

# ON THE POWER OF LINEAR DEPENDENCIES

IMRE BÁRÁNY

ABSTRACT. Simple as they may be, linear dependencies have proved very useful in many ways. In this survey several geometric applications of linear dependencies are discussed, focusing on rearrangements of sums and on sums with  $\pm 1$  signs.

## 1. INTRODUCTION

Linear algebra is a basic and powerful tool in many areas of mathematics. In combinatorics, for instance, there are several cases when the size of a set can be bounded by a number  $n$  because the elements of the set are associated with vectors in  $\mathbb{R}^n$ , and these vectors turn out to be linearly independent. The excellent book [1] by Babai and Frankl (which is unfortunately, still unpublished) contains thousands of beautiful applications of the so-called linear algebra method.

This article describes another kind of use of linear algebra, this time in geometry. The method uses linear dependencies and is often referred to as the *method of floating variables*. The same method is used at other places as well, for instance in discrepancy theory, in the Beck-Fiala theorem [7] or [8], and in probability theory, [9]. Here we focus on rearrangement of sums and on sums with  $\pm 1$  signs.

In what follows the setting is the  $d$ -dimensional Euclidean space  $\mathbb{R}^d$ , together with a (Minkowski) norm,  $\|\cdot\|$  whose unit ball is denoted by  $B$  or  $B^d$ . We write  $\mathbb{N}$  for the set of natural numbers and  $[n]$  for the set  $\{1, 2, \dots, n\}$  where  $n \in \mathbb{N}$ . We assume that  $V \subset B$  is a finite set.

## 2. THE STEINITZ LEMMA

Assume  $V \subset B$  is finite and  $\sum_{v \in V} v = 0$ . The question, due to Riemann and Lévy, is whether there is an ordering,  $v_1, v_2, \dots, v_n$ , of the elements of  $V$  such that all partial sums along this order are bounded by a number that only depends on  $B$ . The answer is yes. An incomplete proof came from Lévy [17] in 1905. The first complete proof, from 1913, is due to Steinitz [20], and that's why it is usually called the Steinitz Lemma.

**Theorem 2.1.** *Given a finite set  $V \subset B$  with  $\sum_{v \in V} v = 0$ , where  $B$  is the unit ball of a norm in  $\mathbb{R}^d$ , there is an ordering  $v_1, v_2, \dots, v_n$  of the elements of  $V$  such that for all  $k \in [n]$*

$$\sum_{i=1}^k v_i \in dB.$$

So all partial sums are contained in a blown-up copy of the unit ball, with blow-up factor  $d$ . We will return to the value of the blow-up factor later.

Let's see first the proof of Theorem 2.1, which is our first application of linear dependencies.

*Proof.* The key step is the construction of sets  $V_{d+1} \subset \dots \subset V_{n-1} \subset V_n = V$  where  $|V_k| = k$ , together with functions  $\alpha_k : V_k \rightarrow [0, 1]$  satisfying

$$\begin{aligned} \sum_{v \in V_k} \alpha_k(v)v &= 0 \\ \sum_{v \in V_k} \alpha_k(v) &= k - d. \end{aligned}$$

So the functions  $\alpha_k(\cdot)$  are *linear dependencies* on  $V_k$ , with coefficients in  $[0, 1]$  that sum to  $k - d$ .

The construction goes by backward induction. The starting case  $k = n$  is easy:  $V_n = V$  and  $\alpha_n = (n - d)/n$  satisfy the requirements. Assume now that  $V_k$  and  $\alpha_k$  have been constructed, and consider the auxiliary system

$$\begin{aligned} \sum_{v \in V_k} \beta(v)v &= 0, \\ \sum_{v \in V_k} \beta(v) &= k - 1 - d, \\ 0 \leq \beta(v) &\leq 1 \text{ for all } v \in V_k. \end{aligned}$$

Write  $P$  for the set of functions  $\beta : V_k \rightarrow [0, 1]$  satisfying this auxiliary system. The elements of  $P$  can and will be regarded as vectors in  $\mathbb{R}^k$  whose components are indexed by the elements of  $V_k$ .

Note that  $P$  is non-empty since  $\beta(v) = \frac{k-1-d}{k-d}\alpha_k(v)$  belongs to  $P$ . Thus  $P$  is a convex polytope, lying in the unit cube of  $\mathbb{R}^k$ . Let  $\beta^*(\cdot)$  be an extreme point of  $P$ .

We claim now that  $\beta^*(v) = 0$  for some  $v \in V_k$ . Indeed, assume  $\beta^*(v) > 0$  for all  $v \in V_k$ . The auxiliary system has  $d + 1$  equations and  $k$  variables, so at least  $k - (d + 1)$  of the inequalities  $\beta^*(v) \leq 1$  are satisfied as equalities (the inequalities  $\beta^*(v) \geq 0$  are all strict). Then  $\sum_{v \in V_k} \beta^*(v) > k - d - 1$  (we use  $k > d + 1$  here), which contradicts one of the conditions defining  $P$ .

Let  $v^* \in V_k$  be an element with  $\beta^*(v^*) = 0$ , and define  $V_{k-1} = V_k \setminus \{v^*\}$  and  $\alpha_{k-1}(v) = \beta^*(v)$  for all  $v \in V_{k-1}$ . All conditions are satisfied for  $V_{k-1}$  and  $\alpha_{k-1}$ . The construction is finished.

Now we ready to order the elements of  $V$ . For  $k = n, n - 1, \dots, d + 2$  we set, quite naturally,

$$v_k = V_k \setminus V_{k-1}.$$

The remaining  $d + 1$  vectors are ordered arbitrarily.

We check, finally, that all partial sums are contained in  $dB$ . This is trivial for the first  $d$  partial sums. Assume now that  $k \geq d + 1$ .

$$\sum_1^k v_i = \sum_{v \in V_k} v = \sum_{v \in V_k} v - \sum_{v \in V_k} \alpha_k(v)v = \sum_{v \in V_k} (1 - \alpha_k(v))v.$$

Taking norms and using that  $1 - \alpha_k(v) \geq 0$  and  $\|v\| \leq 1$  gives

$$\left\| \sum_1^k v_i \right\| \leq \sum_{v \in V_k} (1 - \alpha_k(v)) = k - (k - d) = d. \quad \square$$

This splendid proof, due to Grinberg and Sevastyanov [13], is a streamlined version of an earlier one by Sevastyanov [21]. Steinitz's original proof also used linear dependencies, and gave constant  $2d$  instead of  $d$ . Yet another proof, from Bárány [3], using linear dependencies again, gave blow-up factor  $1.5d$ .

It comes as a surprise that the norm plays a rather marginal role. I describe now another proof, due to Kadets [15], in which the norm is important. The proof works for the Euclidean norm  $B_2$  and gives the weaker blow-up factor  $C_d = \sqrt{(4^d - 1)/3}$ . It goes by induction on dimension and the case  $d = 1$  is very simple: one takes positive elements as long as the sum stays below 1, then one takes negative elements as long as the sum is above  $-1$ , and so on. For the induction step  $d - 1 \rightarrow d$ , let  $V \subset B_2$  be a finite set with  $\sum_{v \in V} v = 0$ . We choose a subset  $W \subset V$  for which  $\|\sum_{v \in W} v\|$  is maximal among all subsets of  $V$ , and set  $a = \sum_{v \in W} v$ . For  $v \in V$  let  $\bar{v}$  denote its orthogonal projection onto the subspace  $A$  orthogonal to  $a$ , and set  $v_a = v - \bar{v}$ . It follows from the maximality of  $W$  that, for all  $v \in W$ ,  $v_a$  points the same direction as  $a$ , and, for all  $v \notin W$ ,  $v_a$  points the opposite direction. Further, each  $\bar{v} \in A$  and has Euclidean norm at most one in  $A$  which is the  $d - 1$ -dimensional Euclidean space. Also,  $\sum_{v \in W} \bar{v} = 0$ . By the induction hypothesis, there is an ordering  $v_1, v_2, \dots$  of the vectors in  $W$  with all partial sums of the  $\bar{v}_i$  having Euclidean length at most  $C_{d-1}$ . The same applies to the set  $V \setminus W$ , so its elements can be ordered as  $w_1, w_2, \dots$  with all partial sums of the  $\bar{w}_j$  shorter than or equal to  $C_{d-1}$ . The sequences  $v_1, \dots$  and  $w_1, \dots$  are then interlaced making sure (using the method given for the case  $d = 1$ ) that the absolute value of the  $a$ -component of each partial sum is at most 1. Then the square of each partial sum of the interlaced sequence is at most  $4C_{d-1}^2 + 1 = C_d^2$  as one can easily see.

This is quite a neat proof, yet the other one is superior: it works for all (even non-symmetric) norms, gives a far better bound, and is much more elegant, as far as I can judge.

### 3. THE STORY OF THE STEINITZ LEMMA

The story actually began with Riemann who showed that a conditionally convergent series (of real numbers) can be made to converge, by a suitable rearrangement of the series, to any real number. What happens with a conditionally convergent series of  $d$ -dimensional vectors? Let  $U = \{u_1, u_2, \dots\}$  be the vectors in the series, and let  $\sigma U$  be the set of points that it can be made to converge by rearrangements. It turns out that  $\sigma U$  is always an affine subspace of  $\mathbb{R}^d$ . For  $d = 1$  this is equivalent to Riemann's result. In higher dimensions the problem quickly reduces to the statement of what is called now Steinitz Lemma, with arbitrary norm and arbitrary constant depending only on dimension. This is what Steinitz proved in [20].

The smallest constant the Steinitz lemma holds with is a number, to be denoted by  $S(B)$ , that depends only on the unit ball  $B$ . It is called the Steinitz constant of  $B$ . Theorem 2.1 says that  $S(B) \leq d$  for all  $B$  in  $\mathbb{R}^d$ . This norm need not be symmetric: a quick look at the last step of the proof shows that  $\|\lambda v\| \leq \lambda \|v\|$  was only used with  $\lambda \geq 0$ . For non-symmetric norms, the estimate  $S(B) \leq d$  is best possible. In the example proving this,

$B$  is a simplex in  $\mathbb{R}^d$  with its center of gravity at the origin, and  $V$  consists of the vectors pointing to the vertices of this simplex.

Yet, of course,  $S(B)$  might be much smaller than  $d$  for any particular norm. Write  $B_p$  for the unit ball of the  $\ell_p$  norm in  $\mathbb{R}^d$ , or  $B_p^d$  if we wish to stress the dimension of the underlying space. It is easy to see that  $S(B_1) \geq d/2$ , so the order of magnitude for symmetric norms cannot be improved. In 1931 Bergström [9] proved that  $S(B_2) = \sqrt{5}/2$  for  $d = 2$ , a surprisingly precise result. The lower bound comes from a construction consisting of  $n/2$  copies of the vector  $(\sqrt{1-t^2}, -t)$ ,  $n/2$  copies of the vector  $(-\sqrt{1-t^2}, -t)$  where  $t = 1/n$ , and the vector  $(0, 1)$ . This is essentially  $n/2$  copies of a slightly modified  $e_1$  and  $-e_1$  compensated by  $e_2$ . In higher dimensions, the analogous example shows that  $S(B_2^d) \geq \sqrt{d+3}/2$ . It has been conjectured that

$$S(B_2^d) = O(d^{1/2}).$$

But even the much weaker  $S(B_2^d) = o(d)$  estimate seems to be out of reach though quite a few mathematicians have tried.

The case of the maximum norm,  $B_\infty$ , is also open. An example can be built from a  $d+1$  by  $d+1$  Hadamard matrix: its first row is the all 1 vector, and the vectors in  $V$  are the  $d+1$  columns of this matrix with the first coordinate deleted. It is not hard to see that the squared Euclidean norm of the sum of  $k$  vectors from  $V$  is  $k(d+1-k)$ . This shows, when  $k = (d+1)/2$ , that one coordinate of the sum is at least  $(d+1)/\sqrt{4d}$  in absolute value, implying that the conjecture,

$$S(B_\infty^d) = O(d^{1/2}),$$

if true, is best possible. But again, there is no proof in sight even for the much weaker  $S(B_\infty^d) = o(d)$  estimate.

The Steinitz Lemma has many applications. It is used, or can be used to prove Lyapunov's theorem stating that the image of an atomless vector valued measure is always convex. In Operations Research, the Lemma has been applied to scheduling problems. In particular, it was used to find optimal flow shop and job shop schedules in polynomial time under some mild conditions, although these scheduling problems are NP-hard in general. See the excellent survey by Sevastyanov [22], or some of the original works [3], [21]. Halperin and Ando [14] cite 290 papers related to the Steinitz Lemma up to 1989, by now the number must be much higher.

#### 4. SIGNED SUM

In this section  $V \subset B$  is a finite set, again, and we want to find signs  $\varepsilon(v) = 1$  or  $-1$  for all  $v \in V$  such that  $\sum_{v \in V} \varepsilon(v)v$  is not too large. In the following theorem, which is from Bárány, Grinberg [4], we work with the Euclidean ball  $B_2$ .

**Theorem 4.1.** *Under the above conditions there are signs  $\varepsilon(v)$  such that*

$$\sum_{v \in V} \varepsilon(v)v \in \sqrt{d}B.$$

The example when  $V$  consists of  $d$  pairwise orthogonal unit vectors shows that the above estimate is best possible.

*Proof.* The proof is in two steps. For the first one, consider the set of linear dependencies  $\alpha : V \rightarrow [-1, 1]$ , that is, functions satisfying

$$\sum_{v \in V} \alpha(v)v = 0 \text{ and } -1 \leq \alpha(v) \leq 1 \text{ for all } v \in V.$$

These functions form a convex polytope in the  $\pm 1$  cube of  $\mathbb{R}^V$ , of dimension at least  $n - d$  (where  $|V| = n$ ). This polytope is non-empty since it contains the origin. At an extreme point,  $\alpha^*$  say, of this polytope many  $\alpha^*(v) \in \{-1, 1\}$ . Precisely, the set of vectors  $v \in V$  with  $-1 < \alpha^*(v) < 1$  are linearly independent since otherwise  $\alpha^*$  is not an extreme point. For simpler notation we assume that this happens for the vectors  $v_1, \dots, v_k$ , where, obviously,  $k \leq d$ . For the rest of the vectors  $v_i$  (with  $i \in \{k+1, \dots, n\}$ ), we have  $\alpha^*(v_i) = \alpha_i \in \{-1, 1\}$ .

Define now  $u = \sum_{k+1}^n \alpha_i v_i$  and set

$$(4.1) \quad Q = \left\{ \sum_1^k \beta_i v_i : \beta_i \in [-1, 1] \right\}.$$

Clearly,  $Q$  is a parallelotope whose sides have Euclidean length at most 2. What we have shown so far is that  $0 \in u + Q$ .

The second step in the proof is geometric. We claim that if a point  $a$  lies in a parallelotope  $Q$  defined by  $k$  linearly independent vectors  $v_1, \dots, v_k$  as in (4.1) with  $\|v_i\| \leq 1$ , then  $Q$  has a vertex at distance at most  $\sqrt{k}$  from  $a$ . The theorem follows from this since  $k \leq d$ .

We prove the claim by induction on  $k$ . The case  $k = 1$  is trivial. In the induction step  $k-1 \rightarrow k$ , we assume that  $a$  is in the interior of  $Q$  as otherwise  $a$  is on a facet of  $Q$  which is itself a parallelotope of dimension  $k-1$  and the induction works. Now put a Euclidean ball,  $B(a)$ , centered at  $a$ , into  $Q$  and increase its radius as long as you can with  $B(a)$  still remaining in  $Q$ . The maximal  $B(a)$  contained in  $Q$  has a point, say  $b$ , on the boundary of  $Q$ . Then  $b$  is contained in a face  $F$  of  $Q$  which is a  $k-1$ -dimensional parallelotope, whose defining vectors are of unit length at most. By induction,  $F$  has a vertex,  $w$  say, at distance at most  $\sqrt{k-1}$  from  $b$ . Of course,  $w$  is a vertex of  $Q$  as well. As  $B(a)$  touches  $F$  at  $b$ ,  $a-b$  is orthogonal to  $F$ . Further,  $\|a-b\| \leq 1$  as otherwise a ball of radius larger than 1 would be contained in  $Q$  which is impossible. Now

$$\|a-w\|^2 = (a-w)^2 = (a-b)^2 + (b-w)^2 \leq 1 + (k-1) = k,$$

finishing the proof.  $\square$

The first step of the proof works for every norm but the second, more geometric step, does not. In general, in the second step one can only guarantee distance  $d$  from a vertex. A point  $a$  in a parallelotope  $Q$  of side length 2 in norm  $B$ , may be far away from all vertices of  $Q$ . The straightforward example of the  $\ell_1$  norm shows that every vertex of the  $\pm 1$  cube in  $\mathbb{R}^d$  is at distance  $d$  from the origin.

The situation is better for the  $B_\infty$  norm, because then every point of the parallelotope  $Q$  is closer than  $6\sqrt{d}$  from some of its vertices. This is a result

due to Spencer [19] and, independently to Gluskin [12], who was relying on earlier work of Kashin [16]. It is an interesting fact that Spencer finds the signs by a combination of the pigeon hole principle and random methods, while Gluskin and Kashin use volume estimates and Minkowski's theorem on lattice points in 0-symmetric convex bodies.

In connection with this we mention a striking question of J Komlós (cf. [2] or [19]). He asks whether there is a universal constant  $C$  such that for every  $d \geq 1$  and for every finite  $V \subset B_2^d$ , there are signs  $\varepsilon(v)$  for each  $v \in V$  such that  $\sum_{v \in V} \varepsilon(v)v \in CB_\infty^d$ . The best result so far in this direction is that of Banaszczyk [2]. He showed the existence of signs such that the signed sum lies in  $C(d)B_\infty^d$  where the constant  $C(d)$  is of order  $\sqrt{\log d}$ .

## 5. SIGNING VECTOR SEQUENCES

In this section  $U$  will denote a sequence,  $u_1, u_2, \dots$  from the unit ball  $B \subset \mathbb{R}^d$ . This time  $B$  is symmetric, and the sequence may be finite or infinite. We wish to find signs  $\varepsilon_i \in \{-1, +1\}$  such that all partial sums  $\sum_1^n \varepsilon_i u_i$  are bounded by a constant depending only on  $B$ . The following result is from Bárány, Grinberg [4].

**Theorem 5.1.** *Under the above conditions there are signs  $\varepsilon_i$  such that for all  $n \in \mathbb{N}$*

$$\sum_1^n \varepsilon_i u_i \in (2d - 1)B.$$

*Proof.* We will only prove that all partial sums are in  $2dB$ . The improvement to  $2d - 1$  is explained in the remark after this proof.

We start again with a construction, which is the prime example of the method of “floating variables”.

Define  $U_k = \{u_1, u_2, \dots, u_{k+d}\}$ ,  $k = 0, 1, 2, \dots$ . We are going to construct mappings  $\beta_k : U_k \rightarrow [-1, 1]$  and subsets  $W_k \subset U_k$  with the following properties (for all  $k$ ):

- (i)  $\sum_{U_k} \beta_k(u)u = 0$ ,
- (ii)  $\beta_k(u) \in \{-1, 1\}$  whenever  $u \in W_k$ ,
- (iii)  $|W_k| = k$  and  $W_k \subset W_{k+1}$  and  $\beta_{k+1}(u) = \beta_k(u)$  if  $u \in W_k$ .

The construction is by induction on  $k$ . For  $k = 0$ ,  $W_0 = \emptyset$  and  $\beta_0(\cdot) = 0$  clearly suffice. Now assume that  $\beta_k$  and  $W_k$  have been constructed and satisfy (i), (ii), and  $|W_k| = k$  from (iii). The  $d + 1$  vectors in  $U_{k+1} \setminus W_k$  are linearly dependent, so there are  $\gamma(u) \in \mathbb{R}$  not all zero such that

$$\sum_{U_{k+1} \setminus W_k} \gamma(u)u = 0.$$

Putting  $\beta_k(u_{k+d+1}) = 0$ , we have

$$\sum_{W_k} \beta_k(u)u + \sum_{U_{k+1} \setminus W_k} (\beta_k(u) + t\gamma(u))u = 0$$

for all  $t \in \mathbb{R}$ . For  $t = 0$  all coefficients lie in  $[-1, 1]$ . Hence for a suitable  $t = t^*$ , all coefficients still belong to  $[-1, 1]$ , and  $\beta_k(u) + t\gamma(u) \in \{-1, 1\}$  for some  $u^* \in U_{k+1} \setminus W_k$ . Set now  $W_{k+1} = W_k \cup \{u^*\}$  and  $\beta_{k+1}(u) = \beta_k(u)$ , if  $u \in W_k$ , and  $\beta_{k+1}(u) = \beta_k(u) + t^*\gamma(u)$ , if  $u \in U_{k+1} \setminus W_k$ . Then  $W_{k+1}$  and

$\beta_{k+1}$  satisfy (i) and (ii) and  $|W_{k+1}| = k+1$  from (iii). Moreover,  $W_k \subset W_{k+1}$  and  $\beta_{k+1}(u) = \beta_k(u)$  for all  $u \in W_k$ .

We now define the signs  $\varepsilon_i$ . Set  $\varepsilon_i = 1$  if  $u_i \in W_k$  and  $\beta_k(u_i) = 1$  for some  $k$ , and set  $\varepsilon_i = -1$  if  $u_i \in W_k$  and  $\beta_k(u_i) = -1$  for some  $k$ . As  $\beta_k(u)$  stabilizes, that is,  $\beta_k(u) = \beta_{k+1}(u) = \beta_{k+2}(u) = \dots$  once  $u \in W_k$ , this definition is correct for all vectors that appear in some  $W_k$ . For the remaining (at most  $d$ ) vectors one can set  $\varepsilon_i = \pm 1$  arbitrarily.

Again, we have to check the partial sums. The first  $d$  (actually, the first  $2d$ ) partial sums lie automatically in  $2dB$ . For  $n > d$  define  $k = n - d$ . Denoting  $\varepsilon_i$  by  $\varepsilon(u_i)$  or simply by  $\varepsilon(u)$  we have

$$\begin{aligned} \sum_1^n \varepsilon_i u_i &= \sum_{u \in U_k} \varepsilon(u) u = \sum_{u \in U_k} \varepsilon(u) u - \sum_{u \in U_k} \beta_k(u) u \\ &= \sum_{u \in U_k} (\varepsilon(u) - \beta_k(u)) u = \sum_{u \in U_k \setminus W_k} (\varepsilon(u) - \beta_k(u)) u. \end{aligned}$$

Note that the last sum has only  $d$  terms, because  $U_k \setminus W_k$  has exactly  $d$  elements. We take the norm:

$$(5.1) \quad \left\| \sum_1^n \varepsilon_i u_i \right\| \leq \sum_{u \in U_k \setminus W_k} |\varepsilon(u) - \beta_k(u)| \leq 2d,$$

since every term in the last sum is at most 2.  $\square$

**Remark.** Where can one get  $2d - 1$  instead of  $2d$  in this proof? Well, when choosing the suitable  $t^*$  which gives  $u^* \in U_{k+1}$  the coefficient 1 or  $-1$ , we can move from  $t = 0$  to both positive or negative values of  $t$ , and this degree of freedom helps. Here is a sketch of how this can be done. For each  $k \geq 1$  one has a *special element*  $v \in U_k \setminus W_k$  with the property that  $\beta_{k+1}(v) \geq \beta_k(v)$  if  $\beta_k(v) > 0$  and  $\beta_{k+1}(v) \leq \beta_k(v)$  if  $\beta_k(v) < 0$ . The special element remains the same as long as  $v \notin W_k$ . What can be reached this way is that  $\beta_k(v)$  has the same sign as long as  $v$  is special. Then  $|\varepsilon(v) - \beta_k(v)| \leq 1$  for the special element in the sum over  $U_k \setminus W_k$  in equation (5.1) and this is where we get  $2d - 1$  instead of  $2d$ . When the special  $v$  enters  $W_k$  we let  $v = u_{k+d}$  be the new special element, and the sign of  $\beta_l(v)$  for  $l > k$  is going to be the same as that of  $\beta_k(v)$ . There is no  $\beta_l(v)$  for  $l < k$  so they can't influence the validity of (5.1) for the previous indices. The choice of the first special element and the case when  $\beta_k(v)$  never reaches  $\pm 1$  needs extra care which we leave to the interested reader.

The above proof gives, in fact, a good algorithm for finding a suitable sign sequence. It is an almost on-line algorithm: it does not have to foresee the whole sequence. At each moment, it only keeps a buffer of  $d$  vectors with undecided signs. In fact, the previous remark shows that a buffer of size  $d - 1$  suffices. But smaller buffer wouldn't do. This was proved by Peng and Yan [18].

The sign sequence constant,  $E(B)$ , of the unit ball  $B$  is the smallest blow-up factor for which Theorem 5.1 holds. By the same theorem,  $E(B) \leq 2d - 1$  always holds for every symmetric norm in  $\mathbb{R}^d$ . For individual norms, of course, much better estimates are possible. The lower bounds for  $S(B_p)$

with  $p = 1, 2, \infty$  apply also to  $E(B)$ . We will return to this question in connection with Chobanyan's remarkable transference theorem in Section 7.

One can set up the problem leading to Theorem 5.1 more generally. Namely, assume we are given a sequence of sets  $V_1, V_2, \dots$ , with  $V_i \subset B$  and  $0 \in \text{conv}V_i$  for each  $i \in \mathbb{N}$ . Can one choose vectors,  $u_i \in V_i$  for each  $i$  such that each partial sum  $\sum_1^n u_i$  lies in  $cB$  with a suitable constant  $c$  that depends only on the norm. The answer is yes, with  $c = 2d$ , and the proof is similar to the one above, see [4]. The case of Theorem 5.1 is when  $V_i = \{u_i, -u_i\}$ . Several other questions treated in this paper have similar generalizations.

## 6. PARTITIONING A SEQUENCE

We now formulate Theorem 5.1 in a different form, suitable for generalization. We need one more piece of notation: if  $U' = \{u_{i_1}, u_{i_2}, \dots\}$  is a subsequence of  $U$  with  $i_1 < i_2 < \dots$ , then let  $\sum_n U'$  denote the sum of all  $u_{i_j}$  with  $i_j \leq n$ . This unusual notation will be very convenient.

With this notation the statement of Theorem 5.1 is that  $U$  can be partitioned into two subsequences  $U^+$  and  $U^-$  such that for every  $n \in \mathbb{N}$

$$\sum_n U^+ - \sum_n U^- \in 2dB, \text{ and also } \sum_n U^- - \sum_n U^+ \in 2dB.$$

The two statements here are equivalent since the norm is symmetric. Further, of course,  $U^+$  ( $U^-$ ) consists of elements of  $U$  for which  $\varepsilon = +1$  ( $\varepsilon = -1$ ).

Adding  $\sum_n U$  to both sides and dividing by 2 gives

$$\sum_n U^+ \in dB + \frac{1}{2} \sum_n U \text{ and } \sum_n U^- \in dB + \frac{1}{2} \sum_n U.$$

The new formulation of the Theorem 5.1 is this. Under the same conditions  $U$  can be partitioned into two subsequences  $U^1$  and  $U^2$  such that for all  $j = 1, 2$  and all  $n \in \mathbb{N}$

$$\sum_n U^j \in dB + \frac{1}{2} \sum_n U.$$

Can one partition  $U$  into  $r$  subsequences with similar properties? The answer is yes. The following theorem is from [5], and is an improved version of a similar result by Doerr and Srivastav [11].

**Theorem 6.1.** *Assume  $U \subset B$  is a sequence, and  $r \geq 2$  is an integer. Then  $U$  can be partitioned into  $r$  subsequences  $U^1, \dots, U^r$  such that for all  $j \in [r]$  and  $n \in \mathbb{N}$*

$$\sum_n U^j \in CdB + \frac{1}{r} \sum_n U,$$

where  $C$  is a universal constant which is smaller than 2.0005.

*Proof.* We only give a sketch. We set  $r = r_0 + r_1$  with  $r_0, r_1 \in \mathbb{N}$  whose values will be chosen later. Then partition  $U$  into two subsequences  $U_0$  and  $U_1$  so that, for all  $n \in \mathbb{N}$ ,

$$(6.1) \quad r_1 \sum_n U_0 - r_0 \sum_n U_1 \in (r_0 + r_1)dB.$$



This is accomplished by the same construction as in the proof of Theorem 5.1, only the bounding interval  $[-1, 1]$  is to be changed to  $[-r_0, r_1]$ . Then we add  $r_0 \sum_n U$  to both sides of (6.1) and divide by  $r = (r_0 + r_1)$  to obtain

$$\sum_n U_0 \in dB + \frac{r_0}{r} \sum_n U.$$

The same way we have

$$\sum_n U_1 \in dB + \frac{r_1}{r} \sum_n U.$$

The proof proceeds from here recursively, by choosing  $r_0 = r_{00} + r_{01}$  and splitting  $U_0$  into subsequences  $U_{00}$  and  $U_{01}$ , just as  $U$  was split into  $U_0$  and  $U_1$ , and then splitting  $U_{00}$  and  $U_{01}$  further.

This recursion gives rise to a recursion tree. It is a binary tree whose root is marked by  $r$ , its two children by  $r_0$  and  $r_1$ , etc. It has to have  $r$  leaves, each marked by 1. For the  $j$ th leaf, let  $b_j$  denote the sum of the reciprocals of the numbers on the nodes from this leaf to the root (including the leaf but excluding the root). The recursion gives then a partition of  $U$  into subsequences  $U^1, \dots, U^r$  such that for all  $j \in [r]$  and  $n \in \mathbb{N}$  we have

$$\sum_n U^j \in b_j dB + \frac{1}{r} \sum_n U.$$

Thus the recursion tree is to be built in such a way that all  $b_j$  be small. This can be achieved, giving  $b_j \leq 2.0005$  for all  $j \in [r]$ , see [11] and [5] for the details.  $\square$

## 7. A TRANSFERENCE THEOREM

The methods for the Steinitz constant  $S(B)$  and the sign-sequence constant  $E(B)$  are similar, and so are the bounds. Is there some deeper connection between them? This is answered by the following beautiful result of Chobanyan [10].

**Theorem 7.1.** *Assume  $B$  is the unit ball of symmetric norm in  $\mathbb{R}^d$ . Then  $S(B) \leq E(B)$ .*

The result shows that the sign-sequence problem is “easier” than the rearrangement problem. One may wonder whether the opposite inequality, that is,  $E(B) \leq CS(B)$ , holds with dimension independent constant  $C$ . It does hold with  $C = 2d - 1$  since  $S(B) \geq 1$  trivially and  $E(B) \leq 2d - 1$  by Theorem 5.1, but this is not interesting.

We mention that Theorem 7.1 holds in any normed space, not necessary finite dimensional. The proof below will show this.

*Proof.* Clearly both  $S(B)$  and  $E(B)$  are at least one. By the definition of  $S(B)$ , for every small  $\eta > 0$  there is a finite  $V \subset B$  with  $\sum_{v \in V} v = 0$ , such that  $V$  has an ordering  $v_1, \dots, v_n$  so that every partial sum along this ordering lies in  $S(B)B$ , but for every ordering of  $V$  there is a partial sum (along that ordering) which is outside of  $(S(B) - \eta)B$ .

Choose a small  $\eta > 0$  together with the finite  $V$  with the above properties. Consider now the sign-sequence  $\varepsilon_1, \dots, \varepsilon_n$  such that for all  $k \in [n]$ ,

$$\sum_1^k \varepsilon_i v_i \in E(B)B.$$

By the definition of  $E(B)$  such a sign sequence exists. The vectors  $v_i$  with  $\varepsilon_i = +1$ , in the same order, form a sequence  $u_1, u_2, \dots, u_m$ , while the vectors  $v_i$  with  $\varepsilon_i = -1$ , in the opposite order, form a sequence  $u_{m+1}, u_{m+2}, \dots, u_n$ . The sequence  $u_1, \dots, u_n$  is a rearrangement of  $V$ . Then one partial sum, the  $k$ th say, has norm greater than  $S(B) - \eta$ .

Assume  $k \leq m$ . Then, clearly,

$$\sum_1^k u_i = \frac{1}{2} \left( \sum_1^k v_i + \sum_1^k \varepsilon_i v_i \right) \in \frac{1}{2} (S(B) + E(B)) B.$$

This shows that  $S(B) - \eta < \frac{1}{2}(S(B) + E(B))$  implying

$$(7.1) \quad S(B) < E(B) + 2\eta.$$

Assume now that  $k > m$ . Then  $\sum_1^k u_i = -\sum_{k+1}^n u_i$  is outside  $(S(B) - \eta)B$ . Consequently,  $\sum_{k+1}^n u_i$  is outside  $(S(B) - \eta)B$  as well. But the last sum is just the sum of the first  $n - k$  elements of the sequence  $v_1, \dots, v_n$  that go with  $\varepsilon_i = -1$ . This sum is equal to

$$\frac{1}{2} \left( \sum_1^{n-k} v_i - \sum_1^{n-k} \varepsilon_i v_i \right) \in \frac{1}{2} (S(B) + E(B)) B,$$

again. This proves inequality (7.1) in all cases.

Finally, since (7.1) holds for all  $\eta > 0$ , we have  $S(B) \leq E(B)$ .  $\square$

## REFERENCES

- [1] L. Babai, P. Frankl, Linear Algebra Methods in Combinatorics, Preliminary version 2. Department of Computer Science, University of Chicago, 1992 (book, 216 pp).
- [2] W. Banaszczyk, Balancing vectors and Gaussian measures of  $n$ -dimensional convex bodies, *Random Structures and Alg.*, **12**(1998), 315–360.
- [3] I. Bárány, A vector-sum theorem and its application to improving flow shop guarantees, *Math. Oper. Res.*, **6**(1981), 445–452.
- [4] I. Bárány, V. S. Grinberg, On some combinatorial questions in finite dimensional spaces, *Linear Alg. Appl.*, **41**(1981), 1–9.
- [5] I. Bárány, B. Doerr, Balanced partitions of vector sequences, *Linear Alg. Appl.*, **414**(2006), 464–469.
- [6] J. Beck, T. Fiala, Roth’s estimate of the discrepancy of integer sequences is nearly sharp, *Combinatorica*, **1**(1981), 319–325.
- [7] J. Beck, T. Fiala, “Integer-making” theorems, *Discrete Appl. Math.*, **3**(1981), 1–6.
- [8] J. Beck, V. T. Sós, Discrepancy theory, in: Handbook of combinatorics, (ed. R. Graham, M. Grötschel, L. Lovász), Elsevier, Amsterdam, 1995, 1405–1446.
- [9] V. Bergström, Zwei Sätze über ebene Vectorpolygone, *Abh. Math. Sem. Univ. Hamburg*, **8**(1931), 148–152.
- [10] S. Chobanyan, Convergence a.s. of rearranged random series in Banach space and associated inequalities, in: Probability in Banach spaces, 9: Proceedings from the 9th international conference on probability in Banach spaces, (ed.: J. Hoffmann-Jorgensen et al.), Prog. Probab. (Birkhäuser Boston), 35, 3–29 (1994).

- [11] B. Doerr, A. Srivastav, Multicolour discrepancies, *Combinatorics, Prob. Comp.*, **12**(2003), 365–399.
- [12] Gluskin, E. D., Extremal properties of orthogonal parallelepipeds and their applications to the geometry of Banach spaces, *Mat. Sbornik*, **136**(1988), 85–96. (in Russian)
- [13] V. S. Grinberg, S. V. Sevastyanov, The value of the Steinitz constant, *Funk. Anal. Prilozh.*, **14**(1980), 56–57. (in Russian)
- [14] Halperin, I., Ando, T., Bibliography: Series of vectors and Riemann sums, Sapporo, Japan, 1989.
- [15] M. I. Kadets, On a property of broken lines in  $n$ -dimensional space, *Uspekhi Mat. Nauk*, **8**(1953), 139–143. (in Russian)
- [16] Kashin, B. S., On some isometric operators in  $L^2(0, 1)$ , *Comptes Rendus Acad. Bulg. Sciences*, **38**(1985), 1613–1615.
- [17] P. Lévy, Sur les séries semi-convergentes, *Nouv. Ann. de Math.*, **64**(1905), 506–511.
- [18] H. Peng, C. H. Yan, Balancing game with a buffer, *Adv. in Appl. Math.*, **21**(1998), 193–204.
- [19] J. Spencer, Six standard deviations suffice, *Transactions AMS*, **289**(1985), 679–706.
- [20] E. Steinitz, Bedingt konvergente Reihen und konvexe Systeme, *J. Reine Ang. Mathematik*, **143**(1913), 128–175, *ibid*, **144**(1914), 1–40, *ibid*, **146**(1916), 1–52.
- [21] S. V. Sevastyanov, On approximate solutions of a scheduling problem, *Metody Diskr. Analiza*, **32**(1978), 66–75. (in Russian)
- [22] S. V. Sevastyanov, On some geometric methods in scheduling theory: a survey, *Discrete Appl. Math.*, **55**(1994), 59–82.

IMRE BÁRÁNY

Rényi Institute of Mathematics  
Hungarian Academy of Sciences  
PO Box 127, 1364 Budapest  
Hungary  
and  
Department of Mathematics  
University College London  
Gower Street, London WC1E 6BT  
England  
e-mail: [barany@renyi.hu](mailto:barany@renyi.hu)