dt \\
&\geq \int (n V(t) - \frac{d}{dt}(h(t) V(t))) \\
&= n \int V(t) dt \\
&= n |B^n|.
\end{aligned}$$

This completes the proof. $\square$

2.2 Proof of the Classical Isoperimetric Inequality via the Brunn-Minkowski Inequality

In this section, we will show another proof of the classical isoperimetric inequality, this time utilizing Brunn-Minkowski theory. For a comprehensive introduction to Brunn-Minkowski theory, see [5].

Definition 2.2 *Let A and B be two arbitrary subsets of $\mathbb{R}^n$. The Minkowski sum of A and B is defined to be*

$$A \oplus B = \{a + b | a \in A, b \in B\}.$$

We will also be needing the definition of a parallelepiped with edges parallel to the coordinate axes:

Definition 2.3 *Let $I_1, I_2, ..., I_n$ be n bounded open intervals of the real line. A subset of $\mathbb{R}^n$ of the form*

$$A = I_1 \times I_2 \times ... \times I_n$$

is called a parallelepiped with edges parallel to the coordinate axes.

Parallelepipeds with edges parallel to the coordinate axes will play an important role in this section, as they will act as building blocks for more complicated sets and, as the following Lemma shows, are closed under the operation of taking a Minkowski sum.

Lemma 2.4 *Let $A = I_1 \times I_2 \times ... \times I_n$ and $B = J_1 \times J_2 \times ... \times J_n$ be two parallelepipeds in $\mathbb{R}^n$ with edges parallel to the coordinate axes. Then*

$$A \oplus B = (I_1 + J_1) \times (I_2 + J_2) \times ... \times (I_n + J_n).$$

If A and B are disjoint, there exists a hyperplane parallel to one of the coordinate hyperplanes separating A and B.

Proof. The first assertion is easy to see by considering the components of $a \in A$ and $b \in B$. Suppose that $A \cap B = \emptyset$. Then for some $k \in \{1, 2, ..., n\}$, $I_k \cap J_k$ is empty. Therefore there exists $a \in \mathbb{R}$ such that the intervals I_k and J_k are on different sides of a. It follows that the hyperplane $x_k = a$ separates A and B. □

Theorem 2 *(The Brunn-Minkowski inequality). Let A and B be two bounded open subsets of Euclidean space $\mathbb{R}^n$. Then*

$$|A|^{\frac{1}{n}} + |B|^{\frac{1}{n}} \leq |A \oplus B|^{\frac{1}{n}}.$$

Proof. The proof proceeds in three stages. The first stage is to prove the inequality for A and B parallelepipeds with edges parallel to the coordinates axes. Then we prove the inequality for A and B disjoint finite unions of bounded open parallelepipeds whose edges are parallel to the coordinate axes. Finally, we apply our previous results and some elementary results in the Lebesgue theory of integration to prove the general case.

Suppose first that $A = I_1 \times I_2 \times ... \times I_n$ and $B = J_1 \times J_2 \times ... \times J_n$, where I_k and J_k, $k = 1, 2, ..., n$ are bounded open intervals of $\mathbb{R}$ with lengths a_i and b_i. Then,

$$\frac{|A|^{\frac{1}{n}} + |B|^{\frac{1}{n}}}{|A \oplus B|^{\frac{1}{n}}} = \frac{(\prod_{k=1}^{n} a_k)^{\frac{1}{n}} + (\prod_{k=1}^{n} b_k)^{\frac{1}{n}}}{(\prod_{k=1}^{n}(a_i + b_i))^{\frac{1}{n}}} = (\prod_{k=1}^{n} \frac{a_k}{a_k + b_k})^{\frac{1}{n}} + (\prod_{k=1}^{n} \frac{b_k}{a_k + b_k})^{\frac{1}{n}}.$$

By the inequality of the arithmetic and geometric means,

$$\frac{|A|^{\frac{1}{n}} + |B|^{\frac{1}{n}}}{|A \oplus B|^{\frac{1}{n}}} \leq \frac{1}{n} \sum_{k=1}^{n} \frac{a_k}{a_k + b_k} + \frac{1}{n} \sum_{k=1}^{n} \frac{b_k}{a_k + b_k} = 1.$$

Now suppose that A and B are disjoint finite unions of bounded open parallelepipeds whose edges are parallel to the coordinate axes:

$$A = \cup_{k=1}^{n} A_k \text{ and } B = \cup_{k=1}^{m} B_k,$$

where A_k and B_k are as in the first case above. We proceed by induction on the total number $n + m$ of parallelepipeds. By the above, the base case of $n + m = 2$ has been established. Suppose that $n + m \geq 3$ and assume that the inequality holds for disjoint finite unions of bounded open parallelepipeds with edges parallel to the coordinate axes and total number less than $n + m$. Since $n + m \geq 3$, either $n > 1$ or $m > 1$. Without loss of generality, assume that $n > 1$. By lemma 2.4, there exists a hyperplane P parallel to the coordinate axes separating A_1 and A_2. Let P^+ and P^- be the two open half-spaces in which P divides $\mathbb{R}^n$ and let $A^+ = A \cap P^+$ and $A^- = A \cap P^-$. Then A^+ and $A^-$are also finite unions of parallelepipeds whose faces are parallel to the coordinate hyperplanes, namely

$$A^+ = \cup_{k=1}^{n^+} A_k^+ \text{ and } A^- = \cup_{k=1}^{n^-} A_k^-$$

where $n^+ < n$ and $n^- < n$, since P separates at least A_1 and A_2. We can find a hyperplane Q parallel to P such that

$$\frac{|A^+|}{|A|} = \frac{|B^+|}{|B|} \tag{2}$$

where we have used the same notation as before. Indeed, the fraction on the left is between 0 and 1, and if one takes Q on the left of all the parallelepipeds of B and displaces it to the right until it surpasses all of them, the fraction on the right side is a continuous function of the position of Q taking values from 0 to 1. By the intermediate value theorem we are guaranteed a position of Q for which equality holds.

Since $|A^+|+|A^-|=|A|$ and $|B^+|+|B^-|=|B|$, by 2,

$$1-\frac{|A^-|}{|A|}=1-\frac{|B^-|}{|B|} \implies \frac{|A^-|}{|A|}=\frac{|B^-|}{|B|}. \tag{3}$$

Moreover, B^+ and B^- are also disjoint finite unions of open parallelepipeds with faces parallel to the coordinate hyperplanes:

$$B^+=\cup_{k=1}^{m^+}B_k^+ \text{ and } B^-=\cup_{k=1}^{m^-}B_k^-,$$

where $m^+\leq m$ and $m^-\leq m$ because Q does not necessarily separate two parallelepipeds of B. Since $n^++m^+<n+m$ and $n^-+m^-<n+m$, we can apply the inductive hypothesis to A^+,B^+ and A^-,B^- to yield

$$|A^+\oplus B^+|\geq[|A^+|^{\frac{1}{n}}+|B^+|^{\frac{1}{n}}]^n, \tag{4}$$

$$|A^-\oplus B^-|\geq[|A^-|^{\frac{1}{n}}+|B^-|^{\frac{1}{n}}]^n. \tag{5}$$

Now, since $A^+\subset P^+$ and $B^+\subset Q^+$, $A^+\oplus B^+\subset P^+\oplus Q^+=(P\oplus Q)^+$ and similarly $A^-\oplus B^-\subset(P\oplus Q)^-$. Indeed, recall that P and Q are parallel so that if $P\doteq\{x\in\mathbb{R}^n : v\cdot x=\alpha_1\}$ then $Q\doteq\{x\in\mathbb{R}^n : v\cdot x=\alpha_2\}$ and therefore $P\oplus Q=\{x\in\mathbb{R}^n : v\cdot x=\alpha_1+\alpha_2\}$. Any element p of P^+ satisfies $v\cdot p>\alpha_1$ and $q\in Q^+$ satifies $v\cdot q>\alpha_2$ implying that $v\cdot(p+q)>\alpha_1+\alpha_2$ and $P^+\oplus Q^+\subset(P\oplus Q)^+$. The converse is similar.

Keeping in mind that $P\oplus Q$ is a hyperplane, we conclude that $A^+\oplus B^+$ and $A^-\oplus B^-$ are disjoint. Therefore, taking (5) into account,

$$|A\oplus B|\geq|A^+\oplus B^+|+|A^-\oplus B^-|$$

$$\geq[|A^+|^{\frac{1}{n}}+|B^+|^{\frac{1}{n}}]^n+[|A^-|^{\frac{1}{n}}+|B^-|^{\frac{1}{n}}]^n.$$

From (2) and (3),

$$|A\oplus B|\geq[|A^+|^{\frac{1}{n}}+|A^+|^{\frac{1}{n}}(\frac{|B|}{|A|})^{\frac{1}{n}}]^n+[|A^-|^{\frac{1}{n}}+|A^-|^{\frac{1}{n}}(\frac{|B|}{|A|})^{\frac{1}{n}}]^n$$

$$\geq|A^+|[1+(\frac{|B|}{|A|})^{\frac{1}{n}}]^n+|A^-|[1+(\frac{|B|}{|A|})^{\frac{1}{n}}]^n\geq|A|[1+(\frac{|B|}{|A|})^{\frac{1}{n}}]^n=[|A|^{\frac{1}{n}}+|B|^{\frac{1}{n}}]^n,$$

completing the induction.

Now consider the general case. Let A and B be any two bounded open sets in $\mathbb{R}^n$. The theory of Lebesgue integration tells us that there are two sequences A_n and B_n, $n \in \mathbb{N}$, of open sets in $\mathbb{R}^n$ of the type that we have considered in the second case, such that $A_n \subset A$ and $B_n \subset B$ and

$$\lim_{n\to\infty} |A_n| = |A| \text{ and } \lim_{n\to\infty} |B_n| = |B|.$$

Hence $A_n \oplus B_n \subset A \oplus B$ $\forall n \in \mathbb{N}$, and so

$$|A \oplus B|^{\frac{1}{n}} \geq |A_n \oplus B_n|^{\frac{1}{n}} \geq |A_n|^{\frac{1}{n}} + |B_n|^{\frac{1}{n}} \ \forall n \in \mathbb{N}.$$

Taking the limit as n tends to infinity completes the proof. $\square$

Definition 2.5 *Let K be a bounded domain with smooth boundary. The surface area of K is defined as the differential rate of volume increase as we add a small Euclidean ball to the body:*

$$|\partial K| = \lim_{\epsilon\to 0} \frac{|K \oplus \epsilon B_2^n| - |K|}{\epsilon}.$$

We now prove the classical isoperimetric inequality via the Brunn-Minkowski inequality [3].

Theorem 3 *(Classical Isoperimetric Inequality). For any body K with n-dimensional volume $|K|$ and surface area $|\partial K|$,*

$$\frac{|K|^{n-1}}{|\partial K|^n} \leq \frac{|B^n|^{n-1}}{|\partial B^n|^n}. \tag{6}$$

Proof. By the Brunn-Minkowski inequality,

$$|K \oplus \epsilon B_2^n| \geq [|K|^{\frac{1}{n}} + \epsilon |B_2^n|^{\frac{1}{n}}]^n$$

$$= |K|[1 + \epsilon(\frac{|B_2^n|}{|K|})^{\frac{1}{n}}]^n$$

$$\geq |K|[1 + n\epsilon(\frac{|B_2^n|}{|K|})^{\frac{1}{n}}]$$

where the second inequality is obtained by keeping the first two terms of the Maclaurin series of $(1+x)^n$.

By definition 2.5,

$$|\partial K| = \lim_{\epsilon\to 0}\frac{|K\oplus\epsilon B_2^n| - |K|}{\epsilon} \geq \lim_{\epsilon\to 0}\frac{|K| + n\epsilon|K|(\frac{|B_2^n|}{|K|})^{\frac{1}{n}} - |K|}{\epsilon} =$$

$$= n|K|^{\frac{n-1}{n}}|B_2^n|^{\frac{1}{n}}.$$

For an n-dimensional unit ball, we have $|\partial B_2^n| = n|B_2^n|$. Therefore

$$\frac{|\partial K|}{|\partial B_2^n|} \geq \frac{n|K|^{\frac{n-1}{n}}|B_2^n|^{\frac{1}{n}}}{n|B_2^n|} = (\frac{|K|}{|B_2^n|})^{\frac{n-1}{n}}.$$

□

3 The Alexandrov Theorem

In this section, we prove Alexandrov's Theorem, which states that a compact connected surface of constant mean curvature is necessarily a sphere. Combined with a short Calculus of Variations argument, this will yield an additional proof of the Isoperimetric Inequality.

3.1 Preliminaries

We begin by stating a few basic facts from differential geometry which will be used in the following sections. We refer the reader to [1] for proofs and further discussion.

Definition 3.1 *If S is a subset of $\mathbb{R}^3$, a differentiable vector field on S is a differentiable vector-valued function $v: S\to\mathbb{R}^3$.*

A *tangent vector field* on a surface S is a differentiable vector-valued function $v: S\to\mathbb{R}^3$ such that $v(x)\in TS_x$ for every $x\in S$, where TS_x is the tangent space to S at x. Similarly, a *normal vector field* to S is a differentiable vector-valued function $v: S\to\mathbb{R}^3$ such that $v(x)\in TS_x^{\perp}$ for every $x\in S$.

Theorem 4 *(Divergence theorem). Let S be a compact connected surface and Ω the inner domain determined by S. If $X : \bar{\Omega} \to \mathbb{R}^3$ is a differentiable vector field, then*

$$\int_\Omega div\ X = -\int_S \langle X, N\rangle,$$

where $N : S \to \mathbb{S}^2$ is the inner unit normal field.

Let S be an orientable surface and N a unit normal vector field on S. Since $|N(x)|^2 = 1$ for all $x \in S$, we have $N(S) \subset \mathbb{S}^2$. Consequently, we can view a unit normal field N on S as a differentiable map $N : S \to \mathbb{S}^2$ of the surface into the unit sphere $\mathbb{S}^2$. This map which takes each point on the surface to a unit vector orthogonal to the surface at this point will be called a *Gauss map* on S.

If $N : S \to \mathbb{S}^2$ is a Gauss map on S, then its differential at a point $p \in S$ is $dN_p : TS_p \to T\mathbb{S}^2_{N(p)}$. Since $N(p) \in TS_p^\perp$, and the tangent space to the sphere $\mathbb{S}^2$ at the point $N(p)$ is the orthogonal complement of the vector $N(p)$, we have $T\mathbb{S}^2_{N(p)} = (TS_p^\perp)^\perp = TS_p$, so that dN_p is an endomorphism of the tangent plane at p.

An endomorphism of a two-dimensional vector space has only two associated invariants: its determinant and its trace. We define the *Gauss curvature* K and *mean curvature* H to the surface at p by

$$K(p) = \det(dN_p)$$

$$H(p) = -\frac{1}{2}\text{trace}(dN_p)$$

for $p \in S$.

It is also the case that the differential dN_p of the Gauss map at each point is self-adjoint, so that it can be diagonalized and its eigenvalues are real. We let $k_1(p)$ and $k_2(p)$ be the two (real) eigenvalues of $-dN_p$ and we will suppose in standard fashion that they are ordered so that $k_1(p) \leq k_2(p)$. These will be called the *principal curvatures* of S at p. Since they are the roots of the characteristic polynomial, k_1 and k_2 satisfy

$$K = k_1 k_2,\ \ H = \frac{1}{2}(k_1 + k_2) \text{ and } k_i^2 - 2Hk_i + K = 0, i = 1, 2.$$

Moreover, since the discriminant of this second degree equation has to be non-negative (because two real roots exist), we have

$$K(p) \leq (H(p))^2, \forall p \in S.$$

Associated to these two eigenvalues are the corresponding eigenspaces, which are two orthogonal lines when $k_1(p) \neq k_2(p)$. We call these the *principal directions of S at the point p.*

We also recall the definitions of tubular and parallel surfaces. Define $F : S \times \mathbb{R} \to \mathbb{R}^3$ by $F(p,t) = p + tN(p)$. The following properties will be important for us:

1. If we restrict the domain of the second argument to $(-\epsilon, \epsilon)$ with ϵ sufficiently small, the map F is a diffeomorphism and the image is open. We call the image the *tubular neighbourhood* of S.

2. If $p \in S$ and $e_1, e_2 \in TS_p$ are principal directions of S at p, we have
$$(dF)_p(e_i) = \begin{pmatrix} 1+tN_x^1 & tN_y^1 & tN_z^1 \\ tN_x^2 & 1+tN_y^2 & tN_z^2 \\ tN_x^3 & tN_y^3 & 1+tN_z^3 \end{pmatrix} (e_i) =$$
$$= e_i + t(dN)_p(e_i) = (1 - tk_i(p))e_i.$$

3. For every $|t| < \epsilon$, the image of the map $F_t(p) = p + tN(p)$ is called the *parallel surface* of S at (oriented) distance t.

4. If $(p,t) \in S \times (0, \epsilon)$, then
$$\begin{aligned} |\mathrm{Jac}(F)|(p,t) &= \det((dF)_{(p,t)}(e_1, 0), (dF)_{(p,t)}(e_2, 0), (dF)_{(p,t)}(0,1)) \\ &= |(1 - tk_1(p))(1 - tk_2(p))| = |1 - 2tH(p) + t^2K(p)|. \end{aligned} \tag{7}$$

5. (Area of a parallel surface) If $t \neq 0$ is small enough, then the image of a compact surface S under the map $F_t(p) = F(p,t)$ is a diffeomorphism onto its image $S_t = F_t(S)$, the parallel surface at distance t. Applying the change of variables formula to the constant function 1 on S_t, we obtain the following expression for the area of the parallel surface at distance t:
$$A(S_t) = A(S) - 2t \int_S H(p)dp + t^2 \int_S K(p)dp. \tag{8}$$

6. (Volume enclosed by a parallel surface) Let $V_t = F(S \times (0, \epsilon)) \subset N_\epsilon(S)$ and let Ω_t be the inner domain determined by S_t. Then
$$|\Omega| - |\Omega_t| = |V_t(S)| = \int_{V_t(S)} 1.$$

Since $F : S \times (0, \epsilon) \to V_\epsilon(S)$ is a differeomorphism,

$$|\Omega| - |\Omega_t| = \int_{S\times(0,t)} |\mathrm{Jac}F|.$$

By the above calculation of the Jacobian, the integral is equal to

$$\int_0^t (\int_S (1 - 2tH(p) + t^2K(p))dp)dt = tA(S) - t^2 \int_S H + \frac{t^3}{3}\int_S K. \quad (9)$$

We also recall the definition of the divergence of a vector field V on a surface S. Let $\{e_1, e_2\}$ be an orthonormal basis of the principal directions of TS_p.

Definition 3.2 *The divergence of a vector field V on a surface S is the function $p \to \langle dV_p(e_1), e_1\rangle + \langle dV_p(e_2), e_2\rangle$ for $p \in S$.*

There is correspondingly a divergence theorem for surfaces:

Theorem 5 *(Divergence theorem for surfaces). Let S be a compact surface and let $V : S \to \mathbb{R}^3$ be a differentiable vector field defined on S. Then*

$$\int_S div\ V = -2\int_S \langle V, N\rangle H.$$

The following result will be used in the proof of Alexandrov's Theorem.

Theorem 6 *(Minkowski's formula). Let S be a compact surface and N its inner Gauss map. Then*

$$\int_S (1 + \langle p, N(p)\rangle H(p))dp = 0,$$

Proof. Let $V : S \to \mathbb{R}^3$ be a vector field given by $V(p) = p$ for all $p \in S$. Since

$$(dV)_p = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

by definition (3.2), div $V = 2$. Applying the divergence theorem for surfaces (Theorem 5), we see that

$$\int_S \mathrm{div}\ V + 2\langle V, N\rangle H = \int_S 2 + 2\langle p, N(p)\rangle H(p)dp = 0.$$

$\square$

We also introduce the so-called *area formula* - an integral formula that generalizes the change of variables formula for the Lebesgue integral when the transformation in question is no longer a diffeomorphism, but only a differentiable map.

Consider a bounded open subset $O \subset \mathbb{R}^3$, a differentiable map $\phi : \bar{O} \to \mathbb{R}^3$, and a function $f : O \to \mathbb{R}$. Let $N = \{x \in \bar{O} | (\mathrm{Jac}(\phi))(x) = 0\}$. We define

$$n(\phi, f) : \mathbb{R}^3 \setminus \phi(N) \to \mathbb{R}$$

via

$$n(\phi, f)(x) = \sum_{p \in \phi^{-1}(x)} f(p),$$

with the convention that the sum is zero when $x \notin \phi(\bar{O})$. The three-dimensional version of Sard's theorem tells us that $\phi(N)$ is a subset of measure zero in $\mathbb{R}^3$, so that $n(\phi, f)$ is defined almost everywhere on $\mathbb{R}^3$. For proofs for the following two theorems, see [1].

Theorem 7 *(Area formula). With the notation above, suppose that $f \cdot |Jac(\phi)|$ is integrable on O. Then $n(\phi, f)$ is integrable on $\mathbb{R}^3$ and*

$$\int_{\mathbb{R}^3} n(\phi, f) = \int_O f(x)|Jac(\phi)|(x)dx.$$

We will employ the following version of the area formula in our proof of Alexandrov's Theorem:

Theorem 8 *(Area formula for products). Let $\phi : \bar{R} \times [a, b] \to \mathbb{R}^3$ be a differentiable map, where R is a region of an orientable surface and $a < b$, and let f be a function on $R \times (a, b)$ such that $f \cdot |Jac(\phi)|$ is integrable on $R \times (a, b)$. Then the function $n(\phi, f)$ given by*

$$n(\phi, f)(x) = \sum_{(p,t) \in \phi^{-1}(x)} f(p, t)$$

is well-defined almost everywhere and is integrable on $\mathbb{R}^3$ with

$$\int_{\mathbb{R}^3} n(\phi, f) = \int_{R \times (a,b)} f(p, t)|Jac(\phi)|(p, t)dpdt.$$

One last preliminary:

Lemma 3.3 *Let S be a compact connected surface with positive mean curvature everywhere and let Ω be its inner domain. Define the map $F : S \times \mathbb{R} \to \mathbb{R}^3$ by $F(p,t) = p + tN(p)$ and define a set*

$$A = \{(p,t) \in S \times \mathbb{R} | 0 \leq t \leq \frac{1}{k_2(p)}\}.$$

Then $\Omega \subset F(A)$.

Proof. Let $q \in \Omega$. By compactness of S, the square of the distance function from q given by $f : S \to \mathbb{R}$, $f(p) = |p-q|^2$, attains a minimum on S. Its differential is

$$(df)_p(v) = 2\langle v, p-q \rangle.$$

Let r be a point at which the minimum is attained. Then

$$\langle v, r-q \rangle = 0.$$

Since v is in the tangent space TS_r of r, it follows that q must lie on the normal to S at r. That is,

$$q = r + tN(r)$$

for some $t \geq 0$. By considering the Hessian matrix at this point, we see that $1 - tk_2(p) \geq 0$. Therefore $t \leq \frac{1}{k_2(p)}$. $\Box$

3.2 Proof of Alexandrov's Theorem

Theorem 9 *(The Heintze-Karcher inequality). Let S be a compact surface whose (inner) mean curvature H is positive everywhere. Then*

$$V(\Omega) \leq \frac{1}{3} \int_S \frac{1}{H(p)} dp,$$

where Ω is the inner domain determined by S. Moreover, equality holds if and only if S is a sphere.

Proof. Let k_2 be the largest principal curvature of S corresponding to the inner normal and observe that it is positive everywhere because $k_2 \geq H$. Define A and F as in Lemma 3.3. By continuity of the function $\frac{1}{k_2(p)}$ and compactness of S, the function attains a maximum M. Let $a > M > 0$ and note that A is a compact set contained in $S \times [0, a)$. Applying Theorem 8 on $S \times (0, a)$ to the map F and the characteristic function χ_A we obtain the integral equality

$$\int_{\mathbb{R}^3} n(F, \chi_A) = \int_{S\times(0,a)} \chi_A |\mathrm{Jac}(F)|.$$

By definition,

$$n(F, \chi_A)(x) = \sum_{(p,t)\in F^{-1}(x)} \chi_A(p, t).$$

Let $x \in F(A)$. Then for some $(p, t) \in A$, we have $F(p, t) = x$. Therefore

$$n(F, \chi_A)(x) = \sum_{(p,t)\in F^{-1}(x)} \chi_A(p, t) \geq 1.$$

By Fubini's theorem, we have

$$|\Omega| \leq \int_S (\int_0^a \chi_A(p, t)|\mathrm{Jac}(F)|(p, t)dt)dp.$$

Substituting the absolute value of the Jacobian from the preliminaries, we have

$$|\Omega| \leq \int_S (\int_0^{\frac{1}{k_2(p)}} |1 - 2tH(p) + t^2K(p)|dt)dp.$$

But when $0 \leq t \leq \frac{1}{k_2(p)}$, the previous integrand is non-negative. Since $K(p) \leq (H(p))^2$,

$$1 - 2tH(p) + t^2K(p) \leq (1 - tH(p))^2$$

with equality only at umbilical points. Hence, since $\frac{1}{k_2(p)} \leq \frac{1}{H(p)}$ for each $o \in S$ and since the function $(1 - tH(p))^2$ is non-negative,

$$|\Omega| \leq \int_S (\int_0^{\frac{1}{H(p)}} (1 - tH(p))^2 dt)dp = \frac{1}{3}\int_S \frac{1}{H(p)} dp.$$

Since the only closed, bounded and totally umbilical surfaces in $\mathbb{R}^3$ are spheres, equality occurs if and only if S is a sphere. □

Theorem 10 *(Alexandrov's theorem). If a compact connected surface has constant mean curvature, then it is a sphere.*

Proof. Let S be a surface satisfying the conditions of the theorem. By the Heintze-Karcher inequality (Theorem 9) and the assumption that H is constant,

$$|\Omega| \leq \frac{A(S)}{3H}.$$

Moreover, if equality occurs, then S is a sphere. Applying the Divergence Theorem (Theorem 4) for the vector field $X(p) = p$, we have

$$\int_\Omega \mathrm{div} X = -\int_S \langle X, N \rangle$$

$$3|\Omega| = -\int_S \langle p, N(p) \rangle dp$$

since $\mathrm{div} X = 3$. By the Minkowski formula (Theorem 6),

$$\int_S (1 + \langle p, N(p) \rangle H(p)) dp = A(s) + H \int_S \langle p, N(p \rangle = A(S) - 3H|\Omega| = 0.$$

Therefore equality is attained, so that S must be a sphere. $\square$

We now present Alexandrov's Moving Plane Method, a beautiful technique which will provide us with another proof of Alexandrov's Theorem. We will be needing a few basic facts from PDE theory. We recall the form of a second-order linear PDE:

$$Lu = \sum_{i,j} a_{i,j}(x) u_{i,j}(x) + \sum_i b_i(x) u_i(x) + c(x) u(x) = f(x).$$

We will need that fact that if $a_{i,j}(x)$ is a positive definite matrix, then the PDE is *elliptic*, which, roughly speaking, implies that if the coefficients are sufficiently smooth, then the solution will also be smooth. We also recall the form of a second-order *quasilinear PDE*:

$$Lu = \sum_{i,j} a_{i,j}(x, u, \nabla u) u_{i,j}(x) + b(x, u, \nabla u) = 0,$$

Since a surface may locally be represented as a graph of a function, we will use $u(x, y) = z$ to reference this. In this local setting, the mean curvature can be written as

$$H = \text{div }(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}) = \sum_{i=1}^{n} D_i(\frac{D_i u}{\sqrt{1+|\nabla u|^2}}).$$

Note that

$$D_i(\frac{D_i u}{\sqrt{1+|\nabla u|^2}}) = (1+|\nabla u|^2)^{-\frac{3}{2}}(D_i^2 u(1+|\nabla u|^2) - D_i u(\sum_j D_j u D_j^2 u)),$$

so that an equation of the form $H = c$ for a constant c is a quasilinear PDE.

In the remainder of this discussion, we restrict to elliptic linear second order PDE unless otherwise stated. Elliptic PDE satisfy a number of nice properties.

Theorem 11 *(Strong Maximum Principle) Suppose that L is a linear elliptic differential operator of second order, and suppose that u is a nonnegative function on a domain Ω (with smooth boundary) satisfying $Lu \leq 0$. Then:*

1. *If $u(x_0) = 0$ at some point in the interior of Ω, then u vanishes identically in a neighborhood of x_0.*
2. *If $u(x_0) = 0$ at some point on the boundary of Ω, then either u vanishes identically in a neighborhood of x_0 or the normal derivative of u at x_0 is strictly negative.*

At first sight, it would appear that the Strong Maximum Principle is not applicable to our situation - after all, we have shown that the equation for constant mean curvature is quasilinear. Nevertheless, the Strong Maximum Principle applies to *differences of solutions* of quasilinear PDE. We illustrate the argument in the case relevant to us.

Consider a hypersurface which can be expressed as the graph of a function u. The mean curvature of this hypersurface can be written as

$$\sum_i D_i(\frac{D_i u}{\sqrt{1+|\nabla u|^2}}) = H$$

Suppose now that we have two surfaces with the same constant mean curvature H which can be expressed as graphs of two functions u and v. This implies

$$0 = \sum_i D_i(D_i(\frac{D_i u}{\sqrt{1+|\nabla u|^2}}) - (\frac{D_i v}{\sqrt{1+|\nabla v|^2}})) = \sum_{i,j} D_i(a_{ij} D_j(u-v)),$$

where

$$a_{ij} = \int_0^1 \frac{(1+|t\nabla u + (1-t)\nabla v|^2)\delta_{ij} - (tD_i u + (1-t)D_i v)(tD_j u + (1-t)D_j v)}{(1+|t\nabla u + (1-t)\nabla v|^2)^{\frac{3}{2}}} dt$$

Indeed,

$$\frac{D_i u}{\sqrt{1+|\nabla u|^2}} - \frac{D_i v}{\sqrt{1+|\nabla v|^2}} = \int_0^1 \frac{d}{dt}(\frac{tD_i u + (1-t)D_i v}{1+|t\nabla u + (1-t)\nabla v|^2})dt = \int_0^1 \sum_{i,j} a_{ij}(D_j u - D_j v)$$

To summarize, we have shown that $\phi := u - v$ is a solution of the equation $\sum_{i,j} D_i(a_{ij} D_j \phi) = 0$. The coefficient matrix a_{ij} is clearly positive denite, so the equation is elliptic (with variable coefficients).

In the general case, the Strong Maximum Principle implies

Corollary 3.4 *Suppose that Q is a quasilinear elliptic differential operator of second order, and suppose that u and v are two functions on a domain Ω with smooth boundary satisfying $u \geq v$ and $Qu = Qv$. Then:*

1. *If $u(x_0) = v(x_0)$ at some point in the interior of Ω, then $u = v$ in a neighborhood of x_0.*

2. *If $u(x_0) = v(x_0)$ at some point on the boundary of Ω, then either $u = v$ in a neighborhood of x_0 or the normal derivative of u at x_0 is strictly smaller than the normal derivative of v at x_0.*

In the following, we will employ the fact that if u is twice-differentiable and solves an elliptic PDE with sufficiently nice coefficients, then u is analytic.

To prove Alexandrov's Theorem using the Moving Plane Method, we prove the following fact, which guarantees that the surface S is a sphere: if for every direction e there is some plane Π_e perpendicular to this direction and such that S is symmetric with respect to Π_e, then S is a sphere.

Theorem 12 *(Alexandrov's Theorem). If a compact connected surface has constant mean curvature, then it is a sphere.*

Proof. We pick a direction e and a plane Π perpendicular to e which does not intersect S. Slide Π along e until it touches S for the first time. At this point, we slide the plane further by a distance ϵ and call the resulting plane Π_ϵ. The plane Π_ϵ intersects S. Reflect S in Π_ϵ and call the image S_ϵ. Now we increase ϵ until one of the following two things happens:

1. The side of S_ϵ that lies, say, to the left of Π_ϵ touches S at some point away from Π_ϵ.

2. At some point $p \in S \cap \Pi_\epsilon$, there is a tangent vector to S at p parallel to e.

In either case, we will show that there is some $p \in S$ and $r > 0$ such that $S_\epsilon \cap B_r(p) = S \cap B_r(p)$.

Case 1: Say the point of contact is p. Then near p we may represent both surfaces by functions u and v such that

$$u \leq v \in \partial\Omega$$

for some domain of the plane, $Lu = Lv \in \Omega$ and $u = v$ at some interior point. Then the second part of the Strong Maximum Principle implies that $S \cap B_r(p) = S_\epsilon \cap B_r(p)$ for some small r.

Case 2: Given the point p, we write both surfaces locally again as graphs of functions u and v, but now p corresponds to a point on the boundary of the patch Ω, and u and v satisfy

$$u \leq v \text{ on } \partial\Omega$$

$$Lu = Lv \text{ in } \Omega$$

$$u = v \text{ at some point of } \partial\Omega$$

$$u_\nu = v_\nu \text{ at some point.}$$

By the first part of the Strong Maximum Principle, we conclude that $u = v$ in all of Ω and thus $S \cap B_r(p') = S_\epsilon \cap B_r(p')$ for some small r and some p' near p.

Since u and v satisfy the elliptic PDE for constant mean curvature, u and v are analytic. We conclude that $S = S_\epsilon$, completing the proof. $\square$

Using Alexandrov's Theorem and basic methods from the Calculus of Variations, we may obtain an additional proof of the Classical Isoperimetric Inequality. The following argument was adopted from [4]. Consider a domain D in $\mathbb{R}^3$ bounded by a surface S.

Lemma 3.5 *If a surface S has minimum area among all surfaces bounding the same volume, then the mean curvature H of S is constant.*

To see why this is so, let $h : S \to \mathbb{R}$ be a smooth real-valued function on S, and let S_t denote the surface obtained by displacing each point of S by the vector thN, where N is the unit exterior normal field to S. If $A(t)$ is the area of S_t and $V(t)$ is the volume enclosed by S_t, then the formulae for the first variation are

$$A'(0) = -\int_S hHdA,$$

$$V'(0) = \int_S hdA.$$

Proof. (proof sketch) It is not hard to see that if there exists a function h for which $V'(0) = 0$ and $A'(0) \neq 0$, then applying a similarity transformation with factor $(V/V(t))^{\frac{1}{3}}$ transforms the surface S_t into a surface $\hat{S}_t$ bounding a volume V and having surface area $\hat{A}(t)$ which for all small values of t will be either strictly greather than or strictly less than A, depending on the sign of t. Thus, in order for S to have minimum area among all surfaces bounding the same volume V, it must be true that whenever $\int_S hdA = 0$, also $\int_S hHdA = 0$. This implies that H must be constant on S. For otherwise, if H had different values at two points then we could select h to be zero everywhere except in small neighbourhoods of those two points, and to have opposite sign at these two neighbourhoods in such a way that $\int_S hdA = 0$ but $\int_S hHdA > 0$. This variation would have the geometric effect of pushing in the neighbourhood having small mean curvature and pulling out the neighbourhood with large mean curvature. The net effect would be a "rounding out" of S which would preserve volume but decrease the surface area. $\square$

Combining Theorem 10 and Lemma 3.5, we obtain another proof of the Classical Isoperimetric Inequality in $\mathbb{R}^3$.

Corollary 3.6 *(Classical Isoperimetric Inequality) In $\mathbb{R}^3$, for a given volume, the surface of minimum area is the sphere.*

We now present an interesting application of the Isoperimetric Inequality to ovaloids.

Definition 3.7 *A compact connected surface with positive Gaussian curvature everywhere is called an ovaloid.*

Theorem 13 *For any ovaloid S,*

$$(\int_S H(p)dp)^2 \geq 4\pi A(S).$$

Proof. Consider an outer parallel surface S_t at distance $t \geq 0$. Since each such S_t is compact and connected, we can apply the isoperimetric inequality (Theorem 3) to S_t to obtain

$$A(S_t)^3 \geq 36\pi|\Omega_t|^2, \ \forall t \geq 0.$$

Using equations (8) and (9) for the area and volume of a parallel surface to obtain the following polynomial inequality:

$$(A(S)+2t\int_S H+t^2\int_S K)^3 \geq 36\pi(|\Omega|+tA(S)+t^2\int_S H+\frac{t^3}{3}\int_S K)^2, \ \forall t>0.$$

Using the Gauss-Bonnet Theorem, we replace $\int_S K$ with 4π. We then divide both sides by t^6 and set $s=\frac{1}{t}$ to yield

$$f(s)=(4\pi+2s\int_S H+s^2A(S))^3-36\pi(\frac{4\pi}{3}+s\int_S H+s^2A(S)+s^3|\Omega|)^2 \geq 0 \text{ for } s>0.$$

At $s=0$,

$$f(s)=(4\pi)^3-36\pi(\frac{4\pi}{3})^2=0.$$

In addition, $f'(s)$ evaluated at 0 is equal to

$$f'(0)=3(4\pi)^2(2\int_S H)-72\pi(\frac{4\pi}{3})(\int_S H)=0.$$

Therefore $f''(0) \geq 0$, which implies that

$$24\pi(\int_S H)^2-96\pi^2|S| \geq 0.$$

□

4 The Isoperimetric Inequality on the Sphere

The Isoperimetric Inequality on the sphere can be proved via the so-called two-point symmetrization, an operation which modifies a set so that it is more similar to a spherical cap. We will encounter similar techniques in section 5, when we discuss the use of "compression" to obtain Discrete Isoperimetric Inequalities. The effectiveness of such techniques is explained by their property that they allow us to compare the perimeter of a given set with that of the "candidate optimal set" directly without us having to first compute the perimeter of an arbitrary set - a formidable, if not intractable, task.

A sketch of the following proof is given in [7]. Let H be a hyperplane through zero in $\mathbb{R}^n$, $S_0 = S^{n-1} \cap H$ and let H^+ and H^- be the two open half spheres in the complement of H. Denote reflection in H by σ_H. Clearly, if $x, y \in H^+$ or H^- then $d(x, y) \leq d(x, \sigma_H(y))$. The *two-point symmetrization* $\tau_H K$ of a set $K \subset S^{n-1}$ is defined as follows (see Fig. 1):

$$\tau_H K = ((K \cap \sigma_H K) \cap H^-) \cup ((K \cup \sigma_H K) \cap H^+) \cup (K \cap S_0).$$

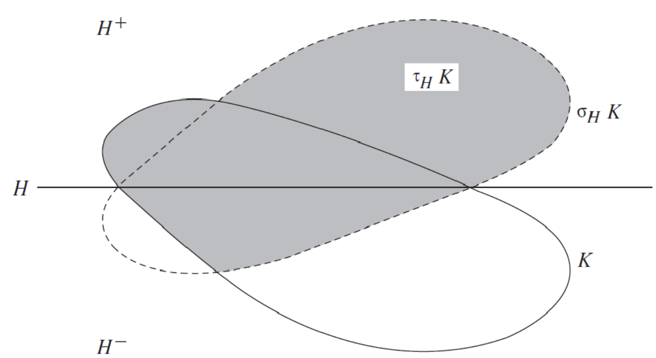

Figure 1: Two-point symmetrization [6].

Let A_ϵ be the ϵ neighbourhood of A, defined as $A_\epsilon = \{x : d(x, A) \leq \epsilon\}$.

Lemma 4.1 *For any subset $A \subset S^{n-1}$ and $\forall \epsilon > 0$,*

$$(\tau_H A)_\epsilon \subset \tau_H(A_\epsilon).$$

Proof. Let $x \in (\tau_H A)_\epsilon$. Then either $d(x, (A \cap \sigma_H A) \cap H^-) \leq \epsilon$, $d(x, (A \cup \sigma_H A) \cap H^+) \leq \epsilon$ or $d(x, A \cap S_0) \leq \epsilon$. In the first case, $d(x, (A \cap \sigma_H A) \cap H^-) \leq \epsilon$ implies that $x \in A_\epsilon \cap \sigma_H(A_\epsilon)$. Clearly $x \in H^+ \cup H^- \cup S_0$, so that

$$x \in \tau_H(A_\epsilon) = ((A_\epsilon \cap \sigma_H(A_\epsilon)) \cap H^-) \cup ((A_\epsilon \cup \sigma_H(A_\epsilon)) \cap H^+) \cup (A_\epsilon \cap S_0).$$

If $d(x, (A \cup \sigma_H A) \cap H^+) \leq \epsilon$, then $x \in A_\epsilon \cup \sigma_H(A_\epsilon)$. If $x \in H^+$ then we are done. If $x \in S_0$ then $x \in A_\epsilon \cap S_0$ since $S_0 \cap (A_\varepsilon \cup \sigma_H(A_\varepsilon)) = S_0 \cap A_\varepsilon$. Finally, if $x \in H^-$ there is a y such that $d(x, y) \leq \epsilon$ and $y \in (A \cup \sigma_H A) \cap H^+$. Since $y \in H^+$, $d(x, y) \leq \epsilon \implies d(x, \sigma_H(y)) \leq \epsilon$. Therefore $x \in A_\epsilon \cap \sigma_H(A_\epsilon) \cap H^-$.

Lastly, if $d(x, A \cap S_0) \leq \epsilon$ then there is some $y \in A \cap S_0$ such that $d(x, y) \leq \epsilon$. Since $y \in S_0$, $\sigma_H(y) = y$. Therefore $x \in A_\epsilon \cap \sigma_H(A_\epsilon)$. It follows that $x \in \tau_H(A_\epsilon)$. $\square$

Lemma 4.1 implies that volume satisfies the following relation under two-point symmetrization:

$$|(\tau_H A)_\epsilon| \leq |\tau_H(A_\epsilon)| \leq |A_\epsilon|.$$

To prove the isoperimetric inequality on the sphere we would like to apply the operation $A \to \tau_H A$ until we reach a set for which application of τ_H does not improve $|A_\epsilon|$ and prove that such a set must be a spherical cap. This method of proof will reappear in section 5.

Recall that the Hausdorff metric D is defined by

$$D(A, B) = \inf\{\epsilon | A \subset B_\epsilon \text{ and } B \subset A_\epsilon\},$$

where A and B are nonempty bounded closed sets.

Also, recall the following property of the Hausdorff metric.

Theorem 14 *Let (X, d) be a metric space and let $\mathcal{H}$ be the collection of all nonempty compact subsets of X. If X is compact in the metric d, then the space $\mathcal{H}$ is compact in the Hausdorff metric D.*

For proof, we refer the reader to [8].

Let $\mathcal{C}$ be the metric space of closed subsets of S^{n-1} with the Hausdorff metric. Fix $A \in \mathcal{C}$ and consider the set $\mathcal{B} \subset \mathcal{C}$ of all sets $B \in \mathcal{C}$ satisfying:

- $\forall \epsilon > 0$, $|B_\epsilon| \leq |A_\epsilon|$
- $|B| = |A|$.

To see that the set $\mathcal{B}$ is closed in $\mathcal{C}$, let B_k, $k = 1, 2, ...$, be a sequence of sets in $\mathcal{B}$ with limit Z. Then $\forall \delta > 0\ \exists k_0$ such that $\forall k > k_0$, $B_k \subset Z_\delta$ and $Z \subset (B_k)_\delta$. Therefore $Z_\epsilon \subset (B_k)_{\epsilon+\delta}$ so that $|Z_\epsilon| \leq |(B_k)_{\epsilon+\delta}| \leq |A_{\epsilon+\delta}|$. We send $\delta \to 0$, and by the dominated convergence theorem, $|A_{\epsilon+\delta}| \to |A_\epsilon|$. Hence $|Z_\epsilon| \leq |A_\epsilon|$. Similarly, $B_k \subset Z_\delta$ so that $|A| = |B_k| \leq |Z_\delta|$, and sending $\delta \to 0$ as before gives us $|Z| = |A|$. Thus $\mathcal{B}$ is closed in $\mathcal{C}$.

Now fix a point $x_0 \in S^{n-1}$ and let C be the closed spherical cap centered at x_0 with volume $|A|$. It is enough to prove that $C \in \mathcal{B}$. For any hyperplane H with $x_0 \notin H$ we denote by H^+ the open half sphere containings x_0. Consider the upper semi-continuous map $\phi : \mathcal{C} \to \mathbb{R}$ given by $\phi(B) = |B \cap C|$. Let B_k be a sequence converging to B. By compactness, ϕ attains a maximum on $\mathcal{B}$, say at M. We will show that $C \subset M$.

Assume by contradiction that this is not the case. Then $|M \setminus C| = |C \setminus M| > 0$. Let $x \in M \setminus C$ and $y \in C \setminus M$ be points of density of the respective sets and let H be a hyperplane perpendicular to segment xy and crossing it at its midpoint. Let $B_r(x) \subset H^-$ and $B_r(y) \subset H^+$ be small balls such that, say, $|B_r(x) \cap (M \setminus C)| > 0.99|B_r(x)|$ and $|B_r(y) \cap (C \setminus M)| > 0.99|B_r(y)|$. Applying τ_H to M, most of $B_r(x)$ will be transferred into $B_y(y)$ while no point of $C \cap M$ will be transferred to a point which is not in C. Thus $|(\tau_H M) \cap C| > |M \cap C|$. Since $\tau_H M$ also belongs to $\mathcal{B}$ we get a contradiction.

5 Discrete Isoperimetric Inequalities[2]

Another variant of the Isoperimetric Inequality is the Discrete Isoperimetric Inequality. In this setting, we work with a graph G, and after defining a suitable notion of perimeter and volume, we find an object for which the perimeter is the smallest of all other objects with the same volume. There are several notions of volume and perimeter used in literature, with choice depending on the particular problem and on its tractability given the choice of volume and perimeter.

One could consider the perimeter of a set of vertices A to be: the number of vertices neighbouring those of A, the number of vertices a distance less than t away, or even the number of edges leaving A. The volume of A can be, for instance, its cardinality, or the sum of the degrees of its vertices. Thus, discrete isoperimetric inequalities can vary significantly even if the underlying

[2]The contents of sections 5.2 and 5.3 have been submitted for publication [30].

graph is the same. The notion of perimeter we will initially use is the vertex neighbourhood:

$$\partial A = \{y \in V(G) : d(A, y) \leq 1\},$$

where $d(A, y)$ is the infimum of $d(x, y)$ over $x \in A$. In particular, note that $A \subset \partial A$.

Another unique aspect of Discrete Isoperimetric Inequalities is that the existence of a minimizer is easy to prove. In particular, no appeal to Geometric Measure Theory is necessary.

5.1 Harper's Vertex Isoperimetric Theorem

We adopt the proofs and images in this section from [9]. The main Theorem in this section, Harper's Theorem, solves the Isoperimetric problem on the graph of the hypercube Q_n defined as follows.

Let $X = \{1, 2, ..., n\}$. We will use the standard shorthand $[n]$ for $\{1, 2, ..., n\}$ and denote the power set of X by $\mathcal{P}(X)$. The vertices of Q_n and the edges are

$$V(Q_n) = \mathcal{P}(X)$$

$$\{x, y\} \in E(Q_n) \iff |x \triangle y| = 1.$$

In other words, the vertices are subsets of X, and they are connected if and only if there is some $i \in X$ such that $x \cup \{i\} = y$ or $y \cup \{i\} = x$.

Another way to represent the hypercube graph Q_n is to view its vertices as binary strings of length n of which two are connected whenever they differ by one bit. This setup is preferable in some situations but in our case we will find the above to be more convenient. Rather than writing, e.g., $\{1, 2, 3\}$, we will use the shorthand version 123. Also, for simplicity, we will write $A \subset Q_n$ instead of $A \subset V(Q_n)$, since no confusion ought to arise.

The Isoperimetric Problem on Q_n is as follows: given $0 \leq m \leq 2^n$, how to choose $A \subset Q_n$ with $|A| = m$ such that $|\partial A|$ is minimal?

To solve this problem, we define an order on the elements of $\mathcal{P}(X)$ called the *simplicial ordering*. Given $x, y \in \mathcal{P}(X)$, let x precede y, written $x < y$, if $|x| < |y|$ or $|x| = |y|$ and $\min(x \triangle y) \in x$. For example, on Q_3 the simplicial ordering is

$$\{\emptyset, 1, 2, 3, 12, 13, 23, 123\}.$$

Harper's Theorem states that initial segments of the simplicial ordering are best for $|\partial A|$:

Theorem 15 *(Harper's Theorem, 1966). Let $A \subset Q_n$ with $|A| = m$ and let $I \subset Q_n$ be the first m elements of Q_n in simplicial order. Then $|\partial A| \geq |\partial I|$. Moreover, if $|A| \geq \sum_{k=0}^{r} \binom{n}{k}$, then $|\partial A| \geq \sum_{k=0}^{r+1} \binom{n}{k}$.*

To prove this theorem, we use the idea of *compressions*, which allow us to avoid direct computation of $|\partial I|$ or $|\partial A|$. Compressing a set of vertices A means replacing A by a set τA such that $|\tau A| = |A|$, $|\partial(\tau A)| \leq |\partial A|$ and τA looks more like I than A does. We would like to keep compression until the resulting set B is easily seen to satisfy $|\partial B| \geq |\partial I|$.

Let $A \subset \mathcal{P}(X)$. The i-sections of A, $i \in \{1, ..., n\}$, are the sets on $X \setminus \{i\}$ given by

$$A_{i-} = \{x \in \mathcal{P}(X \setminus \{i\}) : x \in A\} \subset Q_{n-1}^{(i-1)}$$

$$A_{i+} = \{x \in \mathcal{P}(X \setminus \{i\}) : x \cup \{i\} \in A\} \subset Q_{n-1}^{(i+1)},$$

where $Q_{n-1}^{(i-1)}$, $Q_{n-1}^{(n+1)}$ are copies of Q_{n-1} labeled by sets of $\mathcal{P}(X \setminus \{i\})$. Geometrically, we separate the elements of A into those which lie on one half of the hypercube Q_n, which is a smaller hypercube having dimension $n-1$, and those that lie on the other half, another hypercube of dimension $n-1$. The simplicial ordering on $\mathcal{P}(X \setminus \{i\})$ is just the simplicial ordering on $\mathcal{P}(X)$ restricted to elements in $\mathcal{P}(X \setminus \{i\})$.

The i-compression of A is defined as the set system $C_i(A) \subset \mathcal{P}(X)$ given by its i-sections:

$C_i(A)_{i-} =$ the first $|A_{i-}|$ points in simplicial ordering in $\mathcal{P}(X \setminus \{i\})$,

$C_i(A)_{i+} =$ the first $|A_{i+}|$ points in simplicial ordering in $\mathcal{P}(X \setminus \{i\})$.

The actual set $C_i(A)$ can be obtained by taking the union of $C_i(A)_{i-}$ and $\{x \in \mathcal{P}(X) : x = y \cup \{i\}, y \in C_i(A)_{i+}\}$. We are now ready to prove Harper's Theorem.

Proof. We have

$$|C_i(A)| = |C_i(A)_{i-}| + |C_i(A)_{i+}| = |A_{i-}| + |A_{i+}| = |A|.$$

We would like to show that $|\partial C_i(A)| \leq |\partial A|$. For simplicity, let $C = C_i(A)$. It suffices to show that $|(\partial C)_{i-}| \leq |(\partial A)_{i-}|$ and $|(\partial C)_{i+}| \leq |(\partial A)_{i+}|$ because $|(\partial C)_{i-}| + |(\partial C)_{i+}| = |\partial C|$, and the same applies to A. We have

$$(\partial A)_{i-} = \partial(A_{i-}) \cup A_{i+}.$$

Indeed, let $x \in (\partial A)_{i-}$. Then $i \notin x$. If $x \in A$ then $x \in A_{i-}$. Otherwise, for some j, $x \cup \{j\} \in A$. If $j = i$, then $x \in A_{i+}$. If not, then $x \in \partial(A_{i-})$. Conversely, if $x \in A_{i+}$ then $x \cup \{i\} \in A$ and $i \notin x$. Therefore $x \in (\partial A)_{i-}$. If $x \in \partial(A_{i-})$, then for some j, $x \cup \{j\} \in A_{i-}$. So $i \notin x$, so that $x \in (\partial A)_{i-}$. It follows that the two sets are equal. We will also utilize this equality with A replaced with C.

Since $|C_{i-}| = |A_{i-}|$ and C_{i-} is an initial segment on $\mathcal{P}(X \setminus \{i\})$, one can induct on n to conclude that $|\partial(C_{i-})| \leq |\partial(A_{i-})|$. The base case $n = 1$ is trivial, so the induction does start. We also have $|C_{i+}| = |A_{i+}|$. If C_{i-} is an initial segment of simplicial order, then so is $\partial(C_{i-})$. To see this, note that if C_{i-} is exactly all x with $|x| \leq r$, then this is clearly true. Otherwise, suppose that C_{i-} is all x with $|x| \leq r$ and some set $\{y_1, y_2, ..., y_k\}$ with $|y_j| = r+1$. Proceed by induction on k. For $k = 1$, $y_1 = \{1, 2, ..., r+1\}$, so $y_1 \cup \{r+2, ..., n\} \in \partial(C_{i-})$, and clearly all $|x| \leq r+1$ are in $\partial(C_{i-})$, so $\partial(C_{i-})$ is in simplicial order. Assume the hypothesis for k, and consider the case of $k+1$. Then by the inductive hypothesis, $\partial(C_{i-} \setminus \{y_{k+1}\})$ is in simplicial order. The neighbours of y_{k+1} outside of $\partial(C_{i-} \setminus \{y_{k+1}\})$ are exactly the segment following $\partial(C_{i-} \setminus \{y_{k+1}\})$ because y_{k+1} is the element following y_k since C_{i-} is an initial segment. As a consequence, either $\partial(C_{i-}) \subset C_{i+}$ or $C_{i+} \subset \partial(C_{i-})$.

Since

$$(\partial C)_{i-} = \partial(C_{i-}) \cup C_{i+},$$

it follows that

$$|(\partial C)_{i-}| = \max\{|\partial(C_{i-})|, |C_{i+}|\}.$$

So either $|(\partial C)_{i-}| = |\partial(C_{i-})| \leq |\partial(A_{i-})| \leq |(\partial A)_{i-}|$ or $|(\partial C)_{i-}| = |C_{i+}| = |A_{i+}| \leq |(\partial A)_{i-}|$. Therefore $|(\partial C)_{i-}| \leq |(\partial A)_{i-}|$. A similar argument yields $|(\partial C)_{i+}| = |(\partial A)_{i+}|$, so that $|\partial C| \leq |\partial A|$.

Now define a sequence $\{A_k\}_{k=1,2,...}$ in the following way. If A_j is i-compressed for every i, then terminate the sequence at A_j. Otherwise, there is some i for which A_j is not i-compressed. Define $A_{j+1} = C_i(A_j)$, and continue in this manner inductively. This sequence must terminate because if the compression operator C_i moves A_j to A_{j+1}, then $A_{j+1} < A_j$ in simplicial order. This can only happen a finite number of times, so the sequence is guaranteed to terminate.

The last term of the sequence, call it B, satisfies

$$B = |A_1|,$$

$$|\partial B| \leq |\partial A_1|,$$

and is i-compressed for each i. Interestingly, it is not true that B is an initial segment. For example, $B = \{\emptyset, 1, 2, 12\} \subset \mathcal{P}([3])$ is i-compressed for $i = 1, 2, 3$, but is not an initial segment. The consequence of this observation is that additional work will be needed to prove the result.

Luckily, there is only one such B, as we will now show, so that direct computation shows that its perimeter is not minimal (Lemma 5.1). $\square$

We denote the sphere and ball in the hypercube by

$$X^{(r)} = \{x \in \mathcal{P}(X) : |x| = r\}$$

and

$$X^{(<r)} = \{x \in \mathcal{P}(X) : |x| < r\}.$$

The following seemingly technical, but actually quite elegant, lemma will allow us to conclude the proof of Harper's Theorem.

Lemma 5.1 *Let B be i-compressed for each i but not an initial segment. Then B is of the form*

$$B = \begin{cases} X^{(<\frac{n}{2})} \setminus \{\{(\frac{n+3}{2}), (\frac{n+5}{2}), ..., n\}\} \cup \{\{1, 2, ..., (\frac{n+1}{2})\}\} \\ X^{(<\frac{n}{2})} \cup \{x \in X^{(\frac{n}{2})} : 1 \in x\} \setminus \{\{1, (\frac{n}{2}) + 2, (\frac{n}{2}) + 3, ..., n\}\} \cup \{\{2, 3, ..., (\frac{n}{2}) + 1\}\}, \end{cases}$$

depending on whether n is odd or even (in that order).

Proof. Since B is not an initial segment, there exist elements $x, y \in \mathcal{P}(X)$ such that $x < y$ in simplicial order, $x \notin B$ and $y \in B$. We show that $x \cap y = \emptyset$ and $x^c \cap y^c = \emptyset$, so that by De Morgan's law, $x \cup y = [n]$ and consequently $x = y^c$.

Indeed, suppose that $i \in x \cap y$. Then $y \setminus \{i\} \in B_{i+}$. By assumption, $x < y$ in $\mathcal{P}(X)$, so $x \setminus \{i\} < y \setminus \{i\}$ in $\mathcal{P}(X \setminus \{i\})$, since the simplicial order on $\mathcal{P}(X \setminus \{i\})$ is the restriction of the simplicial order on $\mathcal{P}(X)$. Since B is i-compressed, B_{i+} is an initial segment. Therefore $x \setminus \{i\} \in B_{i+}$, implying that $x \in B$, which is a contradiction. To see that $x \cap y = \emptyset$ is even easier.

This implies that the only $z \notin B$ satisfying $z < y$ is x, since then $z = y^c$, and similarly that the only $z > x$ such that $z \in B$ is y. Therefore x and y are consecutive. We have $B = \{z \in \mathcal{P}(X) : z \leq y\} \setminus \{x\}$, where y is the immediate successor of x and $y^c = x$. If n is odd, $|y| = |x| + 1$ because $|x| + |y| = n$. Thus x is the last element having cardinality $\frac{n-1}{2}$.

If n is even, then $|x| = |y| = \frac{n}{2}$. Since x precedes y, and complements it, $x = \{1, (\frac{n}{2}) + 2, (\frac{n}{2}) + 3, ..., n\}$ and $y = \{2, 3, ..., (\frac{n}{2}) + 1\}$.
□

5.2 Monotonicity of the optimal perimeter in isoperimetric problems on $\mathbb{Z}^k \times \mathbb{N}^d$

We denote vertices of a graph G by $V(G)$, its edges by $E(G)$, and write $G = (V(G), E(G))$. For $x \in V(G)$, we set $N_{V(G)}(x) = \{y \in V(G) : \{x, y\} \in E(G)\}$ to be the vertex-neighbourhood of x and $N_{E(G)}(x) = \{e \in E(G) : e = \{x, y\} \in E(G)$ for some $y \in V(G)\}$ to be its edge neighbourhood. We write $N_G(x)$ for a neighbourhood of x in G when the distinction between vertex and edge does not matter.

We will be needing the following definitions:

Definition 5.2 *A graph $A = (\mathbb{Z}^{k+d}, E(A))$ is locally symmetric if for every $x \in V(A)$, the neighbourhood $N_A(x)$ of x is centrally symmetric about x.*
A graph G with $V(G) = \mathbb{Z}^k \times \mathbb{N}^d$ is locally symmetric if there exists a graph $A = (\mathbb{Z}^{k+d}, E(A))$ with centrally symmetric neighbourhoods such that for every $x \in V(G)$, the neighbourhood $N_G(x)$ of x in G is the intersection of the neighbourhood $N_A(x)$ of x with G:

$$N_G(x) = N_A(x) \cap G.$$

Definition 5.3 *A graph G is induced by a p-norm, $1 \leq p \leq \infty$, if there exists some constant c such that for every $x \in V(g)$, $N_V(x) = \{y \in V(G) : 0 < \|x - y\|_p \leq c\}$.*

Note that if a graph G with $V(G) = \mathbb{Z}^k \times \mathbb{N}^d$ is induced by a p-norm, then it is homogenous.

Recall that colexicographical ordering is defined by

$$(a_1, a_2, \ldots, a_n) < (b_1, b_2, \ldots, b_n) \iff (\exists m > 0)(\forall i > m)(a_i = b_i) \wedge (a_m < b_m)$$

Theorem 16 *Let G be a locally symmetric graph on $\mathbb{Z}^k$. The minimum edge-boundary is a monotonically increasing function of the volume.*

Proof. Let B be a set of cardinality $|A| + 1$ with optimal boundary. We find a point which can be removed without increasing the boundary. This implies that a set of cardinality $|A|$ has minimum boundary $|\partial A|$ less than or equal to $|\partial B|$. Such a point has the property that it has at least as many neighbours in $\mathbb{Z}^k \setminus B$ as it does in B. We derive a contradiction by assuming that such a point does not exist.

Let

$$R_x(y) = (2x_1 - y_1, 2x_2 - y_2, \ldots, 2x_k - y_k)$$

denote the reflection of point y in point x.

Let $H_m = \{x \in \mathbb{Z}^k : \|x\|_\infty = m\}$ and let $r = \max_r H_r \cap B \neq \emptyset$. Let $p \in H_r \cap B$ be greatest in colexicographical order. If p has no neighbour in B, then it can clearly be removed without increasing the boundary. So let y be a neighbour of p in B. We show that the reflection $R_p(y)$ of y in p is not in B. Since p is greater than y colexicographically, $p_i = y_i$ for each $i > m$ and $p_m > y_m$ for some $m > 0$. For each $i > m$, $2p_i - y_i = p_i$. Additionally,

$$2p_m - y_m > p_m,$$

so that $R_p(y) > p$ colexicographically. Therefore $R_p(y) \notin B$. $\square$

Theorem 17 *If G is a graph on $\mathbb{Z}^k \times \mathbb{N}^d$ induced by a p-norm with constant $c < 2$, then the minimum edge-boundary is a monotonically increasing function of the volume.*

Proof. Being induced by a p-norm, G is locally symmetric. Therefore the argument from Theorem 16 shows that for $b \in B$ of greatest colexicographical order, the reflection $R_b(y)$ of a neighbour $y \in B$ is outside of B. A new issue arises, that the image might land outside of the graph. For each index $i > k$ such that $2b_i - y_i < 0$, we apply a reflection in the hyperplane $x_i = 0$. The colexicographical order of the image can only increase, as its entries have increased. Moreover, each entry $2b_i - y_i < 0$ maps to $y_i - 2b_i > 0$, so this image is outside of B and inside the graph. Let z be the image of $R_b(y)$ under these reflections. It remains to see that z is a neighbour of b and that this map is injective, i.e., no two neighbours y and y' map to the same point.

To show that z is a neighbour of b, it suffices to see that for each i, $|z_i - b_i| \leq |y_i - b_i|$. This inequality is clearly true for the indices i for which no reflection in the axes occurs. So consider an index i for which $2b_i - y_i < 0$.

We consider two cases, depending on whether $y_i < 3b_i$ or $y_i \geq 3b_i$.
In the first case, $|z_i - b_i| = |y_i - 3b_i| = 3b_i - y_i$. On the other hand, $|y_i - b_i| = y_i - b_i$, so that

$$|z_i - b_i| \leq |y_i - b_i| \iff y_i \geq 2b_i.$$

In the second case, $|z_i - b_i| = |y_i - 3b_i| = y_i - 3b_i$, so that

$$|z_i - b_i| \leq |y_i - b_i| \iff b_i \geq 0.$$

Therefore $\|z - b\|_p \leq \|y - b\|_p$, so that z is a neighbour of b.

Next we show that if y, y' are neighbours of b then they do not map to the same point. Assume by contradiction that their images after reflections are the same. For each coordinate i, either

$$2b_i - y_i = 2b_i - y'_i$$

or

$$2b_i - y_i = y'_i - 2b_i.$$

Since $y \neq y'$, for at least one coordinate i, $y_i \neq y'_i$ and, consequently, $2b_i - y_i = y'_i - 2b_i \implies 4b_i = y_i + y'_i$. Considering this equation over the nonnegative integers with constraint $y \neq y'$ shows that $|y'_i - b_i| \geq 2$ or $|y_i - b_i| \geq 2$. It follows that one of y and y' is not a neighbour of b, a contradiction. $\square$

Finally, we note that optimal sets are not necessarily nested. Consider the graph $\mathbb{N}^n$ induced by the ∞-norm with $c = 1$. That is, the graph for which $x \sim y$ iff $\max_i |x_i - y_i| \leq 1$. The set of edges is $E(G) = \{\{x, y\} : x, y \in \mathbb{N}^n \text{ and } \|x - y\|_\infty = \max_i |x_i - y_i| = 1\}$. We consider the edge-isoperimetric problem on this graph.

Proposition 5.4 *The optimal sets of $\mathbb{N}^2$ are not nested.*

Proof. Figure 2 shows the uniquely determined up to reflection in $y = x$ sequence of nested optimal sets of $\mathbb{N}^2$. Figure 3 shows a set with $|A| = 11$ which has a smaller boundary than the optimal nested set with $|A| = 11$. $\square$

Nested optimal sets are crucial for the technique of compression and the fact that the optimal sets are not nested means that such an approach will not be possible. However, as we will show, the monotonicity of the optimal boundary established here is a useful tool for obtaining bounds, proving optimality and understanding properties of optimal sets.

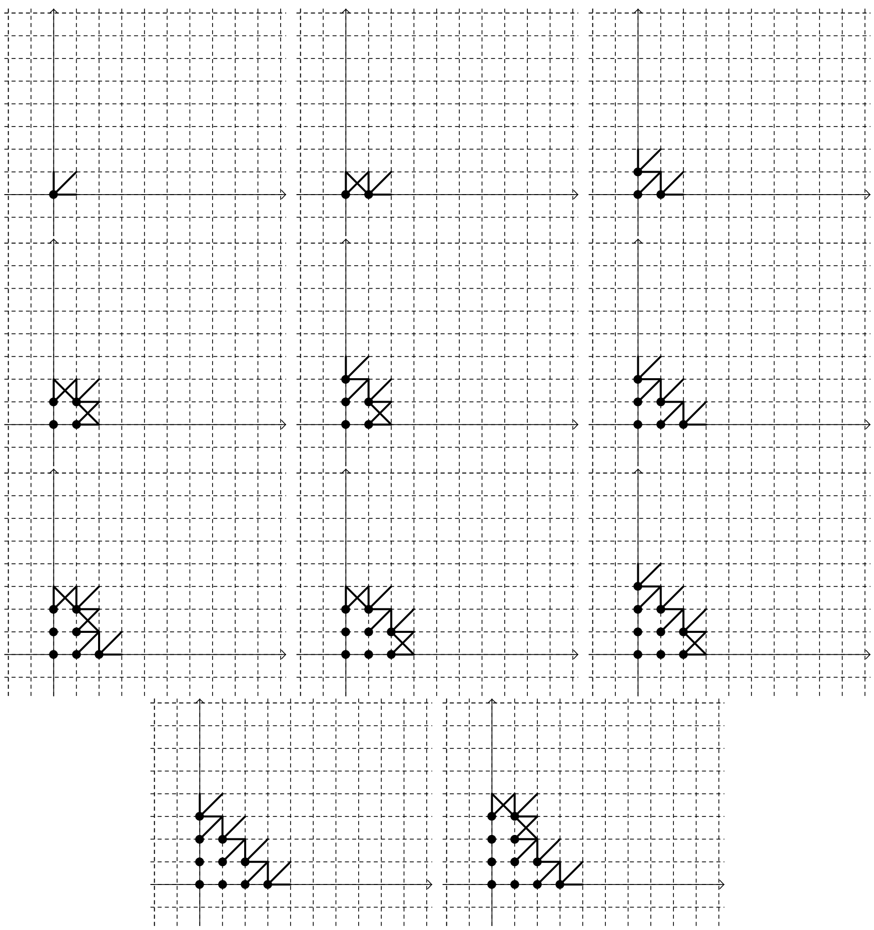

Figure 2: The unique (up to reflection in the line $y = x$) sequence of optimal nested sets for $1 \leq |A| \leq 11$. The optimal nested set with $|A| = 11$ has $|\partial A| = 17$.

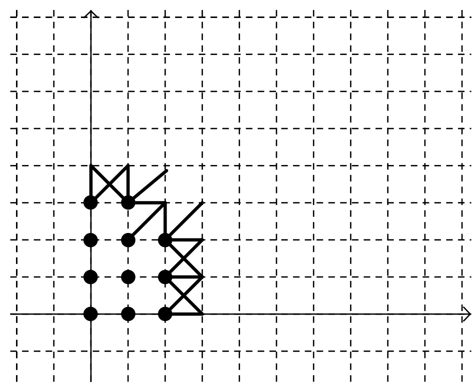

Figure 3: An optimal set having $|A| = 11$ has $|\partial A| = 16$. This set is better than the one of the same volume in Figure 2

5.3 On the Edge-Isoperimetric Problem on $(\mathbb{N}^2, \infty)$

Given a set in $\mathbb{N}^2$, it can be made connected without increasing its volume by translating the connected components towards the origin. Moreover, it can be made to touch both axes. Call A the resulting set. Let X be the point on the x-axis with greatest x-coordinate and let Y be the point on the y-axis of greatest y-coordinate. There is a connected subset C of A containing X and Y. We argue that all points bounded by C and the axes are in A if A is optimal.

Definition 5.5 *Given a connected subset C of $\mathbb{N}^2$ containing at least one point on the x-axis and one point on the y-axis, we say that a point z is bounded by C if point z is bounded by some piecewise linear curve formed by a subset of the edges of C and the axes.*

Lemma 5.6 *If A is an optimal set, then it contains all lattice points bounded by the axes and a connected subset C containing points X and Y.*

Proof. Assume otherwise. Consider the subset $B_C \subset \mathbb{N}^2$ of points lying on or below C and let $B = B_C \cup A$. Let $A_C = B_C \cap A$ be the points of A lying on or below C. By assumption, $|B_C| > |A_C|$, so that $|B| > |A|$. We show that $|\partial B_C| < |\partial A_C|$. This implies that $|\partial B| < |\partial A|$, contradicting Theorem 17.

The perimeter ∂B_C of B_C consists not only of all edges of C which lie outside of B_C, but also of the edges of the outermost layer of points of $B_C \setminus C$ (see

Figure 4) if such points exist. In particular, a point $a \in B_C \setminus C$ contributes whenever the following conditions are met. Point a is the vertex of a unit square $abcd$, the opposite corner c is in the complement of B_C, and the other two corners b and d are on C. In this case, edge ac is added to the perimeter. Suppose that such a point $a \in B_C \setminus A_C$ exists. Then the addition of a to A_C adds a diagonal edge ac to the count but takes away the two edges ab and ad. If a is part of more than one square, it is easy to see that the perimeter will still be strictly improved. If $B_C \setminus A_C$ contains some point which does not meet the conditions, then it does not contribute to the perimeter of B_C. Since there is at least one point of this form or of the prior form, filling in these points strictly improves the perimeter. This completes the proof. $\square$

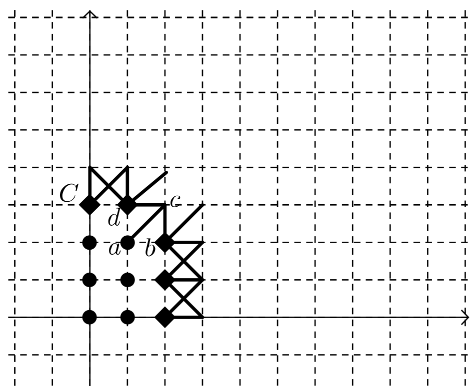

Figure 4: A connected subset C containing X and Y is indicated by diamonds. The perimeter of the set bounded by C receives a contribution not only from the points of C, but also from the layer adjacent to C, e.g., point a.

Conjecture 18 *For every volume, there is an optimal set A which consists of the points bounded by a connected subset $C \subset A$ touching the axes.*

We call sets consisting of the points bounded by a connected subset which touches the axes *bounded*. We will now investigate bounded sets. Though there might be sets which are better, we will still be able to learn much about the problem by considering bounded sets.

Definition 5.7 *Point $g \in \mathbb{Z}^n$ is a j-gap of a set A if $g \notin A$ and there exists some point $p \in A$ with $p_i = g_i$ $\forall i \neq j$ and $g_j < p_j$.*

We will say that a set has no gaps with it has no j-gaps for any $j = 1, 2, \ldots, n$.

Theorem 19 *For every volume, a bounded set can be modified to a bounded set with no gaps without increasing the boundary.*

Proof. We fill in the 1-gaps starting from the lowest gaps by adding points such as x in Figure 5, all the while decreasing the perimeter. By Theorem 17, the optimal perimeter is a monotonic increasing function of the volume, so we see that there can be no 1-gaps. The same argument applies to the 2-gaps. □

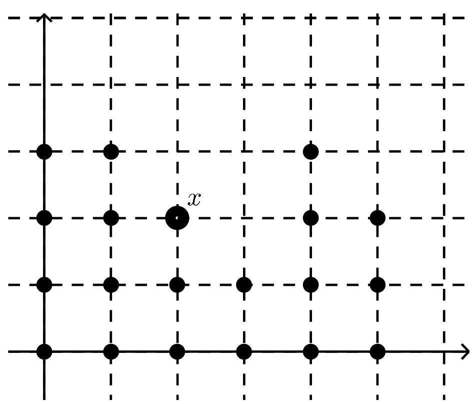

Figure 5: A bounded optimal set A cannot have gaps since filling in the gap with points such as x decreases the boundary.

Let $A_t = \{x \in A : x_1 = t\}$.

Lemma 5.8 *A bounded optimal set can be chosen to have at most one* $t \in \{1, 2, ..., k-1\}$ *for which* $|A_t| - |A_{t+1}| \geq 2$*, and this* t *can be chosen to be* $k-1$*.*

Proof. Let t_0 be the first t for which $|A_t| - |A_{t+1}| \geq 2$ and assume that $t_0 \neq k-1$. We can then transfer points from A_k to A_{t_0+1} until $|A_{t_0}| - |A_{t_0+1}| = 1$ or A_k has no more points left. Throughout, the perimeter does not increase because any point transferred shared at most 8 edges with other points, and after the transfer shares at least 8 edges. If $|A_{t_0}| - |A_{t_0+1}|$ is still greater than

2, then we can transfer points from column A_{k-1}, and continue in this way until either $|A_{t_0}| - |A_{t_0+1}| = 1$ or the number of columns has been reduced so that $t_0 + 1$ is the last column. □

Let us now view the heights of the columns as a function $h : \mathbb{N} \to \mathbb{N}$ given by $h(t) = |A_t|$. For brevity we will say that h is constant whenever we mean that h is constant on its support. We have already shown in Theorem 19 that a bounded optimal set exists which has h non-increasing. We will show that h can be made to take on a specific form. Before we do that, however, we note some special cases that are the only exceptions to the following Lemma. If $|A| = 1$, $|A| = 2$ or $|A| = 4$, then it is easy to see that for the optimal set h is constant. These are the only cases in which h will be constant, as we show next.

Lemma 5.9 *Without increasing the boundary, a bounded set A can be transformed into a bounded set B for which h_B is constant on $\{1, 2, ..., c-1\}$ and strictly decreasing on $\{c, c+1, ..., k\}$. Moreover, if $|A| \neq 1, 2, 4$, we can choose $c < k$.*

Proof. Let $S = \{x, x+1, ..., x+l\}$ be a maximal set of least x for which $h(x+i) < h(x+i-1)\ \forall i = 1, ..., l$. Assume further that $x + l < k$. We consider two cases: when $x+l = k-1$ and when $x+l < k-1$. In the former case, we take the top point from column A_k, which has at most 6 shared edges, and place it on top of column A_{k-1}, and now it has at least 6 shared edges. This reduces us to the second case. In this case, we take a point from column A_k, which can have at most 8 shared edges, and place it at the top of column A_{x+l}. Because $h(x+l) = h(x+l+1)$ and $h(x+l-1) = h(x+l)+1$, the new point has 8 shared edges, so the perimeter is not increased. We have now reduced S to $S \setminus \{x+l\}$, and we continue inductively. This shows that A can be transformed into a set B for which h_B is constant on $\{1, 2, ..., c-1\}$ and strictly decreasing on $\{c, c+1, ..., k\}$.

To see that we can choose $c < k$, assume that $|A| \neq 1, 2, 4$ and that h is constant. If $k = 1$, then we can take the top point of A_k and place it at A_2 without increasing the perimeter. If $k = 2$, we can take the point P at the top of A_k and place it at A_{k+1}. Since h_A was constant, P had at most 3 neighbours. Placing P at A_{k+1} guarantees 2 neighbours and 3 boundary edges, so the perimeter is not increased. For $k \geq 3$, we can take the point P from the top of A_k and place it at the top of A_1, and because $k \geq 3$, P will not have 2 neighbours and 3 boundary edges. □

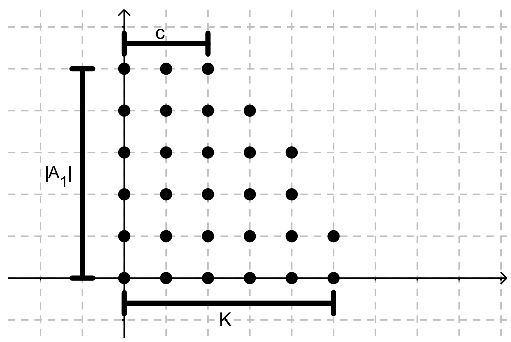

Figure 6: Schematic depicting the form of an optimal bounded set guaranteed by Lemmas 5.8 and 5.9. The first c columns have the same height $|A_1|$, the next $k-c-1$ columns have heights decreasing by 1 at each step, and the last column A_k has height $|A_k|$ less than $|A_{k-1}|$.

We can now find the perimeter of a general set subject to the conditions in the Lemmas above in terms of $|A_1|$, $|A_k|$, k and c. There are $|A_1|$ horizontal edges and k vertical edges. There are $\sum_{t=1}^{k-1}(|A_t| - |A_{t+1}| + 1) + |A_k| = |A_1| + k - 1$ edges parallel to $e_1 + e_2$, $|A_k| - 1 + \max\{|A_{k-1}| - |A_k| - 1, 0\} = |A_k| - 1 + |A_{k-1}| - |A_k| - 1 = |A_{k-1}| - 2$ edges in the $e_1 - e_2$ direction, and $\sum_{t=2}^{k} \delta_{|A_t|,|A_{t-1}|} = c - 1$ edges in the direction $e_2 - e_1$. Consequently, the perimeter is $2|A_1| + |A_{k-1}| + c + 2k - 4 = 2|A_1| + |A_1| - (k-1-c) + c + 2k - 4$, which is equal to

$$|\partial A| = 3|A_1| + 2c + k - 3. \tag{10}$$

We also know that $\sum_{t=1}^{k} |A_t| = |A|$. Therefore $|A| = c|A_1| + \sum_{i=1}^{k-c-1}(|A_1| - i) + |A_k|$, which simplifies to

$$|A| = (k-1)|A_1| + |A_k| - \frac{(k-c-1)(k-c)}{2}. \tag{11}$$

Combining Theorem 17 on the monotonicity of the perimeter and equation 10, we obtain:

Corollary 5.10 *Let A be a bounded optimal set with $|A_k| < |A_{k-1}| - 1$. Then the optimal perimeter of bounded sets of cardinality $|A| + 1$ is $|\partial A|$ and a bounded optimal set of cardinality $|A| + 1$ is given by $A \cup \{(k, |A_k| + 1)\}$.*

Example 5.11 *We show that simplices are not always optimal. Consider the simplex given by $h_i = 15 - i$, $i = 1, 2, ..., 14$. By increasing $|A_1|$ by 1 while* *preserving the shape from Lemma 5.9, we get a truncated simplex given by $h_i = 16 - i$, $i = 1, 2, ..., 10$. The perimeter of the simplex is* 56*, whereas that of the truncated simplex is* 55*.*

Given $|A_1|$, c and $|A|$, the set A is determined. Indeed, we know that

$$|A_i| = \begin{cases} |A_1| & i \leq c \\ |A_1| - i + c & c+1 \leq i \leq k-1 \\ |A| + \frac{(k-c-1)(k-c)}{2} - (k-1)|A_1| & i = k \end{cases}$$

Moreover, k is the unique positive integer such that

$$\sum_{i=1}^{k-c-1} (|A_1| - i) < |A| - c|A_1| \leq \sum_{i=1}^{k-c} (|A_1| - i).$$

Solving for k, we obtain

$$k = \lceil \frac{1}{2}(-1 + 2|A_1| - \sqrt{1 + 8(\binom{|A_1|}{2} - |A| + c|A_1|)} \rceil + c.$$

Therefore the problem is to minimize

$$|\partial A| = 3|A_1| + 3c + \lceil \frac{1}{2}(-1 + 2|A_1| - \sqrt{1 + 8(\binom{|A_1|}{2} - |A| + c|A_1|))} \rceil - 3$$

$$= 4|A_1| + 3c - 3 - \lfloor \frac{1}{2}(1 + \sqrt{1 + 8(\binom{|A_1|}{2} - |A| + c|A_1|))} \rfloor.$$

Lemma 5.12 *Any bounded set A can be transformed into one for which $|A_1| \geq c$ without increasing the boundary.*

Proof. Assume that $|A_1| < c$. We reflect A in the line $y = x$ to obtain a new set B which has $|A_1|$ columns. The first $|A_k|$ are of height k. Columns $|A_k|+1$ through $|A_{k-1}|$ are of height $k-1$. The remaining columns decreased in height by steps of 1, with column $|A_{k-1}|+1$ having height $k-2$ and column $|A_1|$ having height c. Since $c > |A_1| > |A_{k-1}| - |A_k|$, we can take points from column $|A_1|$ of height c and place them on top of columns $|A_k|+1,...,|A_{k-1}|-1$

without increasing the perimeter. The resulting set has the form of Lemma 5.9. The new parameters are $|\tilde{A}_1| = k$, $\tilde{k} = |A_1|$ and $\tilde{c} = |A_{k-1}| - 1 = |A_1| - k + c$. Then

$$c > |A_1| \implies \tilde{c} = |A_1| - k + c < 2c - k.$$

Since $k > c$, $|\tilde{A}_1| > \tilde{c}$. □

In order to obtain a lower bound for the sets considered, we relax our problem to a continuous one:

$$\begin{aligned} \underset{|A_1|, c \in \mathbb{R}}{\text{minimize}} \quad & 4|A_1| + 3c - 3 - \frac{1}{2}\Big(1 + \sqrt{1 + 8\Big(\binom{|A_1|}{2} - |A| + c|A_1|\Big)}\Big) \\ \text{subject to} \quad & 1 \leq |A_1| \leq |A|, \\ & \frac{|A| - \binom{|A_1|}{2}}{|A_1|} \leq c \leq |A_1|. \end{aligned}$$

For any $|A| \geq 2$, we can establish via a direct calculation that the minimum value of the objective is $\sqrt{\frac{7}{2}}\sqrt{8|A| - 1} - 2$ given by the unconstrained minimizer $|A_1| = \frac{3\sqrt{8|A|-1}}{2\sqrt{14}}$ and $c = \frac{1}{28}(14 + \sqrt{14}\sqrt{8|A| - 1})$, and the value is better than the value of the function on the boundary of the feasible region.

To obtain an upper bound, we will utilize the monotonicity of the perimeter from Theorem 17. Let $m \in \mathbb{N}$ and set $|A|^* = 7m^2$, $|A_1|^* = 3m$ and $c^* = m$. It is again a simple calculation to verify that these values give a feasible point. The function

$$g(|A_1|, c) = 4|A_1| + 3c - 2 - \frac{1}{2}\Big(1 + \sqrt{1 + 8\Big(\binom{|A_1|}{2} - |A| + c|A_1|\Big)}\Big)$$

is an upper bound for the perimeter. For $|A|^* = 7m^2$,

$$g(|A_1|^*, c^*) = 15m - \frac{1}{2}\sqrt{4m^2 - 12m + 1} - \frac{5}{2}$$

$$= \frac{15}{\sqrt{7}}\sqrt{|A|^*} - \frac{1}{2}\sqrt{\frac{4}{7}|A|^* - \frac{12}{\sqrt{7}}\sqrt{|A|^*} + 1}.$$

For a general $|A|$, we find m such that $7(m-1)^2 < |A| \leq 7m^2$. Then $|A|^* \leq |A| + 2\sqrt{7|A|} - 8$, so that

$$|\partial A| \leq \frac{15}{\sqrt{7}}\sqrt{|A|+2\sqrt{7|A|}-8} - \frac{1}{2}\sqrt{\frac{4}{7}(|A|+2\sqrt{7|A|}-8) - \frac{12}{\sqrt{7}}\sqrt{|A|+2\sqrt{7|A|}-8}+1}.$$

This complicated expression is asymptotically

$$\frac{15}{\sqrt{7}}\sqrt{|A|} - \frac{1}{\sqrt{7}}\sqrt{|A|} = \sqrt{\frac{7}{2}}\sqrt{8|A|}.$$

Note, however, that the upper bound is real if and only if $|A| \geq 36$.

Theorem 20 *Let the cardinality of a bounded optimal set A be $|A| \geq 36$. Then the perimeter $|\partial A|$ is bounded below by*

$$\lceil \sqrt{\frac{7}{2}}\sqrt{8|A|-1} - 2 \rceil$$

and above by

$$\lfloor \frac{15}{\sqrt{7}}\sqrt{|A|+2\sqrt{7|A|}-8} - \frac{1}{2}\sqrt{\frac{4}{7}(|A|+2\sqrt{7|A|}-8) - \frac{12}{\sqrt{7}}\sqrt{|A|+2\sqrt{7|A|}-8}+1} \rfloor.$$

Moreover, the difference between the upper and lower bound does not exceed the constant $\frac{35}{2}$.

Proof. The upper and lower bounds were shown above. Rather than consider the upper bound as it stands, we consider the slightly weaker but simpler upper bound $u(|A|)$ equal to

$$\frac{15}{\sqrt{7}}\sqrt{|A|+2\sqrt{7|A|}-8} - \frac{1}{2}\sqrt{\frac{4}{7}(|A|+2\sqrt{7|A|}-8) - \frac{12}{\sqrt{7}}\sqrt{|A|+2\sqrt{7|A|}-8}}$$

obtained by dropping the 1 inside of the square root. Then a calculation shows that the difference $d(|A|)$ between the upper bound $u(|A|)$ and lower bound $l(|A|) = \sqrt{\frac{7}{2}}\sqrt{8|A|-1} - 2$ has a non-vanishing derivative. Moreover, at $|A| = 39$, the first point at which this upper bound is defined, the derivative of the difference is positive, so that the difference is an increasing function. Taking the limit, we obtain the value $\frac{35}{2}$. For any $36 \leq |A| \leq 38$, a direct calculation shows that the difference is at most $\frac{35}{2}$. $\square$

Note that the growth of the boundary, even for bounded optimal sets, is slower than linear, though by Theorem 17 the perimeter is an increasing function. Therefore

Corollary 5.13 *There exist arbitrarily long consecutive values of the volume for which the minimum boundary is the same.*

6 Acknowledgements

I would like to thank my adviser Simon Brendle for the support, help and suggestions throughout the writing of this thesis.

References

[1] Montiel, Sebastian and Antonio Ros. Curves and Surfaces. 2nd ed. Providence: American Mathematical Society, 2009. Print.

[2] Lawlor, Gary. *An elementary isoperimetric proof in n-space,* 2010. Print.

[3] Kelner, Jonathan. 18.409 Topics in Theoretical Computer Science: An Algorithmist's Toolkit, Fall 2009. MIT OpenCourseWare. Web. 18 Nov. 2011.

[4] R. Osserman, *The isoperimetric inequality*, Bull. Amer. Math. Soc. **84** (1978), 1182– 1238.

[5] Schneider, Rolf. Convex Bodes: The Brunn-Minkowski Theory. Cambridge University Press, 1993. Print.

[6] G. Auburn and M. Fradelizi. *Two-Point Symmetrization and Convexity.*

[7] G. Schechtman. *Concentration, Results and Applications.*

[8] K. Barich. *Proving Completeness of the Hausdorff Induced Metric Space.*

[9] H. Liu. Discrete Isoperimetric Inequalities.

[10] M. Talagrand, *A new look at independence,* Ann. Probab. **24**, 1 (1996).

[11] P. Levy, *Problmes concrets d'analyse fonctionnelle* (French) 2d ed. (Gauthier-Villars, Paris, 1951).

[12] Béla Bollobás and Imre Leader. An isoperimetric inequality on the discrete torus. *SIAM J. Discrete Math.*, 3(1):32-37, 1990.

[13] Béla Bollobás and Imre Leader. Isoperimetric inequalities and fractional set systems. *J. Combin. Theory Ser. A,* 56(1):63-74, 1991.

[14] Thomas A. Carlson. The edge-isoperimetric problem for discrete tori. *Discrete Math.*, 254(1-3):33-49, 2002.

[15] V. Maz'ya. Lectures on Isoperimetric and Isocapacitary Inequalities in the Theory of Sobolev Spaces. *Contemp. Math.* **338**, 307-340 (2003).

[16] L. H. Harper. Optimal numberings and isoperimetric problems on graphs. *J. Combinatorial Theory*, 1:385-393, 1966.

[17] John H. Lindsey, II. Assignment of numbers to vertices. *Amer. Math. Monthly*, 71:508-516, 1964.

[18] Oliver Riordan. An ordering on the even discrete torus. *SIAM J. Discrete Math.*, 11(1):110-127 (electronic), 1998.

[19] Michel Ledoux. *The concentration of measure phenomenon, volume 89 of Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2001.

[20] Jiri Matousek. *Lectures on discrete geometry*, volume 212 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 2002.

[21] Vitali D. Milman and Gideon Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.

[22] Gilles Pisier. *The volume of convex bodies and Banach space geometry*, volume 94 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge, 1989.

[23] Noga Alon, Itai Benjamini, and Alan Stacey. Percolation on finite graphs and isoperimetric inequalities. *Ann. Probab.*, 32(3A):1727-1745, 2004.

[24] A. Barvinok and A. Samorodnitsky. The distance approach to approximate combinatorial counting. *Geom. Funct. Anal.*, 11(5):871-899, 2001.

[25] Béla Bollobás. Martingales, isoperimetric inequalities and random graphs. In *Combinatorics (Eger, 1987)*, volume 52 of *Colloq. Math. Soc. Janos Bolyai*, pages 113-139. North-Holland, Amsterdam, 1988.

[26] Béla Bollobás. Random graphs revisited. In *Probabilistic combinatorics and its applications (San Francisco, CA, 1991)*, volume 44 of *Proc. Sympos. Appl. Math.*, pages 81-97. Amer. Math. Soc., Providence, RI, 1991.

[27] Wang, D.-L. and Wang, P. (1977) Discrete isoperimetric problems. *SIAM J. Appl. Math.* 32 860-870.

[28] Béla Bollobás and Imre Leader. Edge-isoperimetric inequalities in the grid. *Combinatorica*, 11(4):299-314, 1991.

[29] E. Veomett, A. J. Radcliffe.Vertex Isoperimetric Inequalities for a Family of Graphs on $\mathbb{Z}^k$. The Electronic Journal of Combinatorics **19(2)** (2012), #P45.

[30] E. Tsukerman. Monotonicity of the optimal perimeter in isoperimetric problems on $\mathbb{Z}^k \times \mathbb{N}^d$. Preprint. arXiv: 1303.4139.