Title
High-dimensional tennis balls

Journal
Combinatorial Theory, 2(2)

ISSN
2766-1334

Authors
Gowers, Timothy
Wyczesany, Katarzyna

Publication Date
2022

DOI
10.5070/C62257881

COMBINATORIAL THEORY **2 (2)** (2022), #15 combinatorial-theory.org

High-dimensional tennis balls

Timothy Gowers[1] and Katarzyna Wyczesany*[2]

[1] *Collège de France and the Department of Pure Mathematics and Mathematical Statistics, University of Cambridge*
wtg10@dpmms.cam.ac.uk
[2] *Department of Pure Mathematics, Tel Aviv University, Israel*
kasiawycz@outlook.com

Submitted: Jun 2, 2021; Accepted: May 31, 2022; Published: Jun 30, 2022

Abstract. We show that there exist constants $\alpha, \epsilon > 0$ such that for every positive integer n there is a continuous odd function $f : S^m \to S^n$, with $m \geqslant \alpha n$, such that the ϵ-expansion of the image of f does not contain a great circle. This result is motivated by a conjecture of Vitali Milman about well-complemented almost Euclidean subspaces of spaces uniformly isomorphic to ℓ_2^n.

Keywords. Anti-Ramsey, antipodal subsphere

Mathematics Subject Classifications. 46B09, 60C05

1. Introduction

Let U be a measurable subset of the sphere S^n. How large must the measure of U be in order to guarantee that the ϵ-expansion of U, that is, the set U_ϵ that consists of all points at distance at most ϵ from U, contains the unit sphere of a subspace of dimension k? Such questions have been much studied ever since Milman's famous proof [Mil71] of Dvoretzky's theorem [Dvo61]. Milman's insight was that by the isoperimetric inequality in the sphere, the volume of the $\epsilon/2$-expansion (say) is minimized, for a given measure of U, when U is a spherical cap. But when U is a spherical cap of measure α and n is large, a relatively straightforward calculation shows that $U_{\epsilon/2}$, which is again a spherical cap, has measure very close to 1. From this it follows, again straightforwardly, that under suitable conditions on the parameters, an $\epsilon/2$-net of the sphere of a *random* subspace of dimension k will lie in $U_{\epsilon/2}$ with high probability, and hence that the entire sphere will lie in U_ϵ. This basic argument can be used to prove the surprising result that even if ϵ is quite small, an exponentially small measure for U is sufficient to guarantee that U_ϵ contains the sphere of an k-dimensional subspace for some k of dimension that is linear in n.

If U has measure c^n and c is too small, then the above argument fails, and the conclusion is false. Indeed, if U is a spherical cap of volume c^n, and if its spherical radius is $\pi/2 - \eta$, then

*Supported by the Engineering and Physical Sciences Research Council grant EP/L016516/1.

unless $\epsilon > \eta$ the volume of U_ϵ will not be close to 1, and if $\epsilon < \eta$ then U_ϵ will not even contain two antipodal points, let alone the sphere of a subspace of dimension k.

One could attempt to rule out this simple example by restricting attention to centrally symmetric sets, but that does not achieve much: if U is the cap just discussed, and if its centre is the unit vector u, then $U \cup (-U)$ is centrally symmetric, and $(U \cup (-U))_\epsilon$ is disjoint from the hyperplane orthogonal to u, which implies that $(U \cup (-U))_\epsilon$ does not contain the sphere of any 2-dimensional subspace.

A noticeable feature of this example is that it is in a certain sense zero-dimensional: if we identify antipodal points and write q for the quotient map, then $q(U)$ is a subset of projective n-space, and it is homotopic to a point. Having made this observation, it is natural to wonder what happens if we impose a condition that forces $q(U)$ to have a higher dimension in this topological sense. That motivates the following definition.

Definition 1.1. An m-dimensional *antipodal subsphere* of S^n is the image of a continuous function $f : S^m \to S^n$ that preserves antipodal points.

If an m-dimensional antipodal subsphere of S^n is the unit sphere of an $(m+1)$-dimensional subspace of $\mathbb{R}^{n+1}$, then we shall call it *linear*.

An m-dimensional antipodal subsphere is in a certain sense "genuinely m-dimensional". For instance, if X is such a subsphere and $g : X \to \mathbb{R}^m$ is a continuous function, then $g \circ f$ is a continuous function from S^m to $\mathbb{R}^m$, which implies, by the Borsuk–Ulam theorem, that there is some $x \in S^m$ such that $g(f(x)) = g(f(-x))$, and therefore that $g(f(x)) = g(-f(x))$. Thus, for any continuous map from X to $\mathbb{R}^m$ there will be two antipodal points with the same image. Essentially the same argument shows that in projective n-space there is no homotopy from $q(X)$ (where q is the quotient map defined above) to a lower-dimensional set.

We now ask the following question.

Question 1.2. Let $\epsilon > 0$, let k be a positive integer, and let X be an m-dimensional antipodal subsphere of S^n. How large does m have to be in order to guarantee that X_ϵ contains a linear subsphere of dimension k?

In order to tackle this question, an obvious first step is to see how well one can do using concentration of measure. That is, we consider instead a slightly stronger question.

Question 1.3. Let $\epsilon > 0$, let k be a positive integer, and let X be an m-dimensional antipodal subsphere of S^n. How large does m have to be in order to guarantee that X_ϵ contains almost all linear subspheres of dimension k?

By standard arguments, that is roughly the same as asking for X_ϵ to have measure at least $1 - \epsilon^k$.

The following estimate is well known. See for example [Art02].

Lemma 1.4. *Let $m = \alpha n$ and let $\epsilon(\alpha)$ be such that $\sin \epsilon(\alpha) = \sqrt{1-\alpha}$. Then if X is a linear m-dimensional subsphere, the measure of X_ϵ tends to 1 if $\epsilon > \epsilon(\alpha)$ and to 0 if $\epsilon < \epsilon(\alpha)$.*

This implies that for the second question we need m to be at least αn, where $\sqrt{1-\alpha} = \sin\epsilon$, or $\alpha = \cos^2\epsilon \approx 1 - \epsilon^2/2$.

However, it is not obvious what this observation tells us about the first question, since these examples are linear subspheres, which are good sets to choose for the second question but the worst possible sets to choose for the first. That is, if we wish to find an antipodal subsphere X of dimension m such that X_ϵ contains only a very small proportion of all linear k-dimensional subspaces, then we should take X itself to be linear, but if we would like X_ϵ to contain *no* linear k-dimensional subspace, then obviously we cannot take X to be linear (unless its dimension is less than k).

The main result of this paper is that even when $k = 1$, the dimension of X can be quite large.

Theorem 1.5. *There exist constants $\alpha, \epsilon > 0$ such that for every n there is an $\lfloor \alpha n \rfloor$-dimensional antipodal subsphere X of S^n such that X_ϵ contains no linear subsphere of dimension 1.*

To put this less formally, there is an antipodal subsphere $X \subset S^n$ of dimension linear in n such that the expansion X_ϵ does not contain any 1-dimensional linear subsphere.

We informally call such an antipodal subspace a *tennis ball* because it brings to mind the seam of a genuine tennis ball (though the resemblance is not perfect, since the seam of a genuine tennis ball is not centrally symmetric). We shall also refer to 1-dimensional linear subspheres as *great circles*.

Theorem 1.5 has a simple corollary that can be thought of as an anti-Ramsey theorem for linear subspheres.

Corollary 1.6. *There exist constants $\alpha, \eta > 0$ such that for every n there is a partition of S^n into two subsets M and N such that M_η does not contain the unit sphere of any subspace of codimension less than $\lfloor \alpha n \rfloor$ and N_η does not contain the unit sphere of any 2-dimensional subspace.*

Proof (assuming Theorem 1.5). Let α, ϵ and X be as given by Theorem 1.5. Let $N = X_{\epsilon/2}$, and M its complement on S^n. Then since X_ϵ does not contain a great circle, every great circle contains a point that does not belong to $(X_{\epsilon/2})_{\epsilon/2} = N_{\epsilon/2}$, which implies that this point is at distance at least $\epsilon/2$ from N.

It remains to prove that if $M = S^n \setminus N$ and $r < m$, where $m = \lfloor \alpha n \rfloor$, then $M_{\epsilon/2}$ does not contain the sphere of a subspace of codimension r, since then we will be done with $\eta = \epsilon/2$. Since M_ϵ is disjoint from X, it is sufficient to prove that X intersects every subspace of codimension r.

Let V be such a subspace. We are given that X is the image of some continuous odd function $f : S^m \to S^n$. Hence, after composing f with a rotation, we may assume without loss of generality that $V = \{x : x_1 = \cdots = x_r = 0\}$. Let P_r be the coordinate projection to the first r coordinates. Then $P_r \circ f$ is a continuous map from S^m to $\mathbb{R}^r$, so by the Borsuk–Ulam theorem there exists $x \in S^m$ such that $P_r f(x) = P_r f(-x)$. Since $P_r \circ f$ is also odd, it follows that $P_r f(x) = 0$, and therefore that $f(x) \in V$, as claimed. $\square$

We actually deduce Theorem 1.5 from a stronger result that trivially implies it and is of independent interest.

Theorem 1.7. *There exist constants* $\alpha, \epsilon > 0$ *and a continuous map* $\psi : S^n \to S^n$ *such that* ψ *preserves antipodal points and such that if* X *is a random linear subsphere of dimension* $\lfloor \alpha n \rfloor$ *then with probability* $1 - o(1)$ *the set* $\psi(X)_\epsilon$ *does not contain a great circle.*

1.1. A remark about a question of V. Milman

Let us briefly explain our original motivation for the above result. We were interested in the following question of Vitali Milman, concerning a possible strengthening of Dvoretzky's theorem.

Question 1.8. Let k be a positive integer, let $C \geqslant 1$ and let $\epsilon > 0$. Does there exist n such that if $X = (\mathbb{R}^n, |||\cdot|||)$ is any normed space such that $\|x\| \leqslant |||x||| \leqslant C\|x\|$ for every $x \in X$, then X has a subspace of dimension k that is $(1+\epsilon)$-complemented and has Banach–Mazur distance at most $1 + \epsilon$ from ℓ_2^k?

For the reader unfamiliar with the terminology, the *Banach–Mazur distance* between two isomorphic normed spaces X and Y is the infimum of $\|T\|\|T^{-1}\|$ over all linear isomorphisms $T : X \to Y$, and a subspace $V \subset X$ is α*-complemented* if there is a projection $P : X \to V$ with $\|P\| \leqslant \alpha$.

We initially attempted to obtain a positive answer to a stronger question, namely the following.

Question 1.9. Let k be a positive integer, let $C \geqslant 1$ and let $\epsilon > 0$. Does there exist n such that if $X = (\mathbb{R}^n, |||\cdot|||)$ is any normed space such that $\|x\| \leqslant |||x||| \leqslant C\|x\|$ for every $x \in X$, then X has a subspace V of dimension k such that there exists β with $\beta\|v\| \leqslant |||v||| \leqslant \beta(1+\epsilon)\|v\|$ for every $v \in V$ and such that $\|P\| \leqslant 1 + \epsilon$, where P is the orthogonal projection from X to V?

In relation to this conjecture, we identified a class of points $x \in X$ that we called ϵ*-good*: these are points x such that the orthogonal projection to the 1-dimensional space spanned by x has norm at most $1 + \epsilon$. The second question turns out (by a not very difficult argument) to be equivalent to asking whether for sufficiently large n there is a k-dimensional subspace consisting entirely of ϵ-good points.

A number of examples have led us to believe that for any normed space satisfying the conditions of the question, the set of ϵ-good points should be "genuinely cn-dimensional" for some positive constant c that depends on C and ϵ. (We have not formulated a suitable conjecture, but one possible definition of a set $X \subset S^n$ such that $X = -X$ being "at least m-dimensional" is that when we regard X as a subset of projective n-space, it is not homotopic to a subset of dimension less than m.) Also, a point that is close to an ϵ-good point is 2ϵ-good. Therefore, a positive answer to the question would follow if one could show that the ϵ-expansion of a "genuinely high-dimensional" subset of the sphere contains a k-dimensional linear subsphere.

However, the main result of this paper shows that this is false. In a separate paper we show that the answer to Question 1.9 is also negative [GW21]. However, while the two results arose from the same line of thought, the constructions are somewhat different, and neither result directly implies the other.

2. The construction

Throughout this note we shall write $\|\cdot\|$ for the norm given by the formula

$$\|x\|^2 = n^{-1}\sum_{i=1}^{n} x_i^2. \tag{2.1}$$

The advantage of the factor n^{-1} on the right-hand side is that a typical coordinate of a random vector of norm 1 has order of magnitude 1 rather than order of magnitude $n^{-1/2}$. This norm is often called the L_2^n norm on $\mathbb{R}^n$, and we write $L_2^n = (\mathbb{R}^n, \|\cdot\|)$. It is the Euclidean norm most commonly used in additive combinatorics. Following the standard terminology in that field, we shall sometimes write the right-hand side of the formula above as $\mathbb{E}_i x_i^2$. Moreover, from this point on we shall think of spheres concretely, so S^{n-1} will denote the unit sphere of L_2^n.

2.1. The tennis ball map

We are aiming to prove Theorem 1.7, or in other words to prove that there exists a continuous map (in fact it will be bi-Lipschitz) $\psi : S^{n-1} \to S^{n-1}$ that preserves antipodal points, with the property that if X is a random $\lfloor \alpha n \rfloor$-dimensional subsphere of S^{n-1}, then with high probability $\psi(X)_\epsilon$ contains no linear subsphere of dimension 1. We shall achieve this by identifying a set $\Gamma \subset S^{n-1}$ such that with high probability $\psi(X \cap S^{n-1}) \subset \Gamma$, or equivalently $X \cap S^{n-1} \subset \psi^{-1}(\Gamma)$, and such that every great circle contains a point that does not belong to Γ_ϵ.

These properties are clearly in tension with each other: we need Γ to have small measure, or else its expansion Γ_ϵ will contain a great circle, but on the other hand we also need $\psi^{-1}(\Gamma)$ to have measure very close to 1, or else it will not contain almost all $\lfloor \alpha n \rfloor$-dimensional linear subspheres.

In order to resolve this tension, we define a map φ that takes "typical" vectors to highly "atypical" vectors and let Γ to be the set of "atypical" vectors. One should think of an "atypical" vector as a vector in S^{n-1} whose coordinates belong to a set of small measure., which we call B. Moreover, the set B will have a special structure so that the normalization of φ to a map from S^{n-1} to itself, also mostly has "atypical" coordinates.

To be more precise, let k be a large positive integer to be chosen later, let $\lambda > 1$, and define $s = \lambda^{1/2k}$. (The parameter λ will later be chosen to be 4, but we write most of the arguments in slightly greater generality in order to emphasize a certain flexibility in our construction and to make the role of this parameter more explicit.) Define "wide" sets A_m and "narrow" sets B_m as

$$A_m = [s^{2mk+1}, s^{2(m+1)k-1}] \quad \text{and} \quad B_m = [s^{2mk-1}, s^{2mk+1}]. \tag{2.2}$$

Hence, the "typical" set is given by $A = \bigcup_m (A_m \cup (-A_m))$ and the "atypical" set is $B = \bigcup_m (B_m \cup (-B_m))$. Note that $A \cup B = \mathbb{R}$, the intersection has measure zero and that B is significantly smaller than A. We define a function φ with the property that $\varphi(A_m) = B_{m+1}$ for every m (and then consider its normalization ψ).

Definition 2.1. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a strictly increasing continuous odd function such that for every integer m we have

$$\varphi(s^{2mk-1}) = s^{2mk+1} \text{ and } \varphi(s^{2mk+1}) = s^{2(m+1)k-1}$$

and such that φ is linear on (s^{2mk-1}, s^{2mk+1}) and $(s^{2mk+1}, s^{2(m+1)k-1})$ for every m.

When $x = (x_1, \dots, x_n)$ is a vector in $\mathbb{R}^n$ we shall abuse notation by writing $\varphi(x)$ to denote the vector $(\varphi(x_1), \dots, \varphi(x_n))$. The *tennis ball map* is a function $\psi : S^{n-1} \to S^{n-1}$, which is a normalized version of φ, given by the formula

$$\psi(x) = \frac{\varphi(x)}{\|\varphi(x)\|}.$$

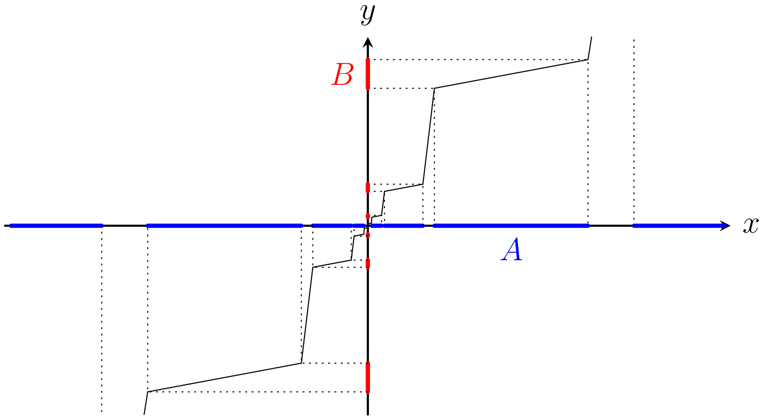

Figure 2.1: The "staircase function" φ.

Note that the graph of φ has a kind of staircase shape with steps of sizes that grow exponentially (see Figure 2.1). Indeed, as x increases from s^{2mk-1} to s^{2mk+1}, which is only a small change proportionately speaking, $\varphi(x)$ increases from s^{2mk+1} to $s^{2(m+1)k-1}$, which is an increase by a factor of almost λ. Similarly, as x increases from s^{2mk+1} to $s^{2(m+1)k-1}$, which is about λ times as big, $\varphi(x)$ increases from $s^{2(m+1)k-1}$ to $s^{2(m+1)k+1}$, which is only a small increase.

Further, observe that if $x, y \in B$, then xy^{-1} belongs to an interval of the form $[s^{2mk-2}, s^{2mk+2}]$, or minus such an interval. Recall that the set BB^{-1} is defined to be $\{xy^{-1} : x, y \in B\}$. As mentioned above, the set B is a "geometric-progression-like" set. This ensures that BB^{-1} is not much larger than B itself. It follows that $B^2B^{-2} = \{x^2y^{-2} : x, y \in B\}$ consists of all points that belong to an interval of the form $[s^{2mk-4}, s^{2mk+4}]$. This fact will be useful later on.

Finally, we have that

$$\|x\| \leqslant \|\varphi(x)\| \leqslant \lambda\|x\| \tag{2.3}$$

for every x. This implies that ψ is a Lipschitz function. Similarly, we observe that the inverse of φ satisfies $\frac{1}{\lambda}\|x\| \leqslant \|\varphi^{-1}(x)\| \leqslant \|x\|$, and hence we get that ψ^{-1} is Lipschitz as well. Thus, the tennis ball map ψ is indeed bi-Lipschitz as claimed.

In what follows we shall show that the tennis ball map ψ takes random linear subspheres of appropriate dimension to tennis balls.

2.2. The definition of the set Γ

Now that we have defined the tennis ball map, let us discuss the set Γ. The rough idea is that Γ is the set of points $x \in S^{n-1}$ with almost all their coordinates in B. Since A has a small complement, one would expect almost all coordinates of a random vector to belong to A, and indeed this is the case. It forms the basis of a probabilistic argument that shows that with high probability every $x \in S^{n-1} \cap X$ has the property that almost every coordinate of x belongs to A, which implies that almost every coordinate of $\varphi(x)$ belongs to B. Thus a "typical" vector (one with almost all coordinates in the large set A) is mapped to a highly "atypical" vector (one with almost all coordinates in the small set B).

This is a slight oversimplification, because of the normalization that replaces φ by ψ. The actual definition of Γ concerns the *ratios* of the coordinates rather than their actual values.

Now let us give some more details. For the purposes of this problem, it is more natural, when talking about a unit vector x, to attach a weight of x_i^2 to the ith coordinate. For example, the statement "almost every ratio $x_i x_j^{-1}$ belongs to BB^{-1}" should be interpreted as meaning that

$$\sum_{x_i x_j^{-1} \in BB^{-1}} x_i^2 x_j^2 \geqslant (1-\epsilon) \sum_{i,j=1}^{n} x_i^2 x_j^2$$

for some small ϵ, and similarly for other statements about coordinates. More precisely, with every vector x we define an associated probability measure as follows.

Definition 2.2. For any $x \in \mathbb{R}^n \setminus \{0\}$ define a measure μ_x on $\{1, 2, \ldots, n\}$ such that for any subset $J \subset \{1, 2, \ldots, n\}$ we have the following formula

$$\mu_x(J) = \frac{\|P_J x\|^2}{\|x\|^2},$$

where P_J is the coordinate projection to the subset J.

Let us now fix some useful notation. For functions

$$f : \{1, 2, \ldots, n\} \to \mathbb{R} \qquad \text{and} \qquad g : \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\} \to \mathbb{R}$$

we write

$$\mathbb{E}_i^x f(i) = \|x\|^{-2} \mathbb{E}_i x_i^2 f(i) \quad \text{and} \quad \mathbb{E}_{i,j}^x g(i,j) = \|x\|^{-4} \mathbb{E}_{i,j} x_i^2 x_j^2 g(i,j),$$

where, as mentioned earlier, $\mathbb{E}_i$ and $\mathbb{E}_{i,j}$ are the usual averages.

Then, given a property Q of integers in the set $\{1, 2, \ldots, n\}$ we define

$$\mathbb{P}_i^x[Q(i)] = \mathbb{E}_i^x \mathbb{1}_Q(i) = \mu_x\{i : Q(i)\}.$$

Similarly, we shall write

$$\mathbb{P}^x_{i,j}[Q(i,j)] = \mathbb{E}^x_{i,j}\mathbb{1}_Q(i,j) = (\mu_x \times \mu_x)\{(i,j) : Q(i,j)\},$$

where $\mu_x \times \mu_x$ is the product measure.

We are now ready to define the set Γ.

Definition 2.3. Let $\Gamma \subset S^{n-1}$ be the set given by

$$\Gamma = \{y \in S^{n-1} :\ \mathbb{P}^y_{i,j}[y_i/y_j \in BB^{-1}] \geqslant 1 - 2\lambda^2\beta\}, \tag{2.4}$$

for a suitable choice of the parameter β.

Also important to us will be a set Δ defined by

$$\Delta = \{x \in S^{n-1} :\ \mathbb{P}^x_i[x_i \in A] < 1 - \beta\}. \tag{2.5}$$

In this section we will be more interested in the complement

$$\Delta^c = \{x \in S^{n-1} :\ \mathbb{P}^x_i[x_i \in A] \geqslant 1 - \beta\}$$

but our choice of notation is more convenient in further sections.

Proposition 2.4. *The image of* Δ^c *under the tennis ball map* ψ *is a subset of* Γ.

Proof. Let $x \in \Delta^c$ and note that if $\mathbb{P}^x_i[x_i \in A] \geqslant 1 - \beta$, then $\mathbb{P}^x_i[\varphi(x_i) \in B] \geqslant 1 - \beta$, by the definition of the function φ. Moreover, from (2.3) it follows that $\|\varphi(x)\|^2$ always lies between $\|x\|^2$ and $\lambda^2\|x\|^2$ and therefore $\mu_{\varphi(x)}(E) \leqslant \lambda^2\mu_x(E)$ for every vector $x \in \mathbb{R}^n$ and every set $E \subset \{1, 2, \ldots, n\}$. From this it follows (considering complements) that

$$\mathbb{P}^{\varphi(x)}_i[\varphi(x_i) \in B] > 1 - \lambda^2\beta.$$

Since this is true for every $x \in \Delta^c$ we get that

$$\varphi(\Delta^c) \subset \{y : \mathbb{P}^y_i[y_i \in B] > 1 - \lambda^2\beta\}. \tag{2.6}$$

In order to show the inclusion for ψ, note that if we have y as above then

$$\mathbb{P}^y_{i,j}[y_i \in B \text{ and } y_j \in B] \geqslant 1 - 2\lambda^2\beta,$$

which in turn gives us

$$\mathbb{P}^y_{i,j}[y_i/y_j \in BB^{-1}] \geqslant 1 - 2\lambda^2\beta.$$

From this follows that such a y belongs to Γ. Since Γ is invariant under positive scalar multiples, we conclude using (2.6) that $\psi(\Delta^c) \subset \Gamma$. $\square$

In the next section we shall prove that the expansion Γ_ϵ contains no great circle, and in the following one we shall prove that for a suitable constant $\alpha > 0$, a random linear subsphere of dimension $\lfloor \alpha n \rfloor$ is contained in Δ^c with high probability.

3. Proving that no great circle is contained in Γ_ϵ

We begin with a couple of lemmas that help us to describe the set Γ_ϵ. Recall that if x is a vector in $\mathbb{R}^n$, then $\|x\|$ is its norm in L_2^n (as defined in (2.1)).

Lemma 3.1. *Let y, z be unit vectors in L_2^n with $\|y-z\| \leqslant \epsilon$, and let E be a subset of $\{1, 2, \ldots, n\}$. Then $\left|\mathbb{P}_i^y[E] - \mathbb{P}_i^z[E]\right| \leqslant 2\epsilon$.*

Proof. The left-hand side is equal to $|\mathbb{E}_i(y_i^2 - z_i^2)\mathbb{1}_E(i)| \leqslant \mathbb{E}_i|y_i^2 - z_i^2|$. By the Cauchy–Schwarz inequality it follows that

$$\begin{aligned}\mathbb{E}_i|y_i^2 - z_i^2| &= \mathbb{E}_i|y_i - z_i|\,|y_i + z_i| \\ &\leqslant \|y-z\|\|y+z\| \\ &\leqslant 2\epsilon,\end{aligned}$$

which proves the result. □

Recall from §2.1 that BB^{-1} is the union of all intervals of the form $[s^{2mk-2}, s^{2mk+2}]$, and B^2B^{-2} is the union of all intervals of the form $[s^{2mk-4}, s^{2mk+4}]$, where $s = \lambda^{1/2k}$ was one of the parameters used to define the "staircase function" φ. It follows that if $t \in BB^{-1}$ and $u \notin B^2B^{-2}$, then $|t/u| \notin [s^{-2}, s^2]$, which implies in particular that $|t/u - 1| \geqslant 1 - s^{-2}$.

Lemma 3.2. *Let $\tau = 1 - s^{-2}$ and $\epsilon < \tau^2$. Then for $z \in \Gamma_\epsilon$ we have*

$$\mathbb{P}_{i,j}^z[z_i/z_j \in B^2B^{-2}] \geqslant 1 - 2\lambda^2\beta - 6\epsilon.$$

Proof. Let $y \in \Gamma$ with $\|y\| = 1$ be such that $\|y - z\| \leqslant \epsilon$. Then $\mathbb{P}_{i,j}^y[y_i/y_j \in BB^{-1}] \geqslant 1 - 2\lambda^2\beta$, or equivalently

$$\mathbb{E}_{i,j}y_i^2y_j^2\mathbb{1}_{[y_i/y_j \in BB^{-1}]} \geqslant 1 - 2\lambda^2\beta.$$

By Lemma 3.1 (and recalling that $\mathbb{E}_i y_i^2 = 1$), it follows that

$$\mathbb{E}_{i,j}(y_i^2 - z_i^2)y_j^2\mathbb{1}_{[y_i/y_j \in BB^{-1}]} \leqslant 2\epsilon$$

and hence combining both inequalities gives

$$\mathbb{E}_{i,j}z_i^2y_j^2\mathbb{1}_{[y_i/y_j \in BB^{-1}]} \geqslant 1 - 2\lambda^2\beta - 2\epsilon. \tag{3.1}$$

If the conclusion that $\mathbb{P}_{i,j}^z[z_i/z_j \in B^2B^{-2}] \geqslant 1 - 2\lambda^2\beta - 6\epsilon$ is not true, then we would have

$$\mathbb{E}_{i,j}z_i^2z_j^2\mathbb{1}_{[z_i/z_j \notin B^2B^{-2}]} \geqslant 2\lambda^2\beta + 6\epsilon,$$

and by Lemma 3.1 again it follows that

$$\mathbb{E}_{i,j}z_i^2y_j^2\mathbb{1}_{[z_i/z_j \notin B^2B^{-2}]} \geqslant 2\lambda^2\beta + 4\epsilon. \tag{3.2}$$

By summing (3.1) and (3.2), and estimating trivially the probability of the union of the events by 1, we deduce that

$$\mathbb{E}_{i,j}z_i^2y_j^2\mathbb{1}_{[y_i/y_j \in BB^{-1} \text{ and } z_i/z_j \notin B^2B^{-2}]} \geqslant 2\epsilon.$$

As remarked before the lemma, if $y_i/y_j \in BB^{-1}$ and $z_i/z_j \notin B^2B^{-2}$, then $|\frac{y_i z_j}{y_j z_i} - 1| > \tau$. It follows that

$$\mathbb{E}_{i,j} z_i^2 y_j^2 \Big(\frac{y_i z_j}{y_j z_i} - 1\Big)^2 > 2\tau^2 \epsilon.$$

But since $\mathbb{E}_i y_i^2 = \mathbb{E}_i z_i^2 = 1$ we have

$$\begin{aligned}\mathbb{E}_{i,j} z_i^2 y_j^2 \Big(\frac{y_i z_j}{y_j z_i} - 1\Big)^2 &= \mathbb{E}_{i,j}(y_i z_j - z_i y_j)^2 \\ &= \mathbb{E}_{i,j}(y_i^2 z_j^2 + y_j^2 z_i^2 - 2y_i y_j z_i z_j) \\ &= 2 - 2\langle y, z\rangle^2.\end{aligned}$$

Furthermore, if $2-2\langle y,z\rangle^2 > 2\tau^2\epsilon$, then $\langle y,z\rangle^2 < 1-\tau^2\epsilon$, which implies that $\langle y,z\rangle < 1-\tau^2\epsilon/2$, which in turn means that

$$\|y-z\|^2 = 2 - 2\langle y,z\rangle > \tau^2\epsilon,$$

and therefore that $\|y-z\| > \tau\sqrt{\epsilon}$. However, since by assumption we have $\tau > \sqrt{\epsilon}$, this is a contradiction. □

Corollary 3.3. *It follows directly from Lemma 3.2 that*

$$\Gamma_\epsilon \subset \{z : \mathbb{P}^z_{i,j}[z_i/z_j \in B^2B^{-2}] \geqslant 1 - 2\lambda^2\beta - 6\epsilon\}.$$

This bigger set resembles Γ but is defined using slightly different parameters. We now turn to the proof that every great circle contains a point that does not belong to this slightly expanded Γ-like set.

3.1. Finding a suitable point in an arbitrary 2-dimensional subspace

Let Y be a 2-dimensional subspace of L_2^n and let $\{u, v\}$ be an orthonormal basis for Y. Then the unit sphere of Y consists of vectors $u\cos\theta + v\sin\theta$. The ith coordinate of such a vector, $u_i\cos\theta + v_i\sin\theta$, can be rewritten as $a_i\sin(\theta+\phi_i)$, where $a_i = \sqrt{u_i^2+v_i^2}$ and ϕ_i is chosen such that $a_i\sin\phi_i = u_i$ and $a_i\cos\phi_i = v_i$. Let $a = (a_1,\dots,a_n)$ and note that $\|a\|^2 = 2$.

We start by proving that there are plenty of pairs (i,j) such that ϕ_i is not close to ϕ_j or $-\phi_j$.

Lemma 3.4. *With $a_1,\dots,a_n$ and $\phi_1,\dots,\phi_n$ as above, we have the inequality*

$$\mathbb{P}^a_{i,j}[\cos(2(\phi_i-\phi_j)) \leqslant 1/2] \geqslant 1/3.$$

Proof. Since u, v are fixed, orthogonal unit vectors (in L_2^n) we have that $\|a\|^2 = \mathbb{E}_i(u_i^2+v_i^2) = 2$ and hence $\mathbb{E}_i^a \sin^2(\theta+\phi_i) = \|a\|^{-2}\mathbb{E}_i a_i^2\sin^2(\theta+\phi_i) = \frac{1}{2}\mathbb{E}_i(u_i^2\sin^2\theta + v_i^2\cos^2\theta) = \frac{1}{2}$ for every θ. Therefore, we find on differentiating with respect to θ that

$$2\mathbb{E}_i^a \sin(\theta+\phi_i)\cos(\theta+\phi_i) = \mathbb{E}_i^a \sin(2\theta+2\phi_i) = 0$$

for every θ, and hence, on differentiating again, that

$$\mathbb{E}_i^a \cos(2\theta+2\phi_i) = 0$$

for every θ as well.

From that it follows that

$$\mathbb{E}^a_{i,j}\Big(\cos(2\theta+2\phi_i)\cos(2\theta+2\phi_j)+\sin(2\theta+2\phi_i)\sin(2\theta+2\phi_j)\Big)=\mathbb{E}^a_{i,j}\cos(2(\phi_i-\phi_j))=0.$$

Let F be the event that $\cos(2(\phi_i-\phi_j))\leqslant 1/2$. We have seen that $\mathbb{E}^a_{i,j}\cos(2(\phi_i-\phi_j))=0$, and we also know that $\cos(2(\phi_i-\phi_j))\in[-1,1]$. So, using the total probability formula we get

$$0=\mathbb{E}^a_{i,j}\cos(2(\phi_i-\phi_j))\geqslant\frac{1}{2}\mathbb{P}^a_{i,j}[F^c]-\mathbb{P}^a_{i,j}[F]=\frac{1}{2}-\frac{3}{2}\mathbb{P}^a_{i,j}[F],$$

from which the desired inequality follows. □

Next, we need a technical lemma that will help us to show that if ϕ_i is not approximately $\pm\phi_j$, then $\sin(\theta+\phi_i)/\sin(\theta+\phi_j)$ is not often close to an element of some given geometric progression.

Lemma 3.5. *Let θ be chosen randomly from $[-\pi,\pi]$ and let $0<a<b$. Then*

$$\mathbb{P}\left[a\leqslant\cot\theta\leqslant b\right]\leqslant\frac{b-a}{\pi(1+a^2)},$$

and the same bound holds for the probability that $\cot\theta\in[-b,-a]$.

Proof. Since $\cot$ is periodic with period π and is decreasing in the interval $(0,\pi)$, the probability in question is $(\cot^{-1}a-\cot^{-1}b)/\pi$. By the mean value theorem, $\cot^{-1}a-\cot^{-1}b$ is at most $|a-b|$ times the absolute value of the derivative of $\cot^{-1}$ at a. Since that derivative is $-1/(1+a^2)$, the first result follows. The second then holds by symmetry. □

Recall once again that B^2B^{-2} is the set of all real numbers x such that

$$|x|\in[\lambda^m s^{-4},\lambda^m s^4]$$

for some positive integer m.

The main point about the bound in the next lemma is not its exact form, but simply that it is $O(\xi)$ except when ϕ_i is close to ϕ_j or $\phi_j+\pi$.

Lemma 3.6. *Let $\xi=s^8-1$ and let $\theta\in[0,2\pi]$ be chosen uniformly at random. Then*

$$\mathbb{P}\left[\frac{a_i\sin(\theta+\phi_i)}{a_j\sin(\theta+\phi_j)}\in B^2B^{-2}\right]\leqslant\frac{4\xi\lambda}{\pi(\lambda-1)}(4+|\cot(\phi_i-\phi_j)|).$$

Proof. Since θ is fixed, by the change of variables $\theta+\phi_j\mapsto\theta$ and $\theta+\phi_i\mapsto\theta+\phi_i-\phi_j$ it follows that the distribution of $\frac{a_i\sin(\theta+\phi_i)}{a_j\sin(\theta+\phi_j)}$ is the same as the distribution of

$$\frac{a_i\sin(\theta+\phi_i-\phi_j)}{a_j\sin(\theta)}=\frac{a_i}{a_j}\Big(\cos(\phi_i-\phi_j)+\sin(\phi_i-\phi_j)\cot\theta\Big).$$

Therefore, we are interested in the probability that $\cot\theta\in\frac{a_j}{a_i\sin(\phi_i-\phi_j)}B^2B^{-2}-\cot(\phi_i-\phi_j)$.

Let $t = |\frac{a_j s^{-4}}{a_i \sin(\phi_i - \phi_j)}|$. Then

$$\frac{a_j}{a_i \sin(\phi_i - \phi_j)} B^2 B^{-2} = \bigcup_m \Big([t\lambda^m, t\lambda^m s^8] \cup [-t\lambda^m s^8, -t\lambda^m] \Big).$$

By Lemma 3.5, we get the bound

$$\mathbb{P}\left[\cot\theta \in \left[t\lambda^m - \cot(\phi_i - \phi_j), t\lambda^m s^8 - \cot(\phi_i - \phi_j)\right]\right] \leqslant \frac{\xi t \lambda^m}{\pi}$$

for all m, and in addition if $t\lambda^m \geqslant 2\cot(\phi_i - \phi_j)$, then since $s^{-2} < 1$ we have an upper bound of

$$\frac{\xi t \lambda^m}{\pi(1 + t^2\lambda^{2m}/4)} \leqslant \frac{4\xi}{\pi t \lambda^m}.$$

If $\cot(\phi_i - \phi_j) \geqslant 0$, then the probability that $\cot\theta + \cot(\phi_i - \phi_j)$ lies in the positive part of $\frac{a_j}{a_i \sin(\phi_i - \phi_j)} B^2 B^{-2}$ is therefore at most ξ/π multiplied by the sum

$$\sum_{t\lambda^m \leqslant S} t\lambda^m + 4 \sum_{t\lambda^m > S} \frac{1}{t\lambda^m},$$

where $S = \max\{2\cot(\phi_i - \phi_j), 1\}$. Let $m_0 = \lfloor \frac{\ln(S/t)}{\ln \lambda} \rfloor$, so that $t\lambda^{m_0} \leqslant S$. By the formula for the sum of a geometric progression, and recalling that $\lambda \geqslant \frac{3}{2}$, the first sum can be estimated by

$$\sum_{t\lambda^m \leqslant S} t\lambda^m = t \sum_{m=-\infty}^{m_0} \left(\frac{1}{\lambda}\right)^m = t \frac{(1/\lambda)^{-m_0}}{1 - \lambda^{-1}} \leqslant S \frac{\lambda}{\lambda - 1},$$

and similarly the second sum is at most $S^{-1} \frac{\lambda}{\lambda - 1}$. Therefore, the total is at most

$$\frac{(S + 4S^{-1})\lambda}{\lambda - 1} \leqslant \frac{(5 + 2\cot(\phi_i - \phi_j))\lambda}{\lambda - 1}.$$

Therefore, we obtain an answer of at most $\xi\lambda(5 + 2\cot(\phi_i - \phi_j))/\pi(\lambda - 1)$.

If $\cot(\phi_i - \phi_j) < 0$, then in the same way we get that the probability that $\cot\theta \in \left[t\lambda^m - \cot(\phi_i - \phi_j), t\lambda^m s^8 - \cot(\phi_i - \phi_j)\right]$ is at most $\xi t\lambda^m/\pi$ for all m but now it is also at most $\xi/\pi t\lambda^m$ for all m. Using the first bound when $t\lambda^m \leqslant 1$ and the second when $t\lambda^m > 1$, we obtain an upper bound of at most

$$\frac{2\xi\lambda}{\pi(\lambda - 1)}.$$

Considering the negative part of B as well and combining these two estimates, we obtain the result stated. □

Let us recall that in Subsection 2.1 we defined $s = \lambda^{1/2k}$, where $\lambda > 1$ and $k \in \mathbb{N}$ large enough. We have further, for convenience, defined parameters $\tau = 1 - s^{-2}$ and $\xi = s^8 - 1$.

Corollary 3.7. *If* $\tau \leqslant 10^{-4}$ *and* $\lambda \geqslant 3/2$*, then in every 2-dimensional subspace of* L_2^n *there is a vector* y *such that*

$$\mathbb{P}^y_{i,j}[y_i/y_j \notin B^2B^{-2}] \geqslant 1/8.$$

Proof. Let a typical unit vector y in the subspace have ith coordinate $y_i = a_i \sin(\theta + \phi_i)$. We will bound the desired probability from below by adding an additional constraint. We will consider the probability that $\frac{y_i}{y_j} \notin B^2B^{-2}$ and $|\cot(\phi_i - \phi_j)| \leqslant 2$, which can be found by calculating the expected probability that $|\cot(\phi_i - \phi_j)| \leqslant 2$ and subtracting from it the probability of the event $\{\frac{y_i}{y_j} \in B^2B^{-2}$ and $|\cot(\phi_i - \phi_j)| \leqslant 2\}$.

We have for each i, j that

$$\begin{aligned}
\mathbb{E}_\theta \sin^2(\theta+\phi_i)\sin^2(\theta+\phi_j) &= \frac{1}{4}\mathbb{E}_\theta\Big(\cos(\phi_i-\phi_j) - \cos(2\theta+\phi_i+\phi_j)\Big)^2 \\
&= \frac{1}{4}\Big(\cos^2(\phi_i-\phi_j) + \mathbb{E}_\theta\cos^2(2\theta+\phi_i+\phi_j)\Big) \\
&= \frac{1}{4}\Big(\cos^2(\phi_i-\phi_j) + \frac{1}{2}\Big) \geqslant \frac{1}{8}.
\end{aligned}$$

Here $\mathbb{E}_\theta$ is just the usual average over $\theta \in [0, 2\pi)$. Recall that y is such that $\|y\| = 1$ and $y_i = a_i \sin(\theta + \phi_i)$ for some such θ. For any event Q that depends on two coordinates i, j, we get

$$\begin{aligned}
\mathbb{E}_y\mathbb{P}^y_{i,j}[Q] &= \mathbb{E}_\theta\mathbb{E}_{i,j}a_i^2a_j^2\sin^2(\theta+\phi_i)\sin^2(\theta+\phi_j)\mathbb{1}_Q(i,j) \\
&\geqslant \frac{1}{8}\,\mathbb{E}_{i,j}a_i^2a_j^2\mathbb{1}_Q(i,j) \\
&= \frac{\|a\|^4}{8}\mathbb{E}^a_{i,j}\mathbb{1}_Q(i,j) \\
&= \frac{1}{2}\mathbb{P}^a_{i,j}[Q],
\end{aligned}$$

where we used that $\|a\|^2 = 2$.

It is easy to check the identity $\cot^2\alpha = \frac{2\cos^2\alpha}{1-\cos(2\alpha)}$, so if $\cos(2(\phi_i - \phi_j)) \leqslant 1/2$, then $|\cot(\phi_i - \phi_j)| \leqslant 2$. Therefore,

$$\begin{aligned}
\mathbb{E}_y\mathbb{P}^y_{i,j}[|\cot(\phi_i-\phi_j)| \leqslant 2] &\geqslant \mathbb{E}_y\mathbb{P}^y_{i,j}[\cos(2(\phi_i-\phi_j)) \leqslant 1/2] \\
&\geqslant \frac{1}{2}\,\mathbb{P}^a_{i,j}[\cos(2(\phi_i-\phi_j)) \leqslant 1/2] \;\geqslant\; \frac{1}{6},
\end{aligned}$$

where the last inequality follows from Lemma 3.4.

Now, by Lemma 3.6, if y is a random such vector, then for each i, j the probability that $y_i/y_j \in B^2B^{-2}$ is at most $\frac{4\xi\lambda}{\pi(\lambda-1)}\Big(4 + |\cot(\phi_i - \phi_j)|\Big) \leqslant 4\xi\Big(4 + |\cot(\phi_i - \phi_j)|\Big)$, where the last inequality uses the fact that $\lambda \geqslant 3/2$, which implies that $\lambda/\pi(\lambda - 1) \leqslant 1$. Note also that since $y_i^2 \leqslant a_i^2$ for each i, and $\mathbb{E}_i y_i^2 = \frac{1}{2}\mathbb{E}_i a_i^2$, we have that $\mathbb{P}^y_{i,j}[Q(i,j)] \leqslant 4\mathbb{P}^a_{i,j}[Q(i,j)]$ for every event $Q(i,j)$ that depends on two coordinates i,j. It follows that

$$\begin{aligned}\mathbb{E}_y\mathbb{P}^y_{i,j}[|\cot(\phi_i-\phi_j)|\leqslant 2 \text{ and } y_i/y_j\in B^2B^{-2}]&\\ &\leqslant 4\mathbb{E}_y\mathbb{P}^a_{i,j}[|\cot(\phi_i-\phi_j)|\leqslant 2 \text{ and } y_i/y_j\in B^2B^{-2}]\\ &\leqslant 16\xi(4+2)\\ &=96\xi.\end{aligned}$$

Together with the estimate in the previous paragraph, this implies that

$$\mathbb{P}^y_{i,j}\big[|\cot(\phi_i-\phi_j)|\leqslant 2 \text{ and } y_i/y_j\notin B\big]\geqslant 1/6-96\xi.$$

It is straightforward to check that our assumption that $\tau\leqslant 10^{-4}$ implies that this is at least $1/8$, and the result follows. □

Corollary 3.8. *Provided that $2\lambda^2\beta+6\epsilon<1/8$, every great circle contains a point that does not belong to Γ_ϵ.*

Proof. The previous corollay, applied to the subspace whose unit sphere is the great circle, gives us a point y such that $\mathbb{P}^y_{i,j}[y_i/y_j\notin B^2B^{-2}]\geqslant 1/8$. Since the event in square brackets is invariant under positive scalar multiples, we may assume that y is a unit vector and thus that it belongs to the great circle.

We showed in Corollary 3.3 that if $z\in\Gamma_\epsilon$, then $\mathbb{P}^z_{i,j}[z_i/z_j\in B^2B^{-2}]\geqslant 1-2\lambda^2\beta-6\epsilon$, where β is a parameter used in (2.4) to define the set Γ. Hence, if $2\lambda^2\beta+6\epsilon<1/8$, this implies that $y\notin\Gamma_\epsilon$, which finishes the proof. □

4. Almost every point has an "atypical" image

In this section we want to show that there exists a subspace X of linear dimension such that $X\cap S^{n-1}\subset\Delta^c=\{x\in S^{n-1}:\ \mathbb{P}^x_i[x_i\in A]\geqslant 1-\beta\}$, so that $\psi(X\cap S^{n-1})\subset\psi(\Delta^c)\subset\Gamma$. Indeed, we shall show that for an appropriate constant $\alpha>0$, almost all subspaces of dimension at most αn have this property. To this end, it will be sufficient to show that Δ has exponentially small measure. Note that

$$\mathbb{P}[x\in\Delta]=\mathbb{P}[\mathbb{P}^x_i[x_i\in A]<1-\beta]=\mathbb{P}[\mathbb{P}^x_i[x_i\in B]\geqslant\beta],$$

where B is, as before, the set

$$\bigcup_m(B_m\cup(-B_m)),$$

and $B_m=[s^{2mk-1},s^{2mk+1}]$ for each integer m. Let $\eta=s-1>0$ and as before let $\lambda=s^{2k}$. Then $B_m=[(1+\eta)^{-1}\lambda^m,(1+\eta)\lambda^m]$.

Let us say that a positive real number t is an *η-approximate power of* λ if there exists an integer m such that

$$(1+\eta)^{-1}\lambda^m\leqslant t\leqslant(1+\eta)\lambda^m.$$

For $\gamma \in [0,1]$ and $\xi \geqslant 0$ define Δ_γ^ξ by

$$\Delta_\gamma^\xi = \Big\{ x \in \mathbb{R}^n : \mathbb{P}_i^x[|x_i| \text{ is a } \xi\text{-approximate power of } \lambda] \geqslant \gamma \Big\}. \tag{4.1}$$

As mentioned before, we shall end up taking $\lambda = 4$. For this reason, although Δ_γ^ξ depends on λ, we suppress this dependence in the notation. We shall be particularly interested in the set Δ_β^η, which, when restricted to S^{n-1}, is equal to the set Δ defined in (2.5).

However, we shall also be interested in the set Δ_β^0, which we shall write simply as Δ_β. That is,

$$\Delta_\beta = \Big\{ x \in \mathbb{R}^n : \mathbb{P}_i^x[|x_i| \text{ is a power of } \lambda] \geqslant \beta \Big\}. \tag{4.2}$$

Lemma 4.1. *If $y \in \Delta_\beta^\eta$ then there exists $x \in \Delta_{\beta(1-2\eta)}$ such that $\|x-y\| \leqslant \eta\|y\|$.*

Proof. We are given that $\mathbb{P}_i^y[|y_i|$ is an η-approximate power of $\lambda] \geqslant \beta$. Let J be the set of all i such that $|y_i|$ is an η-approximate power of λ. For each $i \in J$ let $|x_i|$ be the nearest power of λ to $|y_i|$ and let x_i have the same sign as y_i. For each $i \notin J$ let $x_i = y_i$. Then $|x_i - y_i| \leqslant \eta|y_i|$ for $i \in J$, so, writing P_J for the coordinate projection to J, we have that

$$\|x-y\|^2 = \frac{1}{n}\sum_{i\in J}|x_i-y_i|^2 \leqslant \eta^2\frac{1}{n}\sum_{i\in J}|y_i|^2 = \eta^2\|P_J y\|^2 \leqslant \eta^2\|y\|^2.$$

We now need a lower bound for $\|P_J x\|^2/\|x\|^2$. We know that $\|P_J y\|^2 \geqslant \beta\|y\|^2$, and also that $\|P_J x\|^2 - \|P_J y\|^2 = \|x\|^2 - \|y\|^2$. We also have for each $i \in J$ that

$$(1+\eta)^{-2}y_i^2 \leqslant x_i^2 \leqslant (1+\eta)^2 y_i^2,$$

which implies that

$$(1+\eta)^{-2}\|P_J y\|^2 \leqslant \|P_J x\|^2 \leqslant (1+\eta)^2\|P_J y\|^2.$$

Therefore,

$$\frac{\|P_J x\|^2}{\|x\|^2} = \frac{\zeta\|P_J y\|^2}{\|y-P_J y\|^2 + \zeta\|P_J y\|^2}$$

for some $\zeta \in [(1+\eta)^{-2}, (1+\eta)^2]$. The right-hand side is minimized when $\zeta = (1+\eta)^{-2}$, and then it is at least $\zeta\beta \geqslant (1-2\eta)\beta$, which finishes the proof of the lemma. □

Our next aim is to prove an upper bound for the volume of the η-expansion of $\Delta_{\beta(1-2\eta)}$, which by the above lemma contains Δ_β^η. We shall do this in a series of simple steps.

Corollary 4.2. *Let $\lambda, C > 1$ be real numbers and let m be a positive integer. Then the number of positive integer sequences $(a_1,\dots,a_m)$ such that $\lambda^{a_1}+\cdots+\lambda^{a_m} \leqslant Cm$ is at most $(e\log_\lambda C)^m$.*

Proof. If $m^{-1}(\lambda^{a_1}+\cdots+\lambda^{a_m}) \leqslant C$, then by Jensen's inequality $m^{-1}(a_1+\cdots+a_m) \leqslant \log_\lambda C$, and hence $a_1+\cdots+a_m \leqslant am$, where $a = \log_\lambda C$. The number of sequences with this property is $\binom{\lfloor am\rfloor}{m} \leqslant (ea)^m$, so the result follows. □

Let $\eta > 0$ and consider a set $K \subset \mathbb{R}^n$. We say that a subset $N \subset K$ is an η-net of K if every point in K is within distance η of some point of N. We now find an upper bound on the cardinality of an η-net of a special family of sequences in $\mathbb{R}^m$ described below.

Corollary 4.3. *Let $\lambda > 1$ be a real number, let m be a positive integer, and let $\eta > 0$. Let Ω be the set of all sequences $(x_1, \dots, x_m)$ such that $x_1^2 + \dots + x_m^2 \leqslant C^2 m$ and each $|x_i|$ is a power of λ. Then there is an η-net of Ω of cardinality at most $(2e \log_\lambda(\lambda^2 C/\eta))^m$.*

Proof. Let $x \in \Omega$. For each i such that $|x_i| \leqslant \eta/\lambda$, replace x_i by $\lambda^{-t}\operatorname{sign}(x_i)$, where t is chosen in such a way that $\eta/\lambda \leqslant \lambda^{-t} < \eta$, and let the resulting vector be y. Then $|x_i - y_i| \leqslant \eta$ for every i, so $\|x - y\| \leqslant \eta$. Now let Ω' consist of all vectors $x \in \Omega$ such that each $|x_i|$ is equal to λ^{a_i} for some integer a_i with $a_i \geqslant -t$. We have just shown that Ω' is an η-net of Ω.

The number of points in Ω' with positive coordinates is equal to the number of integer sequences $(a_1, \dots, a_m)$ such that each a_i is at least $-t$ and $\lambda^{2a_1} + \dots + \lambda^{2a_m} \leqslant C^2 m$. Rescaling, we see that is the number of positive-integer sequences $(a_1, \dots, a_m)$ such that $\lambda^{2a_1} + \dots + \lambda^{2a_m} \leqslant \lambda^{2(t+1)} C^2 m$, which by Corollary 4.2 is at most $(e(t + 1 + \log_\lambda C))^m$. Since there are 2^m possible choices of signs, the size of Ω' is at most $(2e(t + 1 + \log_\lambda C))^m$. Noting that $t \leqslant \log_\lambda(\lambda/\eta) = 1 + \log_\lambda(1/\eta)$, we obtain the result. □

The important thing about the bound above is that the number we raise to the power m depends logarithmically on η. This shows that an η-net of Ω is *much* smaller than an η-net of the full sphere of radius C.

We shall need a lemma concerning the sizes of nets of unit balls. It is standard, but the version we give is less commonly used, so for convenience we include a proof. (The argument is essentially due to Rogers [Rog57].)

Lemma 4.4. *Let X be an n-dimensional normed space with unit ball B_X and let $\delta > 0$. If n is sufficiently large, then X contains a δ-net of B_X of cardinality at most $2en\log(n)(1 + \frac{1}{\delta})^n$.*

Proof. Let $\rho > 0$ be a small real number to be chosen later. (It will in fact depend on n.) Then a standard volume estimate shows that there is an ρ-net of B_X of size at most $(3/\rho)^n$. We shall now cover every point of this net with a union of balls of radius $\delta - \rho$ in order to obtain our δ-net, and then we will optimize over ρ.

To do this, let $\zeta = \delta - \rho$ and pick points $x_1, \dots, x_N$ uniformly at random from $(1 + \zeta)B_X$. If y is a point in the ρ-net, then the probability that y is not within any of the balls of radius ζ about the x_i is $(1 - (\frac{\zeta}{1+\zeta})^n)^N \leqslant \exp(-N(\frac{\zeta}{1+\zeta})^n)$. Therefore, we are done as long as

$$\left(\frac{3}{\rho}\right)^n \exp\left(-N\left(\frac{\zeta}{1+\zeta}\right)^n\right) < 1,$$

which is satisfied if $N > n\log(\frac{3}{\rho})(1 + \frac{1}{\delta - \rho})^n$.

It can be checked that $1 + \frac{1}{\delta-\rho} = (1 + \frac{1}{n})(1 + \frac{1}{\delta})$ when $\rho = \delta(\frac{\delta+1}{n+\delta+1})$. For this value of ρ and for n is sufficiently large, we have that $\log(\frac{3}{\rho}) < \log(\frac{3n}{\delta})$, which is at most $\frac{3}{2}\log n$. We also have that $(1 + \frac{1}{n})^n < \frac{4e}{3}$ when n is sufficiently large, and putting these estimates together we find that we can take N to be $2en\log n(1 + \frac{1}{\delta})^n$, as claimed. □

Next, we need a simple technical lemma about the largest proportion of the unit sphere of L_2^n that can be covered by a ball of radius δ.

Lemma 4.5. *Let $B_\delta(x)$ be a closed ball of radius δ about a point x in L_2^n. If n is sufficiently large, then the probability that a random point of the unit sphere of L_2^n lies in $B_\delta(x)$ is at most $2\delta^n$.*

Proof. The intersection of $B_\delta(x)$ with the unit sphere is a spherical cap, and the measure of the spherical cap is maximized when the centre x of $B_\delta(x)$ is a vector of norm $\sqrt{1-\delta^2}$.

Define C to be the set of all y such that $\|y\| \leqslant 1$ and $\left\|x - \frac{y}{\|y\|}\right\| \leqslant \delta$. This is a convex hull of the spherical cap and the origin, and the proportion of its volume to the volume of the entire unit ball, is equal to the probability we are trying to estimate.

We are going to show that $B_\delta(x)$ contains the set $C \setminus (1-2\delta^2)C$. Indeed, we claim now that $B_\delta(x)$ contains all points y such that $\left\|x - \frac{y}{\|y\|}\right\| \leqslant \delta$ and $1 \geqslant \|y\| \geqslant 1-2\delta^2$. By the convexity of $B_\delta(x)$ it is sufficient to prove this when $\|y\| = 1-2\delta^2$. The first assumption on y implies that

$$\|x\|^2 - \frac{2\langle x, y\rangle}{n\|y\|} + 1 \leqslant \delta^2,$$

and therefore, since $\|x\|^2 = 1-\delta^2$, that

$$\langle x, y\rangle \geqslant (1-\delta^2)n\|y\|.$$

This implies that

$$\begin{aligned}\|x-y\|^2 &= \|x\|^2 + \|y\|^2 - \frac{2\langle x, y\rangle}{n} \\ &\leqslant 1-\delta^2 + (1-2\delta^2)^2 - 2(1-\delta^2)(1-2\delta^2) \\ &= \delta^2,\end{aligned}$$

which proves the claim.

We have therefore shown that $B_\delta(x)$ contains the set $C \setminus (1-2\delta^2)C$. Now, since $(1-2\delta^2)C$ has volume $(1-2\delta^2)^n$ times that of C, if n is sufficiently large, then $B_\delta(x)$ contains at least half of C. The result follows, since the volume of $B_\delta(x)$ is δ^n times that of the unit sphere of L_2^n. $\square$

Now let y be a vector in L_2^n supported on $J \subset \{1, \ldots, n\}$ of cardinality m and satisfying the inequality $\|y\|^2 \geqslant \beta(1-2\eta)$. Again let P_J be the coordinate projection to the set J and define

$$V_y = \{x \in S^{n-1} : P_J x = y\}.$$

We will next obtain an upper bound for the spherical volume of $(V_y)_\epsilon$, which is the ϵ-expansion of V_y.

Lemma 4.6. *Let $\delta > \eta > 0$. Then when n is sufficiently large, the probability that a random unit vector belongs to $(V_y)_\eta$ is at most $4e\delta^n n \log n \left(1 + \frac{\sqrt{1-\beta(1-2\eta)}}{\delta-\eta}\right)^{n-m}$.*

Proof. If we cover V_y by N balls of radius $\delta - \eta$, then the balls of radius δ with the same centres cover $(V_y)_\eta$, so by Lemma 4.5 the probability that a random unit vector lies in $(V_y)_\eta$ is at most $2N\delta^n$. But V_y is an $(n-m)$-dimensional sphere of radius at most $\sqrt{1-\beta(1-2\eta)}$, hence Lemma 4.4 implies that it can be covered by at most $N = 2en\log n\Big(1+\frac{\sqrt{1-\beta(1-2\eta)}}{\delta-\eta}\Big)^{n-m}$ balls of radius $\delta - \eta$. This implies the result. □

Theorem 4.7. *There is a choice of parameters λ, δ, η and β such that the probability that a random unit vector belongs to $(\Delta_{\beta(1-2\eta)})_\eta$ is exponentially small.*

Before we prove the theorem we remark that if we choose $\lambda = 4$, $\delta = 10^{-6}, \eta = 10^{-12}$ and $\beta = \frac{1}{257}$, then the probability in question is at most

$$4en^2 \log n(0.9982)^n.$$

Proof. Lemma 4.1 tells us that the set $\Delta_{\beta(1-2\eta)}$ is an η-net of the set of unit vectors in Δ^η_β. Moreover, the same is true if we restrict to vectors of norm at most $1+\eta \leqslant 2$. For each $x \in \Delta_{\beta(1-2\eta)}$ there is a set $J \subset \{1,2,\dots,n\}$ such that $\mu_x(J) \geqslant \beta(1-2\eta)$ and $|x_i|$ is a power of λ for every $i \in J$. If $|J| = m$, then Corollary 4.3 implies that there is an η-net of size at most $\left(2e\log_\lambda\left(\sqrt{\frac{2n}{m}}\frac{\lambda^2}{\eta}\right)\right)^m$ of the set of vectors y such that $|y_i|$ is a power of λ for every $i \in J$, $y_i = 0$ for $i \notin J$, and $\sum_i y_i^2 \leqslant 2n$.

Every unit vector in $\Delta_{\beta(1-2\eta)}$ lies in V_y for some such J and y. Therefore, summing over all J and all y in an η-net for each J and applying Lemma 4.6, we find that the probability that a random unit vector belongs to $(\Delta_{\beta(1-2\eta)})_\eta$ is at most

$$\sum_{m=1}^{n}\binom{n}{m}\Big(2e\log_\lambda\Big(\sqrt{\frac{2n}{m}}\frac{\lambda^2}{\eta}\Big)\Big)^m 4e\delta^n n\log n\Big(1+\frac{\sqrt{1-\beta(1-2\eta)}}{\delta-\eta}\Big)^{n-m}.$$

Now let us set $\lambda = 4$. Using the upper bound $\binom{n}{m} \leqslant (en/m)^m$ and setting $\theta = m/n$, we can bound the previous expression above by

$$4en\log n\sum_{m=1}^{n}\Big(2e^2\delta\frac{1}{\theta}\log_4\Big(\frac{16\sqrt{2}}{\eta\sqrt{\theta}}\Big)\Big)^{\theta n}\Big(\delta+\frac{\delta\sqrt{1-\beta(1-2\eta)}}{\delta-\eta}\Big)^{(1-\theta)n}.$$

To prove that this is exponentially small, it is sufficient to show that

$$\Big(\frac{2e^2}{\log 4}\frac{\delta}{\theta}\log\Big(\frac{16\sqrt{2}}{\eta\sqrt{\theta}}\Big)\Big)^\theta\Big(\delta+\frac{\delta\sqrt{1-\beta(1-2\eta)}}{\delta-\eta}\Big)^{1-\theta} \tag{4.3}$$

is bounded above by a constant less than 1 as θ varies. First of all, note that in §3 we obtain the following bounds on parameters: $\eta \leqslant 10^{-5}$ (since $\eta = s-1$ and we need $\tau = 1-s^{-2} \leqslant 10^{-4}$), $\epsilon \leqslant \tau^2$ and $\beta < \frac{1}{32}\left(\frac{1}{8}-6\epsilon\right)$. We need moreover, that $\delta > \eta$. Let us note that the above expression is a decreasing function of β and since $\epsilon \leqslant 10^{-8}$ we can take $\beta = \frac{1}{257}$.

To begin with, we shall show that there exist constants for which the result holds and then we shall choose some particular values to obtain an upper bound for the maximum. Let us consider the condition $\delta + \frac{\delta\sqrt{1-\beta(1-2\eta)}}{\delta-\eta} < 1$. For $\eta = c\delta$ this becomes

$$\delta + \frac{\sqrt{1-\beta(1-2\eta)}}{1-c} < 1,$$

and we can choose c such that $\frac{\sqrt{1-\beta(1-2\eta)}}{1-c}$ is less than $1-\beta/4$, and then if $\delta \leqslant \beta/8$ the inequality holds. For the first part of expression (4.3), if we assume that $\theta \geqslant \sqrt{\delta}$ (and again take $\eta = c\delta$) we have that

$$\Big(\frac{2e^2}{\log 4}\frac{\delta}{\theta}\log\Big(\frac{16\sqrt{2}}{\eta\sqrt{\theta}}\Big)\Big)^\theta \leqslant \left(2e^2\sqrt{\delta}\Big(\log(16\sqrt{2}) + \log\frac{1}{c\delta} + \frac{1}{4}\log\frac{1}{\delta}\Big)\right)^\theta$$
$$\leqslant 2e^2\sqrt{\delta}\left(C_1 + \frac{5}{4}\log\frac{1}{\delta}\right)^\theta,$$

where C_1 is an absolute constant. Hence, we can choose δ small enough such that this expression is at most, say, $\frac{9}{10}$.

If $\theta < \sqrt{\delta}$, then we need to consider the whole expression (4.3). To begin with note that

$$\Big(\frac{2e^2}{\log 4}\frac{\delta}{\theta}\log\Big(\frac{16\sqrt{2}}{\eta\sqrt{\theta}}\Big)\Big)^\theta \leqslant \Big(2e^2\frac{\delta}{\theta}(C_1 + \frac{5}{2}\log\frac{1}{\theta})\Big)^\theta \leqslant \left(\frac{C_2\delta\log\frac{1}{\theta}}{\theta}\right)^\theta$$
$$\leqslant (C_2\delta)^\theta\left(\frac{1}{\theta^2}\right)^\theta \leqslant \left(\frac{1}{\theta^2}\right)^\theta,$$

for $\delta < 1/C_2$. Moreover, we have that $\log\left((\frac{1}{\theta})^{2\theta}\right) = 2\theta\log\frac{1}{\theta} \leqslant 2\sqrt{\theta} \leqslant 2\delta^{1/4}$. Therefore we can estimate $\left(\frac{1}{\theta^2}\right)^\theta$ by $1+4\delta^{1/4}$. Recalling that the right hand side part in (4.3) is at most $(1-\frac{\beta}{8})^{1-\theta} \leqslant (1-\frac{\beta}{8})^{1-\sqrt{\delta}}$, we deduce that the expression (4.3) is less than 1 if we choose δ such that $(1+4\delta^{1/4})(1-\frac{\beta}{8})^{1-\sqrt{\delta}} < 1$. Finally we choose the smallest δ, so that it fulfils all the inequalities and hence the result follows.

One can check that if we choose $\eta = 10^{-12}, \delta = 10^{-6}$ and $\beta = \frac{1}{257}$, then the maximum is at most 0.9982. Therefore, the desired probability is at most $4en^2\log n(0.9982)^n$. □

This gives us the information we need about the set $\Delta = \Delta_\beta^\eta$ defined in (2.5).

Corollary 4.8. *There exists $\alpha > 0$ such that if n is sufficiently large, then the probability that a random subspace X of dimension at most αn contains a vector $x \in \Delta$ is exponentially small.*

Proof. Let $\sigma < 1$ be such that the probability that a random unit vector belongs to $(\Delta_{\beta(1-2\eta)})_\eta$ is at most σ^n when n is sufficiently large. Now choose $\alpha > 0$ such that $(1+\frac{1}{\eta})^\alpha < \sigma^{-1}$. Then for sufficiently large n, the unit sphere of any subspace of dimension at most αn has an η-net of size τ^{-n} for some τ with $\tau^{-1} < \sigma^{-1}$. If we take such a net and rotate it randomly, then the probability that any element of the net lands within η of $\Delta_{\beta(1-2\eta)}$ is exponentially small.

By Lemma 4.1 we have that $\Delta_\beta^\eta \subset (\Delta_{\beta(1-2\eta)})_\eta$ and hence it follows that the probability that a random subspace intersects Δ is exponentially small. □

Remark 4.9. For the particular choice of constants made in the proof of Theorem 4.7, the condition we obtain in Corollary 4.8 is

$$0.9982(4en^2 \log n)^{1/n} < \sigma.$$

The left hand side is a decreasing function of n with limit 0.9982, so for n large enough the left hand side is less than 0.999. We then have that

$$\alpha < \frac{-\ln(0.999)}{\ln(1+10^{10})} \approx 4.345 \times 10^{-5}.$$

Hence, for n large enough we can take $\alpha = 4.3 \times 10^{-5}$ in the construction.

References

[Art02] S. Artstein. Proportional concentration phenomena on the sphere. *Israel Journal of Mathematics*, 132(1):337–358, Dec 2002. `doi:10.1007/BF02784520`.

[Dvo61] A. Dvoretzky. Some results on convex bodies and Banach spaces. In *Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960)*, pages 123–160. Academic Press, 1961.

[GW21] W. T. Gowers and K. Wyczesany. A counterexample to a strengthening of a question of Milman. 2021. `arXiv:2110.03023`.

[Mil71] V. Milman. New proof of the theorem of A. Dvoretzky on intersections of convex bodies. *Functional Analysis and Its Applications*, 5(4):288–295, 1971.

[Rog57] C. A. Rogers. A note on coverings. *Mathematika*, 4(1):1–6, 1957. `doi:10.1112/S0025579300001030`.