Vector Sum Problems in Convex and Discrete Geometry

PhD Thesis

Rainie Heck

Supervisor: Gergely Ambrus, PhD

Doctoral School of Mathematics
Bolyai Institute
University of Szeged,
Faculty of Science and Informatics

2026
Szeged

Contents

List of Figures

Chapter 1

Introduction

1.1 Definitions and Notations

We begin this dissertation by fixing terminology and notation that will be used. Throughout this work, $\log x$ refers to the natural logarithm. Disjoint unions will be denoted by $\bigsqcup$. We will use the standard notation $[n] := \{1, ..., n\}$. Vectors in $\mathbb{R}^d$ will be understood as column vectors, and when we write the coordinate decomposition $x = (x^{(1)}, \ldots, x^{(d)}) \in \mathbb{R}^d$, $x^{(i)}$ denotes the i^{th} coordinate of x. As usual, $\{e_1, \ldots, e_d\}$ stands for the standard orthonormal basis of $\mathbb{R}^d$. We denote the standard $(d-1)$-dimensional simplex represented in $\mathbb{R}^d$ by

$$\Delta^d := \text{conv}\,\{e_1, \ldots, e_d\} = \Big\{\lambda \in \mathbb{R}^d : \ \lambda^{(i)} \geq 0 \ \forall i \in [d], \ \sum_{i \in [d]} \lambda^{(i)} = 1\Big\}.$$

Given a collection of vectors $V = \{v_1, ..., v_n\}$, we denote its convex hull by

$$\text{conv}\, V := \{\lambda_1 v_1 + \cdots + \lambda_n v_n : \ \lambda \in \Delta^d\}.$$

By convex polytope we always mean a non-empty, bounded intersection of finitely many closed halfspaces (without any requirement on its interior). Let $\mathcal{K}^d$ denote the class of convex bodies in $\mathbb{R}^d$, i.e., compact convex sets with non-empty interior, and let $\mathcal{K}_o^d \subset \mathcal{K}^d$ be the class of convex bodies containing the origin in their interior. For $B \in \mathcal{K}_o^d$, the Minkowski norm generated by B is defined as

$$\|x\|_B := \inf\{r \geq 0 : \ x \in rB\}.$$

Note that this is a norm on $\mathbb{R}^d$ in the classical sense only when B is symmetric about

0, that is $B = -B$; otherwise, $\|.\|_B$ is homogeneous only for positive scalars – in this case $\|.\|_B$ is called an asymmetric norm. In this dissertation the term 'norm' will be used in a general sense that encompasses both cases. Note that B is the unit ball of $\|\cdot\|_B$.

According to standard conventions, B_p^d stands for the unit ball of the ℓ_p-norm on $\mathbb{R}^d$, where

$$\|x\|_p := \Big(\sum_{i\in[d]} |x^{(i)}|^p\Big)^{1/p}, \quad \forall p \geq 1,\ x \in \mathbb{R}^d,$$

and

$$\|x\|_\infty = \max_{i\in[d]} |x^{(i)}|.$$

- The closed positive halfspace orthogonal to u,

$$H_+(u) := \{x \in \mathbb{R}^d : \ \langle x, u\rangle \geq 0\}.$$

- The *spherical cap* of height $(1-t)$ centered at u,

$$C_t(u) := \big\{v \in S^{d-1} : \ \langle v, u\rangle \geq t\big\}.$$

- The *ball-cone* of height t centered at u,

$$K_t(u) := \Big\{v \in B_2^d : \ \Big\langle \frac{v}{|v|}, u\Big\rangle \geq t\Big\} = \mathrm{Conv}(C_t(u) \cup \{0\});$$

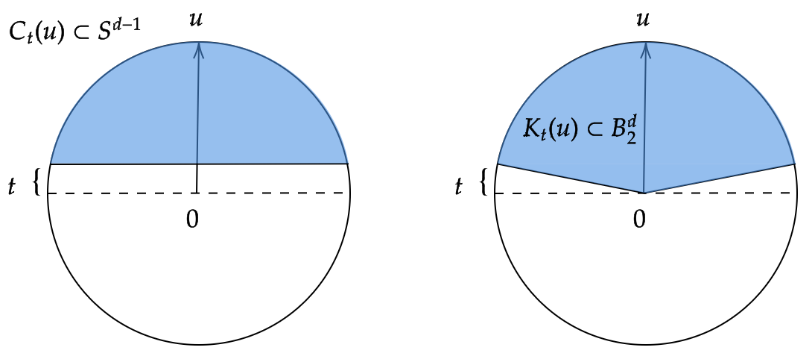

Figure 1.1: *Depiction of $C_t(u)$ and $K_t(u)$.*

Let $\kappa_d := \mathrm{Vol}_d(B_2^d) = \pi^{d/2}/\Gamma(\frac{d}{2}+1)$ denote the volume of the Euclidean unit ball, where

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} \,\mathrm{d}x, \quad x > 0$$

is Euler's gamma function.

Let $\sigma(\cdot)$ be the normalized surface area measure on S^{d-1}. We will need to estimate the measure of spherical caps: to that end, we define

$$\sigma_t := \sigma(C_t(u))$$

for an arbitrary $u \in S^{d-1}$ – note that σ_t is independent of u.

Finally, when useful we denote the sum of a finite collection of vectors $V \subset \mathbb{R}^d$ by $\Sigma(V) := \sum_{v \in V} v$.

1.2 A Brief Overview of History and Results

In this dissertation we will focus on results relating to vector sum problems from convex and discrete geometry, in particular the *vector balancing problem* and the *Steinitz problem*. These two problems, while quite different in nature, are intricately connected by the beautiful transference theorem of Chobanyan [26]. We begin by introducing these two problems and their history, and conclude the section with a brief summary of related work.

The vector balancing question asks the following: given symmetric convex bodies $K, L \in \mathcal{K}_o^d$ with associated Minkowski norms $\|\cdot\|_K, \|\cdot\|_L$ and any collection of vectors $v_1, \ldots, v_n \in K$, select signs $\varepsilon_1, \ldots, \varepsilon_n \in \{\pm 1\}$ so that $\left\| \sum_{i \in [n]} \varepsilon_i v_i \right\|_L$ is minimal. The term *vector balancing* is readily motivated by the following interpretation: placing the vectors into the two plates of a scale according to their associated signs, the problem asks for achieving a nearly equal balance, that is, forcing the sum of the vectors in the plates to be as close as possible.

In order to facilitate the coming work, we introduce the notion of *vector balancing constants* of $K, L \in \mathcal{K}_o^d$. To this end, we define the n-vector balancing constant:

$$\mathrm{vb}(K, L, n) = \max_{v_1, \ldots, v_n \in K} \min_{\varepsilon_1, \ldots, \varepsilon_n \in \{\pm 1\}} \|\varepsilon_1 v_1 + \ldots + \varepsilon_n v_n\|_L. \tag{1.1}$$

The first surprising observation we can make is that even though $\mathrm{vb}(K, L, n)$ depends on the bodies $K, L \in \mathcal{K}_o^d$ and the two parameters $d, n \in \mathbb{N}$, the optimal bounds turn out to be independent of the number of vectors n in the case $n \geq d$

[12, 54, 67], a fact that we will return to in Section 2.2. In light of this observation, we can define the *vector balancing constant* of K and L,

$$\mathrm{vb}(K, L) = \sup_{n \geq d} \mathrm{vb}(K, L, n).$$

When $K = L$, we simply write

$$\mathrm{vb}(K, n) = \mathrm{vb}(K, K, n) \quad \text{and} \quad \mathrm{vb}(K) = \sup_{n \geq d} \mathrm{vb}(K, n).$$

The vector balancing problem is over six decades old, and was first introduced by Dvoretzky [28], who asked for bounds specifically in the ℓ_p setting for $p \geq 1$ and $d \in \mathbb{N}$; that is, on $\mathrm{vb}(B_p^d)$.

The first results for Dvoretzky's question came in the late 1970's, first for the case of the Euclidean norm. The sharp bound

$$\mathrm{vb}(B_2^d, d) = \mathrm{vb}(B_2^d) = \sqrt{d} \tag{1.2}$$

was independently proven by Sevast'yanov [64], Bárány (unpublished at the time, for the proof, see [22]), Spencer [67] and also, perhaps, by V.V. Grinberg [23]. Sevast'yanov and Bárány used linear algebraic techniques, whereas Spencer applied the probabilistic method. In both approaches, the proof reduces to showing that any point of a parallelotope in $\mathbb{R}^d$ can be approximated by a vertex with Euclidean error at most $\sqrt{d}$. This result is the direct predecessor of our Proposition 3.1 in Section 3.1.

The case of the ℓ_∞-norm proved to be much more challenging, but it was later solved by Spencer in 1985 [68], who showed that

$$\mathrm{vb}(B_\infty^d, d) \leq C\sqrt{d} \tag{1.3}$$

and

$$\mathrm{vb}(B_\infty^d) \leq 2C\sqrt{d} \tag{1.4}$$

for a universal constant $C < 6$. These estimates are asymptotically sharp, as one can see using random constructions involving Hadamard matrices. One interesting note about the maximum norm case is that although Spencer's original proof was highly non-constructive, more recent algorithmic approaches have been developed showing that the coloring is actually obtainable in randomized polynomial time [10, 47]. We also note that the weaker bound of $O(\sqrt{n \log d})$ can be shown by applying the probabilistic method, but removing the $\sqrt{\log d}$ factor is not possible using that approach. The upper bound (1.3) was also shown, independently, by Gluskin [33], who applied Minkowski's theorem on lattice points and an argument of Kashin [44]. These results rely on the parallelotope approximation in the maximum norm, which is the predecessor of our Proposition 3.2 in Section 3.2. These vertex approximation results are also closely related another class of vector sum problems known as the Beck-Fiala "integer-making" theorems [15].

In 2022, Reis and Rothvoss [54] proved that there exists a universal constant C' for which $\mathrm{vb}(B_p^d) \leq C'\sqrt{d}$ holds for all $2 \leq p \leq \infty$. This bound can be combined with the following lower bound of Banaszczyk [8] for general norms to fully resolves Dvoretzky's question.

Theorem 1.1 (Lower Bound for General Norms). *Let K, L be two symmetric convex bodies in $\mathbb{R}^d$ and $|K|, |L|$ their d-dimensional volumes. Then there exist vectors $u_1, ..., u_n \in K$ such that for any choice of signs $\varepsilon_1, ..., \varepsilon_d \in \pm 1$,*

$$\|\varepsilon_1 u_d + \cdots + \varepsilon_d u_d\|_L \geq \frac{1}{\sqrt{2\pi e}} \sqrt{n} \big(|K|/|L|\big)^{1/d}.$$

The major open question remaining in this area is the Komlós conjecture (see [9, 68]), which posits that

$$\mathrm{vb}(B_2^d, B_\infty^d, n) \leq C$$

holds for each $n, d \geq 1$ with a universal constant C. Very recently, a new breakthrough was made by Bansal and Jiang, who showed that $\mathrm{vb}(B_2^d, B_\infty^d) = O((\log d)^{1/4})$. This improved on the $O(\sqrt{\log d})$ bound of Banaszczyk from 1998, which was actually a consequence of a more general theorem, stated below.

Theorem 1.2 ([9]). *Let γ_d denote the (standard) Gaussian measure on $\mathbb{R}^d$ with density $\frac{1}{(2\pi)^{-d/2}} \exp(-\|x\|_2^2/2)$, and let $K \subset \mathbb{R}^d$ be a convex body with $\gamma_d(K) \geq 1/2$. Then given $v_1, ..., v_n \in B_2^d$, there exist signs $\varepsilon_1, ..., \varepsilon_n \in \{\pm 1\}$ such that*

$$\|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n\|_K \leq CK,$$

where $C > 0$ is a universal constant.

Similar to the case of Spencer's maximum norm result, the original proof of Banaszczyk was highly non-constructive, but a recent algorithmic proof was given obtaining the colorings in randomized polynomial time [11].

Beyond the unit balls of the ℓ_p norms, this problem has been studied for many other convex bodies, as well as in specific dimensions, in online settings, and through many other variants. We introduce a few here. Giannopoulous provides a nice summary of classical vector balancing results in [32]. Vector balancing of zonotopes, a particular class of convex bodies, was studied in [42] as an extension of Spencer's results in the maximum norm setting. Vector balancing in the specific setting of the plane ($d = 2$) has been studied by Swanepoel [74] and Lund and Magazinov [49]. There are also related online versions of the vector balancing problem, where one is given vectors one at a time, as well as other related combinatorial games; these have been studied extensively by Spencer [66, 69]. One can also ask related anti-balancing questions, see for example [3, 8]. In addition, vector balancing in the maximum norm is intricately connected to *discrepancy theory*, as one can interpret $\mathrm{vb}(B_\infty^d, n)$ as the discrepancy of a set system with d sets on n elements. For more information on discrepancy theory and related results, see Matoušek [50].

An exciting new direction that has driven recent research on vector balancing type problems is a close connection between vector balancing and various problems in machine learning. Applications thus far include, but are not limited to, coresets for kernel density estimation [21, 52, 53, 75], randomized control trials [41], and quantization of neural networks [5].

In Chapters 2 and 3 we will introduce another generalization of the vector balancing problem, called the *colorful vector balancing problem*, and prove that the asymptotically tight bounds in the Euclidean and maximum norm cases extend to this more general setting. The precise results are formulated below.

Theorem ([2], Theorem 1.4). *Given vector families $V_1, ..., V_n \subseteq B_2^d$ with*

$$0 \in \sum_{i \in [n]} \mathrm{Conv}\, V_i,$$

one can select vectors $v_i \in V_i$ for $i \in [n]$ such that $\|v_1 + \cdots + v_n\|_2 \leq \sqrt{d}$.

Theorem 1.3 ([2], Theorem 1.5). *Given vector families $V_1, ..., V_n \subseteq B_\infty^d$ with*

$$0 \in \sum_{i \in [n]} \mathrm{Conv}\, V_i,$$

one can select vectors $v_i \in V_i$ for $i \in [n]$ such that $\|v_1 + \cdots + v_n\|_\infty \leq C\sqrt{d}$, where $C = 22$ suffices.

We now change focus and introduce the *Steinitz problem*, a seemingly unrelated vector sum problem that is in fact closely connected to vector balancing. The Steinitz problem arises in connection with a famous theorem that will be familiar to all mathematicians: the Riemann rearrangement theorem [55]. This theorem, a classic result in analysis, tells us that a conditionally convergent series can be rearranged to converge to any real number. Formulated through a different lens, for any real series, consider the set of all sums of its possible rearrangements. The Riemann rearrangement theorem tells us that this set is either empty, i.e. the series is divergent; a single point, i.e. the series is absolutely convergent; or the entire real line, i.e. the series is conditionally convergent. A natural question is what happens if one studies sequences of complex numbers, or even more generally, sequences of vectors in $\mathbb{R}^d$. This problem was first addressed by Lévy in 1905 [46](at just 19 years old, in his very first article), who proved the following result.

Theorem 1.4 (Lévy-Steinitz Theorem)**.** *Given a series of vectors in $\mathbb{R}^d$, the set of all sums of its rearrangements is empty, or it forms an affine subspace of $\mathbb{R}^d$.*

Recall that an affine subspace of $\mathbb{R}^d$ is of the form $L + x$, where $L \subset \mathbb{R}^d$ is a linear subspace and $x \in \mathbb{R}^d$. The reader may notice that the theorem is also attributed to Steinitz: the reason for this is that Lévy's proof contained serious gaps in dimensions $d \geq 3$, which was pointed out and fixed by Steinitz in a series of works published in three parts [71, 72, 73], which is quite technical and covers much ground. The key step in his proof is the following, which is the birth of what we will call the Steinitz problem.

Theorem 1.5 ([71], p.171)**.** *Given any finite family of vectors $V \subset \mathbb{R}^d$ of Euclidean norm at most 1 summing to 0, one can order the elements of V as $v_1, ..., v_n$ so that for every $k = 1, ..., n$,*

$$\Big\| \sum_{i\in[k]} v_i \Big\|_2 \leq C, \tag{1.5}$$

where C is a constant that depends only on the dimension d.

Steinitz's proof shows that in fact $C \leq 2d$. It is natural to ask for the smallest value of C for which (1.5) holds, in general norms as well. This quantity will be called the *Steinitz constant*, and is defined as follows.

Definition 1.6 (Steinitz constant)**.** *Let $B \in \mathcal{K}_0^d$. The* Steinitz constant of *B, denoted $S(B)$, is the smallest number C for which any finite family of vectors $V \subset B$ with*

$\Sigma(V) = 0$ has an ordering $V = \{v_1, \ldots, v_n\}$ along which each partial sum has norm at most C. That is, for every $k \in [n]$,

$$\Big\| \sum_{i \in [k]} v_i \Big\|_B \leq C.$$

Note that the term 'constant' above refers to the fact that $S(B)$ depends only on the choice of B, but not on the vector family $V \subset B$. We make no reference to the dimension d, as the value of the Steinitz constant is independent of d as long as B can be embedded in $\mathbb{R}^d$.

One can also consider a generalized version of the Steinitz constant, where the zero-sum condition $\Sigma(V) = 0$ on the vector family is dropped:

Definition 1.7 (Relaxed Steinitz constant)**.** *For $B \in \mathcal{K}_0^d$, let $S^*(B)$ denote the smallest constant C for which any finite family of vectors $V \subset B$ has an ordering $V = \{v_1, \ldots, v_n\}$ so that*

$$\Big\| \sum_{i \in [k]} v_i - \frac{k}{n}\Sigma(V) \Big\|_B \leq C \tag{1.6}$$

holds for every $k \in [n]$.

The relationship with the original Steinitz constant is given by the simple chain of inequalities

$$S(B) \leq S^*(B) \leq (1 + \rho(B))S(B), \tag{1.7}$$

where

$$\rho(B) := \max_{v \in B} \| - v\|_B$$

measures the asymmetry of B. Note that $\rho(B) = 1$ if B is symmetric. The lower bound in (1.7) is trivial; to see the upper estimate, one has to observe that starting from any family V of n vectors in B, the triangle inequality implies that $\|\Sigma(V)\|_B \leq n$, hence $\| - \frac{\Sigma(V)}{n}\|_B \leq \rho(B)$. Accordingly, the zero-sum vector family $\big\{v - \frac{\Sigma(V)}{n} : v \in V\big\}$ lies in $(1 + \rho(B))B$, and the estimate readily follows. We note that there are further variants of Definition 1.7 (see e.g. [7]), although these are not directly related to the topics of this dissertation.

Theorem 1.5, proved by Steinitz, justifies that $S(B_2^d)$ and, via (1.7), that $S^*(B_2^d)$ are well-defined. The proof can be extended to any symmetric norm. For asymmetric norms, the justification of Definitions 1.6 and 1.7 is implied by the following general bound, proved in 1978 by Sevastyanov [63] and by Grinberg and Sevastyanov [34] for not necessarily symmetric bodies by a simpler proof.

Theorem 1.8 (The Steinitz Lemma for general norms [34, 63]). *For any convex body $B \in \mathcal{K}_o^d$,*

$$S(B) \leq d. \tag{1.8}$$

The bound is tight for non-symmetric convex bodies, as is shown by taking B to be the regular simplex centered at the origin and choosing V to be the set of its vertices, whereas it is sharp by the order of magnitude for symmetric norms, which is confirmed by the inequality $S(B_1^d) \geq (d+1)/2$, see [34]. For symmetric $B \in \mathcal{K}_o^d$, the estimate in (1.8) can be strengthened to $d - 1 + \frac{1}{d}$, see [65].

Via (1.7), Theorem 1.8 readily implies the bound

$$S^*(B) \leq (1 + \rho(B))d,$$

which also follows from the results in [34]. In particular, $S^*(B) \leq 2d$ holds for symmetric $B \in \mathcal{K}_o^d$.

The following long-standing conjecture of Bergström [18], be it confirmed, would yield a much stronger estimate on the Steinitz constant in the Euclidean case:

Conjecture 1.9. *For all $d \geq 1$, $S(B_2^d) = O(\sqrt{d})$.*

The same bound is expected to hold for the maximum norm. So far, Conjecture 1.9, which is sometimes also called the Euclidean Steinitz problem, has refuted all attempts. An explicit construction [27, 34] shows that $S(B_2^d) \geq \sqrt{d+3}/2$ must hold, meaning that no stronger estimate is possible. The exact value of the planar Euclidean Steinitz constant was determined by Banaszczyk [7], who proved that $S(B_2^2) = \sqrt{5}/2$, matching this lower bound.

Our work in this dissertation focuses on relating the Steinitz constant to the restricted setting of 'nearly unit' vectors: the subscript 'ε' will mean that only families of vectors are considered whose members have norm in the interval $[1-\varepsilon, 1]$.

Definition 1.10 (ε-Steinitz constants). *For $B \in \mathcal{K}_0^d$ and $0 \leq \varepsilon \leq 1$, let $S_\varepsilon^*(B)$ denote the smallest constant C for which any finite family $V \subset \mathbb{R}^d$ consisting of vectors of $\|.\|_B$-norm in $[1-\varepsilon, 1]$ may be ordered as $V = \{v_1, \ldots, v_n\}$ so that*

$$\Big\| \sum_{i \in [k]} v_i - \frac{k}{n} \Sigma(V) \Big\|_B \leq C$$

holds for every $k = 1, \ldots, n$. Furthermore, let $S_\varepsilon(B)$ denote analogous quantity for vector families that satisfy the extra condition $\Sigma(V) = 0$.

Note that for any $0 \leq \varepsilon \leq 1$, $S_0(B) \leq S_\varepsilon(B) \leq S_1(B) = S(B)$, $S_0^*(B) \leq S_\varepsilon^*(B) \leq S_1^*(B) = S^*(B)$, and $S_\varepsilon(B) \leq S_\varepsilon^*(B)$. Thus, (1.7) ensures that

$$S_\varepsilon^*(B) \leq 2S(B) \tag{1.9}$$

for symmetric norms, while

$$S_\varepsilon^*(B) \leq (1 + \rho(B))S(B)$$

holds for arbitrary $B \in \mathcal{K}_o^d$.

Furthermore, observe that setting $\varepsilon = 0$ restricts the problem to families of unit vectors. In the Euclidean case, a construction given by Damsteeg and Halperin [27] implies that

$$\Omega(\sqrt{d}) \leq S_0(B_2^d) \leq S_0^*(B_2^d) \leq S_\varepsilon^*(B_2^d). \tag{1.10}$$

In this dissertation we prove two results establishing reverse estimates of (1.9). The first result is specific to the Euclidean norm.

Theorem 1.11. *For any $0 < \varepsilon < 1$ and all $d \geq 2$,*

$$S(B_2^d) < \frac{1}{\varepsilon}\Big(S_\varepsilon^*(B_2^d) + 200\sqrt{\frac{d}{\log d}}\Big). \tag{1.11}$$

In particular, an $o(d)$ bound on $S_\varepsilon^*(B_2^d)$ for some fixed $0 < \varepsilon \leq 1$ would yield an $o(d)$ estimate on $S(B_2^d)$, hence improving the current strongest bound. Moreover, (1.10) and (1.11) imply that Conjecture 1.9 is equivalent to the statement that $S_\varepsilon^*(B_2^d) = O(\sqrt{d})$ for some constant $\varepsilon \in (0, 1]$.

The second result generalizes and simplifies the techniques of the proof of Theorem 1.11 and yields an even stronger estimate for general norms.

Theorem 1.12 ([4] Theorem 7)**.** *For all $d \geq 2$, for every convex body $B \in \mathcal{K}_o^d$, and $0 < \varepsilon \leq 1$,*

$$S(B) < \frac{1}{\varepsilon}\Big(S_\varepsilon^*(B) + 2\rho(B) + 1\Big). \tag{1.12}$$

In the case that B is symmetric, the bound simplifies to $\frac{1}{\varepsilon}(S_\varepsilon^*(B) + 3)$.

We conclude this section with the Chobanyan transference theorem, a surprising result that connects the vector balancing problem to the Steinitz problem. In order to introduce this theorem, we first describe a variant of the vector balancing problem, called the signed sequence problem. In this problem one is given a symmetric convex body $B \subset \mathbb{R}^d$ and a (potentially infinite) sequence $u_1, u_2, \ldots \in B$. The goal is to find signs $\varepsilon_i \in \{\pm 1\}$ for $i = 1, 2, \ldots$, so that all signed partial sums $\sum_{i \in [k]} \varepsilon_i u_i$ for $k \in \mathbb{N}$ are bounded by a constant C depending only on B. We define the signed sequence constant of B, $E(B)$, to be the smallest constant C that holds for all sequences selected from B. Bárány and Grinberg proved that $E(B) \leq 2d - 1$ for all symmetric convex $B \subset \mathbb{R}^d$ [23].

The Chobanyan transference theorem establishes a close connection between the signed sequence constant and Steinitz constant of any given symmetric convex body.

Theorem 1.13 (Chobanyan Transference Theorem [26]). *Assume B is a symmetric convex body in $\mathbb{R}^d$. Then $S(B) \leq E(B)$.*

In particular, to verify Conjecture 1.9 it would suffice to show that $E(B) = O(\sqrt{d})$. The Chobanyan transference theorem is just one example of deep and beautiful connections between seemingly distinct vector sum problems in discrete and convex geometry.

1.3 Overview of Thesis

To conclude this chapter, we summarize the organization of the thesis. In Chapter 2 we introduce the *colorful vector balancing problem*, a geometric generalization of the original vector balancing problem. In Section 2.1 we discuss the history of the problem and existing results. In Section 2.2 we describe a linear algebraic reduction of the problem which will be key to our proofs. In particular, this aspect of the proofs allows us to prove bounds independent of the number of vectors. In Section 2.3 we extend techniques from [11] to prove that the colorful vector balancing problem can always be bounded in terms of the original vector balancing problem. We also justify the benefits of our more direct, geometric proof of these results.

In Chapter 3 we prove our main results for the colorful vector balancing problem in the Euclidean and maximum norms. In particular, we extend the tight (respectively, asymptotically tight) results for the Euclidean and maximum norms in the vector balancing setting to the colorful setting (see Theorems 1.2 and 1.3). The chapter is structured as follows: in Section 3.1 we prove our result for the Euclidean norm using the probabilistic method. In Section 3.2 we prove our result for the maximum norm, up to the proof of a technical lemma that we defer to Section 3.3 in order to simplify the exposition in Section 3.2. The proof in the maximum norm setting is more involved and is based on analysis of a Gaussian random walk. It is a generalization of the algorithm introduced by Lovett and Meka in [47].

In Chapter 4 we turn our attention to the Steinitz problem and Conjecture 1.9. Our main result is a reduction of the Steinitz problem for arbitrary norms to the setting where the vectors all have norm in $[1-\varepsilon, 1]$ for any fixed constant $0 < \varepsilon < 1$ (which we call "almost-unit vectors"), up to additive $O(1)$ error (see Theorems 1.11 and 1.12). In Section 4.1 we outline the structure of the chapter. In Section 4.3 we present a proof of a slightly weaker result in the specific case of the Euclidean norm, utilizing techniques of independent interest (several technical lemmas are deferred to Section 4.4). Finally, in Section 4.5 we prove our main result, which holds for arbitrary norms.

In Chapter 5 we present a brief conclusion and discuss several potential future extensions of our work.

Chapter 2

Colorful Vector Balancing: a Linear Algebra Reduction

Chapters 2 and 3 of the dissertation are based on the following published paper of the author:

[2] Gergely Ambrus and Rainie Bozzai. Colourful vector balancing. Mathematika, 70(4), August 2024.

2.1 The Colorful Vector Balancing Problem

Recall the vector balancing problem introduced in Section 1, where we are given a norm $\|\cdot\|$ on $\mathbb{R}^d$ with unit ball $B \subset \mathbb{R}^d$ and vectors $v_1, ..., v_n \in B$, and asked to select signs $\varepsilon_1, ..., \varepsilon_n \in \{\pm 1\}$ so that norm of the signed sum, $\|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n\|$, is minimal. This chapter focuses on a natural "colorful" generalization of this problem: again fix a norm $\|\cdot\|$ on $\mathbb{R}^d$ with unit ball $B \subset \mathbb{R}^d$, but now consider vector families $V_1, ..., V_n \subseteq B$ satisfying the condition that $0 \in \text{conv } V_1 + \cdots + \text{conv } V_n$. The goal is to select one vector from each family, $v_i \in V_i$, so that the norm of the sum of the selected vectors, $\|v_1 + \cdots + v_n\|$, is minimal.

We make two remarks about this problem statement. First, to motivate the name "colorful" vector balancing, note that one can interpret the families as color classes, in which case the problem asks for a colorful sum of vectors of minimal norm (see Figure 2.1 for an example. Note that in this example, while $0 \notin \text{conv}V_1$, $\text{conv}V_2$, it is true that $0 \in \text{conv}V_1 + \text{conv}V_2 + \text{conv}V_3$.).

Second, to see that this problem is indeed a generalization of the original vector

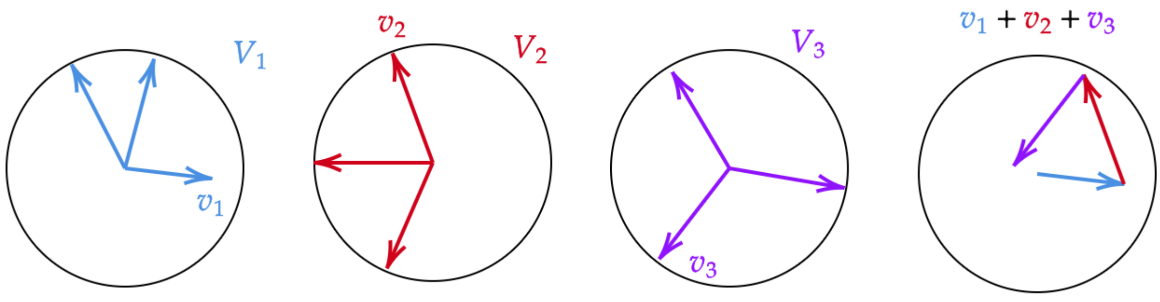

Figure 2.1: *An example of colorful vector balancing with $n = 3$, $d = 2$ in the Euclidean norm.*

balancing problem, note that the original problem is retrieved by setting $V_i = \{\pm v_i\}$ for $i \in [n]$.

The colorful vector balancing problem was first introduced by Bárány and Grinberg, who proved the following result.

Theorem 2.1 (Bárány, Grinberg [23]). *Assume that $B \subset \mathcal{K}_o^d$ is an origin-symmetric convex body, and $V_1, \ldots, V_n \subseteq B$ are vector families so that $0 \in \sum_{i\in[n]} \mathrm{Conv} V_i$. Then there exists a selection of vectors $v_i \in V_i$ for $i \in [n]$ such that*

$$\Big\| \sum_{i\in[n]} v_i \Big\|_B \leq d. \tag{2.1}$$

Taking $B = B_1^d$, $n = d$, and $V_i = \{\pm e_i\}$ for $i \in [n]$ shows that Theorem 2.1 is sharp. Yet, for specific norms, asymptotically stronger estimates may hold. In light of the fact that $\mathrm{vb}(B_2^d) = \sqrt{d}$ and $\mathrm{vb}(B_\infty^d) = O(\sqrt{d})$ (see (1.2) and (1.4)), it is plausible to conjecture that for the Euclidean and the maximum norms, the sharp estimate is of order $O(\sqrt{d})$. For the case of the Euclidean norm, it is mentioned in [23] that V. V. Grinberg proved the sharp bound of $\sqrt{d}$, although this has never been published (or verified) – and 25 years later, the statement was again referred to as a conjecture [14]. Bárány and Grinberg [23] also note that "from the point of view of applications, it would be interesting to know more about" the case of the ℓ_∞-norm.

Recall Theorem 1.2 from Section 1, which for a convex body $K \subset \mathbb{R}^d$ connects $\mathrm{vb}(B_2^d, K)$ to the Gaussian measure of K [9]. This result, as mentioned previously, was recently proven constructively using an algorithm called the *Gram-Schmidt walk* [11]. In this paper, Bansal et al. additionally prove the a colorful generalization of

Banaszczyk's result, albeit in a slightly different colorful setting.

Theorem 2.2 (Bansal, Dadush, Garg, Lovett [11]). *Let $V_1, \ldots, V_n \subseteq B_2^d$ be vector families with $0 \in \operatorname{conv} V_i$ for each $i \in [n]$. Then for any convex body K with $\gamma_d(K) \geq 1/2$, there exist vectors $v_i \in V_i$ such that $\sum_{i=1}^n v_i \in cK$, where $c > 0$ is an absolute constant.*

Note that the condition $0 \in \sum_{i \in [n]} \operatorname{Conv} V_i$ is weaker than requiring $0 \in \operatorname{Conv} V_i$ for each i – by applying a shift of each family, the more general estimate can be derived from the statement under this more restrictive condition, albeit with the loss of a factor 2 compared to the above bound.

Applying Theorem 2.2 to the Euclidean norm, one retrieves a sum of norm at most $C\sqrt{d}$ for some constant $C > 1$, and for the maximum norm one obtains a bound of $O(\sqrt{d \ln d})$ (the latter can also be obtained by a straightforward application of the probabilistic method). We note that their proof method, which is based on the techniques of Lovász, Spencer, and Vesztergombi [48], can be modified to show that the bound in the colorful setting is at most twice the original vector balancing constant, which implies $O(\sqrt{d})$ bounds for both the Euclidean and maximum norm. This asymptotically matches the estimates that we will prove, up to constants; for details, see Section 2.3. In this chapter, we instead provide a direct, constructive approach for proving asymptotically matching, yet tighter estimates for both the Euclidean and maximum norms, which also shed more light on the geometry of the problem and its algorithmic aspects. Our main results show bounds of $\sqrt{d}$ and $O(\sqrt{d})$ for the Euclidean and maximum norm cases, respectively, matching the tight (respectively, tight in order of magnitude) results in the vector balancing setting (see Theorems 1.2 and 1.3).

For an estimate in the dual direction in the Euclidean setting, the following result is well known: if $V_1, \ldots, V_n$ are sets of *unit* vectors with $0 \in \operatorname{Conv} V_i$ for each i, then one may select $v_i \in V_i$ for $i \in [n]$ so that $\left\| \sum_{i \in [n]} v_i \right\|_2 \geq \sqrt{n}$ (for a further generalization, see [1]).

The proofs of Theorems 1.2 and 1.3 are deferred to Chapter 3; in the remainder of this chapter we instead focus on several reductions of the colorful vector balancing problem that motivate our proof technique. In Section 2.2 we will prove an essential linear algebraic reduction of the colorful vector balancing problem that is the key to giving bounds independent of the number of vector families n. The tools developed in Section 2.2 form the basis for the proofs of Theorems 1.2 and 1.3 in Chapter 3. We will also address the above-mentioned reduction of the colorful vector balancing problem to the original vector balancing problem and highlight the advantages of our more geometric approach.

2.2 A Linear Algebraic Reduction

In this section we will use the method of linear dependencies to prove that the number of vector families can be reduced from n to at most d, and moreover that the total number of vectors can also be bounded from above. This approach dates back to the classical work of Shapley and Folkman, and Starr from the 1960's [70]. Several other applications of the method of linear dependencies are well surveyed by Bárány [22].

Recall the setting of the problem: we are given an origin-symmetric convex body B in $\mathbb{R}^d$ and vector families $V_1, ..., V_n \subseteq B$ such that $0 \in \sum_{i\in[n]} \operatorname{conv} V_i$, and our goal is to select vectors $v_i \in V_i$ for $i \in [n]$ such that $\left\| \sum_{i\in[n]} v_i \right\|_B$ is minimal (here $\|\cdot\|_B$ is the Minkowski norm associated to B, see Section 1.1 for the definition).

We first note that by Carathéodory's theorem we may assume that each family V_i is finite (in fact, $|V_i| \leq d+1$ for each $i \in [n]$). Indeed, as we assume that $0 \in \sum_{i\in[n]} \operatorname{conv} V_i$, we know that for each $i \in [n]$ there exists $x_i \in \operatorname{conv} V_i$ so that $x_1 + \cdots + x_n = 0$. By Carathéodory's theorem, for each $i \in [n]$ there exists (up to relabeling) $v_1, ..., v_{d+1} \in V_i$ so that $x_i \in \operatorname{conv}\{v_1, ..., v_{d+1}\}$, thus redefining the families this way still yields a collection of families satisfying the condition that $0 \in \sum_{i\in[n]} \operatorname{conv} V_i$, and without loss of generality we can assume that each family is finite. From now on we will make this assumption.

We identify a set of vectors $U = \{u_1, \ldots, u_m\} \subset \mathbb{R}^d$ with the $d \times m$ matrix

$$U = \begin{pmatrix} u_1 \cdots u_m \end{pmatrix}.$$

Definition 2.3. *Given vector families $V_1, \ldots, V_n \subset \mathbb{R}^d$ with $|V_i| = m_i$ and $\sum_{i\in[n]} m_i = m$, we define the associated* vector family matrix

$$V := \begin{pmatrix} V_1 | V_2 | \cdots | V_n \end{pmatrix} \in \mathbb{R}^{d\times m},$$

which is a partitioned matrix. We also introduce the associated set of convex coefficients

$$\Delta_V := \Delta^{m_1} \times \cdots \times \Delta^{m_n} \subset \mathbb{R}^m \tag{2.2}$$

which is a convex polytope arising as a direct product of simplices.

The relevance of Δ_V is shown by the fact that a vector $v \in \mathbb{R}^d$ is contained in

$\mathrm{Conv}V_i$ if and only if $v = V_i\lambda$ for some $\lambda \in \Delta^{m_i}$. Accordingly,

$$\mathrm{Conv}V_1 + \cdots + \mathrm{Conv}V_n = \{V\lambda :\ \lambda \in \Delta_V\}.$$

In the above scenario, we will usually consider $\mathbb{R}^m$ along with its orthogonal decomposition $\mathbb{R}^m = \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n}$. A collection of vector families $V = \{V_1, \ldots, V_n\}$ will always be identified with its associated vector family matrix V – using the same notation for these two will cause no ambiguity and will be clarified by the context. From now on, U, V and W will always stand for a collection of vector families or their associated vector family matrices.

Throughout this section, Greek letters will be used to denote vectors in the coefficient space $\Delta_V \subset \mathbb{R}^m$, while letters of the Latin alphabet will stand for vectors in $\mathbb{R}^d$. To make the connection between these spaces explicit, coefficient vectors $\beta \in \Delta_V$ will also be indexed by members of V_i as follows:

$$\beta = (\beta(v_i))_{v_i \in V_i,\ i \in [n]} \in \mathbb{R}^m. \tag{2.3}$$

Given a vector family matrix $V \in \mathbb{R}^{m\times d}$ and a set of indices $J \subset [m]$, we naturally define $V|_J$, the restriction of V to the columns indexed by elements of J. This is again a vector family matrix which naturally induces a collection of vector families, the restrictions of the original ones to J. Naturally, $\Delta_{V|_J} \subset \mathbb{R}^{|J|}$ is the set of convex coefficients associated to $V|_J$. By virtue of the indexing (2.3), we may also define the restriction to a subcollection $W \subset V$. In particular, for $\beta \in \Delta_V$ and $V_i \in V$, $\beta|_{V_i} \in \Delta_{V_i}$ consists of the coefficients of vectors in V_i.

Given a partition $I \dot\cup J = [m]$ and vectors $\lambda \in \Delta_{V|_I}$, $\mu \in \Delta_{V|_J}$, we introduce the natural concatenation of λ and μ by $\lambda \vee \mu \in \Delta_V$; that is, $(\lambda \vee \mu)|_I = \lambda$ and $(\lambda \vee \mu)|_J = \mu$.

Definition 2.4. *A number $x \in [0,1]$ is* fractional *if $x \notin \{0,1\}$. Given a vector $\beta \in \Delta_V$, we say that family V_i is* locked by *β if none of the coordinates of $\beta|_{V_i}$ are fractional. Otherwise family V_i is* free under *β. A vector $\beta \in \Delta_V$ is a* selection vector *if every family is locked by β, equivalently, β is a vertex of Δ_V.*

Note that for a selection vector $\beta \in \Delta_V$, $V_i\beta|_{V_i} = v_i$ for some $v_i \in V_i$ for each $i \in [n]$.

The main tool of the section is the following generalization of the Shapley-Folkman lemma [57, 70], a cornerstone result in econometric theory. Alternative versions were proved and used by Grinberg and Sevast'yanov [35] and Bárány and Grinberg [23].

Theorem 2.5. *Given a collection of vector families $V = \{V_1, \ldots, V_n\}$ in $\mathbb{R}^d$ with $0 \in \sum_{i\in[n]} \mathrm{Conv}V_i$, there exists a vector $\alpha \in \Delta_V$ such that* *(i)* $V\alpha = 0$;

(ii) *All but* $k \leq d$ *families* V_i *are locked by* α;

(iii) α *has at most* $k + d$ *fractional coordinates.*

The proof is based on the Shapley-Folkman–style statement below which is related to the geometry of basic feasible solutions of linear programs.

Lemma 2.6 ([70], [35]). *Let* K *be a polyhedron in* $\mathbb{R}^m$ *defined by a system*

$$f_i(x) = a_i, \ i = 1, \ldots, p,$$
$$g_j(x) \leq b_j, \ j = 1, \ldots, q,$$

where f_i, g_j *are linear functions. Let* x_0 *be a vertex of* K *and* $A = \{j : g_j(x_0) = b_j\}$. *Then* $|A| \geq m - p$.

Proof of Theorem 2.5. Given vector families $V_1, \ldots, V_n$ in $\mathbb{R}^d$ with $0 \in \sum_{i \in [n]} \mathrm{Conv} V_i$ and $m = \sum_{i \in [n]} |V_i|$, consider the set

$$\begin{aligned} P &= \big\{\lambda \in \Delta_V : \ V\lambda = 0\big\} \\ &= \Big\{\lambda \in \mathbb{R}^m : \ \sum_{i \in [n]} \sum_{v_i \in V_i} \lambda(v_i) v_i = 0, \ \sum_{v_i \in V_i} \lambda(v_i) = 1 \ \forall i \in [n], \\ &\qquad \lambda(v_i) \geq 0 \ \forall i \in [n], \forall v_i \in V_i\Big\}. \end{aligned} \tag{2.4}$$

By our assumption that $0 \in \sum_{i \in [n]} \mathrm{Conv} V_i$, P is a (non-empty) convex polytope in $\mathbb{R}^m$. Let $\alpha \in P$ be any extreme point of P. Define

$$S := \big\{i \in [n] : \ V_i \text{ is free under } \alpha\big\}$$

and let $k = |S|$. By Lemma 2.6, at most $n + d$ non-negativity inequalities in (2.4) are slack when substituting $\lambda = \alpha$. Each of the $n - k$ families locked by α contribute exactly one slack constraint, arising from the (unique) 1-coordinate. Let f denote the number of fractional coordinates of α; then $f + (n - k)$ is the total number of slack constraints. Thus

$$f + (n - k) \leq n + d$$

which implies that $f \leq k + d$. Since, by definition, $f \geq 2k$, this also shows that $k \leq d$. □

By virtue of allowing us to reduce consideration to at most d families, the following corollary is the main tool for proving upper bounds for the colorful vector balancing problem in arbitrary norms.

Corollary 2.7. *Let $\|\cdot\|$ be a norm on $\mathbb{R}^d$ with unit ball B. Suppose there exists a constant $C(d)$ such that given any collection of $k \leq d$ families $U = \{U_1, \dots, U_k\}$ in B satisfying $|U_1| + \cdots + |U_k| \leq k + d$, and any $\lambda \in \Delta_U$, there exists a selection vector $\mu \in \Delta_U$ such that*

$$\|V\lambda - V\mu\| \leq C(d).$$

Then given any collection of families $V_1, \dots, V_n \subseteq B$ with $0 \in \sum_{i\in[n]} \mathrm{Conv} V_i$, there exists a selection of vectors $v_i \in V_i$ for $i \in [n]$ such that

$$\Big\| \sum_{i\in[n]} v_i \Big\| \leq C(d).$$

Proof. Suppose that the hypothesis of the statement holds. Let $m := |V_1| + \cdots + |V_n|$. Applying Theorem 2.5 to $V = \{V_1, \dots, V_n\}$, we find $\alpha \in \Delta_V$ such that $V\alpha = 0$, all but $k \leq d$ families V_i are locked by α, and α has at most $k + d$ fractional coordinates. Let $F \subset [m]$ be the set of indices of fractional coordinates, and set $L = [m] \setminus F$. Then $|F| \leq k + d$. By hypothesis, there exists a selection vector $\mu \in \Delta_{V|_F}$ such that $\|V|_F \alpha|_F - V|_F \mu\| \leq C(d)$, and so

$$\|V|_L\alpha|_L + V|_F\mu\| \leq \|V|_L\alpha|_L + V|_F\alpha|_F\| + \| - V|_F\alpha|_F + V|_F\mu\| \leq \|V\alpha\| + C(d) = C(d).$$

Taking the selection of vectors given by $\alpha|_L \vee \mu$ completes the proof. □

2.3 A Reduction to Vector Balancing

In this section we describe an alternative approach for proving asymptotic estimates for colorful vector balancing constants matching Theorems 1.2 and 1.3, based on the proof techniques of Lovász, Spencer and Vesztergombi [48] and Bansal, Dadush, Garg, and Lovett [11] [1]. For the following proof we denote the colorful vector balancing constant of two symmetric convex bodies $K, L \subset \mathbb{R}^d$ as

$$\mathrm{colvb}(K, L) := \sup_{n\geq d} \max_{\substack{V_1,\dots,V_n\subseteq K \\ 0\in\sum \mathrm{Conv} V_i}} \min_{v_i\in V_i,\ i\in[n]} \|v_1 + \cdots + v_n\|_L.$$

[1] We thank the anonymous referee for pointing out the argument sketched in this section

Theorem 2.8. *Given any symmetric convex bodies $K, L \subset \mathbb{R}^d$,*

$$\mathrm{colvb}(K, L) \leq 2\mathrm{vb}(K, L).$$

Proof. We are given families $V_1, ..., V_n \subseteq K$ and a vector $\lambda \in \Delta_V$ so that $V\lambda = 0$. Note that by Carathéodory's theorem we may assume that $|V_i| \leq d + 1$ for each $i \in [n]$. Let

$$\rho = \max_{u \in K} \|u\|_L.$$

Fix $\varepsilon > 0$ and take $\ell \in \mathbb{Z}$ so that $n(d + 1)2^{-(\ell-1)}\rho \leq \varepsilon$.

Each coordinate $\lambda(v)$ of λ, for $v \in V_i, i \in [n]$, has a binary expansion, which we truncate at the ℓ^{th} digit after the radix point to obtain the vector μ with coordinates $\mu(v)$ so that $|\mu(v) - \lambda(v)| \leq 2^{-(\ell-1)}$ for each $v \in V_i$, $i \in [n]$. Then

$$\|V\lambda - V\mu\|_L = \Big\| \sum_{i\in[n]} \sum_{v\in V_i} (\lambda(v) - \mu(v))v \Big\|_L \leq \frac{1}{2^{\ell-1}} \sum_{i\in[n]} \sum_{v\in V_i} \|v\|_L \leq \frac{n(d + 1)\rho}{2^{\ell-1}} \leq \varepsilon. \quad (2.5)$$

Denote the j^{th} digit of the binary expansion of $\mu(v)$ by $\mu(v)^{(j)}$. We define the set

$$S_\ell := \{v \in \cup_{i\in[n]} V_i : \ \mu(v)^{(\ell)} = 1\}$$

to be the set of vectors in our collection for which the ℓ^{th} digit of the binary expansion of the corresponding coefficient is 1. Since $\sum_{v\in V_i} \lambda(v) = 1$ for each $i \in [n]$, it follows that $|S_\ell \cap V_i| = 2q_i$ for some $q_i \in \mathbb{Z}$, so we can write $S \cap V_i = \{v^i_1, ..., v^i_{2q_i}\}$ for each $i \in [n]$. We define the auxiliary collections of vectors

$$W_i = \left\{ \tfrac{v^i_{2j} - v^i_{2j-1}}{2} \right\}_{j\in[q_i]} \subseteq K$$

and then balance the collection $W = \cup_{i\in[n]} W_i$, yielding signs $\chi_i(j) \in \{\pm 1\}$ so that

$$\left\| \sum_{i\in[n]} \sum_{j\in[q_i]} \chi_i(j) \frac{v^i_{2j} - v^i_{2j-1}}{2} \right\|_L \leq \mathrm{vb}(K, L).$$

Color the elements of S_ℓ as follows: for each $i \in [n]$, for each $k \in [2q_i]$, we assign

$\beta_i(k) = \chi_i(j)$ for k even and $\beta_i(k) = -\chi_i(j)$ for k odd, so that

$$\left\| \sum_{i\in[n]} \sum_{k\in[2q_i]} \beta_i(k) v_k^i \right\|_L = 2 \left\| \sum_{i\in[n]} \sum_{j\in[q_i]} \chi_i(j) \frac{v_{2j}^i - v_{2j-1}^i}{2} \right\|_L \leq 2\mathrm{vb}(K, L). \tag{2.6}$$

We then update the vector μ as follows: for $v \notin S_\ell$, $\mu_1(v) = \mu(v)$. For $v \in S_\ell$, we know that $v = v_k^i$ for some $i \in [n], k \in [2q_i]$, and we update $\mu_1(v) := \mu(v) + 2^{-\ell}\beta_i(k)$. By construction, $\mu_1 \in \Delta_V \cap 2^{-(\ell-1)}$, and $\|V\mu_1\| \leq 2^{-(\ell-1)}\mathrm{vb}(K, L)$. Iterating the argument for the successive digits leads to a selection vector μ_ℓ for which

$$\|V\lambda\|_L \leq \|V\lambda - V\mu\|_L + \|V\mu\|_L \leq \left\| \sum_{i=0}^{\ell-1} 2^i \mathrm{vb}(\mathrm{K}, \mathrm{L}) \right\|_L \leq \varepsilon + 2\mathrm{vb}(K, L).$$

As $\varepsilon > 0$ was arbitrary, the theorem follows. □

Combining Theorem 2.8 with (1.2) and (1.4) implies our Theorems 1.2 and 1.3 up to constants. We intended to give a direct proof that is more suited to algorithmic applications. Indeed, the computational complexity of finding a solution by the above approach depends heavily on the number of vector families n, whereas our technique illuminates the geometric aspects of the problem and the independence of the number of vector families, including the reduction to $O(d)$ total vectors that is necessary in the maximum norm case. Moreover, it leads to the sharp bound of $\sqrt{d}$ for the Euclidean case as opposed to the asymptotic bound above, and it improves on the constant for the maximum norm given by combining [47] with Theorem 2.3.

2.4 Conclusion

In this chapter we introduced the colorful vector balancing problem as a geometric generalization of the vector balancing problem and discussed its history and existing results. We further used the method of linear dependencies to reduce the number of vector families from n to d, which will be a key tool in our proofs of Theorems 1.2 and 1.3. Finally, we discussed an alternate method of providing bounds on the colorful vector balancing problem by reducing it to the original vector balancing problem, and highlight the benefits of our more direct geometric approach.

Chapter 3

Colorful Vector Balancing: the Euclidean and Maximum Norms

In this chapter we present the proofs of Theorems 1.2 and 1.3. In Section 3.1 we prove Theorem 1.2, and in Section 3.2 we prove Theorem 1.3. Finally, in Section 3.3 we give a detailed proof the our *Skeleton Approximation Lemma*, a technical result that forms the backbone of the proof of Theorem 1.3.

3.1 The Euclidean Norm

To prove Theorem 1.2, we will prove the following vertex approximation property for color classes. Combining Proposition 3.1 with Theorem 2.5 will yield Theorem 1.2.

Proposition 3.1 (Colorful vertex approximation in Euclidean norm)**.** *Given a collection of k vector families $U = \{U_1, \dots, U_k\}$ in B_2^d and any point $\lambda \in \Delta_U$, there exists a selection vector $\mu \in \Delta_U$ such that $\|U\lambda - U\mu\|_2 \le \sqrt{k}$.*

Our proof is inspired by Spencer's argument for the vector balancing case [67] (in particular, Proposition 3.1 generalizes the Lemma in [67], see also Theorem 4.1 of [22]), and it works in any finite dimensional Hilbert space.

Proof of Proposition 3.1. Define $x := U\lambda \in \mathrm{Conv}U_1 + \cdots + \mathrm{Conv}U_k$, so that

$$x = x_1 + \cdots + x_k, \quad x_i = \sum_{u_i \in U_i} \lambda(u_i) u_i \; \forall i \in [k],$$

where $\lambda|_{U_i} \in \Delta_{U_i}$ for each $i \in [k]$. We define a vector-valued random variable $w_i \in \mathbb{R}^d$ for each $i \in [k]$, which takes the value u_i with probability $\lambda(u_i)$ for each $u_i \in U_i$, independently of the other w_j's, $j \in [k] \setminus \{i\}$. Then

$$\mathbb{E}[w_1 + \cdots + w_k] = \sum_{i\in[k]} \mathbb{E}[w_i] = \sum_{i\in[k]} \sum_{u_i \in U_i} \lambda(u_i) U_i = \sum_{i\in[k]} x_i = x.$$

Component-wise this yields

$$\mathbb{E}\big[w_1^{(\ell)} + \cdots + w_k^{(\ell)} - x^{(\ell)}\big] = 0, \quad \ell \in [d]. \tag{3.1}$$

For each $\ell \in [d]$, (3.1) and the independence of the w_i's imply

$$\begin{aligned}\mathbb{E}\big[(w_1^{(\ell)} + \cdots + w_k^{(\ell)} - x^{(\ell)})^2\big] &= \mathbb{E}\big[(w_1^{(\ell)} + \cdots + w_k^{(\ell)} - x^{(\ell)})^2\big] \\ &\quad - \mathbb{E}\big[w_1^{(\ell)} + \cdots + w_k^{(\ell)} - x^{(\ell)}\big]^2 \\ &= \operatorname{Var}\big[w_1^{(\ell)} + \cdots + w_k^{(\ell)} - x^{(\ell)}\big] \\ &= \sum_{i\in[k]} \operatorname{Var}\big[w_i^{(\ell)}\big].\end{aligned} \tag{3.2}$$

Since

$$\|w_1 + \cdots + w_k - x\|_2^2 = \sum_{\ell=1}^{d} \big((w_1^{(\ell)} + \cdots + w_k^{(\ell)}) - x^{(\ell)}\big)^2,$$

by linearity of expectation and (3.2) we conclude

$$\begin{aligned}\mathbb{E}\big[\|w_1 + \cdots + w_k - x\|^2\big] &= \sum_{\ell\in[d]} \mathbb{E}\big[(w_1^{(\ell)} + \cdots + w_k^{(\ell)} - x^{(\ell)})^2\big] \\ &= \sum_{\ell\in[d]} \sum_{i\in[k]} \operatorname{Var}\big[w_i^{(\ell)}\big] \\ &= \sum_{\ell\in[d]} \sum_{i\in[k]} \mathbb{E}\big[(w_i^{(\ell)})^2\big] - \sum_{\ell\in[d]} \sum_{i\in[k]} \mathbb{E}\big[w_i^{(\ell)}\big]^2 \\ &= \sum_{i\in[k]} \mathbb{E}\big[\|w_i\|_2^2\big] - \sum_{\ell\in[d]} \sum_{i\in[k]} \mathbb{E}\big[w_i^{(\ell)}\big]^2.\end{aligned} \tag{3.3}$$

Finally, we note that

$$\mathbb{E}\big[\|w_i\|_2^2\big] = \sum_{u_i\in U_i} \lambda(u_i) \cdot \|u_i\|_2^2 \le \sum_{u_i \in U_i} \lambda(u_i) = 1,$$

hence continuing calculation (3.3),

$$\mathbb{E}\big[\|w_1 + \cdots + w_k - x\|^2\big] \le k - \sum_{\ell\in[d]} \sum_{i\in[k]} \mathbb{E}\big[w_i^{(\ell)}\big]^2 \le k.$$

It follows that for some specific choice of $u_i \in U_i$, $i \in [k]$, we have

$$\|u_1 + \cdots + u_k - x\|_2^2 \leq k.$$

The corresponding selection vector $\mu \in \Delta_U$ satisfies the proposition. □

Theorem 1.2 now follows immediately from Corollary 2.7 and Proposition 3.1.

3.2 The Maximum Norm

To prove Theorem 1.3, we need to show that the vertex approximation property (the analogue of Proposition 3.1) holds for the maximum norm. This result for the original vector balancing problem is due to Spencer [68] and, independently, Gluskin [33]. As in the original vector balancing problem, the challenge is to remove the $\sqrt{\ln d}$ factor. Note that, unlike in the Euclidean case, we need to set an upper bound on the total cardinality of the vector systems.

Proposition 3.2 (Colorful vertex approximation in Maximum norm). *Given a collection of k vector families $U = \{U_1, \ldots, U_k\}$ in B_∞^d satisfying $m := |U_1| + \cdots + |U_k| \leq 2d$, and an arbitrary point $\lambda \in \Delta_U$, there exists a selection vector $\mu \in \Delta_U$ such that $\|U\lambda - U\mu\|_\infty \leq C\sqrt{d}$ for a universal constant $C > 0$.*

Note that applying the probabilistic method directly with the union bound over coordinates results in the weaker upper bound of $O(\sqrt{d}\sqrt{\ln d})$. Thus, in order to reach the bound of $O(\sqrt{d})$, one must apply an alternative argument, just as in the case of the original vector balancing problem in the maximum norm. In [68], Spencer utilized a partial coloring method in order to overcome this difficulty. This technique and the algorithmic argument of Lovett and Meka [47] are the predecessors of our approach described below.

We will prove Proposition 3.2 by iterating the following lemma, which is a close relative of the Partial Coloring Lemma in [47]. We call it the *skeleton approximation lemma*, as it approximates a point in the set of convex coefficients $\Delta_W \subset \mathbb{R}^m$ by a point on the $(m/2)$-skeleton of Δ_W.

Lemma 3.3 (Skeleton Approximation). *Let $W = \{W_1, \ldots, W_k\}$ be a collection of vector families in B_∞^d with $|W_i| \geq 2$ for each i and $m := |W_1| + \cdots + |W_k| \leq 2d$. Then for any point $\lambda \in \Delta_W$, there exists $\mu \in \Delta_W$ such that* *(i)* $\|W\lambda - W\mu\|_\infty \leq \eta\sqrt{m \ln \frac{\xi d}{m}}$ *where* η, ξ *are constants specified as*

$$\eta = \frac{7}{3} \text{ and } \xi = 18; \tag{3.4}$$

(ii) $\mu^{(i)} = 0$ *for at least* $m/2$ *indices* $i \in [m]$.

The proof of Lemma 3.3 is postponed to Section 3.3. We now deduce Proposition 3.2 assuming Lemma 3.3.

Proof of Proposition 3.2. We may assume that $|U_i| \geq 2$ for each i, since any convex coefficient vector corresponding to a 1-element family is necessarily a selection vector.

By an inductive process, we are going to define points $\lambda(s) \in \Delta_U$, sets of indices $F(s), L(s) \subset [m]$, and cardinalities $m(s)$ for $s = 0, 1, \dots$ so that for a suitably large S, $\lambda(S)$ is a selection vector with the desired properties. To initiate the recursive process, take $\lambda(0) = \lambda$, let $F(0) \subset [m]$ be the set of indices of fractional coordinates of $\lambda(0)$, and $L(0) = [m] \setminus F(0)$ be the set of indices of coordinates of $\lambda(0)$ equal to 0 or 1. Introducing $m(0) = |F(0)|$, we have $m(0) \leq m \leq 2d$.

Assuming that iterative step s has been taken, we define step number $s + 1$ as follows. Apply Lemma 3.3 to the vector family matrix $U(s) := U|_{F(s)}$ of total cardinality $m(s) \leq m$ and the point $\lambda(s)|_{F(s)} \in \Delta_{U(s)}$ to find $\mu(s+1) \in \Delta_{U(s)}$ with the prescribed properties. Define $\lambda(s+1) = \lambda(s)|_{L(s)} \vee \mu(s+1)$ to be natural concatenation of these two vectors, obtained by replacing the fractional coordinates of $\lambda(s)$ by the approximating vector $\mu(s+1)$. Let $F(s+1) \subset [m]$ be the set of indices of fractional coordinates of $\lambda(s+1)$, $L(s+1) = [m] \setminus F(s+1)$, and set $m(s+1) = |F(s+1)|$.

By Property (i) of Lemma 3.3 and the definition of $\lambda(s+1)$, for each $s \geq 0$

$$\|U\lambda(s) - U\lambda(s+1)\|_\infty \leq \eta\sqrt{m(s) \ln \tfrac{\xi d}{m(s)}}. \tag{3.5}$$

Also, by property (ii) of Lemma 3.3 we have that

$$m(s) \leq \frac{m}{2^s} \leq \frac{d}{2^{s-1}}. \tag{3.6}$$

Since $m(s) \in \mathbb{N}$, this also yields that after a finite number S of steps, $m(S) = 0$ will hold. Set $\mu = \lambda(S)$. We will show that μ fulfills the criteria of Proposition 3.2.

That $\mu \in \Delta_U$ is a selection vector is shown by $m(S) = 0$. To show the approximation property, note that the function $f(x) = x\ln(1/x)$ is increasing on the interval $[0, 1/4]$.

Combined with (3.5),(3.6), and (3.4), this yields that

$$
\begin{aligned}
\|U\lambda - U\mu\|_\infty &\le \sum_{s=0}^{S-1} \|U\lambda(s) - U\lambda(s+1)\|_\infty \\
&\le \sum_{s=0}^{S-1} \eta\sqrt{\ln\frac{\xi d}{m(s)}}\sqrt{m(s)} \\
&\le \sum_{s=0}^{S-1} \eta\sqrt{\ln\frac{\xi d}{d/2^{s-1}}}\sqrt{\frac{d}{2^{s-1}}} \qquad (3.7)\\
&\le \eta\sqrt{d}\sum_{s=0}^{\infty} 2^{-(s-1)/2}\sqrt{\ln(\xi)+\ln 2\cdot(s-1)} \\
&< 22\sqrt{d}. \qquad \square
\end{aligned}
$$

As in the Euclidean case, Theorem 1.3 now follows from Corollary 2.7 and Proposition 3.2. A simple modification of the proof yields the following version of Proposition 3.2, which provides a significant strengthening for $m \ll 2d$.

Proposition 3.4. *Given a collection of vector families $U = \{U_1, \dots, U_k\}$ in B_∞^d such that $m = |U_1| + \dots + |U_k| \le 2d$ and any point $\lambda \in \Delta_U$, there exists a selection vector $\mu \in \Delta_U$ such that*

$$\|U\lambda - U\mu\|_\infty \le K\sqrt{m}\sqrt{\ln \tfrac{18d}{m}}$$

for a universal constant $K > 0$.

Proof. Take $\mu \in \Delta_V$ as in the proof of Proposition 3.2. Then, substituting (3.6) in (3.7),

$$
\begin{aligned}
\|U\lambda - U\mu\|_\infty &\le \sum_{s=0}^{S-1} \eta\sqrt{\ln\frac{\xi d}{m/2^{s-1}}}\sqrt{\frac{m}{2^{s-1}}} \\
&\le \eta\sqrt{m}\sum_{s=0}^{\infty} \frac{\sqrt{\ln\left(\frac{\xi d}{m}\right)+\ln 2\cdot s}}{\sqrt{2^{s-1}}} \\
&\le \eta\sqrt{m}\sqrt{\ln\tfrac{\xi d}{m}}\sum_{s=0}^{\infty}\frac{\sqrt{1+s}}{\sqrt{2^{s-1}}} \\
&< 9\eta\sqrt{m}\sqrt{\ln\tfrac{18d}{m}}. \qquad \square
\end{aligned}
$$

3.3 The Skeleton Approximation Lemma

Proving Lemma 3.3, the Skeleton Approximation Lemma, requires several standard facts about the behavior of Gaussian random variables. By $\mathcal{N}(\mu, \sigma^2)$ we denote the (1-dimensional) Gaussian distribution with mean μ and variance σ^2. Given a linear subspace $A \subseteq \mathbb{R}^d$, $\mathcal{N}(A)$ denotes the standard multi-dimensional Gaussian distribution on A, i.e. for $G \sim \mathcal{N}(A)$, $G = G_1 a_1 + \cdots + G_m a_m$, where $\{a_1, \ldots, a_m\}$ is any orthonormal basis of A and $G_1, \ldots, G_m \sim \mathcal{N}(0,1)$ are independent Gaussian random variables (for further details, see [20, 29]).

Lemma 3.5. *Let $A \subseteq \mathbb{R}^d$ be a linear subspace with $G \sim \mathcal{N}(A)$. Then given any $u \in \mathbb{R}^d$, $\langle G, u \rangle \sim \mathcal{N}(0, \sigma^2)$, with $\sigma^2 = \|P_A(u)\|^2 \leq \|u\|_2^2$, where $P_A(\cdot)$ denotes the orthogonal projection onto A.*

Corollary 3.6. *Let $A \subseteq \mathbb{R}^d$ be a linear subspace with $G \sim \mathcal{N}(A)$ and define σ_i by $\langle G, e_i \rangle \sim \mathcal{N}(0, \sigma_i^2)$. Then $\sum_{i \in [d]} \sigma_i^2 = \dim A$.*

A proof of Lemma 3.5 can be found in [29, Section III.6] (see also [47]). These results are particularly useful when combined with the following standard tail estimate.

Lemma 3.7. *Given a Gaussian random variable $G \sim \mathcal{N}(\mu, \sigma^2)$, for all $t > 0$,*

$$\mathbb{P}\big[|G - \mu| \geq t\big] \leq \exp\big(-t^2/2\sigma^2\big).$$

This result is a special case of the general version of Hoeffding's inequality (for a proof see e.g. [76]). We will also need a similar bound for *martingales* with Gaussian steps. Recall that a sequence $\{X_i\}_{i \in \mathbb{N}}$ of real-valued random variables is a martingale if $\mathbb{E}[X_{n+1} | X_1, \ldots, X_n] = X_n$.

Lemma 3.8 ([10])**.** *Let $0 = X_0, X_1, \ldots, X_T$ be a martingale in $\mathbb{R}$ with steps $Y_i = X_i - X_{i-1}$ for $i \geq 1$. Suppose that for all $i \in [T]$, $Y_i | X_0, \ldots, X_{i-1}$ is a Gaussian random variable with mean zero and variance at most σ^2. Then for any $c > 0$,*

$$\mathbb{P}\big[|X_T| \geq \sigma c \sqrt{T}\big] \leq 2\exp(-c^2/2).$$

Finally, we will need the following result about sequences of Gaussian random variables. This is a well-known result that can be found for example in [76]; we provide the standard proof for the reader's convenience.

Lemma 3.9. *Let $X_i \sim \mathcal{N}(0, \sigma_i^2)$ with $\sigma_i \leq 1$ for $i = 1, 2, \ldots$ be a sequence of not necessarily independent, jointly Gaussian random variables. Then for any $T \geq 2$,*

$$\mathbb{E} \max_{i \leq T} |X_i| \leq 6\sqrt{\ln T}$$

and

$$\mathbb{E} \max_{i \leq T} |X_i|^2 \leq 10 \ln T.$$

Proof. We define the random variable $Y := \max_{i\in\mathbb{N}} \frac{|X_i|}{\sqrt{1+\ln i}}$. Then by Lemma 3.7, the union bound, and the fact that $\sigma_i \leq 1$ for all i,

$$\begin{aligned}
\mathbb{E}[Y] &= \int_0^\infty \mathbb{P}[Y \geq y] dy \\
&= \int_0^2 \mathbb{P}[Y \geq y] dy + \int_2^\infty \mathbb{P}[Y \geq y] dy \\
&\leq 2 + \int_2^\infty \mathbb{P}\Big[\max_{i\in\mathbb{N}} \tfrac{|X_i|}{\sqrt{1+\ln i}} \geq y\Big] dy \\
&\leq 2 + \int_2^\infty \sum_{i=1}^\infty \mathbb{P}\Big[|X_i| \geq y\sqrt{1+\ln i}\Big] dy \\
&\leq 2 + \int_2^\infty \sum_{i=1}^\infty \exp\big(-y^2(1+\ln i)/2\sigma_i^2\big) dy \\
&\leq 2 + \int_2^\infty \left(\sum_{i=1}^\infty i^{-y^2/2}\right) \exp(-y^2/2) dy \\
&\leq 2 + \tfrac{\pi^2}{6} \cdot 0.06 < 3.
\end{aligned} \tag{3.8}$$

Finally, note that $\sqrt{1+\ln i} \leq \sqrt{1+\ln T}$ for all $i \in [T]$, hence the calculation in (3.8) yields

$$\mathbb{E} \max_{i \leq T} \tfrac{|X_i|}{\sqrt{1+\ln T}} \leq \mathbb{E} \max_{i \leq T} \tfrac{|X_i|}{\sqrt{1+\ln i}} \leq \mathbb{E} \max_{i\in\mathbb{N}} \tfrac{|X_i|}{\sqrt{1+\ln i}} < 3.$$

Then for $T \geq 2$, $\mathbb{E} \max_{i \leq T} |X_i| < 3\sqrt{1+\ln T} \leq 6\sqrt{\ln T}$. The proof for $\max_{i\leq T} |X_i|^2$ follows from an analogous calculation. $\square$

We complete the proof of Theorem 1.3 by proving the crux of the argument, Lemma 3.3. This will be done by means of providing an algorithm that proves the following slightly weaker statement.

Lemma 3.10. *Let $0.01 > \delta > 0$ be arbitrary, and let $W = \{W_1, \ldots, W_k\}$ be a collection of vector families in B_∞^d which satisfies that $|W_i| \geq 2$ for each i, and $m := \sum_{i\in[k]} |W_i| \leq 2d$.*

Define

$$\omega(m) := \eta\sqrt{m \ln \frac{\xi d}{m}} \tag{3.9}$$

where $\eta = \frac{7}{3}$ *and* $\xi = 18$ *as in* (3.4)*. Then for any* $\gamma \in \Delta_W$ *there exists* $\hat{\gamma} \in \Delta_W$ *such that*

(i) $\|W\gamma - W\hat{\gamma}\|_\infty \leq \omega(m)$*;*

(ii) $\hat{\gamma}^{(i)} \leq \delta$ *for at least* $m/2$ *indices* $i \in [m]$.

Lemma 3.3 follows immediately from Lemma 3.10 by standard compactness arguments.

Proof of Lemma 3.10. For each $j \in [d]$, let $W^j \in \mathbb{R}^m$ denote the jth row of the vector family matrix W. The condition $W \subset B^d_\infty$ ensures that $\|W^j\|_\infty \leq 1$ for each $j \in [d]$. Accordingly,

$$\|W^j\|_2^2 \leq m \tag{3.10}$$

for each j.

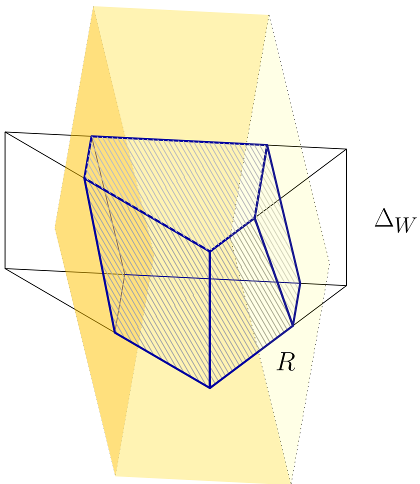

Figure 3.1: *The polytope* R

Consider the polytope

$$R := \Big\{\alpha \in \mathbb{R}^m : \ \alpha \in \Delta_W, \ \|W\alpha - W\gamma\|_\infty \leq \omega(m)\Big\},$$

which is the intersection of Δ_W with d slabs of width $\omega(m)/\|W^j\|_\infty$, $j \in [d]$ (see Figure 3.1). Equivalently, R is defined by the following set of linear equations and inequalities:

$$R = \Big\{ \alpha \in \mathbb{R}^m : \sum_{w_\ell \in W_\ell} \alpha(w_\ell) = 1 \, \forall \ell \in [k],\ \alpha^{(i)} \geq 0 \, \forall i \in [m],\ |\langle \alpha - \gamma, W^j \rangle| \leq \omega(m)\ \forall j \in [d] \Big\}. \tag{3.11}$$

We call the first and second set of constraints *convexity constraints*, as they ensure that $W\alpha \in \mathrm{Conv}(W_1) + \cdots + \mathrm{Conv}(W_k)$ for each $\alpha \in R$. The third set of constraints will be referred to as *maximum constraints*, as they imply that given $\alpha \in R$,

$$\|W\alpha - W\gamma\|_\infty = \max_{j \in [d]} |(W\alpha - W\gamma)_j| = \max_{j \in [d]} |\langle \alpha - \gamma, W^j \rangle| \leq \omega(m).$$

Let Z be the set of normal vectors of the inequality constraints in (3.11):

$$Z = \big\{ e_1, \ldots, e_m, W^1, \ldots, W^d \big\}. \tag{3.12}$$

By the previous remarks, $\|Z\|_\infty = 1$.

The main tool of the argument is to introduce a suitable discrete time Gaussian random walk on $\mathbb{R}^m$, similar to that in [10] and [47]. In order to help the reader navigate through the forthcoming technical details, we first give an intuitive description of the walk, whose position at time $t = 0, 1, \ldots$ will be denoted by $\Gamma_t \in \mathbb{R}^m$.

The walk starts from $\Gamma_0 = \gamma$ and runs in $\mathrm{aff}\, R$ with sufficiently small Gaussian steps as long as Γ_t is in the interior of R, far from its boundary. As Γ_t gets δ-close to crossing a facet of R, we confine the walk to an affine subspace parallel to that facet for the subsequent steps, by intersecting the current range with a hyperplane parallel to the facet. In particular, if any coordinate of Γ_t reaches a value less than δ, we freeze that coordinate for the remainder of the walk.

We show that running the walk long enough, until say time T, with high probability at least half of the coordinates of Γ_T become frozen, while $\Gamma_T \in R$ still holds. This will mean that $\hat{\gamma} = \Gamma_T$ satisfies the criteria of Lemma 3.10. For the proof it is essential that the value of $\omega(m)$ is carefully set (3.9), hence the slabs defining R are sufficiently wide so that the walk is unlikely to escape from them.

Let us turn to the formal definition of the random walk. Let $\varepsilon > 0$ and $T \in \mathbb{N}$ be parameters to be defined later. Define the sets

$$C_t^{conv} := \big\{ i \in [m] : \ \Gamma_t^{(i)} \leq \delta \big\}, \qquad C_t^{max} := \big\{ j \in [d] : \ |\langle \Gamma_t - \Gamma_0, W^j \rangle| \geq \omega(m) - \delta \big\} \tag{3.13}$$

to be the convexity and maximum constraints, respectively, that are at most δ-close to being violated by Γ_t. We will say that coordinate i is *frozen* iff $i \in C_t^{conv}$. Recall that by (2.3), coordinates may be indexed by the vectors, that is, for each $i \in [m]$, $\Gamma_t^{(i)} = \Gamma_t(w_l)$ for some $l \in [k]$ and $w_l \in W_l$. In that case, coordinate w_l is frozen iff $i \in C_t^{conv}$.

Let A be the linear component of $\operatorname{aff} \Delta_W$, that is, $A = \operatorname{lin}(\Delta_W - \Delta_W)$. For each $t \leq 1$, step t is confined to occur in the linear subspace

$$S_t := \left\{\beta \in A : \ \beta^{(i)} = 0 \ \forall i \in C_{t-1}^{conv}, \ \langle \beta - \Gamma_0, W^j \rangle = 0 \ \forall j \in C_{t-1}^{max} \right\}$$

by taking a Gaussian step $\Lambda_t \sim \mathcal{N}(S_t)$ and defining

$$\Gamma_t = \Gamma_{t-1} + \varepsilon \Lambda_t.$$

The walk terminates after T steps: $\hat{\gamma} := \Gamma_T$, where T is to be determined later.

We will show that with certain restrictions on the parameters, Γ_T satisfies properties (i) and (ii) of Lemma 3.10 with probability at least 0.2.

Given $\varepsilon > 0$, we define

$$T := \left\lfloor \frac{0.99^2 \eta^2}{2\varepsilon^2} \right\rfloor. \tag{3.14}$$

Choose $\varepsilon > 0$ small enough so that the following inequalities hold simultaneously:

$$6Td \exp\left(-\frac{\delta^2}{2m\varepsilon^2}\right) < 0.01, \tag{3.15}$$

$$22\varepsilon m^2 \ln T \leq 0.01, \tag{3.16}$$

and

$$10\varepsilon^2 \ln T \leq 1. \tag{3.17}$$

This can indeed be guaranteed since the functions $\exp(-x)/x$, $x \ln \frac{1}{x^2}$ and $x \ln \frac{1}{x}$ converge to 0 as $x \searrow 0$.

We summarize a few useful properties of the random walk.

Lemma 3.11. *Let $\Gamma_0, \ldots, \Gamma_T$ be the steps of the Gaussian random walk defined above and $i \in [m]$, $j \in [d]$. Then:*

(i) Given Γ_{t-1}, $\mathbb{E}[\Lambda_t] = 0$.

(ii) C_t^{conv}, C_t^{max} are nested increasing sets in t.

(iii) S_t is a nested decreasing sequence of linear subspaces in $\mathbb{R}^m$ in t.

(iv) *At any time* $0 \le t \le T$ *and for any* $i \in [k]$, $\sum_{w_i \in W_i} \Gamma_t(w_i) = 1$.

(v) *If the walk leaves the polytope* R *at time* $t \in [T]$, *then* $\Gamma_s \notin R$ *for any* $s \ge t$.

(vi) *If the walk leaves the polytope* R *at time* $t \in [T]$, *then* $|\langle \Lambda_t, z\rangle| \ge \delta/\varepsilon$ *for some* $z \in Z$.

(vii) *If coordinate* i *is frozen at step* t, *that is* $i \in C_t^{conv} \setminus C_{t-1}^{conv}$, *then* $\Gamma_T^{(i)} = \Gamma_t^{(i)} \ge \delta - \varepsilon\big|\Lambda_t^{(i)}\big|$.

Proof. Properties (i)-(v) are straightforward consequences of the definition of Γ_t.

To prove (vi), suppose that the walk leaves the polytope R at time t. Then an inequality constraint in (3.11) with normal vector $z \in Z$ is violated at time t. Suppose that $z = W^j$ for some $j \in [d]$. Since $j \notin C_{t-1}^{max}$,

$$|\langle \Gamma_{t-1} - \Gamma_0, W^j\rangle| < \omega(m) - \delta,$$

while on the other hand,

$$|\langle \Gamma_t - \Gamma_0, W^j\rangle| > \omega(m).$$

Combining these inequalities shows that

$$\varepsilon|\langle \Lambda_t, W^j\rangle| = |\langle \Gamma_t - \Gamma_{t-1}, W^j\rangle| \ge \delta.$$

The proof when $z = e_i$ for $i \in [m]$ is analogous.

To prove (vii), note that $i \notin C_{t-1}^{conv}$ implies that $\Gamma_{t-1}^{(i)} \ge \delta$. Therefore,

$$\Gamma_T^{(i)} = \Gamma_t^{(i)} = \Gamma_{t-1}^{(i)} + \varepsilon\Lambda_t^{(i)} \ge \delta + \varepsilon\Lambda_t^{(i)} \ge \delta - \varepsilon|\Lambda_t^{(i)}|. \qquad \square$$

Equipped with these properties, we are ready to prove that Γ_T satisfies the required conditions of Lemma 3.10 with probability at least 0.01. To show (i), that is

$$\|W\Gamma_0 - W\Gamma_T\|_\infty \le \omega(m), \tag{3.18}$$

it is sufficient to argue that (with high probability) $\Gamma_T \in R$; that is, the walk does not leave the polytope R at any step.

Define the event $E_t := \{\Gamma_t \notin R | \ \Gamma_0, \dots, \Gamma_{t-1} \in R\}$ that the walk steps out of R at time t. If E_t occurs, then by Lemma 3.11(vi), $|\langle \Lambda_t, z\rangle| \ge \delta/\varepsilon$ for some $z \in Z$. By Lemma 3.5 and (3.10), for any $z \in Z$, $\langle \Lambda_t, z\rangle$ is a Gaussian random variable with mean 0 and variance $\sigma^2 \leq m$. Applying Lemma 3.7 to $\langle \Lambda_t, z\rangle$, we find

$$\mathbb{P}\big[|\langle \Lambda_t, z\rangle| \geq \tfrac{\delta}{\varepsilon}\big] \leq 2\exp\big(-\big(\tfrac{\delta}{\varepsilon}\big)^2/2m\big). \tag{3.19}$$

Using the union bound, equations (3.12), (3.19), and the fact that $d \leq 2d$, we derive

$$\begin{aligned}
\mathbb{P}[\exists t \in [T] :\ \Gamma_t \notin R] &= \sum_{t=1}^{T} \mathbb{P}[E_t] \\
&\leq \sum_{t=1}^{T} \sum_{z \in Z} \mathbb{P}\big[|\langle \Lambda_t, z\rangle| \geq \tfrac{\delta}{\varepsilon}\big] \\
&\leq 2T(d+m)\exp\big(-\big(\tfrac{\delta}{\varepsilon}\big)^2/2m\big) \\
&\leq 6Td\exp\Big(-\frac{\delta^2}{2m\varepsilon^2}\Big) \\
&< 0.01
\end{aligned} \tag{3.20}$$

by condition (3.15). This proves that (3.18) holds with probability at least 0.99.

It remains to address (ii) of Lemma 3.10, that (with positive probability) $\Gamma_T^{(i)} \leq \delta$ for at least $m/2$ indices $i \in [m]$. We will reach this by means of proving that

$$\mathbb{E}[|C_T^{conv}|] > 0.51m. \tag{3.21}$$

To this end we derive the following identity, using Lemma 3.11(i):

$$\begin{aligned}
\mathbb{E}\big[\|\Gamma_t\|_2^2\big] &= \mathbb{E}\big[\|\Gamma_{t-1} + \varepsilon\Lambda_t\|_2^2\big] \\
&= \mathbb{E}\big[\|\Gamma_{t-1}\|_2^2\big] + \varepsilon^2\mathbb{E}\big[\|\Lambda_t\|_2^2\big] + 2\varepsilon\mathbb{E}\big[\langle \Gamma_{t-1}, \Lambda_t\rangle\big] \\
&= \mathbb{E}\big[\|\Gamma_{t-1}\|_2^2\big] + \varepsilon^2\mathbb{E}\big[\dim(S_t)\big],
\end{aligned}$$

where in the last equation we use that, by Corollary 3.6,

$$\mathbb{E}\big[\|\Lambda_t\|_2^2\big] = \mathbb{E}\big[\sum_{i\in[m]} \langle \Lambda_t, e_i\rangle^2\big] = \sum_{i\in[m]} \mathbb{E}\big[\langle \Lambda_t, e_i\rangle^2\big] = \dim S_t.$$

Iterating this calculation and using Lemma 3.11(iii),

$$\mathbb{E}\big[\|\Gamma_t\|_2^2\big] \geq \varepsilon^2 \sum_{t\in[T]} \mathbb{E}\big[\dim(S_t)\big] \geq T\varepsilon^2\mathbb{E}\big[\dim S_T\big] = T\varepsilon^2\,\mathbb{E}\big[m - |C_T^{conv}| - |C_T^{max}|\big],$$

and rearranging yields

$$\mathbb{E}\big[|C_T^{conv}|\big] \geq m - \frac{\mathbb{E}\big[\|\Gamma_T\|_2^2\big]}{T\varepsilon^2} - \mathbb{E}\big[|C_T^{max}|\big]. \tag{3.22}$$

The above identity allows us to prove (3.21) by giving upper estimates on $\mathbb{E}\big[\|\Gamma_T\|_2^2\big]$ and $\mathbb{E}\big[|C_T^{max}|\big]$.

We start with the second of these and show that

$$\mathbb{E}\big[|C_T^{max}|\big] \leq \tfrac{2m}{\xi}. \tag{3.23}$$

To this end, we bound the probability that the walk gets close to escaping from a given slab. Note that for fixed $j \in [d]$, $\{\langle \Gamma_t - \Gamma_0, W^j\rangle\}_{t\in[T]}$ for $0 \leq t \leq T$ is a martingale satisfying the conditions of Lemma 3.8. As the step size is $\varepsilon\langle \Lambda_t, W^j\rangle$, by Lemma 3.5 the variance of any step is bounded by $\varepsilon^2\|W^j\|_2^2 \leq \varepsilon^2 m$ (cf. (3.10)).

For any $j \in C_T^{max}$, by (3.13),

$$|\langle \Gamma_T - \Gamma_0, W^j\rangle| \geq \omega(m) - \delta \geq 0.99\,\omega(m),$$

as we have $\delta \leq 0.01$ and $\omega(m) \geq 1$ by (3.9).

Therefore, by Lemma 3.8, (3.9), and (3.14),

$$\begin{aligned}
\mathbb{P}[j \in C_T^{max}] &\leq \mathbb{P}\left[|\langle \Gamma_T - \Gamma_0, W^j\rangle| \geq 0.99\,\omega(m)\right] \\
&\leq 2\exp\left(\frac{-0.99^2 \cdot \eta^2\, \ln(\xi d/m)}{2T\varepsilon^2}\right) \\
&< 2\exp\left(\ln \tfrac{m}{\xi d}\right) = \frac{2m}{\xi d}.
\end{aligned}$$

Thus

$$\mathbb{E}\big[|C_T^{max}|\big] = \sum_{j\in[d]} \mathbb{P}\big[j \in C_T^{max}\big] < \tfrac{2m}{\xi}$$

as desired.

To complete the proof of (3.21) we address the second term in (3.22) and show that

$$\mathbb{E}\big[\|\Gamma_T\|_2^2\big] \leq 1.01m. \tag{3.24}$$

By (2.3), we represent Γ_T in terms of the vector families as

$$\Gamma_T = \big(\Gamma_T(w_i)\big) \quad , \quad w_i \in W_i,\ i \in [k].$$

Then

$$\|\Gamma_T\|_2^2 = \sum_{i\in[k]} \sum_{w_i\in W_i} \big(\Gamma_T(w_i)\big)^2. \tag{3.25}$$

As the above double sum has m terms in total, it suffices to show that the expectation of any of these terms is at most 1.01, that is, $\mathbb{E}\big[\Gamma_T(w_i)^2\big] \le 1.01$ for any $i \in [k]$ and $w_i \in W_i$. By Lemma 3.11(iv), $\sum_{w_i\in W_i} \Gamma_T(w_i) = 1$. Thus, for any fixed $w_i \in W_i$,

$$\Gamma_T(w_i) = 1 - \sum_{w\in W_i\setminus\{w_i\}} \Gamma_T(w). \tag{3.26}$$

Note that in the above sum, $\Gamma_T(w) \ge 0$ unless coordinate w is frozen. In this case, assuming that coordinate w is frozen at step t, by Lemma 3.11(vii) we have $\Gamma_T(w) \ge \delta - \varepsilon\big|\Lambda_t(w)\big|$. Accordingly,

$$\Gamma_T(w) \ge \delta - \max_{t\in[T]} \varepsilon\big|\Lambda_t(w)\big| > -\varepsilon \max_{t\in[T]} \big|\Lambda_t(w)\big|. \tag{3.27}$$

Thus, by (3.26),

$$\Gamma_T(w_i) \le 1 + \varepsilon \sum_{w\in W_i\setminus\{w_i\}} \max_{t\in[T]} \big|\Lambda_t(w)\big|.$$

When $\Gamma_T(w_i) \ge 0$, this leads to

$$\begin{aligned}\big(\Gamma_T(w_i)\big)^2 \le 1 + 2\,\varepsilon \sum_{w\in W_i\setminus\{w_i\}} \max_{t\in[T]} \big|\Lambda_t(w)\big| + \varepsilon^2 \sum_{w\in W_i\setminus\{w_i\}} \max_{t\in[T]} \big|\Lambda_t(w)\big|^2 \\ + \varepsilon^2 \sum_{w\ne u\in W_i\setminus\{w_i\}} \max_{t\in[T]} \big|\Lambda_t(w)\big| \max_{t\in[T]} \big|\Lambda_t(u)\big|.\end{aligned}$$

Note that by Lemma 3.5, for each vector w, $\Lambda_t(w)$ is a Gaussian random variable with variance at most 1. Also, for $a, b \ge 0$ we will use that $ab \le a^2 + b^2$. Therefore, by taking expectations above, and applying Lemma 3.9,

$$\begin{aligned}
&\mathbb{E}\Big(\big(\Gamma_T(w_i)\big)^2 \big| \, \Gamma_T(w_i) \geq 0\Big) \qquad (3.28)\\
&\quad = 1 + 2\,\varepsilon \sum_{w \in W_i \setminus \{w_i\}} \mathbb{E}\big[\max_{t \in [T]} \big|\Lambda_t(w)\big|\big] + \varepsilon^2 \sum_{w \in W_i \setminus \{w_i\}} \mathbb{E}\big[\max_{t \in [T]} \big|\Lambda_t(w)\big|^2\big] \\
&\qquad + \varepsilon^2 \sum_{w \neq u \in W_i \setminus \{w_i\}} \mathbb{E}\big[\max_{t \in [T]} \big|\Lambda_t(w)\big| \max_{t \in [T]} \big|\Lambda_t(u)\big|\big] \\
&\quad \leq 1 + 12 \varepsilon m \sqrt{\ln T} + 10 \varepsilon^2 m \ln T + \varepsilon^2 \sum_{w \neq u \in W_i \setminus \{w_i\}} \mathbb{E}\big[\max_{t \in [T]} \big|\Lambda_t(w)\big|^2 \\
&\qquad + \max_{t \in [T]} \big|\Lambda_t(u)\big|^2\big]
\end{aligned}$$

$$\begin{aligned}
&\leq 1 + 12 \varepsilon m \sqrt{\ln T} + 10 \varepsilon^2 m \ln T + 2 \tfrac{m(m-1)}{2} 10 \varepsilon^2 \ln T \\
&\leq 1 + 12 \varepsilon m \sqrt{\ln T} + 10 \varepsilon^2 m^2 \ln T \\
&\leq 1 + 22 \varepsilon m^2 \ln T \\
&\leq 1.01
\end{aligned}$$

by (3.16) and that $m, \ln T \geq 1, \varepsilon < 1$.

When $\Gamma_T(w_i) < 0$, then coordinate w_i is frozen. Therefore, (3.27) and (3.17) imply that

$$\mathbb{E}\Big(\big(\Gamma_T(w_i)\big)^2 \big| \, \Gamma_T(w_i) < 0\Big) < \varepsilon^2 \, \mathbb{E} \max_{t \in [T]} \big|\Lambda_t(w_i)\big|^2 \leq 10 \varepsilon^2 \ln T < 1.$$

Combining this with (3.28) shows that $\mathbb{E}\left(\big(\Gamma_T(w_i)\big)^2\right) \leq 1.01$ for each $i \in [k]$ and $w_i \in W_i$, and by invoking (3.25), we reach (3.24).

To prove (3.21), we may now combine (3.14), (3.22), (3.23) and (3.24) in order to derive that

$$\begin{aligned}
\mathbb{E}\Big[\big|C_T^{conv}\big|\Big] &\geq m - \frac{1.01m}{T\varepsilon^2} - \frac{2m}{\xi} \\
&\geq \left(1 - \frac{2 \cdot 1.01}{0.99^2 \eta^2} - \frac{2}{\xi}\right) m \\
&> 0.51m.
\end{aligned}$$

Since $\left|C_T^{conv}\right| \leq m$, this leads to

$$\mathbb{P}\Big[\left|C_T^{conv}\right| \geq m/2\Big] \geq 0.02.$$

As the probability of the walk leaving R is less than 0.01 by (3.20), we conclude that the algorithm finds the desired vector Γ_T with probability greater than $0.02 - 0.01 = 0.01$, as claimed. □

Finally, we illustrate how to transform the proof of Proposition 3.2 so as to provide a polynomial time algorithm.

Proposition 3.12. *There exists an algorithm of running time $O(d^7 \ln^2 d)$ which, in the setting of Proposition 3.2, yields the desired selection vector $\mu \in \Delta_U$.*

Proof. Along the course of the proof of Proposition 3.2, we replace the iteration of Lemma 3.3 by that of Lemma 3.10 so as to obtain a vector $\hat{\mu} \in \Delta_U$ such that, for each $i \in [k]$,

$$|\{w_i \in W_i :\ 0 \leq \hat{\mu}(w_i) \leq \delta\}| = |W_i| - 1. \tag{3.29}$$

The existence of such a vector is guaranteed as long as $\delta < 1/|W_i|$ for each $i \in [k]$. At the final step, we take μ to be the closest vertex of Δ_W to $\hat{\mu}$, that is, define

$$\mu(w_i) = \begin{cases} 0 \text{ if } \hat{\mu}(w_i) \leq \delta \\ 1 \text{ if } \hat{\mu}(w_i) > \delta. \end{cases}$$

In particular, taking $\chi := \mu - \hat{\mu}$, we have that by (3.7),

$$|\langle \mu - \lambda, W^j \rangle| = |\langle \hat{\mu} - \lambda, W^j \rangle + \langle \chi, W^j \rangle| \leq 22\sqrt{d} + |\langle \chi, W^j \rangle|$$

for each $j \in [d]$. We show that taking a sufficiently small value of δ ensures that $|\langle \chi, W^j \rangle| \leq O(\sqrt{d})$, accordingly, μ is an appropriate selection vector.

Let $i \in [k]$ be arbitrary, and let $w \in W_i$ be so that $\hat{\mu}(w) > \delta$. Then $\hat{\mu}(w) = 1 - \sum_{w_i \neq w \in W_i} \hat{\mu}(w_i)$. Accordingly,

$$\sum_{w_i \in W_i} |\chi(w_i)| = 2 \sum_{w \neq w_i \in W_i} |\hat{\mu}(w_i)| < 2|W_i|\,\delta.$$

Therefore, as $\|W^j\|_\infty \leq 1$,

$$|\langle \chi, W^j \rangle| \leq \sum_{i\in[k]} \sum_{w_i \in W_i} |\chi(w_i)| \leq \sum_{i\in[k]} (2|W_i|\,\delta).$$

Since $|W_i| \leq m \leq 2d$ for each $i \in [k]$ and $k \leq m \leq 2d$, we conclude that for each $j \in [d]$, $|\langle \chi, W^j \rangle| \leq 8d^2\delta$. Thus, fixing

$$\delta = 0.01d^{-3/2}, \tag{3.30}$$

we indeed obtain

$$\|W\mu - W\lambda\|_\infty = \max_{j\in[d]} |\langle \mu - \lambda, W^j \rangle| \leq 22\sqrt{d} + 8\sqrt{d} = O(\sqrt{d}).$$

Next, we estimate the running time of the algorithm at iteration s of Lemma 3.10. As before, let $m(s)$ be the number of active vectors. For a fixed step $t \in [T]$ of the Gaussian random walk, the calculation of the sets C_t^{var}, C_t^{max} takes time $O(d + m(s))$. An orthonormal basis of the subspace S_t may be determined in $O(d^3)$ time by applying a Gram-Schmidt orthogonalization process, and then the Gaussian vector Λ_{t+1} is sampled in $O(m(s))$ time. Hence, the time complexity of performing a given step of the Gaussian walk is dominated by the calculation of the orthonormal basis of S_t.

By (3.30), (3.6), and the condition that $m \leq 2d$, the maximal ε which satisfies the constraints (3.15), (3.16), and (3.17) simultaneously can be estimated by

$$\frac{1}{\varepsilon^2} = O\big(m(s)d^3 \ln^2 d\big).$$

Since, by (3.14), the number of steps of the random walk is $T = O(1/\varepsilon^2)$, the above estimate shows that the running time of the algorithm within a given iteration s is

$$O\big(m(s)d^6 \ln^2 d\big).$$

Let S be the total number of iterations until reaching the vector $\hat{\mu} \in \Delta_U$ satisfying (3.29). Note that by (3.6), we have $m(s) \leq m/2^s$. Therefore, the total running time of the algorithm is

$$\sum_{s=1}^{S} O\Big(\frac{m}{2^s} d^6 \ln^2 d\Big) = O\big(md^6 \ln^2 d\big) = O\big(d^7 \ln^2 d\big). \qquad \square$$

3.4 Conclusion

In this chapter we extended the optimal results for the Euclidean and maximum norm vector balancing problem to the more general colorful setting. We also discussed the algorithmic complexity of the maximum norm solution, and provided a stronger result for the maximum norm in the setting where the number of vectors is much smaller than the dimension.

Chapter 4

A Reduction of the Steinitz Problem

This chapter of the dissertation is based on the following publication of the author.

[4] Gergely Ambrus and Rainie Heck. A note on the Steinitz constant. Accepted for publication; Mathematika, 2026.

4.1 The Steinitz Problem for 'Almost-Unit' Vectors

In this chapter our goal is to prove Theorems 1.11 and 1.12. Recall that Theorem 1.12 bounds $S(B)$ in terms of $S^*_\varepsilon(B)$ for any $B \in \mathcal{K}^d_o$ and $0 < \varepsilon \leq 1$ (see Definitions 1.6 and 1.10 for details), and that Theorem 1.11 proves a slightly weaker bound for the Euclidean setting using the geometry of the Euclidean ball. The primary motivation for this work is that these results offer a potential approach to resolving Conjecture 1.9. We remark that although Theorem 1.12 is strictly stronger, we present both proofs because the specific techniques used for the geometry of the Euclidean ball are of independent interest.

The chapter is organized as follows: we first delve into the history of the Steinitz problem and briefly summarize existing results in Section 4.2. In Section 4.3 we prove Theorem 1.11, up to a handful of technical lemmas, the proofs of which are deferred to Section 4.4. Finally, in Section 4.5 we prove Theorem 1.12 by generalizing (and thereby simplifying) the proof of 1.11.

4.2 History of the Steinitz Problem

Despite being more than a century old, the story of the Steinitz lemma is still far from complete. In the following, we provide an overview that lists the main, and often forgotten, steps in its development. Recall from Chapter 1 that the Steinitz problem arose from a higher dimensional analog of the Riemann rearrangement theorem, proved by Lévy and Steinitz, restated here for clarity:

Theorem (Lévy-Steinitz theorem). *Given a series of vectors in $\mathbb{R}^d$, the set of all sums of its rearrangements is empty, or it forms an affine subspace of $\mathbb{R}^d$.*

Unfortunately, the proof of Lévy contained serious gaps for dimensions $d \geq 3$, as pointed out by Steinitz [71] in 1913. In turn, he gave the first complete proof of Theorem 4.2, known today as the Lévy-Steinitz theorem. Steinitz's work is quite technical and wide-scoped: it was published in three parts [71, 72, 73], with total length summing well over 100 pages. A key step of his proof is Theorem 1.5 (see [71, p.171]), which he stated with $C = 2d$.

Independently of Lévy and Steinitz, Gross [36] also found a shorter proof for Theorem 4.2 that is reminiscent of Steinitz's method. His approach is again based on the rediscovered Steinitz lemma, yet it yields only the weaker constant $S(B_2^d) \leq 2^d - 1$, which is an inevitable consequence of the induction dimension technique he applies. He also provides a geometric reformulation of Theorem 1.5: given any closed polygonal path in $\mathbb{R}^d$ starting at the origin with side lengths not exceeding 1, it is possible to rearrange the order of its sides so that the resulting polygonal path does not leave the ball of radius C. This later led to the alternate title "polygonal confinement theorem" for Theorem 1.5 (cf. Rosenthal [56]).

Gross was not the last one to rediscover the Steinitz lemma. In 1931, Bergström published two papers on the topic. In the first [19], he gives an alternative proof for Theorem 4.2. The crux of his proof is again Theorem 1.5, which he considers to be of interest on its own, and proves by induction on the dimension, leading to the estimate $S(B_2^d) \leq \sqrt{(4^d-1)/3}$. His second article [18] concentrates solely on the Steinitz lemma. He proves that any family of vectors $V \subset B_2^2$ with $\Sigma(V) = 0$ can be arranged to form a closed polygonal path that fits in a circle of radius $\sqrt{5}/2$ (not necessarily centered at the origin), leading to the upper bound $S(B_2^2) \leq (\sqrt{5}+1)/2$ in the plane. Regarding the question in arbitrary dimensions, he formulates Conjecture 1.9.

In 1936, Hadwiger [38] became the first one to study the Steinitz lemma (which he attributes to Gross and Bergström) for series in general inner product spaces: he manages to bound the norm of partial sums in terms of the number of vectors n. His attention then turned to the extension of the Lévy-Steinitz theorem to abstract Hilbert spaces [39] and finite-dimensional vector spaces [37].

Returning to the Euclidean case, Damsteeg and Halperin [27] provided a construction of Euclidean unit vectors establishing $\frac{1}{2}\sqrt{d+3} \leq S_0(B_2^d) \leq S(B_2^d)$, which implies that the $O(\sqrt{d})$ bound conjectured by Bergström would be optimal by the order of magnitude (this is also shown by considering the vertex set of a centered regular simplex).

In 1954, Behrend entered the scene [16] and by a refinement of Steinitz's original method strengthened the estimate to $S(B_2^d) < d$ for every $d \geq 3$. He also showed that $S(B_2^3) \leq \sqrt{5+2\sqrt{3}}$.

Much of the above information had been blocked by the iron curtain. Although Theorem 1.5 is noted to be a 'well-known lemma of Steinitz' [34], Kadets [43] only rediscovered Bergström's estimate $S(B_2^d) \leq \sqrt{(4^d-1)/3}$ in 1953. Twenty years later, in a series of pioneering works, Sevastyanov studied several variants of the question, introduced the compact vector summation problem, and worked on the algorithmic aspects of the topic, including its connections to scheduling problems. In 1973, he rediscovered and re-proved the planar case of the Steinitz lemma with the bound $S(B_2^2) \leq \sqrt{3}$, see [62]. Turning to the higher dimensional case, he proved [63] that $S(B) \leq d$ for every $B \in \mathcal{K}_o^d$ (this extends a weaker form of Behrend's 1954 bound to arbitrary Minkowksi norms), thus achieving Theorem 1.8. His proof was further simplified in his subsequent joint work with Grinberg [34], where the authors in fact proved that $S^*(B) \leq d$ for arbitrary $B \in \mathcal{K}_o^d$. For further developments related to algorithmic aspects, see [58, 60].

Meanwhile in Hungary, independently of the work in the USSR, Fiala [30] also rediscovered (after Sevastyanov [61] and Belov and Stolin [17]) the connection between the flow shop problem and the Steinitz lemma, and re-proved the latter in the planar case. Inspired by his work, Bárány [12] proved the bound $S(B) \leq 3d/2$ for symmetric B's (note that this is weaker than the estimate in [34, 63]) and solved the flow-shop problem for the maximum norm.

Still in the 1980s, Halperin [40] applied a variation of Lévy's method to obtain an elementary proof of Theorem 4.2, using the Grinberg-Sevastyanov variant of the Steinitz lemma, called here as the 'Polygon Rearrangement Theorem'. He also studies the question in L_p and ℓ_p spaces for $0 < p \leq \infty$. In an expository article, Rosenthal [56] presented the Lévy-Steinitz theorem along the Gross-Steinitz approach.

He cites the Steinitz lemma as the 'Polygonal Confinement Theorem' with the weak bound $S(B_2^d) \leq O(2^d)$, apparently unaware of its stronger forms.

Concentrating on the planar case, Banaszczyk showed in [6] that $S(B) \leq 3/2$ for any symmetric $B \subset \mathbb{R}^2$, and this bound is achieved when B is a square centered at the origin. In [7], he determined the exact value of the planar Euclidean Steinitz constant: $S(B_2^2) = \sqrt{5}/2$.

A possible approach for attacking Conjecture 1.9 is via Chobanyan's transference theorem, which gives an explicit connection between the Steinitz constant and the sign-sequence constant $E(B)$, where one is asked to assign signs to vectors of a (potentially infinite) sequence of vectors selected from B so that all partial sums are bounded by $E(B)$. Chobanyan's result [25, 26] shows that $S(B) \leq E(B)$. For further information about the sign-sequence constant, see the survey article of Bárány [22].

Beyond the results mentioned above, there are many other related problems and results, including coordinate-dependent Steinitz bounds for the maximum norm (see e.g. [59]), various extensions of the Lévy-Steinitz theorem to infinite-dimensional spaces (see e.g. [45]), or colorful versions of the Steinitz lemma (see [13, 51]).

4.3 Proof of Theorem 1.11

In this section, we prove Theorem 1.11, which is specific to the Euclidean norm. For the remainder of the section we denote the Euclidean norm by $|\cdot|$, as it is the only norm that will be used in this section.

Proof of Theorem 1.11. Take $0 \leq \varepsilon < 1$ as in the statement of Theorem 1.11, and fix $0 < t < 1$, whose value we will specify later. Suppose that we are given a finite vector family $V \subset B_2^d$ with $\Sigma(V) = 0$; our goal is to order V in such a way that all partial sums are bounded by the right-hand side of (1.11). The first step of the proof is to partition V as

$$V = \Big(\bigsqcup_{\alpha \in A} V_\alpha \Big) \bigsqcup R,$$

where A is an index set of cardinality m, satisfying the following properties:

(i) For each $\alpha \in A$, there exists $u_\alpha \in S^{d-1}$ such that $V_\alpha \subset K_t(u_\alpha)$.

(ii) For each $\alpha \in A$,

$$\frac{1}{\varepsilon} - 1 \leq |\Sigma(V_\alpha)| < \frac{1}{\varepsilon}.$$

(iii) For any $u \in S^{d-1}$, and any subset $T \subseteq R$,

$$\big|\Sigma(T \cap K_t(u))\big| < 1/\varepsilon.$$

Note that we intentionally use an unordered index set A of cardinality m, rather than $A = [m]$, to emphasize that the vectors are not yet ordered.

We define the families V_α via the following process: initialize $R := V$. As long as there exists $u \in S^{d-1}$ and $T \subseteq R$ such that

$$\Big|\Sigma(T \cap K_t(u))\Big| \geq \frac{1}{\varepsilon} - 1, \tag{4.1}$$

we set $u_\alpha := u$ and select any subfamily $V_\alpha \subseteq T \cap H_+(u_\alpha)$, so that

$$\frac{1}{\varepsilon} - 1 \leq \big|\Sigma(V_\alpha)\big| < \frac{1}{\varepsilon}.$$

Such a set is given e.g. by taking any minimal[1] (with respect to containment) subset $T \subset (T \cap K_t(u))$ that satisfies $|\Sigma(T)| \geq \frac{1}{\varepsilon} - 1$. The family over which we minimize is nonempty by (4.1), and the triangle inequality guarantees that the upper bound holds as well.

We then update $R := V \setminus V_\alpha$ and proceed until there is no choice of a vector $u \in S^{d-1}$ and $T \subseteq R$ satisfying (4.1). Properties (i)-(iii) are immediate consequences of this construction.

For every $\alpha \in A$, let $w_\alpha := \varepsilon \cdot \Sigma(V_\alpha)$, and define $W := \{w_\alpha\}_{\alpha \in A}$. Property (ii) yields that for every $\alpha \in A$,

$$1 - \varepsilon \leq |w_\alpha| < 1.$$

Thus, by the definition of $S^*_\varepsilon(B_2^d)$, there exists an ordering $w_1, ..., w_m$ of W such that for any $j \in [m]$,

$$\Big|\sum_{i \in [j]} w_i - \frac{j}{m}\Sigma(W)\Big| \leq S^*_\varepsilon(B_2^d). \tag{4.2}$$

Further, as $\Sigma(V) = 0$ and $V = (\bigsqcup_{\alpha \in A} V_\alpha) \bigsqcup R$,

$$\Sigma(W) = \varepsilon \Sigma_{\alpha \in A} \Sigma(V_\alpha) = -\varepsilon \Sigma(R), \tag{4.3}$$

[1] Note that here we do not require taking a minimal set with respect to containment, but rather note that such a set satisfies the necessary conditions. In the proof of Theorem 1.12, we will require the set to be minimal containment-wise, which simplifies the proof.

combining (4.2) and (4.3) yields the bound

$$\Big|\sum_{i\in[j]} w_i + \varepsilon\frac{j}{m}\Sigma(R)\Big| \le S^*_\varepsilon(B^d_2). \tag{4.4}$$

We now order our original collection of vectors V as follows: fix the ordering of W as in (4.2), and note that this ordering of $\{w_\alpha\}_{\alpha\in A}$ induces a matching ordering $V_1, \ldots, V_m$ of the sets $\{V_\alpha\}_{\alpha\in A}$. Within each set V_i for $i \in [m]$ we order the vectors $v \in V_i$ arbitrarily. Finally, we also order the remaining vectors in R arbitrarily. Along this ordering, two types of partial sums occur:

(a) $\sum_{i\in[j]} \Sigma(V_i) + \Sigma(U)$ for $0 \le j \le m-1$ and $U \subset V_{j+1}$ (where $U = \emptyset$ is allowed);

(b) $\sum_{i\in[m]} \Sigma(V_i) + \Sigma(T)$ for $T \subseteq R$.

We need to show that any partial sum of type (a) or (b) satisfies the bound in (1.11). To this end, we will need the following two lemmas, the proofs of which we defer to Section 4.4.

Lemma 4.1. *For any $\alpha \in A$ and any $U \subseteq V_\alpha$, and for any $0 < t < 1$,*

$$\big|\Sigma(U)\big| < \frac{1}{\varepsilon t}.$$

Lemma 4.2. *For any $T \subseteq R$, and for any $0 < t < 1$,*

$$\big|\Sigma(T)\big| < \frac{1}{\varepsilon \sigma_t}.$$

We now use Lemmas 4.1 and 4.2 to complete the proof of Theorem 1.11. It will be useful to note that by dividing both sides of (4.4) by ε and applying the triangle inequality, one has

$$\Big|\sum_{i\in[j]} \Sigma(V_i)\Big| \le \frac{j}{m}|\Sigma(R)| + \frac{1}{\varepsilon}S^*_\varepsilon(B^d_2) < \frac{1}{\varepsilon\sigma_t} + \frac{1}{\varepsilon}S^*_\varepsilon(B^d_2), \tag{4.5}$$

where we have used that $j < m$ and Lemma 4.2.

We first consider partial sums of type (a). For any $0 \le j \le m-1$ and $U \subset V_{j+1}$, combining Lemma 4.1 and (4.5) yields

$$\Big|\sum_{i\in[j]} \Sigma(V_i) + \Sigma(U)\Big| \le \Big|\sum_{i\in[j]} \Sigma(V_i)\Big| + \big|\Sigma(U)\big| \le \frac{1}{\varepsilon}\left(\frac{1}{\sigma_t} + S^*_\varepsilon(B^d_2) + \frac{1}{t}\right),$$

where we have interpreted $\Sigma(U) = 0$ for the case $U = \emptyset$. Second, we consider any partial sum of type (b). Fix any $T \subseteq R$, and recall that $\sum_{i\in[m]} \Sigma(V_i) = -\Sigma(R)$. Thus applying Lemma 4.2 to $R \setminus T \subseteq R$,

$$\Big|\sum_{i\in[m]} \Sigma(V_i) + \Sigma(T)\Big| = \Big|\Sigma(T) - \Sigma(R)\Big| = \Big|\Sigma(R \setminus T)\Big| \le \frac{1}{\varepsilon\sigma_t}.$$

We have now shown that every partial sum along the specified ordering has Euclidean norm at most

$$\frac{1}{\varepsilon}\left(S^*_\varepsilon(B_2^d) + \frac{1}{t} + \frac{1}{\sigma_t}\right). \tag{4.6}$$

We finish the proof via the following estimate on the measure of spherical caps, that we prove in the subsequent section:

Lemma 4.3. *For every $d \ge 2$, and for*

$$t = \sqrt{\frac{\log d}{2d}}, \tag{4.7}$$

the estimate

$$\sigma_t \ge \frac{1}{c} \cdot t \tag{4.8}$$

holds with $c = 140$.

Thus, setting t as in (4.7), we have by (4.8) that

$$\frac{1}{t} + \frac{1}{\sigma_t} < 200\sqrt{\frac{d}{\log d}}$$

for every $d \ge 2$, which combined with (4.6) establishes the desired bound. □

We note that by setting $t = \sqrt{\frac{\log d - 2\log\log d}{d}}$, Lemma 4.3 holds with $c = 11$ if d is sufficiently large. Consequently, the constant 200 in Theorem 1.11 can be improved to 12. As this does not constitute an asymptotic improvement, we decided to simplify the calculations by choosing the value of t specified in (4.7).

4.4 Technical Lemmas

In this section, we prove Lemmas 4.1, 4.2, and 4.3.

Proof of Lemma 4.1. Fix $\alpha \in A$ and $U \subset V_\alpha$. By property (i) of our construction, there exists $u_\alpha \in S^{d-1}$ such that $U \subset V_\alpha \subset K_t(u_\alpha)$.

Since $K_t(u_\alpha)$ is convex, and $U \subset V_\alpha \subseteq K_t(u_\alpha)$, we have that

$$\left\langle \frac{\Sigma(U)}{|\Sigma(U)|}, u_\alpha \right\rangle \geq t. \tag{4.9}$$

As $U \subset V_\alpha$ and $0 < \langle v, u_\alpha \rangle$ for all $v \in V_\alpha$,

$$\langle \Sigma(U), u_\alpha \rangle \leq \langle \Sigma(V_\alpha), u_\alpha \rangle \leq \big|\Sigma(V_\alpha)\big| < \frac{1}{\varepsilon}$$

by property (ii) of our construction. Thus, by (4.9),

$$|\Sigma(U)|^2 \leq \frac{1}{t^2} \langle \Sigma(U), u_\alpha \rangle^2 < \frac{1}{\varepsilon^2 t^2}. \qquad \square$$

Proof of Lemma 4.2. Fix an arbitrary subset $T \subseteq R$. The key step of the proof is the observation that for any $u \in S^{d-1}$ and $v \in B_2^d$, $v \in K_t(u)$ if and only if $u \in C_t\big(\frac{v}{|v|}\big)$. To see this, note that

$$v \in K_t(u) \quad \Longleftrightarrow \quad \left\langle \tfrac{v}{|v|}, u \right\rangle \geq t \quad \Longleftrightarrow \quad u \in C_t\big(\tfrac{v}{|v|}\big).$$

Thus, for any $v \in B_2^d$, we have

$$v = \frac{1}{\sigma_t} \int_{S^{d-1}} v \cdot \chi_{K_t(u)}(v) \, \mathrm{d}\sigma(u).$$

Therefore,

$$\begin{aligned}
|\Sigma(T)|^2 &= \langle \Sigma(T), \Sigma(T) \rangle \\
&= \left\langle \Sigma(T), \sum_{v \in S} \frac{1}{\sigma_t} \int_{S^{d-1}} v \cdot \chi_{K_t(u)}(v) \, \mathrm{d}\sigma(u) \right\rangle \\
&= \frac{1}{\sigma_t} \int_{S^{d-1}} \Big\langle \Sigma(T), \Sigma(T \cap K_t(u)) \Big\rangle \, \mathrm{d}\sigma(u) \\
&\leq \frac{1}{\sigma_t} \int_{S^{d-1}} |\Sigma(T)| \cdot |\Sigma(T \cap K_t(u))| \, \mathrm{d}\sigma(u) \\
&< \frac{|\Sigma(T)|}{\sigma_t} \int_{S^{d-1}} \frac{1}{\varepsilon} \, \mathrm{d}\sigma(u) \\
&= \frac{|\Sigma(T)|}{\varepsilon \sigma_t}
\end{aligned}$$

where we used Property (iii) in the penultimate line. Thus we conclude that $|\Sigma(T)| < \frac{1}{\varepsilon\sigma_t}$, as desired. □

Proof of Lemma 4.3. We will use the following bound of Gautschi [31]: for $x > 0$ and $0 < \lambda < 1$,

$$x^{1-\lambda} \leq \frac{\Gamma(x+1)}{\Gamma(x+\lambda)} \leq (x+1)^{1-\lambda}.$$

Combining the above inequality for $x = d/2$ and $\lambda = 1/2$ with the definition of κ_d yields

$$\frac{(d-1)\kappa_{d-1}}{d\kappa_d} = \frac{(d-1)\cdot\pi^{(d-1)/2}\cdot\Gamma(\frac{d}{2}+1)}{d\cdot\pi^{d/2}\cdot\Gamma(\frac{d}{2}+\frac{1}{2})} \geq \frac{\sqrt{d}}{2\sqrt{2\pi}}. \tag{4.10}$$

A standard calculation shows that

$$\sigma_t = \frac{(d-1)\kappa_{d-1}}{d\kappa_d}\int_t^1 (1-x^2)^{\frac{d-3}{2}}\,\mathrm{d}x.$$

Therefore, by (4.10),

$$\sigma_t \geq \frac{\sqrt{d}}{2\sqrt{2\pi}}\int_t^1 (1-x^2)^{\frac{d-3}{2}}\,\mathrm{d}x \geq \frac{\sqrt{d}}{2\sqrt{2\pi}}\int_t^1 (1-x^2)^{d/2}\,\mathrm{d}x = \frac{\sqrt{d}}{2\sqrt{2\pi}}\int_t^1 f(x)\,\mathrm{d}x, \tag{4.11}$$

where $f(x) := (1-x^2)^{d/2}$.

For $d \leq 9$, inequality (4.8) can be verified by directly calculating the above integral; in particular, for $2 \leq d \leq 9$, $\sigma_t \geq 0.05$ and $t/c \leq 0.004$.

Assume that $d \geq 10$. Then (4.7) implies that $t > \frac{1}{\sqrt{d-1}}$. Consequently, since $f(x)$ is convex on the interval $\left[\frac{1}{\sqrt{d-1}}, 1\right]$, we derive that $f(x)$ is convex on $[t, 1]$.

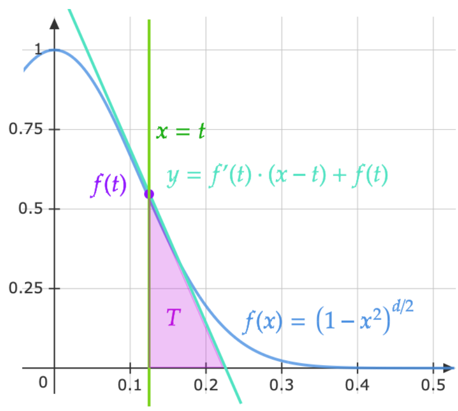

Figure 4.1: *Lower bound on surface area integral using convexity*

Consider the triangle T defined by the x-axis, the tangent line at $f(t)$, and the line $x = t$, as shown in Figure 4.1. By convexity,

$$\int_t^1 f(x)\,\mathrm{d}x \;\geq \mathrm{Area}(T) = \frac{1}{2}\left(-\frac{f(t)^2}{f'(t)}\right) = \frac{(1-t^2)^{d/2+1}}{2dt}. \tag{4.12}$$

Since for $d \geq 4$, $t^2 = \frac{\log d}{2d} < \frac{1}{5}$, by (4.11) we obtain that

$$\sigma_t \geq \frac{1-t^2}{4\sqrt{2\pi}} \cdot \frac{(1-t^2)^{d/2}}{\sqrt{d}t} \geq \frac{1}{5\sqrt{2\pi}} \cdot \frac{(1-t^2)^{d/2}}{\sqrt{d}t} \geq \frac{1}{13} \cdot \frac{(1-t^2)^{d/2}}{\sqrt{d}t}.$$

Thus, the proof of (4.8) boils down to verifying

$$(1-t^2)^{d/2} \geq \frac{13}{c}\sqrt{d}t^2,$$

which, after substituting (4.7), takes the form

$$\left(1 - \frac{\log d}{2d}\right)^{d/2} \geq \frac{13}{2c} \cdot \frac{\log d}{\sqrt{d}}.$$

After taking logarithms, this is equivalent to

$$\frac{d}{2}\log\left(1 - \frac{\log d}{2d}\right) \geq -\log\left(\frac{2c}{13}\right) + \log\log d - \frac{1}{2}\log d. \tag{4.13}$$

By using the standard inequality

$$\log(1-\varepsilon) > -\frac{\varepsilon}{1-\varepsilon}$$

that holds for every $0 < \varepsilon < 1$, and the estimate $1 - \frac{\log d}{2d} \geq \frac{5}{6}$ that is valid for $d \geq 10$, we derive that (4.13) follows from

$$\log d \leq \frac{5}{3}\left(\log d - 2\log\log d + 2\log\left(\frac{2c}{13}\right)\right). \tag{4.14}$$

Let $d_0 := 4 \cdot 10^7$. For $d \geq d_0$, $\log\log d < \frac{1}{5}\log d$ holds and implies (4.14). Finally, for $d \leq d_0$, $\log\log d < 3$ is valid, and since $\log\left(\frac{2c}{13}\right) > 3$, (4.14) holds trivially. □

4.5 An Extension to General Norms

In this section we prove Theorem 1.12. By optimizing the proof of Theorem 1.11, one can see that the strongest bound is given in the case of half-spaces. This generalization not only greatly simplifies the proof and allows us to extend it to arbitrary norms, but also yields a stronger $O(1)$ bound on the additive error introduced. Although the construction is similar to that in the proof of Theorem 1.11, we repeat it for the sake of clarity. We restate the theorem for reference.

Theorem. *For all $d \geq 2$, any convex body $B \in \mathcal{K}_o^d$, and $0 < \varepsilon \leq 1$,*

$$S(B) < \frac{1}{\varepsilon}\Big(S_\varepsilon^*(B) + 2\rho(B) + 1\Big).$$

Proof. Fix $0 < \varepsilon < 1$, $d \geq 2$, and $B \in \mathcal{K}_o^d$, and suppose that we are given a finite vector family $V \subset B$ with $\Sigma(V) = 0$; our goal is to order V in such a way that all partial sums are bounded by the right-hand side of (1.12). For simplicity, we will write $\|.\| := \|.\|_B$ throughout the proof, omitting the dependence on B. Recall also from Chapter 1 that for $u \in S^{d-1}$, we denote the closed positive half-space orthogonal to u by $H_+(u)$.

The first step is to partition V as

$$V = \Big(\bigsqcup_{\alpha \in A} V_\alpha\Big) \bigsqcup R,$$

where A is an index set of cardinality m, satisfying the following properties:

(i) For each $\alpha \in A$, there exists $u_\alpha \in S^{d-1}$ such that $V_\alpha \subset H_+(u_\alpha)$

(ii) For each $\alpha \in A$,
$$\frac{1}{\varepsilon} - 1 \leq \|\Sigma(V_\alpha)\| < \frac{1}{\varepsilon}.$$

(iii) For any $u \in S^{d-1}$, and any subset $T \subseteq R$,
$$\big\|\Sigma(T \cap H_+(u))\big\| < 1/\varepsilon.$$

Note that we intentionally use an unordered index set A of cardinality m, rather than $A = [m]$, to emphasize that the vectors are not yet ordered.

We define the families V_α via the following process: initialize $R := V$. As long as there exist $u \in S^{d-1}$ and $T \subseteq R$ such that

$$\Big\|\Sigma(T \cap H_+(u))\Big\| \geq \frac{1}{\varepsilon} - 1, \tag{4.15}$$

then we set $u_\alpha := u$ and select a containment-wise minimal subfamily $V_\alpha \subseteq T \cap H_+(u_\alpha)$ so that

$$\|\Sigma(V_\alpha)\| \geq \frac{1}{\varepsilon} - 1 \tag{4.16}$$

holds. By the minimality condition, the triangle inequality guarantees that $\|\Sigma(V_\alpha)\| < \frac{1}{\varepsilon}$, therefore V_α satisfies property (ii). We then update $R := V \setminus V_\alpha$ and proceed until there is no choice of a vector $u \in S^{d-1}$ and $T \subseteq R$ satisfying (4.15).

It is an immediate consequence of the construction that properties (i)–(iii) are satisfied for the partition of V obtained above.

Next, for every $\alpha \in A$, let $w_\alpha := \varepsilon \cdot \Sigma(V_\alpha)$, and define $W := \{w_\alpha\}_{\alpha \in A}$. Property (ii) yields that for every $\alpha \in A$,

$$1 - \varepsilon \leq \|w_\alpha\| < 1.$$

Thus, by the definition of $S_\varepsilon^*(B)$, there exists an ordering $w_1, ..., w_m$ of W such that for any $j \in [m]$,

$$\Big\|\sum_{i \in [j]} w_i - \frac{j}{m}\Sigma(W)\Big\| \leq S_\varepsilon^*(B). \tag{4.17}$$

Further, as $\Sigma(V) = 0$ and $V = (\bigsqcup_{\alpha \in A} V_\alpha) \bigsqcup R$,

$$\Sigma(W) = \varepsilon \sum_{\alpha \in A} \Sigma(V_\alpha) = -\varepsilon \Sigma(R),$$

and combining this with (4.17) yields that for any $j \in [m]$,

$$\Big\|\sum_{i \in [j]} w_i + \varepsilon \frac{j}{m}\Sigma(R)\Big\| \leq S_\varepsilon^*(B). \tag{4.18}$$

We now order our original collection of vectors V as follows: fix the ordering of W as above so that (4.17) holds, and note that this ordering of $W = \{w_\alpha\}_{\alpha \in A}$ induces a matching ordering $V_1, ..., V_m$ of the families $\{V_\alpha\}_{\alpha \in A}$. Within each family V_i for $i \in [m]$ we order the vectors $v \in V_i$ arbitrarily. Finally, we also order the remaining vectors in R arbitrarily. Along this ordering, two types of partial sums occur:

(a) $\sum_{i \in [j]} \Sigma(V_i) + \Sigma(U)$ for $0 \leq j \leq m-1$ and $U \subset V_{j+1}$ (where $U = \emptyset$ is allowed);

(b) $\sum_{i \in [m]} \Sigma(V_i) + \Sigma(T)$ with some $T \subseteq R$.

To prove (1.12), we need to show that any partial sum of type (a) or (b) has norm at most $\frac{1}{\varepsilon}\Big(S_\varepsilon^*(B)+2\rho(B)+1\Big)$. Recall that by property (iii) of our construction, for any subset $T' \subseteq R$ and any $u \in S^{d-1}$,

$$\|\Sigma(T' \cap H_+(u))\| < 1/\varepsilon.$$

Fix any direction $u \in S^{d-1}$, and partition T' as

$$T'_+ := T' \cap H_+(u), \quad T'_- := T' \setminus T'_+.$$

Note that these are subfamilies of R, moreover, $T'_- \subset H_+(-u)$. Therefore, property (iii) implies that

$$\|\Sigma(T'_+)\| < \frac{1}{\varepsilon} \text{ and } \|\Sigma(T'_-)\| < \frac{1}{\varepsilon},$$

thus by the triangle inequality

$$\|\Sigma(T')\| < \frac{2}{\varepsilon}. \tag{4.19}$$

We are ready to estimate the norm of partial sums along the ordering specified above. This is simple for sums of type (b): applying (4.19) for the family $T' := R \setminus T$,

$$\Big\|\sum_{i\in[m]}\Sigma(V_i)+\Sigma(T)\Big\| = \Big\|-\Sigma(R)+\Sigma(T)\Big\| = \Big\|-\Sigma(R\setminus T)\Big\| \le \rho(B)\Big\|\Sigma(R\setminus T)\Big\| < \rho(B)\frac{2}{\varepsilon}.$$

Finally, we handle the sums of type (a). Dividing both sides of (4.18) by ε and applying the triangle inequality yields that for any $0 \le j \le m-1$,

$$\begin{aligned}\Big\|\sum_{i\in[j]}\Sigma(V_i)\Big\| &\le \Big\|-\frac{j}{m}\Sigma(R)\Big\| + \Big\|\sum_{i\in[j]}\Sigma(V_i)+\frac{j}{m}\Sigma(R)\Big\| \\ &\le \frac{j}{m}\|-\Sigma(R)\| + \frac{1}{\varepsilon}S_\varepsilon^*(B) \\ &< \rho(B)\frac{2}{\varepsilon} + \frac{1}{\varepsilon}S_\varepsilon^*(B),\end{aligned}$$

where we have used $j < m$ and (4.19) with $T' = R$. Recall that V_{j+1} was chosen as a minimal set (with respect to containment) that satisfies inequality (4.16), that is, $\frac{1}{\varepsilon}-1 \le \|\Sigma(V_{j+1})\|$. In particular, for any $U \subset V_{j+1}$, we know that $\|\Sigma(U)\| < \frac{1}{\varepsilon}-1 < \frac{1}{\varepsilon}$. Combining these estimates, for fixed $0 \le j \le m-1$ and $U \subset V_{j+1}$ we conclude that

$$\Big\|\sum_{i\in[j]}\Sigma(V_i)+\Sigma(U)\Big\| \le \Big\|\sum_{i\in[j]}\Sigma(V_i)\Big\| + \Big\|\Sigma(U)\Big\| < \frac{1}{\varepsilon}\Big(S_\varepsilon^*(B)+2\rho(B)+1\Big).$$

We have now shown that the B-norm of every partial sum along the specified ordering is strictly less than

$$\frac{1}{\varepsilon}\left(S^*_\varepsilon(B) + 2\rho(B) + 1\right). \qquad \square$$

Remark. The proof can not be transformed so as to provide an estimate on $S^*(B)$ instead of $S(B)$: since there is no upper bound on the size of the families V_α, the size of these re-groupings need not be uniform. Hence, the average of the families V_i and the whole family V may differ drastically, which yields that the quantity in (1.6) cannot be estimated in terms of the individual deviations corresponding to V_i.

Chapter 5

Concluding Remarks and Future Work

In this dissertation we have addressed two fascinating problems regarding vector sums in discrete and convex geometry: the vector balancing problem and the Steinitz problem. In Chapters 2 and 3 we prove results about a geometric generalization of the classical vector balancing problem, and in Chapter 4 we prove a reduction of the Steinitz problem to a simpler geometric setting. The nature of these works is clearly different: one generalizes and opens the problem in new directions, whereas the other reduces the problem to a simpler setting and offers a new line of attack on a long-standing conjecture.

In this brief chapter, we discuss future open questions and potential extensions of these works, beginning with the colorful vector balancing problem. As mentioned in Chapter 2, Banaszczyk's famous vector balancing result [9] has already been extended to a colorful setting by Bansal et al. [11], and the strategy that we outline in Section 2.8 shows that in fact up to a constant factor of 2, any bound in the classical vector balancing setting transfers to the colorful setting as well. However, as we argue in Chapter 2, there is value in proofs that show such bounds directly and geometrically, as they shed more light on the problem. Hence a potential direction for future work would be to derive direct proofs of colorful vector balancing bounds for other settings: other ℓ_p norms, to generalize [54], or to any other class of symmetric convex bodies. Further, colorful problems arise in many areas of convex and discrete geometry (as we mentioned, recently the Steinitz problem has also been extended to a colorful setting), and the techniques that we develop in Chapters 2 and 3, particularly understanding the geometry of the direct product of simplices as a parameterization space for the convex hulls, may prove useful in other settings as well.

We now turn to our work on the Steinitz problem. An early motivation for this reduction was the chain of thought that perhaps the Steinitz problem would be easier to solve if one reduces consideration to only unit vectors. Indeed, one can imagine and construct problematic and challenging examples with arbitrarily short vectors that confound certain proof strategies. Our proof falls just short of showing that the Steinitz problem can be reduced to unit vectors (we show that we can reduce to vectors of length at least $1 - \varepsilon$ for any constant ε), and it would be quite interesting to prove that in fact unit vectors suffice. Conversely, it would also be interesting to prove Conjecture 1.9 in the specific setting of unit vectors. Another interesting research direction would be to extend or exploit the "pre-processing" strategy that we use in order to break the vectors into smaller subsets that may be easier to sum, especially in the particular case of spherical caps and the Euclidean norm.

Summary

The PhD thesis presents work on two related problems in discrete and convex geometry: the vector balancing problem and the Steinitz problem.

The majority of the mathematical content of the dissertation is based on the following two publications of the author:

[2] Gergely Ambrus and Rainie Bozzai. Colourful vector balancing. Mathe- matika, 70(4), August 2024.

[4] Gergely Ambrus and Rainie Heck. A note on the Steinitz constant. Accepted for publication; Mathematika, 2026.

In Chapter 1 we introduce all necessary notation and terminology in Section 1.1, followed by a thorough introduction of the vector balancing and Steinitz problems in Section 1.2.

In Chapter 2 we introduce the colorful vector balancing problem, a geometric generalization of the original vector balancing problem. Both Chapters 2 and 3 are based on the results of the paper [2]. To recap, in the vector balancing problem, we are given a symmetric convex body $K \subset \mathbb{R}^d$ and a collection of vectors $v_1, ..., v_n \in K$ and asked to select signs $\varepsilon_1, ..., \varepsilon_n \in \{\pm 1\}$ so that $\|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n\|_K$ is minimal. The colorful vector balancing problem generalizes to the setting where we are given vector families $V_1, ..., V_n \subset K$ satisfying the condition that $0 \in \sum_{i \in [n]} \operatorname{conv} V_i$, and we select one vector from each family to minimize $\|v_1 + \cdots + v_n\|_K$. In the classical vector balancing setting, two well-known results are the following: first, in the Euclidean norm one can always select signs $\varepsilon_i \in \{\pm 1\}$, $i \in [n]$, so that the signed sum has Euclidean norm at most $\sqrt{d}$. Second, in the maximum norm, one can always select signs so that the signed sum has maximum norm at most $O(\sqrt{d})$. Our primary goal in Chapters 2 and 3 is to extend these results to the colorful setting.

In Section 2.1 we introduce the history of the colorful vector balancing problem, including a closely related result of Bansal, Dadush, Garg, and Lovett in a slightly modified colorful setting [11]. In Section 2.2 we prove the following key result (for

the relevant definitions and notation, refer to Section 2.2).

Corollary ([2] Corollary 2.5). *Let $\|\cdot\|$ be a norm on $\mathbb{R}^d$ with unit ball B^d. Suppose that there exists a constant $C(d)$ such that given any collection of $k \leq d$ families $U = \{U_1, ..., U_k\}$ in B^d satisfying $|U_1| + \cdots + |U_k| \leq k + d$, and any $\lambda \in \Delta_V$, there exists a selection vector $\mu \in \Delta_U$ such that $\|V_\lambda - V\mu\| \leq C(d)$. Then given any collection of families $V_1, ..., V_n \subseteq B^d$ with $0 \in \sum_{i\in[n]} \mathrm{Conv} V_i$, there exists a selection of vectors $v_i \in V_i$ for $i \in [n]$ such that*

$$\Big\| \sum_{i\in[n]} v_i \Big\| \leq C(d).$$

In effect, this result allows us to transform the colorful vector balancing problem into a separate problem about vertex approximation in high dimensional direct products of simplices. Furthermore, this result is the key that allows us to prove bounds on the colorful vector balancing problem depending only on the dimension d, and not n, the number of vector families.

Finally, in Section 2.8 we return to the aforementioned result of Bansal, Dadush, Garg, and Lovett and show how one can generalize their techniques to show asymptotically tight bounds in the colorful vector balancing setting based on the vector balancing setting. However, we also justify why our direct geometric approach sheds more light on the problem itself.

In Chapter 3 we continue our study of the colorful vector balancing problem by turning to the specific cases of the Euclidean and maximum norms. In Section 3.1 we prove the following result.

Theorem ([2], Theorem 1.4). *Given vector families $V_1, ..., V_n \subseteq B_2^d$ with*

$$0 \in \sum_{i\in[n]} \mathrm{Conv}\ V_i,$$

one can select vectors $v_i \in V_i$ for $i \in [n]$ such that $\|v_1 + \cdots + v_n\|_2 \leq \sqrt{d}$.

By virtue of our reduction in Chapter 2, it remains to show that we can solve the vertex approximation-style problem introduced in Corollary 2.7. To this end, the key result of Section 3.1 is the following:

Proposition ([2], Proposition 3.1). *Given a collection of k vector families $U_1, ..., U_k \in B_2^d$ and any point $\lambda \in \Delta_U$, there exists a selection vector $\mu \in \Delta_U$ such that $\|U\lambda - U\mu\|_2 \leq \sqrt{k}$.*

The proof follows the probabilistic method, and it is inspired by Spencer's argument in the classical vector balancing setting [67]. In Section 3.2 we turn to the more challenging case of the maximum norm. The main result of Section 3.2 is the following:

Theorem ([2], Theorem 1.5). *Given vector families $V_1, ..., V_n \subseteq B_\infty^d$ with*

$$0 \in \sum_{i\in[n]} \mathrm{Conv}\ V_i,$$

one can select vectors $v_i \in V_i$ for $i \in [n]$ such that $\|v_1 + \cdots + v_n\|_\infty \leq C\sqrt{d}$, where $C = 22$ suffices.

Similarly to Section 3.1, the proof reduces to the following proposition.

Proposition ([2], Proposition 4.1). *Given a collection of k vector families $U_1, ..., U_k \in B_\infty^d$ satisfying $|U_1| + \cdots + |U_d| \leq 2d$ and any point $\lambda \in \Delta_U$, there exists a selection vector $\mu \in \Delta_U$ such that $\|U\lambda - U\mu\|_\infty \leq \sqrt{k}$.*

Our proof, based on the algorithmic proof of Spencer's original result for the vector balancing problem due to Lovett and Meka [47], uses a Gaussian random walk inside of a high dimensional product of simplices to construct "partial colorings" (i.e. assignments of convex coefficients to each family, with the goal of eventually selecting one vector) with high probability. By iterating this algorithm, one can construct a full coloring with bounded error. The key technical aspect of this proof, which we call the *skeleton approximation lemma,* is deferred to Section 3.3 in order to make the exposition of the proof cleaner.

In Chapter 4 we focus our attention on the Steinitz problem. The content of Chapter 4 is based on the results published in [4]. Recall from Chapter 1 that in the Steinitz problem, we are again given a symmetric convex body $B \subset \mathbb{R}^d$ and a collection of vectors $V \subset B$ such that $\sum_{v\in V} v = 0$. The goal is to find an ordering $v_1, ..., v_n$ of V such that every partial sum along this ordering has norm bounded by a constant C depending only on the convex body B. That is, for every $k \in [n]$, $\|v_1 + \cdots + v_k\| \leq C$. The smallest constant C that holds for a given convex body B is called the *Steinitz constant* $S(B)$ of B, and it is a well-known result that $S(B) \leq d$ for any convex body $B \subset \mathbb{R}^d$ [24, 71]. It is a long-standing open conjecture of Bergström [18] that $S(B_2^d) = O(\sqrt{d})$, although not even an $o(d)$ result is known. In Chapter 4 we show that in order to prove this conjecture, one can additionally assume that all of the vectors have length at least $1 - \varepsilon$ for any constant $0 < \varepsilon < 1$, reducing to the setting of "almost-unit vectors". We prove two results, one specifically for the Euclidean norm and one holding for arbitrary norms. We state the precise results below, recalling that $S_\varepsilon(B_2^d)$ denotes the Steinitz constant in the setting where all vectors $v \in V$ satisfy $1 - \varepsilon \le \|v\|_2 \le 1$.

Theorem. *For any* $0 < \varepsilon < 1$ *and all* $d \ge 2$*,*

$$S(B_2^d) \le \frac{1}{\varepsilon}\Big(S_\varepsilon(B_2^d) + 200\sqrt{\frac{d}{\log d}}\Big)$$

In particular, an $o(d)$ estimate on the restricted problem would yield an $o(d)$ estimate for Bergström's conjecture, and an $O(\sqrt{d})$ result in the "almost-unit vectors" setting would resolve the conjecture completely.

The stronger result for arbitrary norms is as follows, where we recall that $\rho(B) := \max_{v \in B} \|-v\|_B$.

Theorem ([4] Theorem 6)**.** *For all* $d \ge 2$*, any convex body* $B \in \mathcal{K}_o^d$*, and* $0 < \varepsilon \le 1$*,*

$$S(B) < \frac{1}{\varepsilon}\Big(S_\varepsilon^*(B) + 2\rho(B) + 1\Big).$$

In Section 4.2 we introduce the interesting and storied history of the Steinitz problem in more detail. In Section 4.3 we prove Theorem 1.11 in broad strokes, deferring the proof of a handful of technical lemmas to Section 4.4. The strategy of the proof is a key pre-processing of the vectors to remove any short (i.e., of norm less than $1 - \varepsilon$) vectors. We do this by summing together short vectors within spherical caps until we get a vector that is sufficiently long; in doing so, we must be careful that all partial sums of the new long vectors remain sufficiently short, and we must deal with a handful of extra vectors that are left after pre-processing; these are the details of the technical lemmas in Section 4.4. In Section 4.5 we generalize this proof technique to arbitrary norms by completing the pre-processing with half spaces in lieu of spherical caps.

Összefoglalás

A disszertációban két, egymáshoz kapcsolódó diszkrét és konvex geometriai témával foglalkozunk: a vektorkiegyensúlyozási feladattal és a Steinitz problémával. A vektorkiegyensúlyozási feladatban legyen $K \in \mathbb{R}^d$ egy szimmetrikus konvex test, és $v_1, ..., v_n$ vektorok K-ban; célunk az $\varepsilon_1, ..., \varepsilon_n \in \{\pm 1\}^n$ együtthatók meghatározása úgy, hogy $\|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n\|_K$ minimális legyen.

Az általánosított, "színezett" vektorkiegyensúlyozási feladatban $V_1, ..., V_n \subset K$ olyan vektorrendszerek, melyekre $0 \in \sum_{i\in[n]} \operatorname{conv}(V_i)$. Célunk kiválasztani minden családból egy $v_i \in V_i$ vektort úgy, hogy $\|v_1 + \cdots + v_n\|_K$ minimális legyen. A klasszikus vektorkiegyensúlyozási probléma két speciális esetére jól ismert korlátok vonatkoznak: az euklideszi normában mindig választhatók olyan előjelek, hogy az előjeles összeg euklideszi normája legfeljebb $\sqrt{d}$ legyen. Továbbá, a maximum normában mindig választhatók olyan előjelek, hogy az előjeles összeg maximum normája legfeljebb $O(\sqrt{d})$ legyen. Mindkét eredményt kiterjesztjük az általánosabb színezett verzióra, éles, illetve aszimptotikusan éles becsléseket igazolva. A bizonyítás kulcslépéseként a színezett vektorkiegyensúlyozási problémát visszavezetjük magas dimenziós szimplexek direkt szorzataiban a csúcsközelítés problémájára.

A Steinitz problémában ismét adott egy $B \subseteq \mathbb{R}^d$ szimmetrikus konvex test, és egy olyan $V \subset B$ vektorrendszer, amelyre $\sum_{v\in V} v = 0$. Célunk V egy olyan $V = \{v_1, ..., v_n\}$ sorbarendezésének meghatározása, melyre a sorrend szerinti parciális összegek normája legfeljebb egy C konstans, amely csak B-től függ: tehát minden $k \in [n]$ esetén $\|v_1 + \cdots + v_k\|_B \leq C$. Adott B-re a C korlát elérhető legkisebb értékét a B Steinitz-konstansának $S(B)$ nevezzük. Jól ismert eredmény, hogy $S(B) \leq d$ bármely $B \subseteq \mathbb{R}^d$ szimmetrikus konvex testre. Bergström régóta fennálló nyílt sejtése szerint $S(B_2^d) = O(\sqrt{d})$, azonban ebben az esetben még egy $o(d)$ becslés sem ismert. A disszertációban a sejtést visszavezetjük arra az esetre, amikor a V vektorrendszer összes elemének normája az $[1 - \varepsilon, 1]$ intervallumban van, tehát a vektorcsalád "közel egységvektorokból" áll. Első eredményünk az euklideszi norma esetére vonatkozik, majd ezt erősítjük és kiterjesztjük tetszőleges, nem feltétlenül szimmetrikus normákra is.

Publications

Journal publications

I declare that as of the submission of this thesis, I have the following two publications, and that both of them are used in the dissertation.

[2] Gergely Ambrus and Rainie Bozzai. Colourful vector balancing. Mathe- matika, 70(4), August 2024.

[4] Gergely Ambrus and Rainie Heck. A note on the Steinitz constant. Accepted for publication; Mathematika, 2026.

Acknowledgments

First of all, I would like to thank my supervisor, Gergely Ambrus. I met Gergely because he taught me discrete and convex geometry in 2019 while I was studying abroad in Budapest. It was his course that captured my interest in this topic, and I am extremely grateful for all of the help and support that he has invested in me since then, beginning with a visiting student position for the 2022-2023 academic year and my PhD studies at SZTE from 2023-2025. I would also like to thank my family for always supporting me in my academic journey, as well as the friends that I have made along the way.

I was also fortunate to be supported by the United States National Science Foundation Graduate Research Fellowship Program (NSF GRFP) from 2020-2023, which made my visiting position at SZTE possible .

Bibliography

[1] Gergely Ambrus. A generalization of Bang's lemma. *Proceedings of the American Mathematical Society*, 151(03):1277–1284, 2023.

[2] Gergely Ambrus and Rainie Bozzai. Colourful vector balancing. *Mathematika*, 70(4), August 2024.

[3] Gergely Ambrus and Bernardo González Merino. Large signed subset sums. *Mathematika*, 67(3):579–595, 2021.

[4] Gergely Ambrus and Rainie Heck. A note on the Steinitz constant. *Accepted for publication; Mathematika*, 2026.

[5] Arturs Backurs, Jerry Chee, Sivakanth Gopi, Rainie Heck, Janardhan Kulkarni, and Thomas Rothvoss. Discquant: A quantization method for neural networks inspired by discrepancy theory. *Proceedings of Machine Learning Research ML Research Press (2025)*, pages 924–951, 2025.

[6] Wojciech Banaszczyk. The Steinitz Constant of the Plane. *Journal für die reine und angewandte Mathematik*, 373:218–220, 1987.

[7] Wojciech Banaszczyk. A note on the Steinitz constant of the Euclidean plane. *CR Math. Rep. Acad. Sci. Canada*, 12(4):97–102, 1990.

[8] Wojciech Banaszczyk. Balancing vectors and convex bodies. *Studia Mathematica*, 106(1):93–100, 1993.

[9] Wojciech Banaszczyk. Balancing vectors and Gaussian measures of n-dimensional convex bodies. *Random Structures & Algorithms*, 12(4):351–360, 1998.

[10] Nikhil Bansal. Constructive algorithms for discrepancy minimization. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 3–10, 2010.

[11] Nikhil Bansal, Daniel Dadush, Shashwat Garg, and Shachar Lovett. The Gram–Schmidt walk: A cure for the Banaszczyk blues. *Theory of Computing*, 15(21): 1–27, 2019.

[12] Imre Bárány. A Vector-sum Theorem and its Application to Improving Flow Shop Guarantees. *Mathematics of Operations Research*, 6(3):445–452, 1981.

[13] Imre Bárány. A matrix version of the Steinitz lemma. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2024(809):261–267, 2024.

[14] Imre Bárány and Benjamin Doerr. Balanced partitions of vector sequences. *Linear algebra and its applications*, 414(2-3):464–469, 2006.

[15] József Beck and Tibor Fiala. “Integer-making” theorems. *Discrete Applied Mathematics*, 3(1):1–8, 1981.

[16] FA Behrend. The Steinitz-Gross theorem on sums of vectors. *Canadian Journal of Mathematics*, 6:108–124, 1954.

[17] I.S. Belov and J.J. Stolin. An Algorithm for the Flow Shop Problem. *Mathematical Economics and Functional Analysis*, 499, 1974.

[18] Viktor Bergström. Zwei Sätze über ebene Vectorpolygone. In *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, volume 8, pages 206–214. Springer, 1931.

[19] Viktor Bergström. Ein neuer Beweis eines Satzes von E. Steinitz. In *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, volume 8, pages 148–152. Springer, 1931.

[20] Patrick Billingsley. *Probability and Measure*. Wiley, 3rd edition, 1995.

[21] Rainie Bozzai and Thomas Rothvoss. Stronger coreset bounds for kernel density estimators via chaining. *Submitted.*, 2023.

[22] Imre Bárány. On the power of linear dependencies. In *Building bridges: Between Mathematics and Computer Science*, volume 19 of *Bolyai Society Mathematical Studies*, pages 31–45. Bolyai Society and Springer, 2010.

[23] Imre Bárány and Victor Grinberg. On some combinatorial questions in finite dimensional spaces. *Linear Algebra and its Applications*, 41:1–9, 1981.

[24] Imre Bárány and Victor Grinberg. A Vector-Sum Theorem in Two-Dimensional Space. *Periodica Mathematica Hungarica*, 16:135–138, 1985.

[25] Sergei Akopovich Chobanyan. Structure of the set of sums of a conditionally convergent series in a normed space. *Mathematics of the USSR-Sbornik*, 56(1): 49, 1987.

[26] Sergej Chobanyan. Convergence A.S. of Rearranged Random Series in Banach Space and Associated Inequalities. In Jørgen Hoffmann-Jørgensen, James Kuelbs, and Michael B. Marcus, editors, *Probability in Banach Spaces, 9*, pages 3–29, Boston, MA, 1994. Birkhäuser Boston.

[27] Ira Damsteeg and Israel Halperin. The Steinitz-Gross theorem on sums of vectors. *Trans. Roy. Soc. Canada Sect. III*, 44:31–35, 1950.

[28] Aryeh Dvoretzky. Problem. In *Convexity*, volume 7 of *Proceedings of Symposia in Pure Mathematics*, page 496. Amer. Math. Soc., 1963.

[29] William Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. Wiley, 1971.

[30] Tibor Fiala. Közelítő Algoritmus a Három Gép Problémára. *Alkalmazott Matematikai Lapok*, 3:389–398, 1977.

[31] Walter Gautschi. Some elementary inequalities relating to the gamma and incomplete gamma function. *Journal of Mathematics and Physics*, 38:77–81, 1959.

[32] Apostolos A Giannopoulos. On some vector balancing problems. *Studia Mathematica*, 122:225–234, 1997.

[33] Efim Davydovich Gluskin. Extremal properties of orthogonal parallelepipeds and their applications to the geometry of Banach spaces. *Mathematics of the USSR-Sbornik*, 64(1):85, 1989.

[34] Victor Grinberg and Sergey Sevast'yanov. Value of the Steinitz Constant. *Functional Analysis and its Applications*, 14:125–126, 1980.

[35] Victor Grinberg and Sergey V. Sevast'yanov. Value of the Steinitz constant. *Functional Analysis and Its Applications*, 14(2):125–126, 1980.

[36] Wilhelm Groß. Bedingt konvergente Reihen. *Monatshefte für Mathematik und Physik*, 28:221–237, 1917.

[37] H Hadwiger. Ein Satz über bedingt konvergente Vektorreihen. *Mathematische Zeitschrift*, 47(1):663–668, 1942.

[38] Hugo Hadwiger. Ein Satz über geschlossene Vektorpolygone des Hilbertschen Raumes. *Mathematische Zeitschrift*, 41(1):732–738, 1936.

[39] Hugo Hadwiger. Über das Umordnungsproblem im Hilbertschen Raum. *Mathematische Zeitschrift*, 46(1):70–79, 1940.

[40] Israel Halperin. Sums of a series, permitting rearrangements. *CR Math. Rep. Acad. Sci. Canada*, 8(2):87–102, 1986.

[41] Christopher Harshaw, Fredrik Sävje, Daniel A Spielman, and Peng Zhang. Balancing covariates in randomized experiments with the gram–schmidt walk design. *Journal of the American Statistical Association*, pages 1–13, 2024.

[42] Laurel Heck, Victor Reis, and Thomas Rothvoss. The vector balancing constant for zonotopes. *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1292–1300, 2022.

[43] Mikhail Kadets. On a Property of Broken Lines in n-Dimensional Space. *Uspekhi Mat. Nauk.*, 8(1):139–143, 1953.

[44] Boris Sergeevich Kashin. On one isometric operator in $L^2(0,1)$. *Comptes Rendus de l'Academie Bulgare des Sciences*, 38(12):1613–1615, 1985.

[45] Yitzhak Katznelson and O Carruth McGehee. Conditionally convergent series in $\mathbb{R}^\infty$. *Michigan Mathematical Journal*, 21(2):97–106, 1974.

[46] Paul Lévy. Sur les séries semi-convergentes. *Nouvelles annales de Mathématiques: journal des candidats aux Écoles Polytechnique et Normale*, 5:506–511, 1905.

[47] Shachar Lovett and Raghu Meka. Constructive discrepancy minimization by walking on the edges. *SIAM Journal on Computing*, 44(5):1573–1582, 2015.

[48] L. Lovász, J. Spencer, and K. Vesztergombi. Discrepancy of set-systems and matrices. *European Journal of Combinatorics*, 7(2):151–160, 1986.

[49] Ben Lund and Alexander Magazinov. The sign-sequence constant of the plane. *Acta Mathematica Hungarica*, 151:117–123, 2017.

[50] Jiri Matousek. *Geometric discrepancy: An illustrated guide*, volume 18. Springer Science & Business Media, 1999.

[51] Timm Oertel, Joseph Paat, and Robert Weismantel. A Colorful Steinitz Lemma with Application to Block-Structured Integer Programs. *Mathematical Programming*, 204(1):677–702, 2024.

[52] Jeff Phillips and Wai Tai. Near-optimal coresets of kernel density estimates. *Discrete and Computational Geometry*, 63, 06 2020.

[53] Jeff M. Phillips. ε-samples for Kernels. In *Proceedings of the Twenty-Fourth Annual Symposium on Discrete Algorithms*, pages 1622–1632. SIAM, 2013.

[54] Victor Reis and Thomas Rothvoss. Vector balancing in Lebesgue spaces. *Random Structures & Algorithms*, 08 2022.

[55] Bernhard Riemann. *Über die Darstellbarkeit einer Function durch eine trigonometrische Reihe*, volume 13. Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, 1868.

[56] Peter Rosenthal. The Remarkable Theorem of Lévy and Steinitz. *The American Mathematical Monthly*, 94(4):342–351, 1987.

[57] Rolf Schneider. *Convex Bodies: The Brunn–Minkowski Theory*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2 edition, 2013.

[58] Sergey Sevastianov. Nonstrict vector summation in multi-operation scheduling. *Annals of Operations Research*, 83(0):179–212, 1998.

[59] Sergey Sevast'janov and Wojciech Banaszczyk. To the Steinitz Lemma in Coordinate Form. *Discrete Mathematics*, 169(1):145–152, 1997.

[60] S.V. Sevast'janov. On some geometric methods in scheduling theory: a survey. *Discrete Applied Mathematics*, pages 59–82, 1994.

[61] Sergey V. Sevastyanov. Asymptotical Approach to Some Scheduling Problems. In *Third All-Union Conf. Problems of Theoretical Cybernetics*, pages 67–69, Thes. Dokl., Novosibirsk, 1974. (in Russian).

[62] Sergey V. Sevastyanov. Asymptotical Approach to Some Scheduling Problems. *Upravlyaemye Sistemy*, 14:40–51, 1975. (in Russian).

[63] Sergey V. Sevast'yanov. On Approximate Solutions of Scheduling Problems. *Metody Discretnogo Analiza*, 32:66–75, 1978.

[64] Sergey V. Sevast'yanov. On the approximate solution of the problem of calendar planning. *Upravlyaemye Systemy*, 20:49–63, 1980. (in Russian).

[65] Sergey V. Sevastyanov. On a Compact Vector Summation. *Diskretnaya Matematika*, 3:66–72, 1991. (in Russian).

[66] Joel Spencer. Balancing games. *Journal of Combinatorial Theory, Series B*, 23(1): 68–74, 1977.

[67] Joel Spencer. Balancing unit vectors. *Journal of Combinatorial Theory*, 30: 349–350, 1981.

[68] Joel Spencer. Six standard deviations suffice. *Transactions of the American Mathematical Society*, 289(2):679–706, 1985.

[69] Joel Spencer. Balancing vectors in the max norm. *Combinatorica*, 6:55–65, 1986.

[70] Ross M Starr. Quasi-equilibria in markets with non-convex preferences. Appendix 2: The Shapley–Folkman theorem, pp. 35–37). *Econometrica: Journal of the Econometric Society*, pages 25–38, 1969.

[71] Ernst Steinitz. Bedingt Konvergente Reihen und Konvexe Systeme. *Journal für die reine und angewandte Mathematik*, 143:128–176, 1913.

[72] Ernst Steinitz. Bedingt Konvergente Reihen und Konvexe Systeme. *Journal für die reine und angewandte Mathematik*, 144:1–40, 1914.

[73] Ernst Steinitz. Bedingt Konvergente Reihen und Konvexe Systeme. *Journal für die reine und angewandte Mathematik*, 146:1–52, 1916.

[74] Konrad Swanepoel. Balancing unit vectors. *Journal of Combinatorial Theory*, 89: 105–112, 2000.

[75] Wai Ming Tai. Optimal Coreset for Gaussian Kernel Density Estimation. In Xavier Goaoc and Michael Kerber, editors, *38th International Symposium on Computational Geometry (SoCG 2022)*, volume 224 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 63:1–63:15, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[76] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018.