THE BIDIRECTIONAL BALLOT POLYTOPE

Steven J. Miller
Department of Mathematics and Statistics, Williams College, Williamstown, Massachusetts
sjm1@williams.edu; Steven.Miller.MC.96@aya.yale.edu

Carsten Peterson
Department of Mathematics, University of Michigan, Ann Arbor, Michigan
carstenp@umich.edu

Carsten Sprunger
Department of Mathematics, Stanford University, Palo Alto, California
csprun@stanford.edu

Roger Van Peski
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts
rvp@mit.edu

Received: , Revised: , Accepted: , Published:

Abstract

A bidirectional ballot sequence (BBS) is a finite binary sequence with the property that every prefix and suffix contains strictly more ones than zeros. BBS's were introduced by Zhao, and independently by Bosquet-Mélou and Ponty as $(1,1)$-culminating paths. Both sets of authors noted the difficulty in counting these objects, and to date research on bidirectional ballot sequences has been concerned with asymptotics. We introduce a continuous analogue of bidirectional ballot sequences which we call bidirectional gerrymanders, and show that the set of bidirectional gerrymanders form a convex polytope sitting inside the unit cube, which we refer to as the bidirectional ballot polytope. We prove that every $(2n-1)$-dimensional unit cube can be partitioned into $2n-1$ isometric copies of the $(2n-1)$-dimensional bidirectional ballot polytope. Furthermore, we show that the vertices of this polytope are all also vertices of the cube, and that the vertices are in bijection with BBS's. An immediate corollary is a geometric explanation of the result of Zhao and of Bosquet-Mélou and Ponty that the number of BBS's of length n is $\Theta(2^n/n)$.

1. Introduction

In [11], Zhao introduced a family of combinatorial objects called bidirectional ballot sequences, defined as follows.

Definition 1. A finite 0-1 sequence is a **bidirectional ballot sequence** (**BBS**) if every prefix and every suffix contains strictly more ones than zeros. Let B_n denote the number of bidirectional ballot sequences of length n.

Bidirectional ballot sequences have a natural interpretation in terms of lattice paths. Suppose we start at $(0, 0)$ and take a finite number of steps either of the form $(1, 1)$ or $(1, -1)$. We call such a path a **standard lattice path**. We define the length of the path to be the number of steps it contains. We define the height of a point in the lattice path to be its y-coordinate. Bidirectional ballot sequences of length n are in bijection with standard lattice paths of length n whose unique minimum height is attained at the first point in the path, and whose unique maximum height is attained at the last point in the path. The bijection is given by identifying the digit '0' in a BBS with a step of the form $(1, -1)$ and the digit '1' with a step of the form $(1, 1)$ (for an example of this, see Section 4).

From this perspective, bidirectional ballot sequences were independently introduced by Bosquet-Mélou and Ponty [2] as a special type of what they call culminating paths. In particular, an (a, b)-culminating path is a sequence of lattice points starting at $(0, 0)$ such that each step is of the form $(1, a)$ or $(1, -b)$ and such that the unique minimum height is achieved at the first point and the unique maximum height is achieved at the last point. Thus bidirectional ballot sequences are in bijection with $(1, 1)$-culminating paths. In [2] it is noted that $(1, 1)$-culminating paths had been used in [5] with connections to theoretical physics, and general (a, b)-culminating paths had been used in [1], [4], and [10] with connections to bioinformatics.

In both [11] and [2], it is noted that unlike other easy to define classes of lattice paths (e.g. Dyck paths), the enumeration of BBS's is tricky; there is no obvious recursive structure to such paths. Both authors focused on the asymptotics of B_n. In particular, [2] obtained a generating function in n for the number of (a, b)-culminating paths of length n with fixed height k (the generating function for the $(1, 1)$ case was found in [5]). Furthermore, they showed that $B_n \sim 2^n/4n$. Independently, [11] showed that $B_n = \Theta(2^n/n)$ and stated without detailed proof that $B_n \sim 2^n/4n$. Additionally in [11], the author conjectured an even finer asymptotic expression for B_n. This conjecture was later proved by Hackl, Heuberger, Prodinger and Wagner [6], who refined the asymptotic expression even further using techniques from analytic combinatorics.

The motivation for the study of culminating paths in [2] was the observation that such paths had been independently introduced and utilized in disparate contexts (theoretical physics and bioinformatics) as well as a general interest in understanding subfamilies of lattice paths. However, the motivation in [11], as well as our original motivation for studying BBS's, arises from additive combinatorics. Let $A \subset \mathbb{Z}$ be a finite set of integers. We define the sumset $A + A$ as those elements in $\mathbb{Z}$ expressible as $a + b$ with $a, b \in A$. Simi-

larly, the difference set $A - A$ is those elements expressible as $a - b$ with $a, b \in A$. We say that A is a **more sums than differences** (**MSTD**) set if $|A + A| > |A - A|$. Because of the commutativity of addition, one may intuitively expect that in general $|A - A| \geq |A + A|$. This intuition turns out to be correct in some contexts (see [7]), in particular if each element in $[n] := \{1, 2, \ldots, n\}$ is independently chosen to be in A with some probability $p(n)$ tending to zero). Let ρ_n be the proportion of subsets of $[n]$ which are MSTD. In [8], it was shown that $\rho_n > 2 \times 10^{-7}$ for $n \geq 15$, and in [12] it was shown that $\lim_{n\to\infty} \rho_n$ converges to a positive number; experimental data suggests this limit to be of order 10^{-4}. Thus, in this sense, a positive proportion of sets are MSTD. However, the techniques in [8] are probabilistic, and to date no known constant density family of MSTD subsets of $[n]$ as $n \to \infty$ is known.

The best density explicit construction of MSTD sets is due to Zhao in [11] using BBS's. Let B be a binary sequence of length n. We can associate to B the set $A \subseteq [n]$ defined as $A := \{i : B_i = 1\}$. For example if $B = 01101$, then $A = \{2, 3, 5\}$. Those subsets A of $[n]$ arising from BBS's have the property that $A + A = \{i : 2 \leq i \leq 2n\}$, which is to say that the sumset is as large as possible (similarly it turns out that the difference set is also as large as possible). Using this property, Zhao was able to translate those subsets of $[n]$ arising from BBS's and append extra elements to the fringes to obtain an MSTD set for each set arising from a BBS. From this, one immediately gets a density $\Theta(1/n)$ family of MSTD sets.

Motivated by the use of BBS's in additive combinatorics, in this paper we study the natural analgoue of BBS's in a continuous setting, which we call **bidirectional gerrymanders**; in the related paper [9], we use similar ideas as in this paper to study the analogue of MSTD sets in a continuous setting.

We first set some notation and then describe our main results. Let $\mathbb{I}_n$ denote the set of all subsets of $\mathbb{R}$ consisting of exactly n disjoint open intervals such that the leftmost interval starts at 0. Suppose $\mathcal{A} \in \mathbb{I}_n$. If we translate $\mathcal{A}$, then the sumset and difference set merely translate as well. Thus, when studying additive behavior, we do not lose any generality by restricting our attention to collections of intervals such that the leftmost interval starts at zero. We can topologize $\mathbb{I}_n$ by identifying it with $\mathbb{R}_{\geq 0}^{2n-1}$, the non-negative orthant: let $\mathcal{A} = I_1 \cup I_2 \cup \cdots \cup I_n \in \mathbb{I}_n$ with I_i to the left of I_j for $i < j$. Suppose $I_j = (a_j, b_j)$. We then identify $\mathcal{A}$ with the vector $v_{\mathcal{A}} = [b_1 - a_1, a_2 - b_1, b_2 - a_2, a_3 - b_2, \ldots, b_n - a_n]$. Thus the first entry is the length of the first interval, the second entry is the size of the gap between the first and second intervals, the third entry is the length of the second interval, etc. We shall find it convenient to restrict our attention to the following set: let $\mathbb{J}_n \subset \mathbb{I}_n$ be the set of collections of n non-overlapping intervals such that the leftmost interval starts at zero, the length of each interval is between 0 and 1, and the gap between adjacent intervals is between 0 and 1 (if we scale $\mathcal{A} \in \mathbb{I}_n$ by $\alpha \neq 0$, then the sumset and difference set scale by α as well, so $\alpha\mathcal{A}$ has the same essential additive behavior as $\mathcal{A}$; note that up to scaling, every element of $\mathbb{I}_n$ is an element of $\mathbb{J}_n$). We can topologize $\mathbb{J}_n$ by identifying it

with $C_{2n-1} = [0,1]^{2n-1}$, the $2n-1$ dimensional unit cube[1]. For other ways to topologize $\mathbb{I}_n$ and related spaces, see [9].

The bidirectional gerrymanders in $\mathbb{J}_n$ form a convex, compact polytope contained in C_{2n-1} which we call the **bidirectional ballot polytope**, $\mathcal{P}_n$. This polytope has a number of extraordinary combinatorial features. In Section 2 we formally define this polytope and show that C_{2n-1} can be partitioned into $2n-1$ disjoint isometric copies of $\mathcal{P}_n$, which in particular shows that the volume of $\mathcal{P}_n$ is $1/(2n-1)$. In Section 3 we show that the vertices of $\mathcal{P}_n$ are vertices of C_{2n-1}. Finally in Section 4 we show that the vertices of $\mathcal{P}_n$ are in bijection with B_{2n+3}, and that a particular subset of the vertices are in bijection with B_{2n-1}. From this we are able to immediately rederive geometrically that $|B_n| = \Theta(2^n/n)$, i.e., there are positive constants α and β such that for all n sufficiently large we have $\alpha 2^n/n \leq |B_n| \leq \beta 2^n/n$.

2. The Bidirectional Ballot Cone and Polytope

We first set some notation. Let $m = 2n-1$ for some $n \in \mathbb{N}$.

Definition 2. Let the set of **left ballot vectors**, L_n, and the set of **right ballot vectors**, R_n, be the following sets of vectors in $\mathbb{R}^m$:

$$L_n := \{[1,-1,0,\ldots,0],[1,-1,1,-1,0,\ldots,0],\ldots,[1,-1,\ldots,1,-1,0]\}, \quad (1)$$

$$R_n := \{[0,\ldots,0,-1,1],[0,\ldots,0,-1,1,-1,1],\ldots,[0,-1,1,\ldots,-1,1]\}. \quad (2)$$

We define V_n, the set of **ballot vectors**, as $V_n = L_n \cup R_n$.

Definition 3. The **bidirectional ballot cone**, $\mathcal{B}_n$, is the set of $x \in \mathbb{R}^m$ such that $x \cdot w \geq 0$ for all $w \in V_n$. When the value of n is obvious, we simply refer to it as $\mathcal{B}$.

We now define the continuous analogue of BBS's, and show in Proposition 1 that it is the right generalization.

Definition 4. Let $\mathcal{A} \in \mathbb{I}_n$. We call $\mathcal{A}$ a **bidirectional gerrymander** if $v_{\mathcal{A}} \in \mathcal{B}$.

Proposition 1. *Suppose $\mathcal{A} = I_1 \cup \cdots \cup I_n \in \mathbb{I}_n$ with endpoints ordered as before. Suppose the right endpoint of I_n is b. Then, $\mathcal{A}$ is a bidirectional gerrymander if and only if $\mu(\mathcal{A} \cap [0,t]) \geq t/2$ and $\mu(\mathcal{A} \cap [b-t,b]) \geq t/2$ for all $t \in [0,b]$.*

Proof. Clearly if these measure conditions hold, then $\mathcal{A}$ is a bidirectional gerrymander, as setting t to be left and right endpoints of the I_i yields the nonnegativity conditions of pairing

[1]Because the endpoints of an open interval cannot be equal, strictly speaking we are taking $\mathbb{I}_n$ to be the set of all weakly increasing $2n$-tuples of points on the real line and identifying these with collections of n intervals by treating them as endpoints (and correspondingly for $\mathbb{J}_n$). However, in the edge case when $a_j = b_j$, we still allow an 'empty' interval at a_j, which is included in the data of an element of $\mathbb{I}_n$. Including these degenerate cases allows us to indeed identify $\mathbb{J}_n$ with the closed unit cube.

with the ballot vectors. The condition $\mu(\mathcal{A}\cap[0,t]) \geq t/2$ is equivalent to the non-negativity of $\mu(\mathcal{A}\cap[0,t]) - \mu((\mathbb{R}\setminus\mathcal{A})\cap[0,t])$. For $t \in [0,b]$, $\mu(\mathcal{A}\cap[0,t]) - \mu((\mathbb{R}\setminus\mathcal{A})\cap[0,t])$ takes a local minimum only if t is a left endpoint of an interval I_i. Hence if $v_{\mathcal{A}} \cdot w \geq 0$ for all $w \in L_n$, then the function is nonnegative at its minima and so the first measure condition holds. Similarly, the second measure condition holds as well by the nonnegativity of pairing with the right ballot vectors. □

A BBS in the sense of [11] is a binary sequence for which any subsequence truncated on the left or right contains more 1's than 0's, and Proposition 1 shows that a bidirectional gerrymander is a subset of $\mathbb{R}$ contained in $[0,a]$ for which any subset obtained by truncating on the left or right contains "more" points (in a measure theoretic sense) in the original set than points not in this set. It is thus clear that they are a natural analogue, but, as we shall see, what is surprising is that they can be used to prove results about standard (discrete) BBS's.

Definition 5. The **bidirectional ballot polytope** $\mathcal{P}_n$, is defined as $\mathcal{B}_n \cap C_m$. Equivalently, it is those vectors $v_{\mathcal{A}}$ such that $\mathcal{A} \in \mathbb{J}_n$ is a bidirectional gerrymander. When the value of n is obvious, we shall refer to it simply as $\mathcal{P}$.

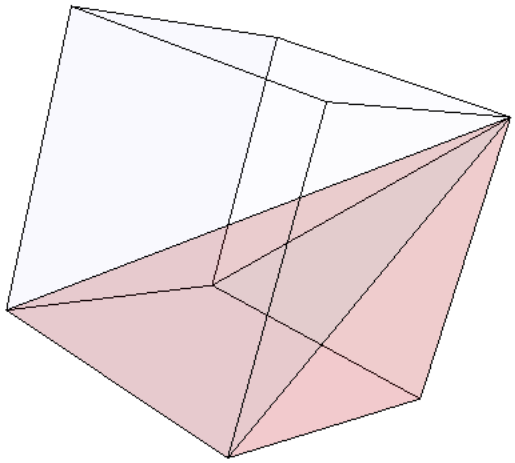


Figure 1: The polytope $\mathcal{P}_2$ (red) sitting inside C_3. Notice that adding two additional copies of $\mathcal{P}_2$, rotated about the main diagonal of the cube by $2\pi/3$ and $4\pi/3$ respectively, would result in a partition of C_3 (neglecting overlap of boundaries).

Definition 6. Let Z_m be the cyclic group of order m with generator ρ. Let Z_m act on $\mathbb{R}^m$ by cyclically permuting the entries (e.g. $\rho^2([0,1,2,3,4]) = [3,4,0,1,2]$). For a given set

of vectors V and $\sigma \in Z_m$, let $\sigma(V) := \{\sigma(v) : v \in V\}$ with $\sigma \in Z_m$. For each $\sigma \in Z_m$, define $\mathcal{B}_\sigma$ by

$$\mathcal{B}_\sigma \;:=\; \{v \in \mathbb{R}^m_{\geq 0} : v \cdot w \geq 0 \text{ for all } w \in \sigma(V_n)\}, \tag{3}$$

and $\mathcal{P}_\sigma$ likewise. Note that $\mathcal{B}_\sigma = \sigma^{-1}(\mathcal{B})$, and that $\mathcal{B} = \mathcal{B}_{\text{ld}}$ and $\mathcal{P} = \mathcal{P}_{\text{ld}}$.

Theorem 1. *The non-negative orthant, $\mathbb{R}^m_{\geq 0}$, is contained in $\bigcup_{\sigma \in Z_m} \mathcal{B}_\sigma$. Furthermore, for $\sigma_1 \neq \sigma_2$, the interiors of $\mathcal{B}_{\sigma_1}$ and $\mathcal{B}_{\sigma_2}$ are disjoint.*

Proof. Let $\tau = \rho^2 \in Z_m$ be the cyclic shift by two places. Because m is odd, τ generates Z_m. In particular, we see that the set of left and right ballot vectors V_n as defined in Definition 2 is equal to

$$V_n \;=\; \left\{ \sum_{i=0}^{k} \tau^i(w) : 0 \leq k \leq 2n-3 \right\}, \tag{4}$$

where $w = [1, -1, 0, \ldots, 0]$. If $\ell < k \leq 2n-3$ then

$$\sum_{i=0}^{k} \tau^i(w) - \sum_{i=0}^{\ell} \tau^i(w) \;=\; \sum_{i=\ell+1}^{k} \tau^i(w) \;=\; \tau^{\ell+1} \sum_{i=0}^{k-\ell-1} \tau^i(w), \tag{5}$$

and since $\sum_{i=0}^{2n-2} \tau^i(w) = [0, \ldots, 0]$ we have similarly that, for $0 \leq k \leq \ell$,

$$\sum_{i=0}^{k} \tau^i(w) - \sum_{i=0}^{\ell} \tau^i(w) \;=\; \tau^{\ell+1} \sum_{i=0}^{(2n-2)+(k-\ell)} \tau^i(w). \tag{6}$$

Then for each ℓ we have that

$$\left\{ \sum_{i=0}^{k} \tau^i(w) - \sum_{i=0}^{\ell} \tau^i(w) : 0 \leq k \leq 2n-2, k \neq \ell \right\} \;=\; \tau^{\ell+1}(V_n). \tag{7}$$

Now let $w_k = \sum_{i=0}^{k} \tau^i(w)$, take any $v \in [0,1]^m$, and choose $0 \leq \ell \leq 2n-2$ minimizing $v \cdot w_\ell$ (this ℓ may not be unique). Then

$$v \cdot \left(\sum_{i=0}^{k} \tau^i(w) - \sum_{i=0}^{\ell} \tau^i(w) \right) \;\geq\; 0 \tag{8}$$

for all $0 \leq k \leq 2n-2$. Therefore $v \cdot r \geq 0$ for all $r \in \tau^{\ell+1}(V_n)$, so $v \in \mathcal{B}_{\tau^{\ell+1}}$. This shows that $\mathbb{R}^m_{\geq 0} = \bigcup_{\sigma \in Z_m} \mathcal{B}_\sigma$. Intersecting with C gives the corresponding result for $\mathcal{P}$.

Conversely, if $v \in \text{Int}(\mathcal{B}_{\tau^{\ell+1}}) \cap \text{Int}(\mathcal{B}_{\tau^{k+1}})$ and $\tau^{\ell+1} \neq \tau^{k+1}$, then (because taking the interior simply changes the inequalities defining $\mathcal{B}_{\tau^{\ell+1}}$ to strict ones) we have both

$$v \cdot \left(\sum_{i=0}^{k} \tau^i(w) - \sum_{i=0}^{\ell} \tau^i(w) \right) \;>\; 0$$

$$v \cdot \left(\sum_{i=0}^{\ell} \tau^i(w) - \sum_{i=0}^{k} \tau^i(w) \right) \;>\; 0.$$

This is a contradiction, so the interiors distinct regions $\mathcal{B}_{\tau^{\ell+1}}$ are disjoint, and it follows immediately that the interiors of distinct regions $\mathcal{P}_{\tau^{\ell+1}}$ are disjoint. □

Corollary 1. *The unit cube C_m equals $\bigcup_{\sigma\in Z_m}\mathcal{P}_\sigma$. Furthermore, for $\sigma_1 \neq \sigma_2$, the interiors of $\mathcal{P}_{\sigma_1}$ and $\mathcal{P}_{\sigma_2}$ are disjoint. Consequently, the volume of $\mathcal{P}$ is exactly $1/m$.*

Proof. Intersecting the nonnegative orthant and the translates $\mathcal{B}_\sigma$ with C_m, Theorem 1 yields that C_m is partitioned into m regions produced by permuting the coordinates of $\mathcal{P}$. Because the matrix representing $\tau = \rho^2$ has determinant 1 it leaves volume invariant. Therefore, $\text{Vol}(\mathcal{P}_\sigma) = \text{Vol}(\mathcal{P})$ for all $\sigma \in Z_m$, so $\text{Vol}(\mathcal{P}) = 1/m$. □

Corollary 2. *For any vector $v \in \mathbb{R}^m_{\geq 0}$, there exists $\sigma \in Z_m$ such that the vector $v' = (v'_1, v'_2, \ldots, v'_m) = \sigma(v)$ has the following property: For all $1 \leq k \leq n$,*

$$\sum_{i=1}^{k}(v'_{2i-1} - v'_{2i}) \;\geq\; 0 \tag{9}$$

and

$$\sum_{i=1}^{k}(v'_{2n-(2i-1)} - v'_{2n-2i}) \;\geq\; 0. \tag{10}$$

If furthermore these are all positive, then σ is unique.

One interpretation of the above corollary is as follows. Suppose you have a necklace with an odd number of beads. On each bead you write a non-negative number. Then there exists some place where you can cut the necklace such that when you lay out the necklace and think of the sequence of values on the beads as a vector in $\mathbb{R}^m$, this vector is a bidirectional gerrymander. Furthermore, if the numbers you write on the beads are "generic", in the sense that the inequalities corresponding to (9) and (10) are strict, then there is exactly one such place you can cut the necklace.

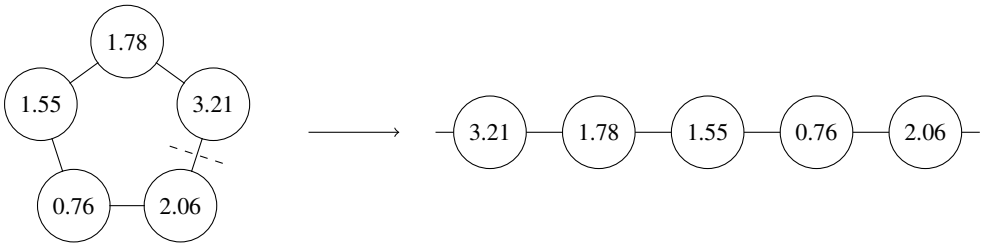


Figure 2: An example "cut" of a necklace as in Corollary 2.

3. Vertices of the Bidirectional Ballot Polytope are Vertices of the Cube

In this section we show that the vertices of $\mathcal{P}_n$ are also vertices of C_m, the unit cube. We had previously defined $\mathcal{P}_n$ as the intersection of the unit cube with the ballot cone, which is equivalent to the set of vectors $[\ell_1, g_1, \ldots, g_{n-1}, \ell_n]$ satisfying the below inequality:

$$\begin{array}{r} \text{cube vectors} \\ \\ \\ \\ \\ \\ \text{left ballot vectors} \\ \\ \\ \text{right ballot vectors} \\ \\ \end{array} \left\{ \begin{array}{c} \left\{ \begin{array}{c} \left[\begin{array}{cccccccccc} 1 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right] \\ \\ \left[\begin{array}{cccccccccc} 1 & -1 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right] \\ \\ \left[\begin{array}{cccccccccc} 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & \ldots & 0 & -1 & 1 & -1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right] \end{array} \right. \end{array} \left[\begin{array}{c} \ell_1 \\ g_1 \\ \ell_2 \\ g_2 \\ \vdots \\ g_{n-1} \\ \ell_n \end{array} \right] \geq \left[\begin{array}{c} 0 \\ -1 \\ 0 \\ -1 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \end{array} \right]. \tag{11}$$

The first collection of rows in the above matrix is necessary to ensure that we only deal with points inside of the unit cube. Thus we call any vector of the form $[0, \ldots, 0, \pm 1, 0, \ldots, 0]$ a **cube vector**.

Before proving the main result of this section, we must review a few concepts related to convex polytopes. We follow the terminology of [3].

Definition 7. Let P be a polytope in $\mathbb{R}^n$ defined by the inequalities $a_i^T x \geq b_i$ for $i \in \{1, 2, \ldots, k\}$. Let x^* be such that for some i, $a_i^T x^* = b_i$. Then, we say that the i^{th} constraint is **active** at x^*.

Definition 8. A vector $x^* \in \mathbb{R}^n$ is called a **basic solution** if out of all of the constraints that are active at x^*, there is some collection of n of them which is linearly independent. If x^* is a basic solution that satisfies all of the constraints, then it is called a **basic feasible solution**.

Part of what makes the study of convex polytopes interesting is that there are several equivalent but strikingly different ways of defining what the vertices of a polytope are. In particular, one definition is that a point v is a vertex if and only if it is a basic feasible solution.

The following shorthand will be helpful in the proof of the main theorem of this section.

Definition 9. A matrix/vector is called **flat** if all of its entries are 0, 1, or -1.

Let Q_n denote the set of vertices in the polytope $\mathcal{P}_n$. Let S_n denote the set of vertices of the unit cube C_m. The main result of this section is the following.

Theorem 2. *All of the vertices of the bidirectional ballot polytope $\mathcal{P}_n$ are also vertices of the unit cube C_m; i.e., $Q_n \subset S_n$.*

Proof. By the above discussion, we know that we must show that all basic feasible solutions are vertices of the cube. Throughout this proof, we let n be fixed, and let $m = 2n-1$. Thus we unambiguously let $\mathcal{P} = \mathcal{P}_n$, $C = C_{2n-1}$, $Q = Q_n$, and $S = S_n$. Notice that $\mathbb{Z}^m \cap \mathcal{P} \subset S$. From this observation, we now describe the strategy for proving the theorem. Suppose x^* is a basic solution whose corresponding constraints are $a_{i_1}, \ldots, a_{i_m}$. Then x^* satisfies

$$\begin{bmatrix} —a_{i_1}— \\ \vdots \\ —a_{i_m}— \end{bmatrix} x^* = \begin{bmatrix} b_{i_1} \\ \vdots \\ b_{i_m} \end{bmatrix}. \tag{12}$$

Let A be the matrix in (12). Let b^* be the vector on the right hand side in (12). Thus $x^* = A^{-1}b^*$. Note that $b^* \in \mathbb{Z}^m$ since it is some subset of the entries in the vector on the right hand side of (11). If we can show that $\det(A) = \pm 1$, it will imply that A^{-1} has integer entries, and thus that $A^{-1}b^* \in \mathbb{Z}^m$. From the earlier observation, if x^* is a basic feasible solution, then we must have that $A^{-1}b^* = x^* \in S$, which would prove the theorem.

Now we must show that if A is invertible, then it has determinant ± 1. In order to show this, we keep track of what happens to the determinant in the process of carrying out Gaussian elimination, which converts A into the identity matrix. In particular, we show that at every step, the determinant changes by a factor of ± 1. Since the identity matrix has determinant 1, we could then conclude that A has determinant ± 1. The only elementary row operation which potentially changes the absolute value of the determinant of a matrix is multiplying a row by a scalar. Thus it suffices to show that when Gaussian elimination is performed on A, no row is ever multiplied by a scalar other than ± 1. In Gaussian elimination, a row is multiplied by a scalar to convert some non-zero entry in that row to a one. If every non-zero entry in that row is ± 1, then we would simply need to multiply by ± 1. Thus, we shall instead prove the stronger hypothesis that at every step of Gaussian elimination, the intermediate matrix is flat, and hence all of its non-zero entries are ± 1. This is the content of Lemma 1. □

Before proving Lemma 1, we include an example to illustrate the method. Here we omit row swapping for clarity, and we obtain a permutation matrix, which has determinant ± 1. At each step, the leading nonzero term in the bolded row is used to clear the corresponding

column.

$$A_0:\begin{bmatrix}0&1&0&0&0\\0&0&0&0&1\\\mathbf{1}&\mathbf{-1}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\0&-1&1&-1&1\\0&0&0&-1&1\end{bmatrix}\rightarrow A_1:\begin{bmatrix}\mathbf{0}&\mathbf{1}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\0&0&0&0&1\\1&-1&0&0&0\\0&-1&1&-1&1\\0&0&0&-1&1\end{bmatrix}\rightarrow A_2:\begin{bmatrix}0&1&0&0&0\\0&0&0&0&1\\1&0&0&0&0\\\mathbf{0}&\mathbf{0}&\mathbf{1}&\mathbf{-1}&\mathbf{1}\\0&0&0&-1&1\end{bmatrix} \tag{13}$$

$$\rightarrow A_3:\begin{bmatrix}0&1&0&0&0\\0&0&0&0&1\\1&0&0&0&0\\0&0&1&-1&1\\\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{-1}&\mathbf{1}\end{bmatrix}\rightarrow A_4:\begin{bmatrix}0&1&0&0&0\\\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{1}\\1&0&0&0&0\\0&0&1&0&0\\0&0&0&1&-1\end{bmatrix}\rightarrow A_5:\begin{bmatrix}0&1&0&0&0\\0&0&0&0&1\\1&0&0&0&0\\0&0&1&0&0\\0&0&0&1&0\end{bmatrix}. \tag{14}$$

Lemma 1. *In carrying out Gaussian elimination on the matrix A as in Theorem 2, all intermediate matrices are flat.*

Proof. We proceed by induction. Let A_k denote the matrix resulting from the k^{th} step of Gaussian elimination (i.e. the matrix obtained after "clearing" the first k columns). We shall show that for each k, every row of the matrix A_k is of exactly one of six types depending on the form of the first k entries of that row and the last $m-k$ entries of that row (in the sequel, we will refer to this as saying that every row is one of the six types with respect to k).

We now describe these six types. Let α_n denote any sequence of length n consisting of alternating plus ones and minus ones (e.g. $\alpha_3=[-1,1,-1]$ or $\alpha_1=[1]$). Let β_n denote the sequence of length n consisting of all zeros. Let γ_n denote any binary sequence of length n containing exactly one one (e.g. $\gamma_4=[0,0,1,0]$). Let $\oplus$ refer to the operation of vector concatenation (e.g. $[1,2,3]\oplus[4,5]=[1,2,3,4,5]$). The six types (with respect to k) are listed in Table 1f.

Type	First k	Last $m-k$	Example ($k=3$, $m=7$)
1	β_k	$\beta_{\ell\geq 1}\ \oplus\ \alpha_{j\geq 1}\ \oplus\beta_{m-k-\ell-j\geq 1}$	$[0,0,0\mid 0,1,-1,0]$
2	β_k	$\alpha_{\ell\geq 1}\ \oplus\ \beta_{n-k-\ell\geq 0}$	$[0,0,0\mid 1,-1,1,0]$
3	β_k	$\beta_{\ell\geq 1}\ \oplus\ \alpha_{m-k-\ell\geq 0}$	$[0,0,0\mid 0,0,1,-1]$
4	γ_k	$\beta_{\ell\geq 1}\oplus\alpha_{j\geq 0}\ \oplus\beta_{m-k-\ell-j\geq 1}$	$[0,1,0\mid 0,0,0,0]$
5	γ_k	$\alpha_{\ell\geq 1}\ \oplus\ \beta_{n-k-\ell\geq 0}$	$[0,1,0\mid 1,-1,1,0]$
6	γ_k	$\beta_{\ell\geq 1}\ \oplus\ \alpha_{m-k-\ell\geq 0}$	$[0,1,0\mid 0,0,1,-1]$

Table 1: The six types with respect to k

We now go through the inductive argument. For the base case, notice that when $k=0$, the cube vectors are type 1, the left ballot vectors are type 2, and the right ballot vectors are type 3. Thus the claim is proven in the base case.

Now for the inductive step, we shall show that if all rows of A_k are of one of the above types with respect to k, then all rows of A_{k+1} are of one of the above types with respect to $k+1$. As described in the proof of Theorem 2, at step k we must first find some row whose first k entries are zero, and whose $k+1$ entry is ± 1. We see then that we must select some row of type 2, call it T. We then subtract T from all other rows whose $k+1$ entry is non-zero. Thus the only types we must worry about are types 2 and 5. Notice that when we subtract T from a row of type 2, we get a row either or type 1, type 2, or type 3 with respect to $k+1$. When we subtract T from a row of type 5, we get a row either of type 4, 5, or 6 with respect to $k+1$. All other rows remain the same. Thus when we catalog the new rows with respect to $k+1$, we get that those of type 1 become either type 1 or type 2. As mentioned before, those of type 2 become those of type 1, 2, or 3, except for row T which becomes of type 4 or 5. Type 3 becomes type 2 or 3. Type 4 remains type 4 or becomes type 5. As mentioned before, type 5 becomes type 4, 5, or 6. Lastly, type 6 becomes type 5 or type 6. Thus, by induction, we have proven the desired statement, implying in particular that the matrix is flat at every step. □

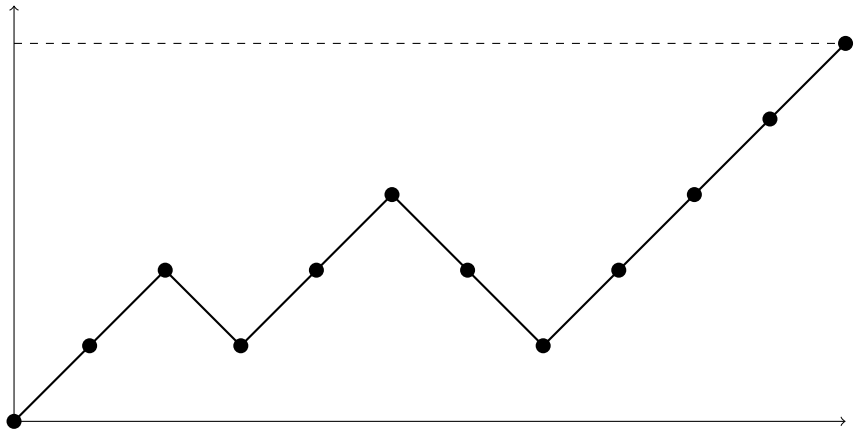
4. Vertices of the cube in the ballot region

In this section, we demonstrate that bidirectional ballot sequences of length $2n-1$ correspond in a natural way to Q_n, and we rederive the growth rate given in [11] and [2].

Definition 10. A **slope vector** is a vector $\lambda = [\lambda_1, \ldots, \lambda_m] \in \mathbb{R}^m$ with $m \in \mathbb{N}$. To a slope vector λ, we associate the unique continuous piecewise linear function $f_\lambda : [0, m] \to \mathbb{R}$ such that $f(0) = 0$ and $f'_\lambda(x) = \lambda_i$ for $x \in (i-1, i)$ for each $1 \leq i \leq m$.

Given any binary sequence $b = b_1 \cdots b_m$, we associate to this sequence the graph of the function f_λ where $\lambda = (\lambda_1, \ldots, \lambda_m)$ with $\lambda_i \coloneqq (-1)^{b_i - 1}$.

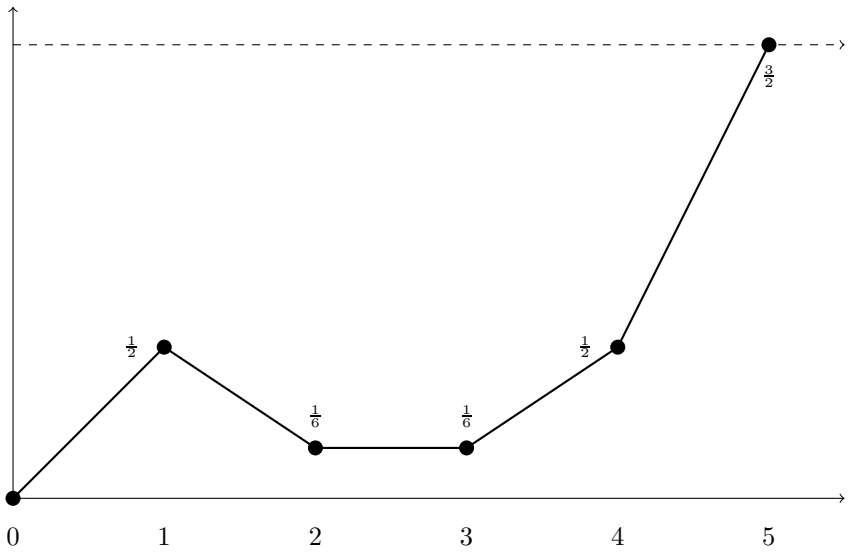
Example 1. The bidirectional ballot sequence 11011001111 corresponds to the path



This is a bijection from binary sequences of length m to graphs of functions f_λ with $\lambda \in \{\pm 1\}^m$. Recall from Section 1 that the graphs which correspond to bidirectional ballot sequences are those of functions f_λ where $f_\lambda(0) < f_\lambda(t) < f_\lambda(m)$ for all $0 < t < m$.

Now we will draw a correspondence between Q_n and B_{2n+3} through these graphs, as well as a correspondence between a certain subset of Q_n and B_{2n-1}, by describing a way to interpret vectors $v \in C_{2n-1} = [0,1]^{2n-1}$ as paths as in the discrete case in such a way that the vertices of the ballot polytope are realized as exactly the graphs above. Given a vector $v = [v_1, \ldots, v_{2n-1}] \in C_{2n-1}$, define the slope vector $\lambda_v = [\lambda_1, \ldots, \lambda_{2n-1}]$ by $\lambda_i := (-1)^{i-1}(2v_i - 1)$, and associate to v the graph of the function f_{λ_v}.

Example 2. The gap-parametrization vector $v = \left[\frac{3}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\right] \in [0,1]^5$ gives the slope vector $\lambda_v = \left[\frac{1}{2}, -\frac{1}{3}, 0, \frac{1}{3}, 1\right]$, which gives the following graph of the function f_{λ_v}, where the values next to the points indicate the distance above the x-axis:



Although the function f_{λ_v} in Example 2 has the property that it achieves global minimum and maximum values at it left and right endpoints (respectively), we will see that this is not always the case (see Example 3). We determine this behavior more precisely now.

If $v = [v_1, \ldots, v_{2n-1}] \in C_{2n-1}$, then for $0 \le k \le 2n-1$ we have

$$f_{\lambda_v}(k) \;=\; \sum_{j=1}^{k} (-1)^{j-1}(2v_j - 1) = \begin{cases} 2\sum_{j=1}^{k}(-1)^{j-1}v_j & k \text{ is even} \\ -1 + 2\sum_{j=1}^{k}(-1)^{j-1}v_j & k \text{ is odd,} \end{cases} \tag{15}$$

and similarly

$$f_{\lambda_v}(2n-1-k) \;=\; \begin{cases} f_{\lambda_v}(2n-1) - 2\sum_{j=1}^{k}(-1)^{j-1}v_{2n-j} & k \text{ is even} \\ f_{\lambda_v}(2n-1) + 1 - 2\sum_{j=1}^{k}(-1)^{j-1}v_{2n-j} & k \text{ is odd.} \end{cases} \tag{16}$$

One can see now that, even if $v \in \mathcal{P}_n$, it is possible for the graph to fail the property stated above, i.e., to achieve a global maximum or minimum at a point in the interior of its interval of definition (again, see Example 3 for an explicit example). However, one can also see that if $v \in \mathcal{P}_n$, it cannot fail this property to a great extent; namely, the values at the left and right endpoints will be within a distance of 1 from the maximum and minimum values, since the large sums in the RHS of (15) and (16) will be non-negative. Nonetheless, we would like the graphs of the functions f_{λ_v} with $v \in Q_n$ to match the graphs of bidirectional ballot sequences in B_{2n+3}, and for that reason we give a way to modify a vector $v \in Q_n$ before associating it to a graph. Namely, we will add a sort of buffer to each side of the vector, so that the left and right endpoints get a leg up.

Definition 11. If $v = [v_1, \ldots, v_{2n-1}] \in C_{2n-1}$, we define

$$\alpha(v) \coloneqq [1, 0, v_1, v_2, \ldots, v_{2n-2}, v_{2n-1}, 0, 1].$$

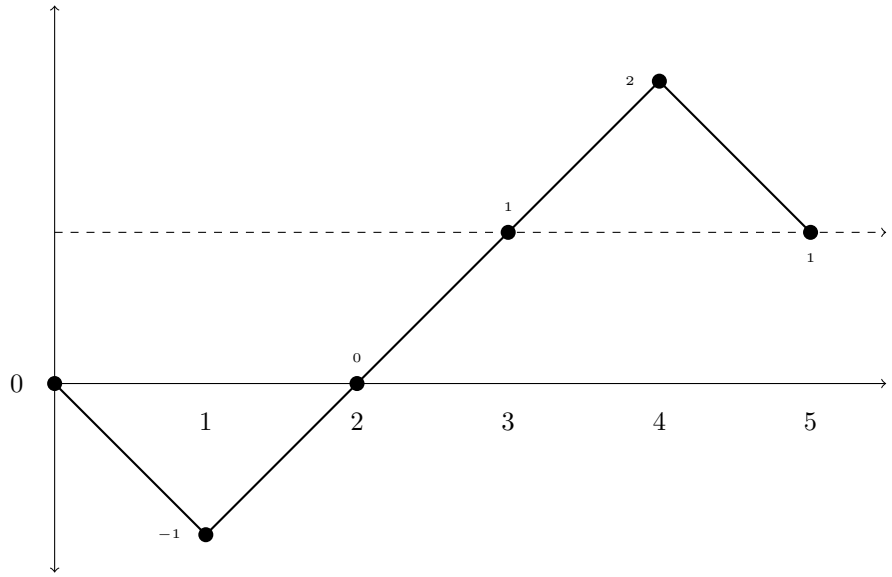
We now present two correspondences, the first stated more naturally, and the second proven more naturally, which are nonetheless very closely related. The first correspondence is as follows.

Theorem 3. *The set Q_n is in bijection with B_{2n+3}, induced by the map*

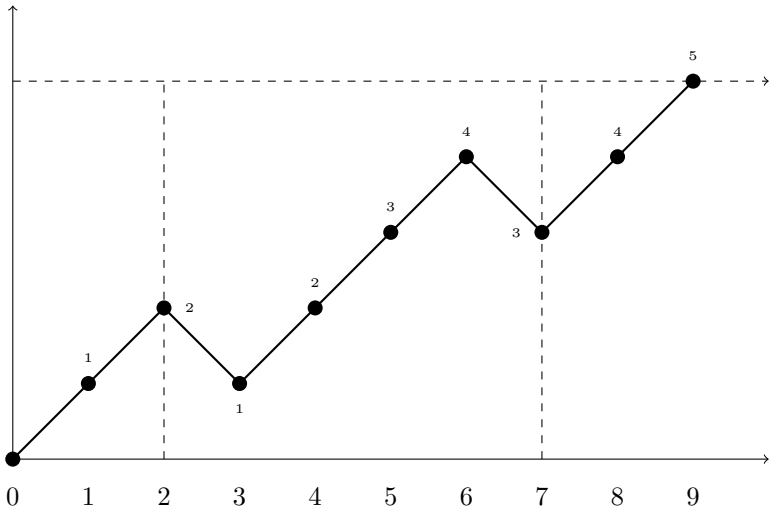
$$v \mapsto f_{\lambda_{\alpha(v)}}. \tag{17}$$

Before we prove Theorem 3, we give an example of the process that induces the bijection.

Example 3. Consider the gap-parametrization vector $v = [0, 0, 1, 0, 0] \in [0, 1]^5$, an element of Q_3. We shall obtain a bidirectional ballot sequence from v. *We see that v gives the slope vector $\lambda_v = [-1, 1, 1, 1, -1]$. The graph of f_{λ_v} is the following, where the values next to the points indicate the distance above the x-axis:*



This is not *the graph of a bidirectional ballot sequence. Namely, the graph passes below the x-axis and above the line $y = f_{\lambda_v}(5)$. Let's now consider $\alpha(v) = [1,0,0,0,1,0,0,0,1] \in [0,1]^9$, which gives slope vector $\lambda_{\alpha(v)} = [1,1,-1,1,1,1,-1,1,1]$ and leads to the following graph of $f_{\lambda_{\alpha(v)}}$.*



The portion of the graph between the vertical dotted lines is simply the graph of f_{λ_v} translated in the plane by the vector $[2,2]$. This graph does *correspond to a bidirectional ballot sequence, namely* 110111011. *We now prove that this process gives a bijection as in the statement of the theorem.*

Proof of Theorem 3. By the correspondence between bidirectional ballot sequences and graphs of certain functions given in Example 1, it suffices to show that the map of (17) puts Q_n in bijection with

$$F \;=\; \{f_\mu : \mu \in \{\pm 1\}^{2n+3},\; f_\mu(0) \;<\; f_\mu(t) \;<\; f_\mu(2n+3) \;\text{ for all }\; t \in (0, 2n+3)\}. \tag{18}$$

If $v \in C_{2n-1}$ is any gap-parametrization vector, then, in light of (15), (16), and the fact that f_{λ_v} achieves maxima and minima only at integer values, we have that $f_{\lambda_v}(0) - 1 \leq f_{\lambda_v}(t) \leq f_{\lambda_v}(2n-1)+1$ for $t \in [0, 2n-1]$ if and only if v is a bidirectional gerrymander. Furthermore, if v is a vertex of the cube C_{2n-1}, then $\alpha(v)$ is a vertex of $C_{2n+3} = [0,1]^{2n+3}$ so that $f_{\lambda_{\alpha(v)}}$ takes integers to integers. Since for any $v \in C_{2n-1}$ we have $f_{\lambda_{\alpha(v)}}(k+2) = f_{\lambda_v}(k) + 2$ for $0 \leq k \leq 2n-1$, $f_{\lambda_{\alpha(v)}}(i) = i$ for $i = 0, 1, 2$, and $f_{\lambda_{\alpha(v)}}(2n+1+i) = f_{\lambda_{\alpha(v)}}(2n+1)+i$ for $i = 1, 2$. Thus if v is a vertex of C_{2n-1} then $f_{\lambda_{\alpha(v)}}(0) < f_{\lambda_{\alpha(v)}}(t) < f_{\lambda_{\alpha(v)}}(2n+3)$ for all $t \in (0, 2n+3)$ if and only if $v \in Q_n$. It follows then that, since $\lambda_{\alpha(v)} \in \{\pm 1\}^{2n+3}$ when $v \in Q_n$, we indeed have that $f_{\lambda_{\alpha(v)}} \in F$, and so the map in (17) does indeed take Q_n to graphs of bidirectional ballot sequences in B_{2n+3}.

Injectivity of the map is clear. To show that the map is surjective, we provide an inverse. For a bidirectional ballot sequence $b = b_1 \cdots b_{2n+3}$ of length $2n+3$, we define the vector $w = [w_1, \ldots, w_{2n-1}]$, where

$$w_j \;\coloneqq\; \begin{cases} 1 & \text{if } j \equiv b_{j+2} \pmod 2 \\ 0 & \text{if } j \not\equiv b_{j+2} \pmod 2. \end{cases} \tag{19}$$

It is easily verified that the graph of $f_{\lambda_{\alpha(w)}}$ is the one associated to b. Moreover, the two statements directly following (18) imply that, since $w \in \{\pm 1\}^{2n-1}$ and the graph of $f_{\lambda_{\alpha(w)}}$ is that of a bidirectional ballot sequence, we must have that $w \in Q_n$. It is clear that this map is both a right- and left-inverse of the map given by (17). □

We now give the second correspondence. Let $\mathcal{I}_n$ denote the interior of $\mathcal{B}_n$ in $\mathbb{R}^{2n-1}$. Let $T_n = \mathcal{I}_n \cap Q_n$, i.e. those vertices of $\mathcal{P}_n$ in the interior of $\mathcal{B}_n$.

Corollary 3. *We have T_n is in bijection with B_{2n-1}, induced by the map*

$$v \;\mapsto\; f_{\lambda_v}. \tag{20}$$

Proof. The proof here is essentially the same as that of Theorem 3. The point here is that, when $v \in T_n$, we already have $f_{\lambda_v}(0) < f_{\lambda_v}(t) < f_{\lambda_v}(2n-1)$, following similar reasoning as in the statements directly following (18). □

Lastly, we use these correspondences along with our previous analysis of $\mathcal{P}_n$ and its translates to obtain the growth rate in [11].

Corollary 4. *For ℓ odd,*

$$B_\ell \;\geq\; \frac{2^\ell}{16(\ell-4)}. \tag{21}$$

Proof. The inequality is trivial if $\ell \in \{1, 3\}$, so assume $\ell \geq 5$. Let $m = \ell - 4$; this is $2n - 1$ for some $n \in \mathbb{N}$. By Theorem 3, we know that the vertices of $\mathcal{P}_n$ are in bijection with B_{m+4}. From Corollary 1, we know that every vertex of C_{2n-1} is contained in $\mathcal{P}_\sigma$ for some $\sigma \in Z_m$. Since there are m such copies of $\mathcal{P}$, we have

$$mB_{m+4} \geq 2^m. \tag{22}$$

By rearrangement we get

$$B_\ell \geq \frac{2^\ell}{16(\ell - 4)}. \tag{23}$$

□

Corollary 5. *For ℓ odd,*

$$B_\ell \leq \frac{2^\ell}{\ell}. \tag{24}$$

Proof. Suppose $\ell = 2n - 1$. From Corollary 3, we know that the vertices of $\mathcal{P}_n$ which are in the interior of $\mathcal{B}_n$, namely T_n, are in bijection with B_m. Since the interiors of $\mathcal{B}_{\sigma_1}$ and $\mathcal{B}_{\sigma_2}$ are disjoint if $\sigma_1 \neq \sigma_2$, we have that $\sigma_1(T_n) \cap \sigma_2(T_n) = \emptyset$ for $\sigma_1 \neq \sigma_2$. Therefore, summing over all the vertices in $\sigma(T)$ for each $\sigma \in Z_\ell$, we at most get every vertex of the cube once. That is,

$$\ell B_\ell \leq 2^\ell. \tag{25}$$

Rearranging yields

$$B_\ell \leq \frac{2^\ell}{\ell}. \tag{26}$$

□

Corollary 6. *For all ℓ, the growth rate of B_ℓ is $\Theta(2^\ell/\ell)$.*

Proof. By Corollaries 4 and 5, we know that for ℓ odd, the growth rate is $\Theta(2^\ell/\ell)$. The only additional insight needed is that for all ℓ, $B_{\ell+1} \geq B_\ell$. To see this, note that given a BBS of length ℓ, by appending a 1 to the end of it, we obtain a BBS of length $\ell + 1$. Thus up to fixed constants, the inequalities in Corollaries 4 and 5 are correct for even ℓ as well. Thus, for all ℓ, B_ℓ grows like $\Theta(2^\ell/\ell)$. □

5. Conclusion

Our methods reveal a rich combinatorial structure underlying bidirectional ballot sequences. In previous papers on BBS's ([11], [2], [6]), analytic techniques were used to obtain asymptotics, but our techniques reveal a geometric interpretation for the $\Theta(2^n/n)$ growth rate.

Interestingly, in the final section of [11], Zhao states without detailed proof that $nB_n/2^n$ goes to $1/4$, but claims his proof is "calculation-heavy". He then posits that "[t]here should be some natural, combinatorial explanation, perhaps along the lines of grouping all possible walks into orbits of size mostly n under some symmetry, so that almost every orbit contains exactly one walk with the desired property." Zhao's statement is strikingly similar to the ideas presented in our paper. Though we have made some effort, we have not been able to derive that $nB_n/2^n \to 1/4$ using the techniques of our paper, but we feel that there is hope for such a proof.

The second, more general takeaway from this paper is the potential for the ideas originally presented in [9]. The ideas in this paper in fact evolved from the ideas in [9]. In passing to the continuous setting, several additive number theory and combinatorial problems reveal a rich structure which was not otherwise visible. We believe that there is even greater potential still in such ideas and techniques.

References

[1] S. F. Altshul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, Basic local alignment search tool, *J. Molecular Biology* **215** (1990), no. 3, 403–410.

[2] M. Bosquet-Mélou and Y. Ponty, Culminating Paths, *Discrete Math. Theor. Comput. Sci.* **10** (2008), no. 2, 125–152.

[3] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, Belmont, MA, 1997.

[4] A. Califano and I. Rigoutsos, Flash, A fast look-up algorithm for string homology, *In Proceedings of the 1st International Conference on Intelligent Systems for Molecular Biology*, pages 56–64. AAAI Press, 1993.

[5] P. Di Franeso, E. Guitter and C. Kristjansen, Integrable 2D Lorentzian gravity and random walks, *Nuclear Phys. B* **567** (2000), no. 3, 515–553.

[6] B. Hackl, C. Heuberger, H. Prodinger and S. Wagner, Analysis of Bidirectional Ballot Sequences and Random Walks Ending in their Maximum, *Ann. Comb.* **20** (2016), 775–797.

[7] P. Hegarty and S. J. Miller, When almost all sets are difference dominated, *Random Structures Algorithms* **35** (2009), no. 1, 118–136.

[8] G. Martin and K. O'Bryant, Many sets have more sums than differences, *Additive combinatorics*, 287-305, CRM Proc. Lecture Notes **43**, Amer. Math. Soc., Providence, RI, 2007.

[9] S. J. Miller and C. Peterson, A geometric perspective on the MSTD question, *ArXiv e-prints* (Sept. 2017). https://arxiv.org/pdf/1709.00606.

[10] W. R. Pearson and D. J. Lipman, Improved tools for biological sequence comparison, *Proc. Natl. Acad. Sci. USA* **85** (1998), 2444–2448.

[11] Y. Zhao, Constructing MSTD sets using bidirectional ballot sequences, *J. Number Theory* **130** (2010), no. 5, 1212–1220.

[12] Y. Zhao, Sets characterized by missing sums and differences, *J. Number Theory* **131** (2011), no. 11, 2107–2134.