ISRAEL JOURNAL OF MATHEMATICS 112 (1999), 1–27

# A HIGHLY NON-SMOOTH NORM ON HILBERT SPACE

BY

JIŘÍ MATOUŠEK*

*Department of Applied Mathematics, Charles University*
*Malostranské nám. 25, 118 00 Praha 1, Czech Republic*
*and*
*Institut für Theoretische Informatik, ETH Zürich, Switzerland*
*e-mail: matousek@kam.mff.cuni.cz*

AND

EVA MATOUŠKOVÁ**

*Mathematical Institute, Czech Academy of Sciences*
*Žitná 25, 11567 Praha, Czech Republic*
*and*
*Institut für Mathematik, Johannes Kepler Universität, A-4040 Linz, Austria*
*e-mail: eva@caddo.bayou.uni-linz.ac.at*

ABSTRACT

We show that there exists a family $\mathcal{F}$ of unit balls in the Hilbert space $\ell_2$ such that $\bigcup\mathcal{F}$ is dense in $\ell_2$ but the complement of $\bigcup\mathcal{F}$ is large in the sense of measure. In an appendix, we present a considerable simplification of the proof due to Preiss. As a corollary, we prove that there is an equivalent norm $p$ on $\ell_2$ such that the set of points where $p$ is Fréchet differentiable is Aronszajn null. This disproves a conjecture of Borwein and Noll in a very strong sense.

* Research supported by ETH Zürich, by Czech Republic Grant GAČR 0194/1996 and by Charles University grants No. 193,194/1996.

** This author was supported by J. Kepler Universität and by the Czech Academy of Sciences grant GAAV-A1019705.

Received July 15, 1997 and in revised form October 24, 1997

## 1. Introduction

Numerous results deal with the smoothness properties of convex continuous functions or, more generally, of locally Lipschitz functions on Banach spaces. A theorem of Mazur says that every convex continuous function on a separable Banach space is Gâteaux differentiable on a dense $G_\delta$-set. If we confine ourselves to Banach spaces with a separable dual, any convex continuous function is even Fréchet differentiable on a dense $G_\delta$-set.

A natural question is whether the set of points of differentiability has to be large also in a sense of measure. In a finite dimension, the answer is positive; a classical theorem of Rademacher says that any locally Lipschitz function on $\mathbb{R}^n$ is Fréchet differentiable almost everywhere. In infinite-dimensional Banach spaces, there is no measure analogous to the Lebesgue measure on $\mathbb{R}^n$, and also no canonical notion of a null set analogous to the family of sets of Lebesgue measure zero in $\mathbb{R}^n$. One possible notion of a null set in infinite-dimensional Banach spaces was defined by Aronszajn [Aro76]. He proved that every locally Lipschitz function on a separable Banach space is Gâteaux differentiable everywhere except for such a null set. However, this is not the case for the Fréchet differentiability. Preiss and Tišer [PT95] even showed that on every separable infinite-dimensional Banach space, a Lipschitz function $f$ exists such that the set of points where $f$ is Fréchet differentiable is Aronszajn null.

In this paper, we establish an analogous result for *convex continuous* functions on the separable Hilbert space $\ell_2$. In fact, we prove that there is an equivalent norm $p$ on $\ell_2$ such that the set of points where $p$ is Fréchet differentiable is Aronszajn null (Aronszajn null sets are defined in Section 2). To do so, we modify a method of Preiss and Zajíček [PZ84] and combine it with a result on finite-dimensional coverings by unit balls inspired by a ball covering construction of Rogers [Rog57].

Having read a preliminary version of this manuscript and employing some of its ideas, David Preiss came up with a much simpler proof. His main new insight is that instead of proving the relatively complicated finite-dimensional ball covering result, one can do a simple inductive construction directly in the Hilbert space. In the interest of the readers and with Preiss' permission, we reproduce his proof in an appendix. The original proof is included as well, since we find the finite-dimensional problem of some independent interest and since the techniques could perhaps be useful in other situations.

Borwein and Noll conjecture in [BN94] that the set of points where a convex continuous function on $\ell_2$ is Lipschitz smooth is never Aronszajn null. Our result shows that this is not the case, since Lipschitz smoothness implies Fréchet differentiability.

## 2. Notation and preliminaries

Let $X$ be a Banach space. We let $B_X(x,r)$ denote the closed ball in $X$ with center $x$ and radius $r$; the subscript $X$ will be often omitted where clear from the context. For a set $C \subset X$, we write $B(C,r) = \bigcup_{x\in C} B(x,r)$ for the $r$-neighborhood of $C$.

Let $X$ be a Banach space and let $f\colon X \to \mathbb{R}$ be a function. A continuous linear map $f'(x)\colon X \to \mathbb{R}$ is a **Gâteaux derivative** of $f$ at a point $x \in X$ if

$$f'(x)(h) = \lim_{t\to 0} \frac{f(x+th) - f(x)}{t}$$

for every $h \in X$. If, moreover, the above limit is uniform for $\|h\| \le 1$, then $f'(x)$ is the **Fréchet derivative** of $f$ at $x$.

NEGLIGIBLE SETS. There is no nontrivial translation-invariant Borel measure in infinite-dimensional Banach spaces. Several authors have defined various classes of "null sets" in infinite-dimensional Banach spaces, trying to mimic the basic properties of Lebesgue null sets in $\mathbb{R}^n$ (a countable union of null sets is null; a translate and a subset of a null set is null; no nonempty open set is null; each class restricted to $\mathbb{R}^n$ gives the null sets for the $n$-dimensional Lebesgue measure). Such classes of null sets are not necessarily induced by a single measure on the considered space.

The following notion of a null set was introduced by Aronszajn [Aro76].

*Definition 2.1:* Let $X$ be a separable Banach space and let $A$ be a subset of $X$. The set $A$ is called **Aronszajn null** if for every sequence $(x_i)_{i=1}^{\infty}$ in $X$ whose closed linear span is $X$ there exist Borel sets $A_i \subset X$ such that $A \subset \bigcup_{i=1}^{\infty} A_i$ and the intersection of $A_i$ with any line in the direction $x_i$ has the one-dimensional Lebesgue measure zero, for each $i \in \mathbb{N}$.

As an easy consequence of Fubini's theorem, the following can be shown (see [Aro76], Proposition 1): If $n \in \mathbb{N}$ and $A$ is a Borel subset of $X$ such that the intersection of $A$ with any $n$-dimensional affine subspace of $X$ is of $n$-dimensional measure zero, then $A$ is Aronszajn null. We will also need the following straightforward modification.

LEMMA 2.2: *Let $X$ be a separable Banach space, let $A \subset X$ be a Borel set, and let $Y$ be a closed subspace of $X$ of a finite codimension. Let $n \in \mathbb{N}$ be such that the intersection of $A$ with any $n$-dimensional affine subspace of $X$ parallel to $Y$ is of $n$-dimensional measure zero. Then $A$ is Aronszajn null.*

*Proof:* Let $k \in \mathbb{N}$ be the codimension of $Y$. Let $Z$ be an $(n+k)$-dimensional subspace of $X$. Then $Z = Z_1 \oplus Z_2$, where $Z_1$ is an $n$-dimensional subspace of $Y$ and $Z_2$ is a subspace of $X$. Let $x \in X$ be given. All $n$-dimensional slices of $A \cap (Z + x)$ parallel to $Z_1$ are of $n$-dimensional Lebesgue measure zero, hence $A \cap (Z+x)$ is of $(n+k)$-dimensional Lebesgue measure zero by Fubini's theorem. The set $A$ is Aronszajn null by the remark above the Lemma. ∎

HILBERT SPACE. Let $\ell_2$ denote the separable Hilbert space and let $(e_1, e_2, \ldots)$ be its orthonormal basis. We identify the Euclidean space $\mathbb{R}^n$ with the linear span of $\{e_1, e_2, \ldots, e_n\}$ in $\ell_2$. For a point $x$ in $\mathbb{R}^n$ or in $\ell_2$, $\|x\| = \left(\sum_i x_i^2\right)^{1/2}$ denotes the Euclidean norm. For a point $x = (x_1, x_2, x_3, \ldots) = \sum x_i e_i$, we will write $x[p..q] = \sum_{i=p}^{q} x_i e_i$, and $x[p\ldots] = \sum_{i=p}^{\infty} x_i e_i$.

GEOMETRIC CONSIDERATIONS. Let $V_n(r)$ denote the volume of the $n$-dimensional ball of radius $r$. A well-known formula says

$$V_n(r) = \frac{\pi^{n/2}}{\Gamma(n/2+1)} r^n.$$

We will need the following simple (and rough) corollary:

LEMMA 2.3: *For natural numbers $m$ and $n$, $m < n$, we have $V_{n-m}(1) \leq n^m V_n(1)$.*

We also need a standard estimate on the number of grid points in a set depending on the volume of a suitable neighborhood.

LEMMA 2.4: *Let $X \subset \mathbb{R}^n$ be a bounded set, and let $\mathbb{Z}^n$ denote the grid of integer points in $\mathbb{R}^n$. Then we have*

$$|X \cap \mathbb{Z}^n| \leq \lambda^n \left(B(X, \sqrt{n}/2)\right)$$

*and also*

$$\left|B(X, \sqrt{n}/2) \cap \mathbb{Z}^n\right| \geq \lambda^n(X).$$

*Proof:* To each grid point $g \in \mathbb{Z}^n$, assign the axis-parallel unit cube centered at $g$. For any $g \in X$, this cube is completely contained in $B(X, \sqrt{n}/2)$ and this gives the first inequality. On the other hand, if the cube of some $g$ intersects $X$ then $g \in B(X, \sqrt{n}/2)$, and this gives the second inequality. ∎

PROBABILITY THEORY RESULTS. In probability estimates, we will mostly follow the rule "whenever you see an expression $1 - x$ (with $x$ small), estimate $1 - x \leq e^{-x}$." We also need a tail estimate for the probability that at least $a$ events among $m$ very rare independent events occur:

LEMMA 2.5 (Poisson approximation to binomial distribution): *Let $X_1, X_2, \ldots, X_m$ be mutually independent random variables, each attaining value 1 with probability $p$ and value 0 with probability $1-p$, where $mp < 1$. Let $a \geq 1$ be a parameter. Then*

$$\text{Prob}\left[\sum_{i=1}^{m} X_i \geq a\right] < (emp)^a.$$

*Proof sketch:* This follows easily, e.g., from Theorem A.12 in Alon and Spencer [AS93]. In our situation, that Theorem says that the probability we are considering is below $\left[e^{\beta-1}\beta^{-\beta}\right]^{pm}$ with $\beta = a/pm$. Using $e^{\beta-1} \leq e^{\beta}$ and $1/a \leq 1$ gives the form in Lemma 2.5. ∎

Next, we recall the so-called Lovász Local Lemma about events with a bounded dependence (see [AS93] for a proof):

LEMMA 2.6 (Lovász Local Lemma): *Let $A_1, A_2, \ldots, A_n$ be events in some probability space. For each $i = 1, 2, \ldots, n$, let $D(i)$ be a set of indices such that the event $A_i$ is independent of all the events $A_j$ with $j \in \{1, 2, \ldots, n\} \setminus D(i)$ (note that $i \in D(i)$). Suppose that numbers $x_1, x_2, \ldots, x_n \in (0, 1)$ exist such that*

$$\text{Prob}[A_i] \leq x_i \prod_{j \in D(i)} (1 - x_j)$$

*holds for all $i = 1, 2, \ldots, n$. Then the probability that none of the events $A_1, A_2, \ldots, A_n$ occurs is strictly positive (in symbols,* $\text{Prob}\left[\bigwedge_{i=1}^{n} \overline{A}_i\right] > 0$*).*

We will use the following consequence.

COROLLARY 2.7: *Let the events $A_i$ and the sets $D(i)$ be as in Lemma 2.6, and suppose that*

$$\sum_{j \in D(i)} \text{Prob}[A_j] \leq 1/2e$$

*holds for each $i$. Then we can conclude* $\text{Prob}\left[\bigwedge_{i=1}^{n} \overline{A}_i\right] > 0$.

*Proof:* Put $x_i = e\text{Prob}[A_i]$. We have, in particular, $x_i \leq \frac{1}{2}$ for all $i$, and so the inequality $1 - x_i \geq e^{-2x_i}$ holds (elementary calculus). Hence $\prod_{j \in D(i)}(1 - x_j) \geq \exp\left(-2\sum_{j \in D(i)} x_j\right) \geq e^{-1}$, and we can use Lemma 2.6. ∎

## 3. Convex functions and coverings

Let $f$ be a convex continuous function on $\ell_2$, and let $D$ be the set of points where $f$ is Fréchet differentiable. Our original aim was to find $f$ so that the complement of $D$ is not Haar null. (The class of Haar null sets [Chr74] is bigger than the class of Aronszajn null sets, hence it is, of course, "easier" to find $f$ with $\ell_2 \setminus D$ not Haar null than to find $f$ with $D$ Haar null, and this is "easier" than to find $f$ with $D$ Aronszajn null.)

If we considered functions defined on a nonreflexive Banach space instead of on $\ell_2$, we could get such an example as follows. According to [MS96], there exists a closed convex set $K \subset X$ with empty interior which is not Haar null. The function $f$ on $X$ defined as the distance from $K$ is convex and continuous, and it is Fréchet differentiable at no point of $K$. However, we want to construct an example on the separable Hilbert space, and each closed convex set in $\ell_2$ with empty interior is Haar null [Mat97], [Mat]. (Let us remark that this result holds also in a considerably more general setting.) Therefore, we cannot simply use a distance function of a convex set and we have to proceed differently.

Our approach is based on suitable low-density packings of unit balls. Suppose $\mathcal{F}$ is a collection of balls of radius 1 in $\ell_2$ such that $\bigcup \mathcal{F}$ is dense in $\ell_2$. Put

$$N = \ell_2 \setminus \left(\bigcup \mathcal{F} \cap B(0,5)\right).$$

Let $C$ be the closed convex hull of the set $\{(y,t) \in N \times \mathbb{R}\colon t \geq \|y\|^2\}$. As observed by Preiss and Zajíček [PZ84], the function $f(x) = \inf\{t \in \mathbb{R}\colon (x,t) \in C\}$ is a well-defined convex continuous function on $\ell_2$, and it is not Fréchet differentiable at any point of $B(0,3) \setminus \bigcup \mathcal{F}$. The latter can be proved by showing that the subdifferential $\partial f$ has oscillation 1 at each point of $B(0,3) \setminus \bigcup \mathcal{F}$ (we will recall the definition of the subdifferential in Section 5).

How to ensure that the complement of $\bigcup \mathcal{F}$ is large? First, we present a heuristic consideration which doesn't quite work but might perhaps be helpful for understanding the actual proof.

Let $\mathcal{B}$ be a covering of $\mathbb{R}^n$ by unit balls (that is, $\mathcal{B}$ is a set of unit balls in $\mathbb{R}^n$ with $\bigcup \mathcal{B} = \mathbb{R}^n$). The **upper density** of $\mathcal{B}$, denoted by $\overline{d}(\mathcal{B})$, is defined by

$$\overline{d}(\mathcal{B}) = \lim\sup_{R\to\infty} \frac{\sum_{B\in\mathcal{B}\colon B\subset B(0,R)} \lambda^n(B)}{\lambda^n(B(0,R))},$$

where $\lambda^n$ denotes the $n$-dimensional Lebesgue measure. Rogers [Rog57] established the existence of coverings with a relatively small upper density. Namely, he proved the existence of a covering $\mathcal{B}_n$ of $\mathbb{R}^n$ by unit balls such that

$$\overline{d}(\mathcal{B}_n) \leq n \ln n + n \ln\ln n + 5n.$$

Let $Z_n$ denote the set of centers of the balls in such a covering $\mathcal{B}_n$. Set $r_n = 1 + 1/\sqrt{n}$, and put $C_n = r_n Z_n$. Hence $C_n$ determines a low-density covering of $\mathbb{R}^n$ by balls of radius $r_n$. At the same time, setting $\mathcal{F}_n = \{B(c,1)\colon c \in C_n\}$, one can calculate that the upper density of $\mathcal{F}_n$ decreases to 0 very quickly as $n \to \infty$ (it is roughly of the order $e^{-\sqrt{n}}$). If we now identify $\mathbb{R}^n$ with the linear span of $\{e_1, e_2, \dots, e_n\}$ in $\ell_2$ and define $\mathcal{F} = \{B_{\ell_2}(c,1)\colon c \in C_n, n \in \mathbb{N}\}$, then clearly $\bigcup \mathcal{F}$ is dense in $\ell_2$. Since the density of the coverings $\mathcal{F}_n$ decreases to 0 (this means that the fraction of the volume of $\mathbb{R}^n$ covered by the balls of radius 1 centered in $C_n$ decreases to zero with the dimension), we can hope that there will be still "enough" space in $\ell_2$ left after removing all the balls in $\mathcal{F}$. This is roughly how the proof of Theorem 3.1 below goes, but instead of the result of Rogers we use Lemma 3.3 below and choose the increasing sequence of subspaces of $\ell_2$ more carefully.

THEOREM 3.1: *There exists a convex continuous function $f$ on the separable Hilbert space such that the set of points where $f$ is Fréchet differentiable is Aronszajn null.*

In Section 5 we show that $f$ can even be an equivalent norm.

In order to control the measure properties of sets, we will use 12-dimensional **test cubes***. We let $U_0$ be the unit cube $[0,1]^{12}$. Since we consider each $\mathbb{R}^n$ canonically embedded in $\ell_2$, $U_0$ is also a subset of $\ell_2$. By a test cube, we mean any congruent copy $U$ of $U_0$ in $\ell_2$. In other words, if $x_0 \in \ell_2$ is a translation vector and $\mathbf{u} = (u_1, u_2, \dots, u_{12})$ is a 12-tuple of orthonormal vectors in $\ell_2$, we set

$$U = \left\{ x_0 + \sum_{i=1}^{12} a_i u_i \colon 0 \le a_i \le 1, i = 1, 2, \dots, 12 \right\}.$$

We denote the 12-dimensional Lebesgue measure on $U$ by $\lambda_U$. Theorem 3.1 is a consequence of the following:

PROPOSITION 3.2: *Let $\varepsilon > 0$ be given. There exist a number $r > 0$ and a countable set $C \subset \ell_2$ such that*

(A) *for any $\delta > 0$, $B(C, r + \delta) = \ell_2$, and*
(B) *$\lambda_U(U \cap B(C,r)) \le \varepsilon$ for any test cube $U$.*

---

* Why just 12-dimensional? This is the smallest dimension where certain technical calculation goes through. With some more effort, the proofs could be made to work also with a somewhat smaller dimension as well, but, interestingly enough, it seems that the current proof method cannot work for a cube of dimension smaller than 3.

*Proof of Theorem 3.1:* For each $m \in \mathbb{N}$, apply Proposition 3.2 with $\varepsilon = 1/m$, obtaining $r_m > 0$ and a set $C_m$. Put $F = \bigcap_{m=1}^{\infty} B(C_m, r_m)$. To see that $F$ is Aronszajn null, it is enough to show that the intersection of $F$ with any 12-dimensional affine subspace $Z$ of $\ell_2$ has 12-dimensional Lebesgue measure zero. The space $Z$ can be covered by countably many 12-dimensional test cubes. If $U$ is any such cube and $m \in \mathbb{N}$ then $\lambda_U(U \cap B(C_m, r_m)) \leq 1/m$; consequently $\lambda_U(U \cap F) = 0$.

By a result of Preiss and Zajíček [PZ84], there exists a convex, continuous function $f$ defined on $\ell_2$ such that $f$ is Fréchet differentiable only at points of $F$. (Construct countably many functions $f_m$ similarly as was described in the beginning of this section, and put $f = \sum_{m=1}^{\infty} a_m f_m$, where the $a_m > 0$ are sufficiently small.) ∎

We are going to prove Proposition 3.2 from a somewhat technical Lemma 3.3 below on ball coverings in $\mathbb{R}^n$. As was mentioned in the introduction, there is a much simpler and direct proof due to Preiss. Readers interested (naturally enough) in this simplified proof may read the appendix and skip the rest of this section and the next section.

Let $K$ be some (yet unspecified) large natural number. This time we will use $K$-times enlarged test cubes. So we let $T_0 = [0, K]^{12}$, and for a vector $x_0 \in \ell_2$ and an orthonormal family $\mathbf{u} = (u_1, u_2, \ldots, u_{12})$ in $\ell_2$, we put

$$T = T(x_0, \mathbf{u}) = \left\{ x_0 + \sum_{i=1}^{12} a_i u_i \colon 0 \leq a_i \leq K, i = 1, 2, \ldots, 12 \right\}.$$

We let aff$T$ denote the affine span of $T$ in $\ell_2$, that is,

$$\mathrm{aff}T = \left\{ x_0 + \sum_{i=1}^{12} \alpha_i u_i \colon \alpha_i \in \mathbb{R} \right\}.$$

Let $\mu = \mu_T$ be the uniform probability measure on $T$ (obtained by re-scaling the 12-dimensional Lebesgue measure).

Let $v \in \ell_2$ be a point. We say that $v$ is **of type** $j$ **with respect to** $T$ if $\mathrm{dist}(v, T) \leq 1$ and the distance of $v$ from aff$T$ is at least $1 - 2^{-j+1}$ and at most $1 - 2^{-j}$. (The type of a point $v$ essentially determines how large a part of $T$ does a unit ball centered at $v$ cover.) Now we are ready to formulate the key lemma.

LEMMA 3.3: *Let a number $K_0$ be given. Then we can choose numbers $K = K(K_0) \geq K_0$ and $n_0 = n_0(K_0)$ so that for any natural number $\ell \geq n_0$, a natural number $n = n(\ell) \geq \ell$ and a countable set $C = C(n) \subset \mathbb{R}^n$ exist with the following properties:*

(i) *We have* $B_{\mathbb{R}^n}(C, 1+\delta) = \mathbb{R}^n$, *where* $\delta = \delta(n) \geq 0$ *is a function of* $n$ *tending to 0 with* $n \to \infty$.

(ii) *Let* $T = T(x_0, \mathbf{u})$ *be any congruent copy of the cube* $T_0$ *in* $\ell_2$. *We let* $C^{T,j}$ *denote the set of all points* $c \in C$ *that are of type* $j$ *with respect to* $T$. *Then we have, for any* $T$ *and any* $j$,

$$\mu_T\Big(T \cap B_{\ell_2}(C^{T,j}, 1)\Big) \leq \frac{1}{K^4 2^{4j}}.$$

(iii) *The distance of* $C$ *from the subspace of* $\mathbb{R}^n$ *spanned by the first* $\ell$ *coordinate axes is at least 1, that is,* $\|c[\ell+1..n]\| \geq 1$ *for each* $c = (c_1, c_2, \ldots, c_n) \in C$.

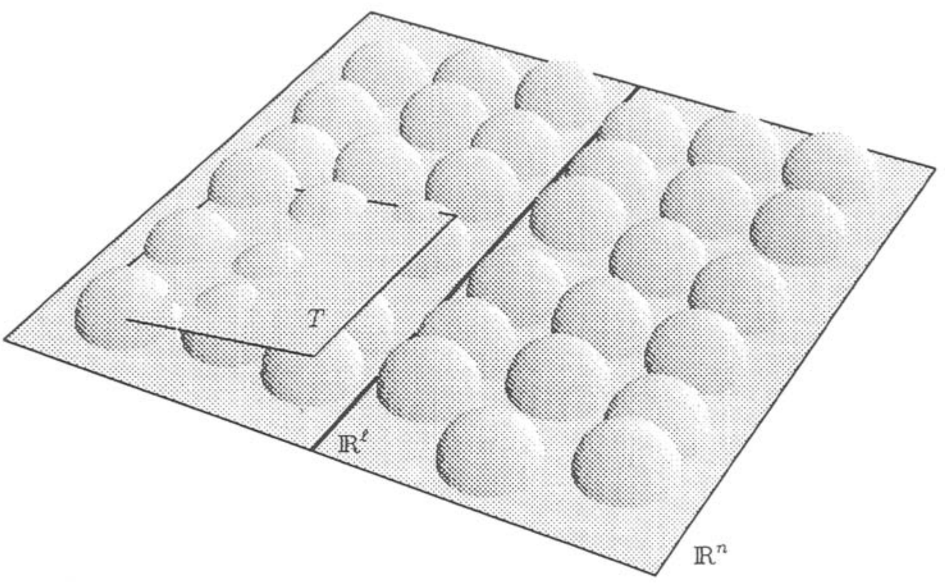


Figure 1. Illustration to the statement of Lemma 3.3.

Figure 1 illustrates the situation the Lemma talks about; for obvious reasons, we had to reduce the dimensions somewhat, and so $\ell_2$ is pictured 3-dimensional, $n = 2$, $\ell = 1$ and $T$ is shown 2-dimensional. Lemma 3.3 will be proved in Section 4.

*Proof of Prcposition 3.2:* Let $\varepsilon > 0$ be given, and let us set $K_0 = \max\{1/\varepsilon, 10^6\}$. Let $K \geq K_0$ and $n_0$ be as in Lemma 3.3. We will show that there exists a subset $C'$ of $\ell_2$ so that

(A′) for any $\delta > 0$, $B(C, 1+\delta) = \ell_2$, and

(B′) $\mu_T(T \cap B(C,1)) \le \varepsilon$ for any congruent copy $T$ of $T_0$.

Once we prove this, the Proposition is obtained by putting $r = 1/K$ and $C = (1/K)C'$.

For $k = 1, 2, \ldots$, put $n_k = n(n_{k-1})$, where $n(\ell)$ is the function as in Lemma 3.3. Let $C_k = C(n_k) \subset \mathbb{R}^{n_k} \subset \ell_2$ and $\delta_k = \delta(n_k)$ be as in Lemma 3.3. We put $C' = \bigcup_{k=1}^{\infty} C_k$.

The condition (A′) is straightforward to check. Consider any point $x = (x_1, x_2, \ldots) \in \ell_2$ and an arbitrarily small number $\gamma > 0$. Let $k$ be large enough so that $\delta_k \le \gamma/2$ and $\|x[n_k+1\ldots]\| \le \gamma/2$. Then, by Lemma 3.3(i), there exists a point $c \in C_k$ with $\|c - x[1..n_k]\| \le 1 + \gamma/2$ and we get $\|c - x\| \le 1 + \gamma$.

To verify condition (B′), let $T = T(x_0, \mathbf{u})$ be a congruent copy of the cube $T_0$. For $j = 1, 2, \ldots$, let $I_j$ be the set of all the indices $k \ge 1$ such that $C_k$ contains at least one point of type $j$ with respect to $T$. We claim that $|I_j| \le K^3 2^{2j}$.

We may assume $I_j \ne \emptyset$. We let $m = \min I_j$. Since all points of $C_m$ lie in the subspace spanned by the first $n_m$ coordinates and some point of $C_m$ lies at distance at most 1 from $T$, we have $\|x_1[n_m+1\ldots]\| \le 1$ for some point $x_1 \in T$, and consequently

$$\|x_0[n_m+1\ldots]\| \le \|x_1[n_m+1\ldots]\| + \operatorname{diam}(T) \le 1 + \sqrt{12}K < 4K. \tag{1}$$

Next, consider an index $k \in I_j$. Let $c \in C_k$ be such that $\operatorname{dist}(T, c) \le 1$ and $\operatorname{dist}(\operatorname{aff}T, c) \le 1 - 2^{-j}$. Let $x \in \operatorname{aff}T$ be the point attaining the distance of $\operatorname{aff}T$ to this $c$. We have $\|x - c\| \le 1 - 2^{-j}$ and also $\operatorname{dist}(T, x) \le \|x - c\| + \operatorname{dist}(c, T) \le 2$. We can write $x = x_0 + \sum_{i=1}^{12} b_i u_i$ for some numbers $b_1, \ldots, b_{12}$ with $|b_i| \le K + 2 < 2K$.

By Lemma 3.3(iii), we have $\|c[n_{k-1}+1..n_k]\| \ge 1$, and hence

$$\begin{aligned}\|x[n_{k-1}+1..n_k]\| &\ge \|c[n_{k-1}+1..n_k]\| - \|(c-x)[n_{k-1}+1..n_k]\| \\ &\ge 1 - (1 - 2^{-j}) = 2^{-j}.\end{aligned}$$

Therefore, at least one of the following inequalities holds:

$$\|x_0[n_{k-1}+1..n_k]\| \ge \frac{1}{13} \cdot 2^{-j} \tag{2}$$

or, for some $i \in \{1, 2, \ldots, 12\}$,

$$\|u_i[n_{k-1}+1..n_k]\| \ge \frac{1}{2K} \cdot \frac{1}{13} \cdot 2^{-j}. \tag{3}$$

Let $J_0 \subseteq I_j$ be the set of those indices $k \ne m$ for which (2) holds, and let $J_i \subseteq I_j$ be the set of indices $k \ne m$ for which (3) holds with $i$, $i = 1, 2, \ldots, 12$. There exists an $i_0 \in \{0,1,\dots,12\}$ with $|J_{i_0}| \geq (|I_j|-1)/13$ (pigeonhole). First, we consider the case $i_0 = 0$. Then we have, by (1), by the theorem of Pythagoras, and by (2),

$$(4K)^2 \geq \|x_0[n_m+1\dots]\|^2 \geq \sum_{k\in I_0} \|x_0[n_{k-1}+1..n_k]\|^2 \geq \frac{|I_j|-1}{13}\cdot\frac{1}{13^2}\cdot 2^{-2j}$$

and $|I_j| \leq 16K^2\cdot 13^3\cdot 2^{2j}+1 < 2^{2j}K^3$ follows. A similar calculation works in the case $i_0 \in \{1,2,\dots,12\}$, using the unit vector $u_i$.

Using Lemma 3.3(ii), we can now calculate

$$\begin{aligned}\mu_T(T\cap B(C',1)) &\leq \sum_{k=1}^{\infty}\mu_T\Big(T\cap B(C_k,1)\Big) \leq \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\mu_T\Big(T\cap B(C_k^{T,j},1)\Big)\\ &= \sum_{j=1}^{\infty}\sum_{k\in I_j}\mu_T\Big(T\cap B(C_k^{T,j},1)\Big) \leq \sum_{j=1}^{\infty}|I_j|\frac{1}{K^4 2^{4j}}\\ &\leq \varepsilon\sum_{j=1}^{\infty}2^{-2j} = \frac{\varepsilon}{3} < \varepsilon\end{aligned}$$

Proposition 3.2 is proved. ∎

## 4. Proof of Lemma 3.3

In the proof, we assume that $K = K(K_0)$ is a large enough natural number (just how large can in principle be determined by an inspection of the calculations below), and that $n_0$ is still much larger than $K$. The dimension $\ell \geq n_0$ is given, and $n = n(\ell)$ is chosen large enough in terms of $\ell$ (e.g., $n = \ell^3$ will work). We think of $\ell$ and $n$ as tending to infinity, while $K$ is very large but fixed. We set $\delta = \delta(n) = 4/\sqrt{n}$.

To construct the set $C$, we set $L = K^2$, we choose a suitable finite set $D \subset [0,L)^n$, and we replicate $D$ periodically with period $L$ along each axis; in other words, we set $C = D + L\mathbb{Z}^n$. Since we need to replicate also other sets periodically in this manner, let us write $X^\# = X + L\mathbb{Z}^n$ for an arbitrary set $X \subseteq \mathbb{R}^n$ (so that $C = D^\#$).

Let us set $\eta = 1/n$, and let $G_0$ be the points of the grid with spacing $\eta$ within the cube $[0,L)^n$, that is, $G_0 = \eta\mathbb{Z}^n \cap [0,L)^n$. The set $D$ will be chosen as a suitable subset of $G_0$. In order to satisfy condition (iii) of the lemma (distance of $C$ from the subspace spanned by the first $\ell$ basis vectors), we define the "forbidden region"

$$F = \{x \in \mathbb{R}^n\colon \|x[\ell+1..n]\| < 1\}$$

and we set $G = G_0 \setminus F^{\#}$.

Let $p \in (0,1)$ be a real parameter; its value will be fixed later. Let $D \subset G$ be a random* subset of $G$, where we include each point $x \in G$ into $D$ with probability $p$, this choice being mutually independent for distinct points $x$. From such a random $D$ we construct the set $C = D^{\#}$ as above. We will show that the probability of obtaining a set $C = D^{\#}$ satisfying the properties required in the Lemma is nonzero, and consequently the required $C$ exists. Note that condition (iii) of the Lemma will be automatically satisfied for any $C = D^{\#}$, with $D \subset G$. A two-dimensional picture, with $\ell = 1$ and $n = 2$, can perhaps be slightly helpful (although it is misleading too); see Figure 2.

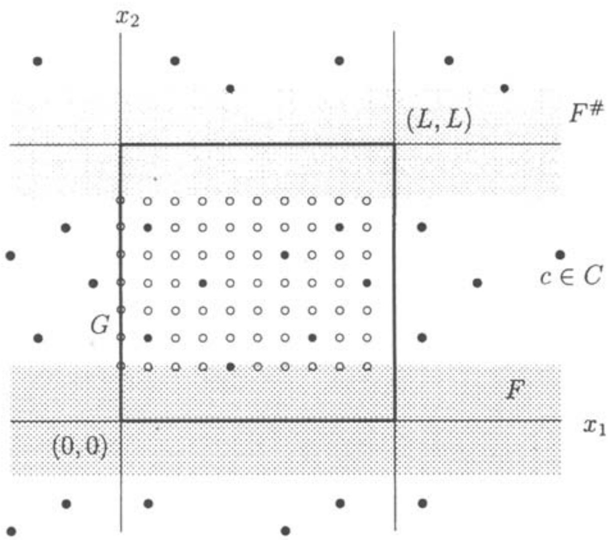


Figure 2. A 2-dimensional illustration to the proof of Lemma 3.3.

For each $x \in G_0$, let $A_x$ denote the event "$x \notin B(C, 1 + \delta/2)$". Since each point $x \in [0, L)^n$ lies at distance at most $\eta\sqrt{n} < \delta/2$ from some point of $G_0$, if none of the events $A_x$ for $x \in G_0$ occurs then $B(C, 1+\delta) = \mathbb{R}^n$ and condition (i) holds.

Let us estimate the probability of the event $A_x$. A given point $x \in G_0$ is covered by $B(C, 1+\delta/2)$ if and only if some point of the set $B(x, 1+\delta/2) \cap \eta\mathbb{Z}^n$ falls into $C$. There is a slight complication since the grid points in the forbidden

* The underlying finite probability space is $(2^G, 2^{2^G}, \mathrm{Prob})$, where for each atom $D \subset G$ we have $\mathrm{Prob}(D) = p^{|D|}(1-p)^{|G \setminus D|}$.

region $F^{\#}$ are not chosen into $D$. The probability of $x$ not being covered by $B(C, 1+\delta/2)$ is thus $\text{Prob}\,[A_x] = (1-p)^{\nu_x}$, where

$$\nu_x = \left|B(x, 1+\delta/2) \cap G^{\#}\right|.$$

A simple calculation, which we postpone, gives the following:

CLAIM 4.1: *For any $x \in G_0$, we have*

$$\nu_x \geq \nu = \frac{V_n\left(1+\frac{4}{3}n^{-1/2}\right)}{\eta^n}$$

*(recall that $V_n(r)$ denotes the volume of the $n$-dimensional ball of radius $r$).*

Let us put $p = (2n \ln n)/\nu$. We show that for this setting, there is a fairly small probability that any of the events $A_x$ occurs. We have

$$\begin{aligned}\text{Prob}\,[\text{some } A_x \text{ occurs}] &\leq \sum_{x \in G_0} \text{Prob}\,[A_x] \leq |G_0|(1-p)^{\nu} \leq (L/\eta)^n \exp(-p\nu) \\ &\leq \exp(n \ln L + n \ln n - 2n \ln n) = (L/n)^n.\end{aligned}$$

The last expression quickly tends to 0 for $n \to \infty$, and so we can assume

$$\sum_{x \in G_0} \text{Prob}\,[A_x] \leq \frac{1}{100}, \tag{4}$$

say. Hence for the given choice of $p$, the covering condition (i) in Lemma 3.3 is typically satisfied.

Next, we are going to deal with condition (ii) (sparse covering of all the 12-dimensional cubes $T$). First we note that although the Lemma considers all cubes $T$ in $\ell_2$, we may restrict ourselves to cubes $T \subset \mathbb{R}^{\bar{n}}$, where we write $\bar{n} = n + 13$. This is because if $T \subset \ell_2$ is an arbitrary 12-dimensional cube, there exists an isometry of the linear span of $T \cup \mathbb{R}^n$ onto $\mathbb{R}^{\bar{n}}$ fixing $\mathbb{R}^n$.

For a given congruent copy $T$ of $T_0$, we now estimate the probability that our random set $C = D^{\#}$ contains many points that are close to $T$. This amounts to a volume computation plus an application of a large deviation tail estimate for the binomial distribution.

CLAIM 4.2: *For any congruent copy $T$ of $T_0$ and any real number $\rho \in (0, \frac{1}{2}]$, we have*

$$\begin{aligned}\text{Prob}\Big[|\{x \in C\colon\ \text{dist}(x, \text{aff}T) \leq 1-\rho/2 \textit{ and } \text{dist}(x, T) \leq 2\}| > K\rho^{-2}\Big] \\ < \exp\Big(-Kn/10\rho\Big).\end{aligned}$$

This claim, whose proof we again postpone, allows us to estimate the probability that a given fraction of some fixed cube $T$ is covered by the unit balls centered at points of $C$ that are of type $j$ (because we can estimate the area covered by one such ball). But we need to handle all possible $T$'s at the same time. Similarly as we did for the covering property (i), we replace the set of all $T$'s by a suitable discrete approximation. This time we need one family $\mathcal{T}_j$ of cubes for each $j$, and also we need a bit more sophisticated choice of these cubes than just taking translates and rotations in a sufficiently fine grid.

For a given $j$, set $\rho = 2^{-j}$, and let $N$ be a $\rho/4$-net in the large cube $[0, L)^{\bar{n}}$ (recall that $\bar{n} = n + 13$). That is, $N$ is an inclusion-maximal subset of $[0, L)^{\bar{n}}$ such that every two points of $N$ have distance at least $\rho/4$. Further let $M$ be a set of orthonormal families as in the following Claim (whose proof is, as usual, postponed):

Claim 4.3: *Let $\rho \in (0,1)$. There exists a set $M$ consisting of orthonormal 12-tuples $\mathbf{v} = (v_1, \ldots, v_{12})$ in $\mathbb{R}^{\bar{n}}$ such that given any orthonormal family $\mathbf{u} = (u_1, u_2, \ldots, u_{12})$ in $\mathbb{R}^{\bar{n}}$, there is a $\mathbf{v} \in M$ with $\|v_i - z_i\| \leq \rho/K^2$ for all $i = 1, 2, \ldots, 12$, and moreover we have*

$$|M| \leq \left(K^3/\rho\right)^{12\bar{n}}.$$

We define $\mathcal{T}_j = \{T(x_0, \mathbf{v})\colon x_0 \in N, \mathbf{v} \in M\}$. It will be notationally convenient to assume that the families $\mathcal{T}_j$ are all disjoint. For a $T \in \mathcal{T}_j$, let $A_T$ be the event

> "for more than $K\rho^{-2}$ points $x$ of $C$, both $\operatorname{dist}(x, \operatorname{aff} T) \leq 1 - \rho/2$ and $\operatorname{dist}(x, T) \leq 2$ holds, where $\rho = 2^{-j}$".

To establish Lemma 3.3, it is sufficient to prove the following two claims:

Claim 4.4: *If the set $D$ is chosen at random in the above-described manner, then with a positive probability, none of the events $A_x$ for $x \in G_0$ (grid points uncovered) and $A_T$ for $T \in \mathcal{T}_j$, $j = 1, 2, \ldots$, occurs.*

Claim 4.5: *If none of the events $A_T$ occurs, $T \in \mathcal{T}_j$, $j = 1, 2, \ldots$, then the condition* (ii) *in Lemma* 3.3 *is satisfied.*

Before proving these Claims, we formulate a simple geometric statement (whose proof is omitted). This is the only point in the proof where the dimension of $T$ really plays a role.

Claim 4.6: *Let $T$ be a congruent copy of $T_0$ and $x \in \ell_2$ be a point at distance at least $1 - \rho/2$ from $\operatorname{aff} T$. Then $B(x, 1) \cap \operatorname{aff} T$ is contained in a 12-dimensional* *ball of radius at most $\sqrt{\rho}$, whose Lebesgue measure is thus at most $\beta\rho^6$, where $\beta$ is an absolute constant.*

*Proof of Claim 4.5:* Let $T$ be a congruent copy of $T_0$. As was remarked above, it suffices to consider the case $T \subset \mathbb{R}^{\bar{n}}$, $\bar{n} = n + 13$. Also, since $C$ is $L$-periodic, we can assume that $T = T(x, \mathbf{u})$ with $x \in [0, L)^{\bar{n}}$. Again write $\rho = 2^{-j}$. By the choice of $\mathcal{T}_j$, we can find a $T_1 \in \mathcal{T}_j$ such that if $y \in \operatorname{aff}T$ and $\operatorname{dist}(y, T) \leq 1$ then there is a $z \in \operatorname{aff}T_1$ with $\operatorname{dist}(z, T_1) \leq 1$ such that $\|y - z\| \leq \rho/2$. Therefore, if $c \in C$ is any point of type $j$ with respect to $T$ then $\operatorname{dist}(c, \operatorname{aff}T_1) \leq 1 - \rho/2$ and $\operatorname{dist}(c, T_1) \leq 2$.

Since we suppose that the event $A_{T_1}$ doesn't occur, there are at most $K\rho^{-2}$ such points $c$. By Claim 4.6, the unit ball around each such point $c$ swallows no more than $\beta\rho^6$ of the Lebesgue measure sitting on $\operatorname{aff}T$. Hence

$$\mu_T\left(T \cap B(C^{T,j}, 1)\right) \leq \frac{\beta\rho^6}{K^{12}} \cdot \frac{K}{\rho^2} \leq \frac{\rho^4}{K^4}$$

if $K$ is large enough. This establishes Claim 4.5. ∎

*Proof of Claim 4.2:* Put $r = 1 - \rho/2$. Let $m$ denote the number of points of the grid $G^\#$ that are at distance at most $r$ from $\operatorname{aff}T$ and distance at most 2 from $T$. Since the diameter of $T$ is $\sqrt{12}K$ and the period of $G^\#$ is $L = K^2 > \operatorname{diam}(T) + 4$, no two of these points are a periodic replication of the same point in $G$. Hence the number of the relevant points in $C$ is the sum of $m$ independent random variables, each of them attaining value 1 with probability $p$ and value 0 with probability $1 - p$. By Lemma 2.5, the probability we seek is at most $(epm)^{K/\rho^2}$; it remains to estimate $m$.

Recall that we assume $T = T(x_0, \mathbf{u}) \subset \mathbb{R}^{\bar{n}}$. The number $m$ is no bigger than the number of points of the grid $\eta\mathbb{Z}^{\bar{n}}$ in $B(T, 2) \cap B(\operatorname{aff}T, r)$, with $r = 1 - \rho/2$ and the balls being in the $\bar{n}$-dimensional space. Let $Y$ be the orthogonal complement of $\mathbf{u}$ in $\mathbb{R}^{\bar{n}}$. Then $B(T, 2) \cap B(\operatorname{aff}T, r)$ is contained in the set

$$Z = \{x + y \colon x \in \operatorname{aff}T \cap B(x_0, 4K), y \in Y, \|y\| \leq r\}.$$

This is a Cartesian product of an $(n+1)$-dimensional $r$-ball and a 12-dimensional ball of radius $4K$. By the first inequality in Lemma 2.4, we get

$$m \leq \frac{\lambda^{\bar{n}}(B(Z, \frac{1}{2}\eta\sqrt{\bar{n}}))}{\eta^{\bar{n}}} \leq \frac{V_{12}(5K)V_{n+1}\left(r + \frac{1}{2}\eta\sqrt{\bar{n}}\right)}{\eta^{\bar{n}}}.$$

We have $V_{12}(5K) \leq (10K)^{12}$, and we also get

$$V_{n+1}\left(r + \frac{\eta}{2}\sqrt{\bar{n}}\right) = V_{n+1}\left(r + \frac{\sqrt{n+13}}{2n}\right) \leq V_n\left(r + n^{-1/2}\right).$$

By substituting $p = 2n\ln n/\nu$, with $\nu = V_n\left(1+\frac{4}{3}n^{-1/2}\right)\eta^{-n}$ as in Claim 4.1, we get

$$\begin{aligned} pm &\le \frac{(2n\ln n).\eta^n}{V_n\left(1+\frac{4}{3}n^{-1/2}\right)}\cdot\frac{(10K)^{12}V_n\left(1-\frac{1}{2}\rho+n^{-1/2}\right)}{\eta^{n+13}} \\ &\le 2n^{14}(10K)^{12}\ln n\left(\frac{1-\frac{1}{2}\rho+n^{-1/2}}{1+\frac{4}{3}n^{-1/2}}\right)^n. \qquad (5) \end{aligned}$$

We have $2n^{14}(10K)^{12}\ln n \le n^{15}$. We distinguish two cases depending on the value $\rho$. For $\rho > n^{-1/2}$, we get

$$\begin{aligned} pm &\le n^{15}\left(\frac{1+n^{-1/2}-\rho/2}{1+\frac{4}{3}n^{-1/2}}\right)^n \le n^{15}\left(1-\frac{\rho}{2(1+\frac{4}{3}n^{-1/2})}\right)^n \\ &\le n^{15}\Big(1-\rho/4\Big)^n \le n^{15}\exp(-n\rho/4) \\ &\le \exp(-n\rho/8) \end{aligned}$$

(recall that $n$ is large and $\rho n \ge \sqrt{n}$). For $\rho \le n^{-1/2}$, we ignore the $\frac{1}{2}\rho$ term in (5) and we calculate

$$\begin{aligned} pm &\le n^{15}\left(\frac{1+n^{-1/2}}{1+\frac{4}{3}n^{-1/2}}\right)^n \le n^{15}\left(1-\frac{\frac{1}{3}n^{-1/2}}{1+\frac{4}{3}n^{-1/2}}\right)^n \\ &\le n^{15}\Big(1-\frac{1}{4}n^{-1/2}\Big)^n \le n^{15}\exp(-\sqrt{n}/4) \\ &< \exp(-\sqrt{n}/5). \end{aligned}$$

In both cases, simple estimates lead to $(epm)^{K/\rho^2} \le \exp(-Kn/(10\rho))$. Claim 4.2 is proved. ∎

*Proof of Claim 4.4:* In order to show that the probability of none of the events $A_x$ and $A_T$ occurring is nonzero, we want to apply the Lovász Local Lemma in the form of Corollary 2.7.

First we note that although the number of the events $A_T$ is formally infinite ($j$ can be any natural number), all but finitely many of them are impossible. Namely, the event $A_T$ requires in particular that $|C| \ge K\rho^{-2}$ with $\rho = 2^{-j}$. Since no two points of $C$ interacting with $T$ can be periodic replications of the same point of $D$, we also have $|D| \ge K\rho^{-2}$. Since $|D| \le |G|$ is bounded by some function of $n$, $A_T$ is impossible for $j$ too large.

For each of the events $A_x$ and $A_T$, we need to find all other events it might possibly depend on and sum up their probabilities. We need not care about the dependence on the events $A_x$, since we have calculated in (4) that all these events together have probability at most $\frac{1}{100}$.

Let us consider an event $A_x$, and let us look which events $A_T$ may possibly affect $A_x$. Clearly, if $A_x$ is not independent of $A_T$ then there must be a point $y \in G$ such that the event "$y \in D$" influences both $A_x$ and $A_T$. If $A_x$ should depend on "$y \in D$" then some periodic copy $y_1 \in y + L\mathbb{Z}^n$ of $y$ must lie in $B(x, 1 + \delta/2) \subset B(x, 2)$. Similarly, "$y \in D$" interacting with $A_T$ means that some $y_2 \in y + L\mathbb{Z}^n$ lies in $B(T, 2) \subseteq B(x_0, 4K)$ where $T = T(x_0, \mathbf{u})$. Putting this together yields that $x_0 \in B(x, 5K)^{\#} = B(x, 5K) + L\mathbb{Z}^n$ for any $T = T(x_0, \mathbf{u})$ with $A_T$ affecting $A_x$.

We recall that the set $\mathcal{T}_j$ was defined as $\{T(x_0, \mathbf{v})\colon x_0 \in N, \mathbf{v} \in M\}$, where $N$ is a $\rho/4$-net in $[0, L)^{\bar{n}}$, $\rho = 2^{-j}$, and $M$ is as in Claim 4.3. By a standard volume argument, we get that the number of points of $N$ in any ball of radius $5K$ is at most

$$\frac{V_{\bar{n}}(5K + \rho/8)}{V_{\bar{n}}(\rho/8)} \leq \left(\frac{50K}{\rho}\right)^{\bar{n}}.$$

Moreover, since $L > \operatorname{diam}(B(x, 5K))$, at most $3^n$ periodic copies of $B(x, 5K)$ in $B(x, 5K)^{\#}$ may intersect the cube $[0, L)^{\bar{n}}$. Therefore, the number of events $A_T$ with $T \in \mathcal{T}_j$ that may possibly influence $A_x$ is bounded by

$$3^n \left(\frac{50K}{\rho}\right)^{\bar{n}} |M| \leq \left(\frac{K}{\rho}\right)^{bn}$$

for an absolute constant $b$. Using Claim 4.2, we get that the sum of probabilities of these $A_T$'s is bounded by

$$\left(\frac{K}{\rho}\right)^{bn} \exp\left(-\frac{Kn}{10\rho}\right) = \exp\left(-n\left[\frac{K}{10\rho} - b\ln(K/\rho)\right]\right).$$

If $K$ is large, the expression in the exponent is at most $-Kn/(20\rho) = -Kn2^j/20$. By summing over all $j = 1, 2, \ldots$, we conclude that the sum of the probabilities of all events that may possibly influence our event $A_x$ is small (smaller than any prescribed constant). A similar reasoning gives the same estimate for the events some $A_T$ may depend on. Claim 4.4 thus follows from Corollary 2.7. ∎

*Remark:* It seems that sum of the probabilities of all the events $A_T$ together (not only of those that some other among our events depends on) cannot be bounded. The reason is that we need too many events $A_T$. Namely, for $j$ being a small constant, the probability $\operatorname{Prob}[A_T]$ with $T \in \mathcal{T}_j$ can only be bounded by a function $\exp(-\alpha n)$ for some positive constant $\alpha$, but the number of points of a

$2^{-j-1}$-net in the cube $[0, L]^n$ grows superexponentially with $n$ (because the ratio of the volume of a cube and its inscribed ball grows superexponentially with the dimension). Somewhat ironically, this is the only reason for applying the Lovász Local Lemma (instead of simply summing up the probabilities). This, in turn, forces us to choose the points of $D$ from a discrete set by independent trials (instead of the perhaps more natural way of choosing $D$ as $\nu$ independent points uniformly distributed in $[0, L)^n$).

*Proof of Claim 4.1:* First, we count the points of the full grid $G_0^{\#} = \eta\mathbb{Z}^n$ falling into $B(x, 1 + \delta/2)$. By the second inequality in Lemma 2.4, this number is at least $V_n(1 + \delta/2 - \eta\sqrt{n}/2)/\eta^n = V_n(1 + \frac{3}{2}n^{-1/2})/\eta^n$.

Next, we estimate the number of points in $\eta\mathbb{Z}^n \cap B(x, 1 + \delta/2)$ falling into the forbidden region $F^{\#}$. The region $F^{\#}$ consists of translated copies of the cylinder $F = \{y \in \mathbb{R}^n \colon \|y[\ell + 1..n]\| < 1\}$, and the ball $B(x, 1 + \delta/2)$ may only intersect one of these copies; so we may as well assume it intersects $F$ itself. We have

$$B(x, 1 + \delta/2) \cap F \subseteq \{z \in \mathbb{R}^\ell \colon \|z - x\| \leq 1 + \delta/2\} \times \{z' \in \mathbb{R}^{n-\ell} \colon \|z'\| < 1\}.$$

(Here we consider $\mathbb{R}^{n-\ell}$ as the span of $e_{\ell+1}, \ldots, e_n$.) If this last region is denoted by $R$, the number of points of the grid $\eta\mathbb{Z}^n$ in $R$ is no more than the volume of the $\frac{1}{2}\eta\sqrt{n}$-neighborhood of $R$, by Lemma 2.4. The volume of $B(R, \eta\sqrt{n})$ is bounded by

$$V_\ell(1 + 3n^{-1/2})V_{n-\ell}(1 + n^{-1/2}).$$

Using Lemma 2.3, one can check that if $n$ is large enough in terms of $\ell$ ($n = \ell^3$ will do), then the last displayed expression is smaller than $\frac{1}{2}V_n(1 + \frac{3}{2}n^{-1/2})$ (say). Therefore, $\nu_x \geq \frac{1}{2}V_n(1 + \frac{3}{2}n^{-1/2})/\eta^n \geq V_n(1 + \frac{4}{3}n^{-1/2})/\eta^n$. This finishes the proof of Claim 4.1 and thus also of Lemma 3.3. ∎

*Proof of Claim 4.3:* In the Claim, we have used the maximum metric for measuring the distance of two 12-tuples. For the proof of the Claim, it will be more convenient to use the Euclidean metric, that is, we consider the metric space $U$ of all orthonormal 12-tuples $\mathbf{u}$ in $\mathbb{R}^{\bar{n}}$ with metric given by $\mathrm{dist}(\mathbf{u}, \mathbf{u}') = \left(\sum_{i=1}^{12} \|u_i - u_i'\|^2\right)^{1/2}$. This metric space can be isometrically identified (as a subset of the $\ell_2$-sum of 12 copies of $\mathbb{R}^{\bar{n}}$) with a subset of $\mathbb{R}^{12\bar{n}}$; it even lies in the ball $B(0, \sqrt{12}) \subset B(0, 4)$ in $\mathbb{R}^{12\bar{n}}$. We choose $M$ as a $\rho_1$-net in $U$, with $\rho_1 = \rho/K^2$. In $\mathbb{R}^{12\bar{n}}$, the balls of radius $\rho_1/2$ around the points of $M$ are disjoint and they are also contained in the ball $B(0, 5)$, say, and so we get

$$|M| \leq \frac{V_{12\bar{n}}(5)}{V_{12\bar{n}}(\rho_1/2)} \leq \left(\frac{10K^2}{\rho}\right)^{12\bar{n}} \leq \left(\frac{K^3}{\rho}\right)^{12\bar{n}}. \quad ∎$$

## 5. An almost nowhere Fréchet smooth norm

In this section, we strengthen Theorem 3.1 as follows:

THEOREM 5.1: *There exists an equivalent norm $p$ on $\ell_2$ such that the set of points where $p$ is Fréchet differentiable is Aronszajn null.*

The proof is not conceptually difficult but a bit technical. The idea how to get the points where a convex function is not Fréchet smooth is the same as in [PZ84]. In order to prove easily that a certain set is Aronszajn null, we intersect a sphere by cones instead of intersecting a paraboloid by cylinders as in [PZ84]. Proving the non-differentiability then requires more computations.

We begin with some notation and preliminaries; see, for example, the book [Ph89] for more details. If $f$ is a convex continuous function on $\ell_2$, we define the **subdifferential of** $f$ at a point $x \in \ell_2$ by

$$\partial f(x) = \{u \in \ell_2 \colon \langle u, y - x\rangle \leq f(y) - f(x) \text{ for all } y \in \ell_2\}.$$

(The elements of the subdifferential are thought of as hyperplanes supporting the graph of $f$ at $(x, f(x))$.) The **oscillation** of $\partial f$ at the point $x$ is given by

$$\operatorname{osc}(\partial f, x) = \lim_{t\to 0} \sup\{\|u - v\| \colon \|x - y\| \leq t, u \in \partial f(x), v \in \partial f(y)\}.$$

The function $f$ is Fréchet differentiable at a point $x$ exactly when $\operatorname{osc}(\partial f, x) = 0$ (see e.g. [Ph89], p. 19).

When we try to construct many points of non-smoothness, sums of convex functions have the advantage that none of the functions in the sum can destroy the "bad" points of the other functions.

LEMMA 5.2: *Let $f$ and $f_1, f_2, f_3, \ldots$ be convex continuous functions on $\ell_2$ such that $f = \sum_{i=1}^{\infty} f_i$. Let $D_i$ be the set of points where $f_i$ is not Fréchet differentiable. Then $f$ is Fréchet differentiable at no point of the set $\bigcup_{i=1}^{\infty} D_i$.*

This seems to be a folklore result but we know no explicit reference, so we include a short proof.

*Proof:* Since $f = f_k + \sum_{i=1}^{k-1} f_i + \sum_{i=k+1}^{\infty} f_i$, it is enough to show the statement for $f = f_1 + f_2$. So suppose that $x \in D_1$. If $f_1$ or $f_2$ are not Gâteaux differentiable at $x$ then $f_1 + f_2$ is also not Gâteaux differentiable at $x$ either since $\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x)$, and Gâteaux differentiability of convex functions is equivalent to single-valuedness of the subdifferential. So suppose both $f_1$ and $f_2$ are Gâteaux differentiable at $x$ and denote by $u_i$ the unique element of $\partial f_i(x)$. Since $f_1$ is not Fréchet differentiable at $x$, there is an $\alpha > 0$ such that for $t > 0$ arbitrarily small we have some $y \in \ell_2$ with $\|x-y\| < t$ and $f_1(y) - f_1(x) - \langle u_1, y-x\rangle \geq \alpha\|x-y\|$. Since $f_2(y) - f_2(x) - \langle u_2, y-x\rangle \geq 0$, we get $f_1(y) + f_2(y) - (f_1(x) + f_2(x)) - \langle u_1 + u_2, y-x\rangle \geq \alpha\|x-y\|$, and $f_1 + f_2$ is not Fréchet differentiable at $x$. ∎

Let $K \subset \ell_2$ be a symmetric, closed, convex, and bounded set containing the origin in its interior. Let $\rho > 0$ be such that $B(0,\rho) \subset K$. The **Minkowski functional** $p\colon \ell_2 \to \mathbb{R}$ of $K$ is given by

$$p(x) = \inf\{t > 0 \colon x \in tK\}.$$

Such a $p$ defines an equivalent norm on $\ell_2$. The subdifferential $\partial p(x)$ is related to the supporting hyperplanes of $K$ as follows (see, e.g., [Ph89], p. 78):

LEMMA 5.3: *Let $p$ and $\rho > 0$ be as above, and let $0 \neq x \in \ell_2$. Then $v \in \partial p(x)$ holds for a $v \in \ell_2$ if and only if $\langle v, x\rangle = p(x)$ and $v$ supports $K$ at the point $x/p(x)$. This means that*

$$1 = \left\langle v, \frac{x}{p(x)} \right\rangle = \max_{y \in K} \langle v, y\rangle.$$

*In particular, any $v \in \partial p(x)$ satisfies $\|v\| \leq 1/\rho$.*

For a set $A \subseteq \ell_2$, let dcone$(A)$ denote the **double cone** of $A$ with apex at the origin, that is,

$$\operatorname{dcone}(A) = \bigcup_{t \in \mathbb{R}} tA,$$

and similarly

$$\operatorname{cone}(A) = \bigcup_{t \geq 0} tA.$$

Theorem 5.1 is an easy consequence of the results presented in the preceding sections and of the following:

PROPOSITION 5.4: *Let $u \in \ell_2$ be a unit vector and let $H$ be the hyperplane $u + \operatorname{Ker} u$ (where $\operatorname{Ker} u$ stands for $\{x \in \ell_2 \colon \langle u, x\rangle = 0\}$). Let $r > 0$ and let $\mathcal{F}$ be a family of balls in $H$ (relative to $H$) of radius $r$ such that $\bigcup \mathcal{F}$ is dense in $H$. Then there exists an equivalent norm $p$ on $\ell_2$ such that $p$ is Fréchet differentiable at no point of the set $D = \operatorname{dcone}(H \setminus \bigcup \mathcal{F})$.*

*Proof:* If $\bigcup \mathcal{F} = H$ the result is obvious (put $p = \|.\|$), so we may assume that there exists a point $a \in H \setminus \bigcup \mathcal{F}$. Let

$$K = \overline{\operatorname{conv}}(B(0,1) \setminus \bigcup \mathcal{F}).$$

The set $K$ is closed, convex, bounded, and symmetric. Since $K$ contains the points $a/\|a\|$ and $-a/\|a\|$ and the set $B(0,1) \cap \operatorname{Ker} u$, there exists a $\rho > 0$ with $B(0,\rho) \subset K$. Therefore $p$, the Minkowski functional of $K$, defines an equivalent norm on $\ell_2$. It remains to show that $p$ is not Fréchet differentiable at any point $x \in D$, $x \neq 0$.

Without loss of generality we can suppose that $\|x\| = 1$ and $\langle x, u\rangle > 0$. Let us set $\alpha = \langle x, u\rangle$. This is a fixed positive number (depending on $x$ only). We will show that

$$\operatorname{osc}(\partial p, x) \geq \beta = \beta(\alpha)$$

for a certain positive $\beta$ depending on $\alpha$ only, and consequently $p$ is not Fréchet differentiable at $x$.

Clearly $x \in \operatorname{bdr} K$ and hence $p(x) = 1$. Since $x$ supports $B(0,1) \supset K$ at the point $x$ and $\langle x, x\rangle = 1$, we get $x \in \partial p(x)$. Next, we want to exhibit an element $v$ of the subdifferential $\partial p(y)$ at a point $y$ arbitrarily near to $x$ such that $\|x - v\| \geq \beta$.

Let $\varepsilon > 0$ be an arbitrarily small number (going to 0 while $\alpha$ and anything depending on $\alpha$ only are fixed). Let $\pi_H\colon H \to \operatorname{bdr} B(0,1)$ be the central projection of the hyperplane $H$ to the unit sphere, given by $\pi_H(w) = w/\|w\|$. Put $x_H = \pi_H^{-1}(x)$, and choose a ball $B_H \in \mathcal{F}$ at distance at most $\varepsilon_1$ from $x_H$, where $\varepsilon_1$ is chosen small enough in terms of $\varepsilon$ and $\alpha$. Let $\tilde{x}_H$ be the point of $B_H$ nearest to $x_H$, and put $\tilde{x} = \pi_H(\tilde{x}_H)$; see Figure 3. Since $\pi_H$ is continuous, $\varepsilon_1$ can be chosen in such a way that $\|x - \tilde{x}\| \leq \rho\varepsilon$.

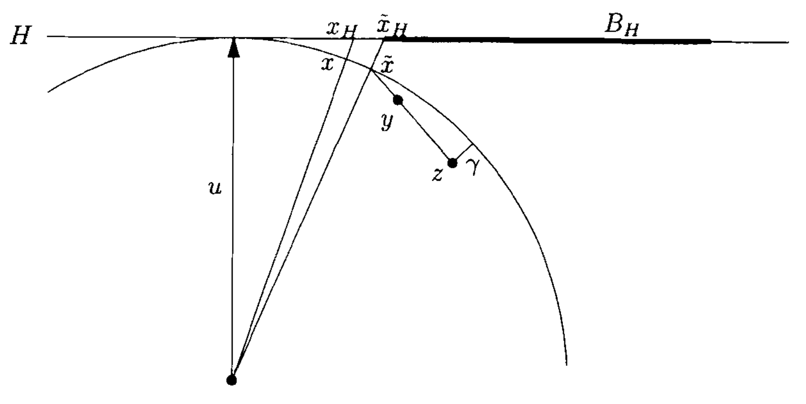


Figure 3. The situation in Claim 5.5.

We need the following geometric claim:

CLAIM 5.5: *There exists a point $z$ such that the segment $\tilde{x}z = \mathrm{conv}\{\tilde{x}, z\}$ does not intersect the interior of $K$, and $\|z\| \le 1-\gamma$, where $\gamma > 0$ depends on $\alpha$ and $r$ but not on $\varepsilon$.*

We prove this claim later. Assuming its validity, we first finish the proof of Proposition 5.4. Let $y$ be the point on the segment $\tilde{x}z$ given by $y = \tilde{x} + \sqrt{\varepsilon}(z-\tilde{x})$. Let $v$ lie in the subdifferential $\partial p(y)$. By Lemma 5.3, we have $\|v\| \le 1/\rho$. The same Lemma further gives $\langle v, y\rangle = p(y) \ge 1$ (because $y \notin \mathrm{int}\, K$), and also $\langle v, x\rangle \le 1$ since $x \in K$. We calculate

$$\langle v, \tilde{x}\rangle = \langle v, \tilde{x} - x\rangle + \langle v, x\rangle \le \|v\| \cdot \|\tilde{x} - x\| + 1 \le 1 + \varepsilon.$$

Since $z = \frac{1}{\sqrt{\varepsilon}} y - (\frac{1}{\sqrt{\varepsilon}} - 1)\tilde{x}$,

$$\langle v, z\rangle = \frac{1}{\sqrt{\varepsilon}}\langle v, y\rangle - (\frac{1}{\sqrt{\varepsilon}} - 1)\langle v, \tilde{x}\rangle \ge \frac{1}{\sqrt{\varepsilon}} - (\frac{1}{\sqrt{\varepsilon}} - 1)(1+\varepsilon) \ge 1 - \sqrt{\varepsilon}.$$

Then we have

$$\|v - x\| \ge \left\langle v - x, \frac{z}{\|z\|}\right\rangle \ge \langle v - x, z\rangle \ge 1 - \sqrt{\varepsilon} - \|x\| \cdot \|z\| \ge 1 - \sqrt{\varepsilon} - (1-\gamma) \ge \frac{\gamma}{2}.$$

Proposition 5.4 is proved; it remains to prove Claim 5.5.

*Proof of Claim 5.5:* The idea is to show that the portion of the unit ball "bitten off" by the cone $C = \mathrm{cone}(B_H)$ is "sufficiently deep". Namely, we want to choose a suitable $z$ in such a way that the hyperplane $Z = z + \mathrm{Ker}\, z$ contains $\tilde{x}$ and satisfies $Z \cap B(0,1) \subset C$ (Figure 4).

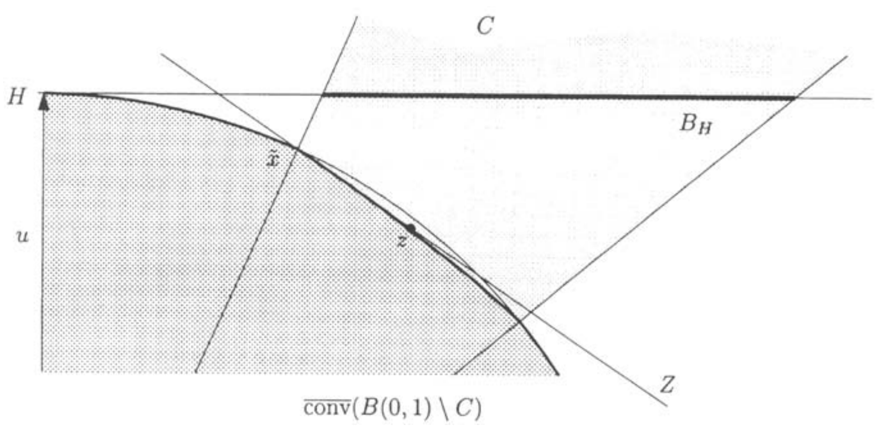


Figure 4. Illustration to the proof of Claim 5.5.

The situation is not as simple as a two-dimensional picture might suggest, since $C$ is in general an elliptic cone rather than a circular one. But if we show that $C$ can be written as a union of circular cones with opening angle bounded from below (i.e. cones of the form $\mathrm{cone}(B(q, r_1))$ with $\|q\| = 1$ and with some fixed $r_1 > 0$ depending on $r, \alpha$) we are done: we can take a circular cone $C' = \mathrm{cone}(B(q, r_1)) \subseteq C$ having $\tilde{x}$ on its boundary, and let $Z$ be the hyperplane cutting off exactly the cap of the unit sphere contained in $C'$.

To show that $C$ is a union of suitable circular cones, we first consider the cone $C_0 = \mathrm{cone}(B(u, r) \cap H)$; see Figure 5. This cone also equals $\mathrm{cone}(B_0)$, where $B_0 = B(u, r')$ is a ball of a suitable radius $r' = r'(r)$ (somewhat smaller than $r$).

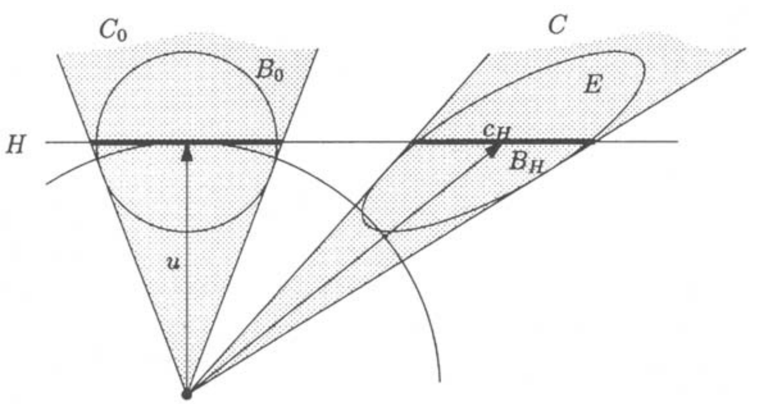


Figure 5. The affine map $F$ sending $C_0$ to $C$.

Let $c_H$ denote the center of the ball $B_H$ in $H$. Consider the linear map $F\colon \ell_2 \to \ell_2$ given by

$$F(x) = x + \langle u, x\rangle(c_H - u).$$

Within $H$, $F$ acts as the translation by the vector $c_H - u$, hence $B_H = F(B(u, r) \cap H)$, and consequently $C = F(C_0)$. The ball $B_0$ is mapped to an ellipsoid $E$; for our purposes, it suffices that the intersection of $E$ with any 2-dimensional affine subspace containing $c_H$ is an ellipse. If $a_E$ denotes the supremum and $b_E$ the infimum of lengths of the semiaxes of these ellipses, then

$$a_E = r'\|F\| \quad \text{and} \quad \frac{1}{b_E} = \frac{1}{r'}\|F^{-1}\|.$$

Since both the norms $\|F\|$ and $\|F^{-1}\|$ are bounded by functions of $\|c_H\|$, both $1/b_E$ and $a_E$ can also be bounded by functions of $\|c_H\|$ and $r$.

It remains to show that $E$ can be expressed as a union of balls of a fixed radius $r_2 > 0$ (depending on $a_E$ and $b_E$). In fact, we only need to show that the points on the boundary of $E$ are contained in such balls. Let $x \in \operatorname{bdr} E$, and let $B_x$ be a ball of some radius $r_2$ containing $x$ at its boundary and having the same tangent hyperplane at $x$ as $E$ does (and lying on the same side of this hyperplane as $E$). If $B_x$ is not completely contained in $E$ then there is a 2-dimensional plane $\tau$ containing $x$ and the center of $E$ such that $B_x \cap \tau$ is not contained in $E \cap \tau$. Since $B_x \cap \tau$ is a circular disc of radius at most $r_2$ and $E \cap \tau$ is an ellipse with semiaxes lengths lying in the interval $[b_E, a_E]$, it suffices to check the following statement in the plane: If $E$ is an ellipse with semiaxes lying in an interval $[b, a]$, $0 < b \leq a$, then there exists a radius $r_2 = r_2(a, b)$ such that for any point $x \in \operatorname{bdr} E$, the circle of radius $r_2$ touching $E$ at $x$ from inside is completely contained in $E$. This can be checked by elementary arguments. In fact, the value $r_2$ is the reciprocal the maximal curvature of an ellipse with semiaxes $a$ and $b$, and we can set $r_2 = b^2/a$. ∎

*Proof of Theorem 5.1:* We fix a unit vector $u \in \ell_2$ and set $H = u + \operatorname{Ker} u$. By Proposition 3.2 there exists, for each $m \in \mathbb{N}$, a family $\mathcal{F}_m$ of congruent balls in $H$ (relative to $H$) such that $\bigcup \mathcal{F}_m$ is dense in $H$ and $\lambda_U(U \cap \bigcup \mathcal{F}_m) \leq 1/m$ for each test cube $U$ in $H$. By Proposition 5.4, there exists an equivalent norm $p_m$ on $\ell_2$ which is Fréchet differentiable at no point of the set $\operatorname{dcone}(H \setminus \bigcup \mathcal{F}_m)$. Put $p = \sum a_m p_m$, where $a_m$ are sufficiently small positive numbers (if $p_m \leq c_m \|\cdot\|$ then put $a_m = 1/2^m c_m$, say). Then $p$ is an equivalent norm on $\ell_2$ which can be Fréchet differentiable only at points of $\bigcap_{m=1}^{\infty} \operatorname{dcone}(\bigcup \mathcal{F}_m)$, by Lemma 5.2.

Let $x \in \ell_2$ and let $X$ be a 12-dimensional subspace of $\operatorname{Ker} u$. Write $x = tu + x_0$, with $t \in \mathbb{R}$ and $x_0 \in \operatorname{Ker} u$. If $t \neq 0$, we have

$$\begin{aligned}(x + X) \cap \bigcap_{m=1}^{\infty} \operatorname{dcone}\Big(\bigcup \mathcal{F}_m\Big) &= t(u + \frac{1}{t}x_0 + X) \cap \bigcap_{m=1}^{\infty} \operatorname{dcone}\Big(\bigcup \mathcal{F}_m\Big) \\ &= t\left((u + \frac{1}{t}x_0 + X) \cap \bigcap_{m=1}^{\infty} \bigcup \mathcal{F}_m\right).\end{aligned}$$

Similarly as in the proof of Theorem 3.1, we see that the 12-dimensional measure of the set in the parentheses is zero for each $t \neq 0$. Since the hyperplane $\operatorname{Ker} u$ (corresponding to $t = 0$) is Aronszajn null, the set $\bigcap_{m=1}^{\infty} \operatorname{dcone}(\bigcup \mathcal{F}_m)$ is Aronszajn null by Lemma 2.2. This proves Theorem 5.1. ∎

### A simple proof of Proposition 3.2 according to David Preiss

Here it is convenient to take 5-dimensional test cubes. So by a test cube, we mean any congruent copy $U$ of the 5-dimensional unit cube $[0,1]^5$ in $\ell_2$. We denote the 5-dimensional Lebesgue measure on $U$ by $\lambda_U$. In the proof we will again use the $K$-times enlarged test cubes congruent to $T_0=[0,K]^5$; for an orthonormal family $\mathbf{u}=(u_1,\ldots,u_5)$ in $\ell_2$ and $x\in\ell_2$, we put

$$T=T(x,\mathbf{u})=\left\{x+\sum_{i=1}^{5}a_iu_i\colon 0\le a_i\le K, i=1,\ldots,5\right\}.$$

Let $\mu=\mu_T$ be the uniform probability measure on $T$ obtained by re-scaling the 5-dimensional Lebesgue measure.

Instead of Claim 4.6 we use the following.

CLAIM 4.6$'$: *Let $Z$ be a 5-dimensional subspace of $\ell_2$ and let $x\in\ell_2$ be so that $\operatorname{dist}(x,Z)\ge 1-\rho$ for some $0<\rho<1$. Then $B(x,1)\cap Z$ is contained in a 5-dimensional ball of radius at most $2\sqrt{\rho}$, whose 5-dimensional Lebesgue measure is thus at most $\beta\rho^{5/2}$, where $\beta$ is an absolute constant.*

For the reader's convenience, we recall the statement being proved.

PROPOSITION 3.2: *Let $\varepsilon>0$ be given. There exist a number $r>0$ and a countable set $C\subset\ell_2$ such that*
(A) *for any $\delta>0$, $B(C,r+\delta)=\ell_2$, and*
(B) *$\lambda_U(U\cap B(C,r))\le\varepsilon$ for any test cube $U$.*

*Proof:* Let $(e_n)$ be the orthonormal basis of $\ell_2$. For $x\in\ell_2$ we define the support of $x$ as $\operatorname{spt}x=\{i\in\mathbb{N}\colon \langle x,e_i\rangle\ne 0\}$. Let $(x_k)_{k=1}^{\infty}$ be a dense sequence in $\ell_2$ with each $x_k$ finitely supported. Choose $n_1<n_2<\cdots$ such that $\max\operatorname{spt}(x_k)<n_k$. Define $c_k=x_k+e_{n_k}$ and $C=\{c_k\colon k\in\mathbb{N}\}$. Let $\varepsilon>0$ be given. We will show that for $K>0$ large enough
(A$'$) for any $\delta>0$, $B(C,1+\delta)=\ell_2$, and
(B$'$) $\mu_T(T\cap B(C,1))\le\varepsilon$ for any congruent copy $T$ of $T_0$.
Once we prove this, the Proposition is obtained by putting $r=1/K$ and replacing $C$ by $(1/K)C'$.

The condition (A$'$) is satisfied since $x_k\in B(c_k,1)$ for all $k$.

To verify condition (B$'$), let $T=T(x,\mathbf{u})$ be a congruent copy of the cube $T_0$. Put $D_T=B(x,R)\cap\operatorname{aff}(T)$, where $R=10K$, and for $j=0,1,\ldots$ let

$$I_j=\{k\in\mathbb{N}\colon 1-2^{-j}\le\operatorname{dist}(D_T,c_k)<1-2^{-j-1}\}.$$

Further, let $w_k \in D_T$ denote a point attaining the distance of $D_T$ to $c_k$, i.e. with $\|w_k - c_k\| = \text{dist}(D_T, c_k)$. If $\text{dist}(c_k, T) < 1$, then $k \in I_j$ for some $j$, and we have $\mu(T \cap B(c_k, 1)) \leq \beta 2^{-5j/2}/K^5$ by Claim 4.6′. Hence

$$\mu(B(C,1) \cap T) \leq \frac{\beta}{K^5} \sum_{j=0}^{\infty} 2^{-5j/2} |I_j|. \tag{6}$$

To estimate $|I_j|$, we first observe that for all but few $k$'s in $I_j$, $e_{n_k}$ is near-orthogonal to $\text{span}\,\mathbf{u}$. Namely, set $\eta = 1/5R \cdot 2^{j+2}$ and define

$$I'_j = \{k \in I_j \colon |\langle u_i, e_{n_k} \rangle| < \eta \text{ for all } i = 1, 2, \ldots, 5\}.$$

Since each $u_i$ is a unit vector, we have

$$|\{k \in \mathbb{N} \colon |\langle u_i, e_{n_k} \rangle| \geq \eta\}| \leq \eta^{-2},$$

and hence $|I_j \setminus I'_j| \leq \beta_1 K^2 2^{2j}$, $\beta_1$ a constant.

Next, we bound $|I'_j|$. We have

$$\|w_k - c_k\| \geq |\langle w_k, e_{n_k} \rangle - \langle c_k, e_{n_k} \rangle| = |\langle w_k, e_{n_k} \rangle - 1|,$$

and since $\|w_k - c_k\| < 1 - 2^{-j-1}$, we derive $\langle w_k, e_{n_k} \rangle > 2^{-j-1}$. Writing $w_k = x + \sum_{i=1}^{5} \alpha_i u_i$, where $|\alpha_i| \leq R$, we get

$$\langle x, e_{n_k} \rangle \geq \langle w_k, e_{n_k} \rangle - 5R\eta \geq 2^{-j-2}.$$

On the other hand, for all $k \in I'_j$ we have $\|x[n_k + 1 \ldots]\| \leq \|x - c_k\| < 2R$, and so if $p$ is the first index with $\|x[p+1 \ldots]\| < 2R$, we find

$$4R^2 > \|[x[p+1 \ldots]\|^2 \geq \sum_{k \in I'_j, n_k > p} \langle x, e_{n_k} \rangle^2 \geq (|I'_j| - 1) \cdot 2^{-2j-4}.$$

Estimating $|I'_j|$ from this inequality and combining with the bound for $|I_j \setminus I'_j|$ derived above, we get that $|I_j| \leq \beta_2 2^{2j} K^2$, where $\beta_2$ is a constant. Finally, by substituting into (6), we arrive at

$$\mu(B(C,1) \cap T) \leq \frac{\beta}{K^5} \beta_2 K^2 \sum_{j=0}^{\infty} 2^{-\frac{j}{2}} \leq \beta_3 \frac{1}{K}.$$

This is at most $\varepsilon$ for $K$ large enough. ∎

ACKNOWLEDGEMENT: We would like to thank David Preiss for explaining his simplified proof to us and Luděk Zajíček for turning our attention to [PZ84].

## References

[AS93] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley, Chichester, 1993.

[Aro76] N. Aronszajn, *Differentiability of Lipschitzian mappings between Banach spaces*, Studia Mathematica **57** (1976), 147–190.

[BN94] J. M. Borwein and D. Noll, *Second order differentiability of convex functions in Banach spaces*, Transactions of the American Mathematical Society **342** (1994), 43–81.

[Chr74] J. P. R. Christensen, *Topology and Borel Structure*, North-Holland, Amsterdam, 1974.

[MS96] E. Matoušková and C. Stegall, *A characterization of reflexive Banach spaces*, Proceedings of the American Mathematical Society **124** (1996), 1083–1090.

[Mat97] E. Matoušková, *Convexity and Haar null sets*, Proceedings of the American Mathematical Society **125** (1997), 1793–1799.

[Mat] E. Matoušková, *The Banach–Saks property and Haar null sets*, Commentationes Mathematicae Universitatis Carolinae **39** (1998), 71–80.

[Ph89] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, Springer-Verlag, Berlin, 1989.

[PT95] D. Preiss and J. Tišer, *Two unexpected examples concerning differentiability of Lipschitz functions on Banach spaces*, in *Geometric Aspects of Functional Analysis*, GAFA Israel Seminar 92–94 (V. D. Milman and J. Lindenstrauss, eds.), Birkhäuser, 1995, pp. 219–238.

[PZ84] D. Preiss and L. Zajíček, *Stronger estimates of smallness of sets of Fréchet nondifferentiability of convex functions*, in *Proceedings of the 11th Winter School*, Supplement to Rendiconti del Circolo Matematico di Palermo, Serie II, No. 3, 1984, pp. 219–223.

[Rog57] C. A. Rogers, *A note on coverings*, Mathematika **4** (1957), 1–6.