CHAPTER 1

The Ramsey–Dvoretzky–Milman phenomenon

1.1. Finite oscillation stability

Uniform — rather than metric — spaces provide the most natural setting for studying the Ramsey–Dvoretzky–Milman phenomenon.

Definition 1.1.1. Recall that a *uniform space* is a pair $(X,\mathcal{U})$, consisting of a set X and a *uniform structure*, $\mathcal{U}$, on X, that is, a collection of subsets of $X \times X$ (binary relations on X), called *entourages of the diagonal*, satisfying the following properties.

(1) The family $\mathcal{U}$ is closed under finite intersections and supersets (if $V \in \mathcal{U}$ and $V \subseteq U \subseteq X \times X$, then $U \in \mathcal{U}$).
(2) Every $V \in \mathcal{U}$ contains the diagonal $\Delta = \{(x, x) : x \in X\}$.
(3) If $V \in \mathcal{U}$, then $V^{-1} \equiv \{(x, y) : (y, x) \in V\}$ is in $\mathcal{U}$.
(4) If $V \in \mathcal{U}$, there exists a $U \in \mathcal{U}$ such that $U \circ U \equiv \{(x, z) : \exists y \in X,\ (x, y) \in U,\ (y, z) \in U\}$ is a subset of V.

Definition 1.1.2. A subfamily $\mathcal{B} \subseteq \mathcal{U}$ is called a *basis* of the uniformity $\mathcal{U}$ if for every $U, V \in \mathcal{B}$ there is a $C \in \mathcal{B}$ with $C \subseteq A \cap B$, and every entourage $V \in \mathcal{U}$ contains, as a subset, an element $U \in \mathcal{B}$.

A family $\mathcal{B}$ of subsets of a set X serves as a basis for a uniform structure if and only if it satisfies the following conditions:

(1) For every $V, U \in \mathcal{B}$ there is a $C \in \mathcal{B}$ with $C \subseteq A \cap B$ (that is, $\mathcal{B}$ forms a *prefilter*).
(2) Every $V \in \mathcal{B}$ contains the diagonal Δ.
(3) If $V \in \mathcal{B}$, then for some $U \in \mathcal{B}$, $U \subseteq V^{-1}$.
(4) If $V \in \mathcal{B}$, there exists a $U \in \mathcal{B}$ such that $U \circ U \subseteq V$.

For every element V of the uniform structure $\mathcal{U}$ and each $x \in X$, denote $V[x] = \{y \in X : (x, y) \in V\}$. This is the V-*neighbourhood* of x. The sets $V[x]$, where $V \in \mathcal{U}$, form a neighbourhood basis for x with regard to a certain topology on X, called the topology *determined* by, or *associated* to, $\mathcal{U}$. If $\mathcal{U}$ is a uniform structure determining the topology of a given topological space X, then $\mathcal{U}$ is said to be *compatible*.

If $\mathcal{U}$ is *separated* (as we will usually assume), that is, $\cap\mathcal{U} = \Delta$, then the associated topology is Tychonoff. There is a converse statement as well.

Exercise 1.1.3. Prove that every Tychonoff topological space admits a compatible uniform structure, and that every such structure is necessarily separated. (*Hint:* if f is a real-valued function on a set, then $\{(x, y) : |f(x) - f(y)| < \varepsilon\}$ is an entourage of the diagonal in $X \times X$.)

To every uniform space $(X,\mathcal{U})$ one can associate a *separated replica*, that is, a quotient set of X by the equivalence relation $\cap\mathcal{U}$, equipped with the uniformity consisting of images of entourages from $\mathcal{U}$ under the quotient map. We will usually assume that all our uniform spaces are separated.

EXERCISE 1.1.4. Show that every element of a compatible uniformity $\mathcal{U}$ on a topological space X is a neighbourhood of the diagonal in the product topology on $X \times X$. Deduce that every compact space X admits a *unique* compatible uniformity, consisting of all neighbourhoods of the diagonal.

EXAMPLE 1.1.5. The *additive uniform structure* on a topological vector space E has, as a basis, the set of all entourages of the diagonal of the form

$$\{(x,y) \in E \times E \colon x - y \in V\},$$

where V runs over the neighbourhood basis of zero.

EXAMPLE 1.1.6. The *left uniform structure*, $\mathcal{U}_L(G)$, on a topological group G has as basic entourages of the diagonal the sets

$$V_L = \{(x,y) \colon x^{-1}y \in V\},$$

where V is a neighbourhood of identity.

EXAMPLE 1.1.7. Every metric space (X,d) supports a natural uniform structure, whose basic entourages are of the form

$$\{(x,y) \colon d(x,y) < \varepsilon\},\ \varepsilon > 0.$$

For instance, if d is a left-invariant metric generating the topology of a topological group G, then the corresponding uniform structure is the left uniform structure on G.

Let A be a subset of a uniform space $(X,\mathcal{U})$, and let $V \in \mathcal{U}$ be an entourage of the diagonal. Say that A is *V-small* if $A \times A \subseteq V$, that is, for each $a, b \in A$ one has $(a,b) \in V$.

A *Cauchy filter* on a uniform space X is a filter $\mathscr{F}$ containing a V-small set for every $V \in \mathcal{U}$. A uniform space is complete if all Cauchy filters converge. Two Cauchy filters are *equivalent* if their intersection is again a Cauchy filter. It is now clear how to define the completion of a separated uniform space, along the same lines as it is being done for metric spaces. The uniform structure $\mathcal{U}$ admits a unique extension to the completion $\hat{X}$ of a separated uniform space $X = (X,\mathcal{U})$, the space $\hat{X}$ is separated and contains X as a uniform subspace (which concept is defined in an obvious way).

A uniform structure $\mathcal{U}$ is *totally bounded* if for every $V \in \mathcal{U}$ there is a finite $F \subseteq X$ with $V[F] = X$. Here $V[F]$ is the *V-neighbourhood* of F, given by

$$V[F] = \cup_{x\in F} V[x] = \{y \in X \colon \exists x \in F,\ (x,y) \in V\}.$$

Totally bounded uniform spaces are exactly those whose completions are compact.

A pseudometric d on a uniform space $(X,\mathcal{U})$ is *uniformly continuous* if for every $\varepsilon > 0$ there is a $V \in \mathcal{U}$ with the property $d(x,y) < \varepsilon$ whenever $(x,y) \in V$. Every uniform space admits a family of uniformly continuous bounded pseudometrics that determine the uniform structure, in the following sense: for every $V \in \mathcal{U}$ there is a bounded uniformly continuous pseudometric d such that

$$\{(x,y) \in X \times X \colon d(x,y) < 1\} \subseteq V.$$

We will only consider such families of pseudometrics that are *directed,* that is, for every two pseudometric d_1, d_2 there is a third, d_3, that is greater than any of the two.

A mapping f between uniform spaces $(X,\mathcal{U}_X)$ and $(Y,\mathcal{U}_Y)$ is *uniformly continuous* if for every $V \in \mathcal{U}_Y$ there is a $U \in \mathcal{U}_X$ with $(f \times f)(U) \subseteq V$.

For a more detailed review of uniform spaces, the reader may consult, for instance, [**56**] (or its 1989 edition, revised and completed: Berlin, Heldermann Verlag), and [**27**].

Now comes one of the main concepts of the present theory.

DEFINITION 1.1.8. Let a group G act on a set X. A function f on X, taking values in a uniform space $(Y,\mathcal{U}_Y)$, is called *finitely oscillation stable* if for every finite subset $F \subset X$ and every entourage $V \in \mathcal{U}_Y$ there is a transformation $g \in G$ such that the image $f(gF)$ is V-small:

$$f(gF) \times f(gF) \subseteq V.$$

EXERCISE 1.1.9. Assume a group G acts on a uniform space X uniformly equicontinuously. (That is, for all $V \in \mathcal{U}_X$ there is a $U \in \mathcal{U}_X$ such that if $(x,y) \in U$, then for all $g \in G$ one has $(gx, gy) \in V$.) Let $f \colon X \to Y$ be uniformly continuous. Show that one can replace in the above definition 1.1.8 finite subsets $F \subset X$ with compact ones.

If (Y,d) is a metric space (equipped with the corresponding uniformity), the definition can be conveniently restated. For a function f on a set A with values in a metric space (Y,d_Y), define *oscillation* of f on A as the value

$$\operatorname{osc}(f|A) = \sup_{x,y \in A} d_Y(f(x), f(y)).$$

PROPOSITION 1.1.10. *Let a group G act on a set X. A function f on X, taking values in a metric space (Y,d_Y), is finitely oscillation stable if and only if for every finite subset $F \subset X$ and every $\varepsilon > 0$ there is a transformation $g \in G$ such that*

$$\operatorname{osc}(f|gF) < \varepsilon.$$

In other words, for every $\varepsilon > 0$ the function f is constant to within ε on a suitable translate of F by an element of G. □

The concept is illustrated in Fig. 1.1.

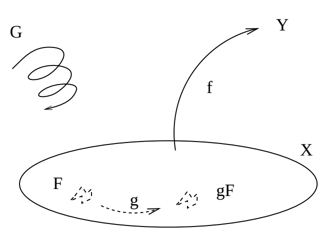

FIGURE 1.1. A finitely oscillation stable function.

Now let G be a group of (not necessarily all) uniform isomorphisms of a uniform space $(X,\mathcal{U}_X)$. We will say that (G,X) is a *uniform G-space.* At this stage, we do not equip G with topology; however, whenever G is a topological group, we will assume that the action of G on X is continuous.

DEFINITION 1.1.11. Say that a uniform G-space X has the *Ramsey–Dvoretzky–Milman property*, or is *finitely oscillation stable,* if every bounded uniformly continuous real-valued function f on X is finitely oscillation stable.

Notice that the above property is dynamical, that is, it depends not only on the uniform space X in question, but also on the selected group G of uniform isomorphisms of X. In many, but not all, important concrete situations, X will be a metric space, and G will be a group of isometries of X.

REMARK 1.1.12. The above concept was introduced (in the case where X is the unit sphere in the Hilbert space and G is the group of unitary operators) by V.Milman [**142, 143**], cf. also [**151**], though the terminology we use is of later origin (cf. [**169**]). According to Milman, the *spectrum* of a function f as above is the set $\gamma(f)$ of all $a \in Y$ such that for every $\varepsilon > 0$ and every finite $F \subset X$ there is a $g \in G$ with $d(f(x), a) < \varepsilon$ whenever $x \in gF$. If $f(X)$ is relatively compact in Y (which is the case that matters), then clearly $\gamma(f) \neq \emptyset$ if and only if f is finitely oscillation stable.

It will be useful, to reformulate the concept in several equivalent forms using the language of uniform topology.

Let $(X,\mathcal{U})$ be a (separated) uniform space. By $\mathcal{C}^*\mathcal{U}$ we will denote the *totally bounded replica* of $\mathcal{U}$, that is, the finest totally bounded uniformity contained in $\mathcal{U}$. Equivalently, $\mathcal{C}^*\mathcal{U}$ is the coarsest uniformity on X with regard to which every bounded uniformly continuous function on X remains uniformly continuous. If $\mathcal{U}$ is separated, then $\mathcal{C}^*\mathcal{U}$ is separated as well.

The completion of the uniform space $(X, \mathcal{C}^*\mathcal{U})$ is a compactification of X, called the *Samuel compactification* or else the *universal uniform compactification* of X. It is denoted by σX. This is the maximal ideal space of the commutative C^*-algebra UCB$\,(X)$ of all bounded uniformly continuous complex-valued functions on X. In other words, elements of σX are multiplicative functionals on UCB$\,(X)$ (morphisms of unital involution algebras), which are automatically continuous, and the set σX is equipped with the weak*-topology.

EXERCISE 1.1.13. Prove that X is everywhere dense in the maximal ideal space of UCB$\,(X)$.

EXAMPLE 1.1.14. If X is a Tychonoff topological space equipped with the finest compatible uniformity, then $\sigma X = \beta X$ is the Stone–Čech compactification of X. In particular, for $X = \mathbb{N}$ equipped with the discrete uniformity (the one containing every entourage of the diagonal $\Delta_{\mathbb{N}}$), one has $\sigma\mathbb{N} = \beta\mathbb{N}$.

We are going to give a convenient description of the totally bounded replica of a given uniformity $\mathcal{U}$ on a set X. If γ is a cover of X, denote by $\tilde{\gamma}$ the corresponding entourage of the diagonal in $X \times X$, that is,

$$\tilde{\gamma} = \cup_{A\in\gamma} A \times A.$$

For an entourage of the diagonal $V \in \mathcal{U}$ and a γ as before, denote by γ_V the V-*thickening* of γ,

$$\gamma_V = \{V[A] \colon A \in \gamma\}.$$

It is easy to verify that $\widetilde{\gamma_V} = V \circ \tilde{\gamma} \circ V^{-1}$.

Lemma 1.1.15. *Let* $(X,\mathcal{U})$ *be a uniform space. A basis for the totally bounded replica of* $\mathcal{U}$ *is given by all entourages of the diagonal of the form* $\widetilde{\gamma_V}$*, where* γ *is a finite cover of* X *and* $V \in \mathcal{U}$*.*

Proof. Given a finite cover γ and an $V \in \mathcal{U}$, choose a bounded uniformly continuous pseudometric d on X with the property

$$(d(x,y) < 1) \implies (x,y) \in V.$$

For each $A \in \gamma$ and every $x \in X$, set $d_A(x) := \inf\{d(x,a)\colon a \in A\}$. The functions d_A are uniformly continuous bounded. Let $x, y \in X$ be such that $|d_A(x) - d_A(y)| < 1$ for all $A \in \gamma$. For some $A' \in \gamma$ one has $x \in A'$, therefore $d_{A'}(x) = 0$, $d_{A'}(y) < 1$ and $y \in V[A']$. Consequently, $(x,y) \in \widetilde{\gamma_V}$. We conclude: $\widetilde{\gamma_V} \in \mathcal{C}^*\mathcal{U}$.

It remains to prove that every $W \in \mathcal{C}^*\mathcal{U}$ contains an entourage of the form $\widetilde{\gamma_V}$. By definition of the totally bounded replica, there exist $n \in \mathbb{N}$, $f \in \mathrm{UCB}\,(X;\ell^\infty(n))$ and $\varepsilon > 0$ such that $(x,y) \in W$ whenever $\|f(x) - f(y)\|_\infty < \varepsilon$. Find a $V \in \mathcal{U}$ with $(x,y) \in V$ implying $\|f(x) - f(y)\|_\infty < \varepsilon/3$, and cover the set $f(X) \subset \ell^\infty(n)$ with a finite family α of sets of diameter $< \varepsilon/3$ each. Let $\gamma = \{f^{-1}(A)\colon A \in \alpha\}$. If $(x,y) \in \widetilde{\gamma_V}$, there are x', y' and an $A \in \alpha$ with $(x,x') \in V$, $f(x'), f(y') \in A$, and $(y,y') \in V$. By the triangle inequality, $\|f(x) - f(y)\|_\infty < \varepsilon$, meaning that $\widetilde{\gamma_V} \subseteq W$. □

Corollary 1.1.16. *If* X *is a metric space, then a basis for the totally bounded replica of the metric uniformity on* X *is given by entourages* $\widetilde{\gamma_\varepsilon}$*,* $\varepsilon > 0$*, where* γ_ε *is the* ε*-thickening of a finite cover* γ*.* □

Example 1.1.17. Let $\mathbb{S}^\infty$ be the unit sphere of the separable Hilbert space ℓ^2. The *norm uniformity*, $\mathcal{U}_{\|\cdot\|}$, on $\mathbb{S}^\infty$ is generated by the usual norm distance, $d_{\|\cdot\|}$. The *weak uniformity*, $\mathcal{U}_w$, is the coarsest uniformity making the restriction to $\mathbb{S}^\infty$ of each continuous linear functional on ℓ^2 uniformly continuous. Equivalently, $\mathcal{U}_w$ is the restriction to the sphere of the additive uniformity on the space ℓ^2 with the weak topology. As is easy to see, $\mathcal{U}_w$ is totally bounded and contained in the norm uniformity. It is well known and easily seen that the uniform completion of the sphere with the weak uniformity is the compact unit ball $\mathbb{B}^\infty$ of ℓ^2 with the weak topology (and the unique compatible uniformity).

The totally bounded replica $\mathcal{C}^*\mathcal{U}_{\|\cdot\|}$ of the norm uniformity is strictly finer than the weak uniformity $\mathcal{U}_w$. To see this, let X be any infinite uniformly discrete subset of $\mathbb{S}^\infty$, for example, the set $X = \{e_1, e_2, \ldots\}$ of all standard basic vectors. Every $\{0,1\}$-valued function f on X extends to a bounded uniformly continuous, indeed 1-Lipschitz, function $\tilde{f}$ on $(\mathbb{S}^\infty, \mathcal{U}_{\|\cdot\|})$, e.g. through the Katětov construction:

$$\mathbb{S}^\infty \ni \xi \mapsto \tilde{f}(\xi) = \inf\{\|\xi - x\| + f(x)\colon x \in X\}.$$

(More generally, every uniformly continuous bounded function can be extended over a uniform space from a subspace, cf. Exercise 8.5.6 in [**56**].) Consequently, the Stone-Čech compactification $\beta\mathbb{N}$ embeds into $\sigma\mathbb{S}^\infty$ as a topological subspace. But as we have seen, the weak completion of $\mathbb{S}^\infty$ is separable metrizable.

For a natural number k, we denote, following a convention usual in combinatorics, $[k] = \{1, 2, \ldots, k\}$.

The following result, although purely technical in appearance, is central to the entire theory, because it shows why the Ramsey-Dvoretzky-Milman property is closely linked both to Ramsey theory and to the existence of fixed points.

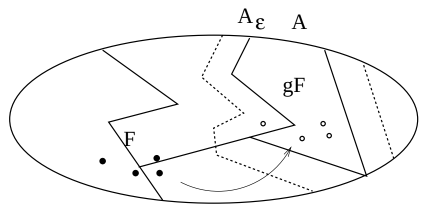

Figure 1.2. Finite oscillation stability (condition (7)).

Theorem 1.1.18. *For a uniform G-space $(X,\mathcal{U})$ the following are equivalent.*

(1) *X has the Ramsey–Milman–Dvoretzky property, that is, every uniformly continuous bounded function $f\colon X \to \mathbb{R}$ is finitely oscillation stable.*
(2) *Let $\mathcal{R}$ be a directed collection of bounded uniformly continuous pseudometrics generating the uniformity $\mathcal{U}$. Then for every $d \in \mathcal{R}$, every real-valued bounded 1-Lipschitz function on (X,d) is finitely oscillation stable.*
(3) *Every bounded uniformly continuous function f from X to each finite-dimensional Euclidean space $\mathbb{R}^n$, $n \in \mathbb{N}$ is finitely oscillation stable.*
(4) *Every uniformly continuous mapping from X to a compact space is finitely oscillation stable.*
(5) *The canonical mapping from X to the Samuel compactification σX is finitely oscillation stable.*
(6) *For every entourage $W \in \mathcal{C}^*\mathcal{U}$ and every finite $F \subseteq X$, there is a $g \in G$ such that gF is W-small: $(gF \times gF) \subseteq W$.*
(7) *For every finite cover γ of X ("colouring of X with finitely many colours"), every $V \in \mathcal{U}$, and every finite $F \subseteq X$ there are a $g \in G$ and an $A \in \gamma$ such that $gF \subseteq V[A]$ (that is, gF is "monochromatic to within V").*
(8) *For every binary cover $\{A,B\}$ of X, every $V \in \mathcal{U}$ and every finite $F \subseteq X$, there is a $g \in G$ such that gF is contained either in $V[A]$ or in $V[B]$.*
(9) *For every entourage $V \in \mathcal{U}$, every $k \in \mathbb{N}$ and every finite $F \subseteq X$, there is a finite $K \subseteq X$ with the property that for every colouring $c\colon K \to [k]$ with k colours, there are $g \in G$ and $i \in [k]$ such that $gF \subseteq K$ and $gF \subseteq V[c^{-1}(i)]$, that is, gF is monochromatic to within V.*
(10) *For every $V \in \mathcal{U}$ and every finite $F \subseteq X$, there is a finite $K \subseteq X$ with the property that for every binary cover $\gamma = \{A,B\}$ of K there is a $g \in G$ such that $gF \subseteq K$ and gF is contained either in $V[A]$ or in $V[B]$.*

If in addition X is a metric space and $\mathcal{U}$ is the metric uniformity, then each of the above is equivalent to any of the following:

(11) *Every bounded 1-Lipschitz function $f\colon X \to \mathbb{R}$ is finitely oscillation stable.*
(12) *For every $n \in \mathbb{N}$, each bounded 1-Lipschitz function from X to $\ell^\infty(n)$ is finitely oscillation stable.*

Proof. (1) $\Longrightarrow$ (2): trivial.

(2) $\Longrightarrow$ (8): for a $V \in \mathcal{U}$, choose a uniformly continuous bounded pseudometric $d \in \mathcal{R}$ such that $(d(x,y) < 1) \implies ((x,y) \in V)$. Let $\{A,B\}$ be a binary cover of X. The distance function d_A from A is a real-valued, bounded, and 1-Lipschitz function on (X,d), and by assumption, for every finite $F \subseteq X$ there is $g \in G$ such that $\operatorname{osc}(d_A|gF) < 1$. Unless $gF \subseteq B$ (in which case we are of course done), the

function d_A assumes value zero at some point of gF, and consequently $gF \subseteq V[A]$.

(8) $\implies$ (10): if X is finite, there is nothing to prove. Suppose X is infinite, and assume that (10) does not hold, that is, there exist a $V \in \mathcal{U}$ and a finite $F \subseteq X$ such that every finite $K \subseteq X$ admits a binary cover $\{A_K, B_K\}$ with the property that, whenever $g \in G$ and $gF \subseteq K$, one has $gF \not\subseteq V[A_K]$ and $gF \not\subseteq V[B_K]$. Choose an ultrafilter ξ on the collection $\mathcal{P}_{\text{fin}}(X)$ of all finite subsets of X with the property that for every $F' \in \mathcal{P}_{\text{fin}}(X)$, the set $\{K \in \mathcal{P}_{\text{fin}}(X) \colon K \supseteq F'\}$ is in ξ. Define $A = \lim_{K\in\xi} A_K$, that is, $x \in A$ if and only if $\{K \in \mathcal{P}_{\text{fin}}(X) \colon x \in A_K\} \in \xi$. Similarly, set $B = \lim_{K\in\xi} B_K$. Then $\{A, B\}$ is a binary cover of X. Let $g \in G$ be arbitrary. For every finite $K \supseteq gF$, there are elements $a_K, b_K \in gF$ such that for every $x \in A_K$, $y \in B_K$ one has $(a_K, y) \notin V$ and $(b_K, x) \notin V$. Since the set of all finite supersets of gF is in ξ, one can define $a = \lim_{K\to\xi} a_K \in gF$ and $b = \lim_{K\to\xi} b_K \in gF$. It follows that for every $x \in A$, $y \in B$ one has $(a, y) \notin V$ and $(b, x) \notin V$. In other words, $gF \not\subseteq V[A]$ and $gF \not\subseteq V[B]$, meaning that (8) does not hold either.

(10) $\implies$ (9): finite induction in the number of colours, k. The case $k = 2$ forms the base of induction and holds by assumption. Suppose the statement is proved for $k = \ell$ (for all finite F and all entourages $V \in \mathcal{U}$). Let F' be chosen so that for every $c \colon F' \to [\ell]$ there is a $g \in G$ such that $gF \subseteq F'$ and gF is monochromatic to within V. Choose a finite $K \subseteq X$ in such a way that for every binary cover $\gamma = \{A, B\}$ of K there is an $h \in G$ such that $hF' \subseteq K$ and hF' is contained either in $V[A]$ or in $V[B]$. Let now $c \colon K \to [\ell + 1]$. By property (10), applied to the binary cover of K with $A = c^{-1}[\ell]$ and $B = c^{-1}\{\ell + 1\}$, there exists $h \in G$ such that $hF' \subseteq K$ and either $hF' \subseteq V[c^{-1}\{\ell + 1\}]$ (in which case we are done, as for some $g \in G$ one has $hgF \subseteq hF'$), or else $hF' \subseteq V[c^{-1}[\ell]]$. In the latter case, choose a $W \in \mathcal{U}$ so that $(x, y) \in W$ implies $(hx, hy) \in V$ (uniform continuity of translations). Apply the induction hypothesis to this W and the colouring $c_1 = c \circ h \colon F' \to [\ell]$ (that is, $c_1(x) = c(hx)$) to find a $g \in G$ such that $gF \subseteq F'$ and gF is contained in $W[c_1^{-1}(i)]$ for some $i \in [\ell]$. Then $hgF \subseteq K$ and $hgF \subseteq V[c^{-1}(i)]$, completing the step.

(9) $\implies$ (7): obvious, if one applies (9) to V, F, and $k = |\gamma|$.

(7) $\implies$ (6): if $W \in \mathcal{C}^*\mathcal{U}$, then, by Lemma 1.1.15, there exist a finite cover γ of X and a $V \in \mathcal{U}$ such that $\widetilde{\gamma_V} \subseteq W$. By assumption, for every finite $F \subseteq X$ there are $g \in G$ and $A \in \gamma$ with $gF \subseteq V[A]$. This means exactly that gF is $\widetilde{\gamma_V}$-small.

(6) $\implies$ (5): let W be an element of the unique compatible uniformity $\mathcal{U}_{\sigma X}$ on σX, and let $F \subseteq X$ be finite. The restriction W' of W to X is an element of $\mathcal{C}^*\mathcal{U}$, and by assumption there is $g \in G$ such that gF is W'-small, and consequently W-small, meaning that the map $i \colon X \to \sigma X$ is oscillation stable.

(5) $\implies$ (4): follows from the fact that every uniformly continuous function f from X to a compact space C factors through the canonical map $i \colon X \to \sigma X$.

(4) $\implies$ (3): apply (4) to the closure of the image $f(X)$ in $\mathbb{R}^n$.

(3) $\implies$ (1): trivial.

If in addition (X, d) is a metric space, then, in a trivial way, (3) $\implies$ (12) $\implies$ (11) $\implies$ (2) (with $\mathcal{R} = \{d\}$). □

Next we will discuss two examples of oscillation stable G-spaces of fundamental importance.

1.2. First example: the sphere $\mathbb{S}^\infty$

The first example is the unit sphere $\mathbb{S}^\infty$ of the infinite dimensional separable Hilbert space ℓ^2 with the usual norm distance, equipped with the standard action of the unitary group $U(\ell^2)$ by rotations. We will prove that the pair $(\mathbb{S}^\infty, U(\ell^2))$ is finitely oscillation stable.

The proof of this fact, belonging to V. Milman, is based on three ingredients:

- every finite subset F of the sphere $\mathbb{S}^\infty$ is contained in a sphere $\mathbb{S}^N$ of an arbitrarily high finite dimension N;
- there is a compact subgroup of $U(\ell^2)$, acting transitively on $\mathbb{S}^N$;
- concentration of measure on the spheres $\mathbb{S}^N$, $N \to \infty$.

The first two observations are obvious, while the third one requires explanation.

The *phenomenon of concentration of measure on high-dimensional structures* says, intuitively speaking, that the geometric structures – such as the Euclidean spheres – of high finite dimension typically have the property that an overwhelming proportion of points are very close to every set containing at least half of the points. Technically, the phenomenon is dealt with in the following framework.

Definition 1.2.1 (Gromov and Milman [**95**]). A *space with metric and measure,* or an *mm-space*, is a triple, (X, d, μ), consisting of a set X, a metric d on X, and a probability Borel measure on the metric space (X, d).

For a subset A of a metric space X and an $\varepsilon > 0$, denote by A_ε the ε-neighbourhood of A in X.

Definition 1.2.2 (*ibid.*). A family $\mathscr{X} = (X_n, d_n, \mu_n)_{n\in\mathbb{N}}$ of mm-spaces is a *Lévy family* if, whenever Borel subsets $A_n \subseteq X_n$ satisfy

$$\liminf_{n\to\infty} \mu_n(A_n) > 0,$$

one has for every $\varepsilon > 0$

$$\lim_{n\to\infty} \mu_n((A_n)_\varepsilon) = 1.$$

Remark 1.2.3. The concept of a Lévy family can be reformulated in many equivalent ways. For example, a family $\mathscr{X}$ as above is Lévy if and only if for every $\varepsilon > 0$, whenever A_n, B_n are Borel subsets of X_n satisfying

$$\mu_n(A_n) \geq \varepsilon,\ \ \mu_n(B_n) \geq \varepsilon,$$

then $d(A_n, B_n) \to 0$ as $n \to \infty$. (Exercise.)

This is formalized using the notion of *separation distance,* proposed by Gromov ([**93**], Section $3\frac{1}{2}$.30). Given numbers $\kappa_0, \kappa_1, \dots, \kappa_N > 0$, one defines the invariant

$$\mathrm{Sep}\,(X; \kappa_0, \kappa_1, \dots, \kappa_N)$$

as the supremum of all δ such that X contains Borel subsets X_i, $i = 0, 1, \dots, N$ with $\mu(X_i) \geq \kappa_i$, every two of which are at a distance $\geq \delta$ from each other. Now a family $\mathscr{X} = (X_n, d_n, \mu_n)_{n\in\mathbb{N}}$ is a Lévy family if and only if for every $0 < \varepsilon < \frac{1}{2}$, one has

$$\mathrm{Sep}\,(X; \varepsilon, \varepsilon) \to 0 \text{ as } n \to \infty.$$

The reader should consult Ch. $3\frac{1}{2}$ in [**93**] for numerous other characterisations of Lévy familes of mm-spaces.

EXAMPLE 1.2.4. The Euclidean spheres $\mathbb{S}^n$, $n \in \mathbb{N}_+$ of unit radius, equipped with the Haar measure (translation-invariant probability measure) and Euclidean (or geodesic) distance, form a Lévy family.

Before proving this fact in Theorem 1.3.15 below, we will deduce from it the oscillation stability of the sphere $\mathbb{S}^\infty$.

THEOREM 1.2.5. *The $U(\ell^2)$-space $\mathbb{S}^\infty$ is finitely oscillation stable.*

PROOF. Let $F \subset \mathbb{S}^\infty$ be finite, let γ be a finite cover of $\mathbb{S}^\infty$, and let $\varepsilon > 0$. Denote $n = |F|$ and $m = |\gamma|$. Choose a natural N so large that $N \geq n$ and whenever $A \subset \mathbb{S}^N$ has the property $\mu_N(A) \geq 1/m$, one has $\mu_N(A_\varepsilon) > 1 - 1/n$. Now fix a Euclidean sphere $\mathbb{S}^N \subset \mathbb{S}^\infty$ of dimension N containing the set F. The group of isometries of $\mathbb{S}^N$ is isomorphic to $U(N)$ and naturally embeds into $U(\ell^2)$: if an orthonormal basis of ℓ^2 is so chosen that $\mathbb{S}^N$ is the unit sphere of the linear span of the first $N+1$ basic vectors, then the embedding is given by

$$U(N) \ni u \mapsto \begin{pmatrix} u & 0 \\ 0 & \mathbb{I} \end{pmatrix} \in U(\ell^2).$$

Equipped with the compact-open topology with regard to its action on $\mathbb{S}^N$, the group $U(N)$ is compact and acts on the sphere continuously.

For some $A \in \gamma$, one has $\mu_N(A \cap \mathbb{S}^N) \geq 1/m$, and therefore $\mu_N(A_\varepsilon \cap \mathbb{S}^N) > 1 - 1/n$. Denote by ν the Haar measure on $U(N)$ normalised to one. Using the uniqueness of a rotation-invariant Borel probability measure on $\mathbb{S}^N$, one has for every $\xi \in F$

$$\nu\{u \in U(N) \colon u\xi \in A_\varepsilon\} = \mu_N(A_\varepsilon) > 1 - \frac{1}{n},$$

and consequently

$$\bigcap_{\xi \in F} \{u \in U(N) \colon u\xi \in A_\varepsilon\} \neq \emptyset.$$

Any element u belonging to the intersection on the left hand side has the property $uF \subset A_\varepsilon$. Since $u \in U(\ell^2)$, we are done by Theorem 1.1.18(7). □

The same argument *verbatim* gives us, at no extra cost, a more general result (Th. 1.2.10 and 1.2.12). It is convenient to ever so slightly extend the concept of a Lévy family by replacing a metric with a uniformity.

DEFINITION 1.2.6. We say that a net (μ_α) of probability Borel measures on a uniform space $(X, \mathcal{U})$ has the *Lévy concentration property,* or simply *concentrates*, if for every family of Borel subsets $A_\alpha \subseteq X$ satisfying $\liminf_\alpha \mu_\alpha(A_\alpha) > 0$ and every entourage $V \in \mathcal{U}_X$ one has $\mu_\alpha(V[A_\alpha]) \to_\alpha 1$.

EXERCISE 1.2.7. Suppose a net of measures (μ_α) on a uniform space X concentrates, and consider each μ_α as a measure on the Samuel compactification σX (by pushing the measure forward along the canonical map $X \to \sigma X$). Equip the space $P(\sigma X)$ of probability measures on σX with the usual weak* topology, regarding it as a subspace of the dual space to $C(\sigma X)$. With this topology, $P(\sigma X)$ is compact. Show that all cluster points of the net (μ_α) in $P(\sigma X)$ are Dirac measures corresponding to singletons.

The converse implication is false: weak convergence of measures to a point mass in general does not imply concentration.

Example 1.2.8. Let $X = \mathbb{S}^\infty$ be the sphere with the norm distance. For every finite collection $\mathcal{F}$ of bounded uniformly continuous functions on $\mathbb{S}^\infty$ and each $\varepsilon > 0$ there is a pair of points $a_{\mathcal{F},\varepsilon}, b_{\mathcal{F},\varepsilon} \in \mathbb{S}^\infty$ at a norm distance $\sqrt{2}/2$ from each other such that the oscillation of every $f \in \mathcal{F}$ on this pair is $< \varepsilon$ (finite oscillation stability of the sphere). As a consequence, the net of probability measures

$$\mu_{a_{\mathcal{F},\varepsilon}} = \frac{1}{2}(\delta_{a_{\mathcal{F},\varepsilon}} + \delta_{b_{\mathcal{F},\varepsilon}}),$$

indexed by pairs $(\mathcal{F}, \varepsilon)$ as above, concentrates with regard to the totally bounded replica of the norm uniformity on $\mathbb{S}^\infty$. By Exercise 1.2.7, the measures $(\mu_{a_{\mathcal{F},\varepsilon}})$ converge to a point mass in the space of probability measures on the Samuel compactification $\sigma\mathbb{S}^\infty$.

At the same time, the net $(\mu_{a_{\mathcal{F},\varepsilon}})$ does not concentrate with regard to the norm uniformity: it is enough to consider the family of Borel subsets $A_{\mathcal{F},\varepsilon} = \{a_{\mathcal{F},\varepsilon}\}$ and the element of the norm uniformity $V = \{(x, y) \in \mathbb{S}^\infty \times \mathbb{S}^\infty \colon \|x - y\| < \sqrt{2}/2\}$.

Remark 1.2.9. A family $\mathscr{X} = (X_n, d_n, \mu_n)_{n\in\mathbb{N}}$ of mm-spaces is a Lévy family if and only if the measures (μ_n) concentrate in the sense of Def. 1.2.6 if considered as probability measures on the disjoint union $\oplus_{n=1}^\infty X_n$, equipped with a metric d inducing the metrics d_n on each X_n and making X_n into an open and closed subset.

Here is a result that we have in fact proved while establishing Theorem 1.2.5.

Theorem 1.2.10. *Let a topological group G act continuously by isometries on a metric space (X, d). Assume there is a family of compact subgroups (G_α), directed by inclusion and such that for some $\xi \in X$ the orbits $G_\alpha \cdot \xi$*

(1) *have an everywhere dense union in X, and*
(2) *form a Lévy family with regard to the restriction of the metric d and the probability measures $\mu_\alpha \cdot \xi$, where μ_α is the normalized Haar measure on G_α.*

Then the G-space X is finitely oscillation stable. □

Here $\mu_\alpha \cdot \xi$ stands for the push-forward of the normalized Haar measure μ_α on the compact group G_α under the orbit map $G_\alpha \ni g \mapsto g \cdot \xi \in X$.

Remark 1.2.11. In fact, one can get rid of the compact subgroups from the definition altogether, using two facts: (i) the group of isometries, $\mathrm{Iso}\,(K)$, of a compact metric space K, equipped with the topology of simple convergence, is compact, and (ii) the orbit map $\mathrm{Iso}\,(K) \to K$ is open. (As a continuous map between compact spaces, it is closed, hence a quotient map onto a homogeneous space, hence open; the statement remains true for Polish groups acting on Polish spaces, this is the Effros theorem [**52**], but the proof becomes much less trivial.) Thus, K is a homogeneous factor-space of the compact group and supports a unique probability measure invariant under isometries. The groups $\mathrm{Iso}\,(K_\alpha)$ will no longer be topological subgroups of G, in fact they will be continuous homomorphic images of suitable subgroups of G, but since their role is auxiliary, like that of the unitary groups $U(N)$ in the proof of 1.2.5, this is enough. We arrive at the following result.

Corollary 1.2.12. *Let a group G act on a metric space X by isometries. Suppose there exists a family (K_α) of compact subspaces of X such that*

- *(K_α) are directed by inclusion,*
- *the union of (K_α) is everywhere dense in X,*

- *for every α, the isometries of X, stabilizing K_α setwise, act on K_α transitively, and*
- *the (unique) normalized probability measures on K_α, invariant under isometries, concentrate.*

Then the G-space X is finitely oscillation stable. □

1.3. Concentration of measure on spheres

EXERCISE 1.3.1. Show that in Definition 1.2.2 of a Lévy family it is enough to consider the case where the values $\mu_n(A_n)$ are bounded away from zero by an apriori chosen constant strictly between zero and one, for instance 1/2. More precisely, prove that a family $\mathscr{X} = (X_n, d_n, \mu_n)$ of mm-spaces is a Lévy family if and only if, whenever Borel subsets $A_n \subseteq X_n$ satisfy $\mu_n(A_n) \geq 1/2$, one has for every $\varepsilon > 0$

$$\lim_{n\to\infty} \mu_n(A_n)_\varepsilon = 1.$$

This reformulation takes us to the following concept [**148, 153**], providing convenient quantitative bounds on the rate of convergence of $\mu_n(A_n)_\varepsilon$ to one.

DEFINITION 1.3.2. Let (X, d, μ) be a space with metric and measure. The *concentration function* of X, denoted by $\alpha_X(\varepsilon)$, is a real-valued function on the positive axis $\mathbb{R}_+ = [0, \infty)$, defined by letting

$$\alpha_X(\varepsilon) = \begin{cases} \frac{1}{2}, & \text{if } \varepsilon = 0, \\ 1 - \inf\left\{\mu\left(B_\varepsilon\right) : B \subseteq X,\ \mu(B) \geq \frac{1}{2}\right\}, & \text{if } \varepsilon > 0. \end{cases}$$

REMARK 1.3.3. A family $\mathscr{X} = (X_n, d_n, \mu_n)_{n\in\mathbb{N}}$ of mm-spaces is a Lévy family if and only if

$$\alpha_{X_n} \to 0 \text{ pointwise on } (0, +\infty) \text{ as } n \to \infty.$$

DEFINITION 1.3.4. A Lévy family $\mathscr{X}$ as above is called *normal* if for suitable constants $C_1, C_2 > 0$,

$$\alpha_{X_n}(\varepsilon) \leq C_1 e^{-C_2\varepsilon^2 n}.$$

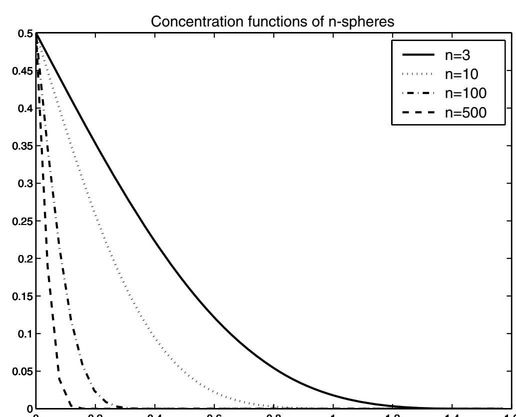

FIGURE 1.3. The concentration functions of spheres in various dimensions.

As we will see shorlty, the spheres $(\mathbb{S}^n)$ form a normal Lévy family. Fig. 1.3 shows the concentration functions of spheres $\mathbb{S}^n$ in dimensions $n = 3, 10, 100, 500$.

REMARK 1.3.5. The concept of a Lévy family was introduced in [**95**], and that of a normal Lévy family appears in [**4**].

Next we will establish the Brunn–Minkowski inequality. By a *box* in the Euclidean space $\mathbb{R}^n$ we mean an (open or closed) coordinate parallelepiped.

LEMMA 1.3.6. *If Π_1 and Π_2 are two disjoint boxes in $\mathbb{R}^n$, then there are an $i = 1, 2, \dots, n$ and $c \in \mathbb{R}$, such that for all $x \in \Pi_1$, $x_i \leq c$, while for all $x \in \Pi_2$, $x_i \geq c$.* □

Geometrically, the result means that any two boxes are separated by a coordinate hyperplane, and immediately follows from the obvious one-dimensional case.

LEMMA 1.3.7. *The geometric mean of a finite collection of positive reals never exceeds the arithmetic mean:*

$$(a_1 a_2 \dots a_n)^{1/n} \leq \frac{1}{n}(a_1 + a_2 + \dots + a_n), \ a_1, a_2, \dots, a_n \geq 0.$$

□

THEOREM 1.3.8 (Brunn–Minkowski inequality — the additive form). *Let A and B be non-empty subsets of $\mathbb{R}^n$ such that A, B and $A + B$ are all measurable. Then*

$$\mu(A + B)^{1/n} \geq \mu(A)^{1/n} + \mu(B)^{1/n}.$$

PROOF. Without loss in generality, we can assume that A and B are unions of finite families of disjoint open boxes. (Approximate A and B by compact sets from the inside and then by unions of open boxes from the outside.) The proof is by induction in k, the total number of such boxes in the two sets.

Basis of induction. Let A and B be open boxes with side lengths $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots, b_n$ respectively. The Minkowski sum of A and B is again an open box, and one has

$$\mu(A) = a_1 a_2 \dots a_n, \ \mu(B) = b_1 b_2 \dots b_n,$$

and

$$\mu(A + B) = (a_1 + b_1)(a_2 + b_2) \dots (a_n + b_n).$$

Applying Lemma 1.3.7 twice, first to the non-negative numbers $a_i/(a_i + b_i)$ and then to $b_i/(a_i + b_i)$, $i = 1, 2, \dots, n$, and adding up the resulting inequalities, one gets

$$\begin{aligned} \prod_{i=1}^n \left(\frac{a_i}{a_i + b_i}\right)^{\frac{1}{n}} + \prod_{i=1}^n \left(\frac{b_i}{a_i + b_i}\right)^{\frac{1}{n}} &\leq \frac{1}{n}\sum_{i=1}^n \frac{a_i}{a_i + b_i} + \frac{1}{n}\sum_{i=1}^n \frac{b_i}{a_i + b_i} \\ &= 1, \end{aligned} \tag{1.1}$$

whence it follows that

$$\begin{aligned} \mu(A)^{\frac{1}{n}} + \mu(B)^{\frac{1}{n}} &= \prod_{i=1}^{n}(a_i)^{\frac{1}{n}} + \prod_{i=1}^{n}(b_i)^{\frac{1}{n}} \\ &\leq \prod_{i=1}^{n}(a_i+b_i)^{\frac{1}{n}} \\ &= \mu(A+B)^{\frac{1}{n}}. \end{aligned} \tag{1.2}$$

Step of induction. Suppose that the inequality has been proved for all pairs of sets A, B representable each as a finite union of disjoint open boxes, $\leq k-1$ in total. Let A and B be finite unions of disjoint open boxes, $k \geq 3$ in total. At least one set, say A, must be made up of at least 2 boxes. (Then B consists of at most $k-2$ boxes.)

Using Lemma 1.3.6, find an $x \in \mathbb{R}^n$ such that the translates by x of at least two open boxes making up A are on the opposite sides of the coordinate hyperplane $x_i = 0$. By cutting boxes up with the hyperplane $x_i = 0$ and throwing away the negligible set $A \cap \{x \colon x_i = 0\}$, we can assume that in fact each of the newly-diminished boxes is entirely on one side of the hyperplane.

Denote by A^L the union of all boxes to the 'left' from the hyperplane (that is, for each $a \in A^L$ one has $a_i < x_i$), and by A^R the union of all boxes to the 'right' of the hyperplane.

Now denote $\lambda = \mu(A^L)/\mu(A)$ and for every $x \in \mathbb{R}$, set

$$f(x) = \mu\{b \in B \colon b_i < x\}.$$

The function $f \colon \mathbb{R} \to [0, \mu(B)]$ is piecewise-linear and assumes both values 0 and $\mu(B)$. By the Intermediate Value Theorem, there is a $\xi \in \mathbb{R}$ with $f(\xi) = \lambda\mu(B)$.

Choose now an arbitrary vector $y \in \mathbb{R}^n$ whose i-th coordinate is $-\xi$, and observe that the translate $B+y$ has the following property. Denote by $B^L = \{b \in B \colon b < y\}$ the part of $B+y$ that is to the 'left' from the same coordinate hyperplane $x_i = 0$, and by $B^R = \{b \in B \colon b > y\}$ the part to the 'right' of it. Then

$$\frac{\mu(B^L)}{\mu(B)} = \lambda \ \text{ and } \ \frac{\mu(B^R)}{\mu(B)} = 1 - \lambda = \frac{\mu(A^R)}{\mu(A)}.$$

The total number of boxes in each of the pairs A^L, B^L and A^R, B^R is strictly less than k, and the inductive hypothesis applies. Since the sets $A^L + B^L$ and $A^R + B^R$ are on the opposite sides of the hyperplane $x_i = 0$ and so disjoint,

$$\mu[(A^L + B^L) \cup (A^R + B^R)] = \mu(A^L + B^L) + \mu(A^R + B^R).$$

Clearly, $(A^L + B^L) \cup (A^R + B^R) \subseteq A + B + (x+y)$, and therefore

$$\begin{aligned} \mu(A+B) &\geq \mu(A^L + B^L) + \mu(A^R + B^R) \\ &\geq \left[\mu(A^L)^{\frac{1}{n}} + \mu(B^L)^{\frac{1}{n}}\right]^n + \left[\mu(A^R)^{\frac{1}{n}} + \mu(B^R)^{\frac{1}{n}}\right]^n \\ &= \lambda\left[\mu(A)^{\frac{1}{n}} + \mu(B)^{\frac{1}{n}}\right]^n + (1-\lambda)\left[\mu(A)^{\frac{1}{n}} + \mu(B)^{\frac{1}{n}}\right]^n \\ &= \left[\mu(A)^{\frac{1}{n}} + \mu(B)^{\frac{1}{n}}\right]^n, \end{aligned} \tag{1.3}$$

finishing the step of induction. □

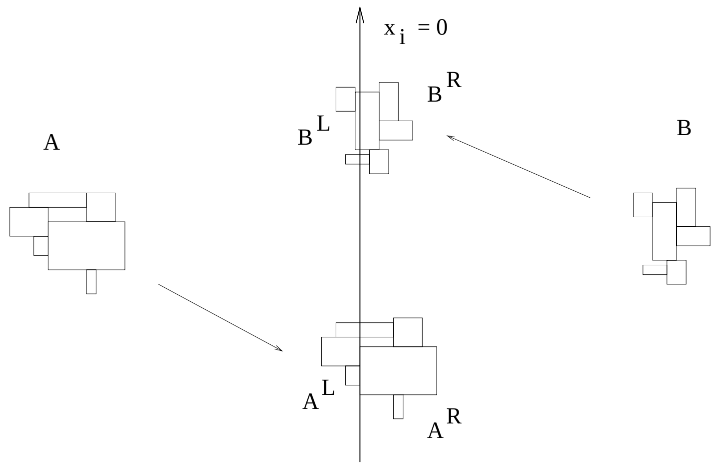

FIGURE 1.4. To the proof of the Brunn–Minkowski inequality.

THEOREM 1.3.9 (Brunn–Minkowski inequality — the multiplicative form). *Let A and B be non-empty Borel subsets of $\mathbb{R}^n$. Then*

$$\mu\left(\frac{A+B}{2}\right) \geq (\mu(A)\mu(B))^{1/2}.$$

PROOF. Applying the obvious inequality $a^2+b^2 \geq 2ab$ to $a=\mu(A)^{1/2n}$ and $b=\mu(B)^{1/2n}$, one derives from Theorem 1.3.8:

$$\begin{aligned}\mu\left(\frac{A+B}{2}\right)^{\frac{1}{n}} &\geq \mu\left(\frac{A}{2}\right)^{\frac{1}{n}}+\mu\left(\frac{B}{2}\right)^{\frac{1}{n}} \\ &= \frac{1}{2}\cdot\left(\mu(A)^{1/n}+\mu(B)^{1/n}\right) \\ &\geq (\mu(A)\mu(B))^{1/2n},\end{aligned}$$

now it is enough to raise to the power of n on both sides. □

We need the following geometric property.

PROPOSITION 1.3.10. *The unit ball $\mathbb{B}$ in the Hilbert space is* uniformly convex *in the sense that for every $x, y \in \mathbb{B}$ at a distance $\varepsilon = d(x,y)$ from each other, the distance from the mid-point $(x+y)/2$ to the complement of $\mathbb{B}$ is at least*

$$\delta(\varepsilon) = 1-\sqrt{1-\varepsilon^2/4},$$

which number is called the modulus of convexity *of $\mathbb{B}$.*

PROOF. It is rather clear that the minimal value of the distance between $(x+y)/2$ and the boundary of $\mathbb{B}$ is achieved when x and y are themselves elements of the sphere. Consider the section of the ball by the plane spanned by x and y (Fig. 1.5). Since $|By| = \varepsilon/2$ and OBy forms a right triangle, $|OB| = \sqrt{1-\varepsilon^2/4}$ and finally

$$|BA| = 1-\sqrt{1-\varepsilon^2/4}.$$

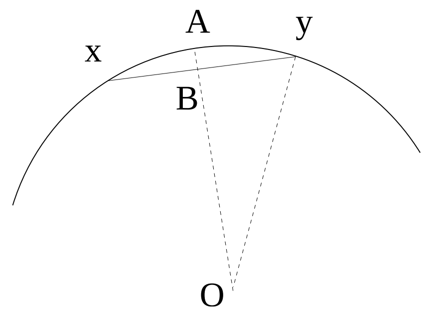

FIGURE 1.5. Uniform convexity of the Euclidean ball.

□

LEMMA 1.3.11. $\delta(\varepsilon) \geq \frac{\varepsilon^2}{8}$.

PROOF. The desired inequality

$$1 - \left(1 - \frac{\varepsilon^2}{4}\right)^{\frac{1}{2}} \geq \frac{\varepsilon^2}{8}$$

is equivalent to

$$1 - \frac{\varepsilon^2}{4} \leq \left(1 - \frac{\varepsilon^2}{8}\right)^2,$$

and is clearly correct once the brackets on the r.h.s. are opened. □

REMARK 1.3.12. Let A be a subset of $\mathbb{S}^n$. Denote by

$$\tilde{A} = \{ta \colon a \in A,\ t \in [0,1]\}$$

the *cone* over A. If $A \subseteq \mathbb{S}^n$ is a Borel subset, then $\tilde{A}$ is a Borel subset of $\mathbb{B}^{n+1}$. Besides,

$$\nu_n(A) \equiv \nu(A) = \frac{\mu(\tilde{A})}{\mu(\mathbb{B}^{n+1})},$$

where ν_n denotes the Haar measure on the sphere and $\mu = \mu_{n+1}$ stands for the Lebesgue measure in $\mathbb{R}^{n+1}$.

LEMMA 1.3.13. *Let A be a Borel subset of $\mathbb{S}^{n-1}$, let $\varepsilon > 0$, let $B = \mathbb{S}^{n-1} \setminus A_\varepsilon$, and let $a \in \tilde{A}$ and $b \in \tilde{B}$ belong to the respective cones. Then*

$$\left\| \frac{a+b}{2} \right\| \leq 1 - \delta(\varepsilon).$$

□

REMARK 1.3.14. In addition to the Euclidean distance on the sphere, there is also the geodesic distance. Since for every $x, y \in \mathbb{S}^n$,

$$d_{eucl}(x,y) \leq d_{geo}(x,y) \leq \frac{\pi}{2} d_{eucl}(x,y),$$

it makes no difference (apart from an extra constant in a formula), which of the two distances we choose. In the next result it will be more convenient for us to use the Euclidean distance.

THEOREM 1.3.15. *The unit spheres $\mathbb{S}^n$, equipped with the rotation-invariant probability Borel measures and the Euclidean distance, form a normal Lévy family:*

$$\alpha_{\mathbb{S}^{n-1}}(\varepsilon) \leq 2e^{-n\varepsilon^2/4}.$$

PROOF. Let A be a Borel subset of the $(n-1)$-dimensional Euclidean sphere $\mathbb{S}^{n-1}$, containing at least half of all points: $\nu(A) \geq \frac{1}{2}$. We will show that for every $\varepsilon > 0$ one has

$$\begin{aligned} \nu(A_\varepsilon) &\geq 1 - 2e^{-2n\delta(\varepsilon)} \\ &\geq 1 - 2e^{-n\varepsilon^2/4}, \end{aligned} \tag{1.4}$$

where the second inequality follows easily from Lemma 1.3.11.

Denote $B = \mathbb{S}^{n-1} \setminus A_\varepsilon$. Lemma 1.3.13 implies that the set $(\tilde{A}+\tilde{B})/2$ is contained inside the closed ball of radius $1-\delta(\varepsilon)$ and therefore, modulo the Brunn–Minkowski inequality in the multiplicative form (Theorem 1.3.9),

$$\begin{aligned} (1-\delta(\varepsilon))^n &\geq \tilde{\mu}\left(\frac{\tilde{A}+\tilde{B}}{2}\right) \\ &\geq \left(\tilde{\mu}(\tilde{A})\tilde{\mu}(\tilde{B})\right)^{1/2} \\ &\geq \left(\frac{1}{2}\tilde{\mu}(\tilde{B})\right)^{1/2}, \end{aligned} \tag{1.5}$$

where $\tilde{\mu}$ denotes the Lebesgue measure normalized so that the volume of the unit ball is 1. This formula yields

$$\nu(B) \equiv \tilde{\mu}(\tilde{B}) \leq 2(1-\delta(\varepsilon))^{2n}.$$

It remains to notice that

$$1 - \delta(\varepsilon) \leq e^{-\delta(\varepsilon)} = 1 - \delta(\varepsilon) + \frac{\delta(\varepsilon)^2}{2} - \ldots,$$

from which one concludes

$$\nu(B) \leq 2e^{-2n\delta(\varepsilon)},$$

which is the desired result in a slight disguise. □

REMARK 1.3.16. The upper bound in Theorem 1.3.15 is not the best possible and can be improved, using the technique of isoperimetric inequalities, to at least this:

$$\alpha_{\mathbb{S}^{n-1}}(\varepsilon) \leq \sqrt{\frac{\pi}{8}} e^{-n\varepsilon^2/2}.$$

However, even in this case the Gaussian estimate falls short of closely approximating the true concentration function, as can be seen from Fig. 1.6.

REMARK 1.3.17. There are at least five different known proofs of concentration of measure in the Euclidean spheres, and probably many more. In presenting a proof based on Brunn–Minkowski inequality, we followed [**153, 96**]. The original proof by Lévy [**127**], based on the isoperimetric inequality, was made completely rigorous by Gromov in the general setting of Riemannian manifolds, cf. his apprendix in [**153**], and another well-known proof uses the spectral theory of the Laplacian, cf.

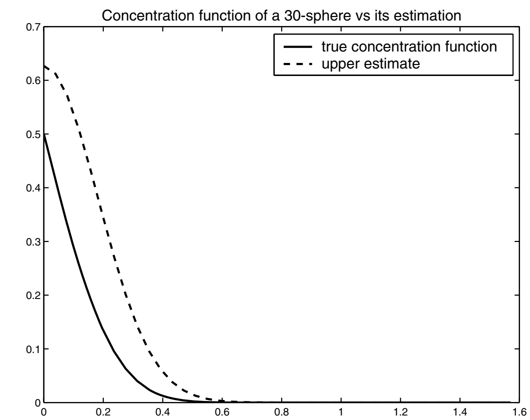

FIGURE 1.6. The concentration function of 30-sphere versus a Gaussian upper bound.

[**95**]. There is an interesting proof based on two-point symmetrization [**60**]. For more, see [**153, 20, 126**].

1.4. Dvoretzky's theorem

A deduction of Dvoretzky's theorem (Th. 0.0.1) from the finite oscillation stability of the pair $(U(\ell^2), \mathbb{S}^\infty)$ can be roughly subdivided into three stages.

(1) The Dvoretzky theorem applies to a finite-dimensional space of some high finite dimension N, meaning that we need to consider the sphere in $\ell^2(N)$ rather than in ℓ^2 and essentially apply finite oscillation stability in the equivalent form (11) or (12) as per Theorem 1.1.18. Concrete bounds on the finite cardinalities involved become of importance. With this purpose, we are now going to refine Theorem 1.2.5.

Every real-valued measurable function f on a finite measure space has a *median* (or *Lévy mean*), that is, a number M with the property

$$\mu\{x \in X\colon f(x) \geq M\} \geq \frac{1}{2} \text{ and } \mu\{x \in X\colon f(x) \leq M\} \geq \frac{1}{2}.$$

In general, the median value of a function need not be unique, though it is uniquely defined for continuous functions on connected mm-spaces (X, d, μ) (where "connected" should be of course understood in the sense that the support of the measure μ in X is topologically connected).

For a uniformly continuous function f on a metric space X, denote by $\delta = \delta(\varepsilon)$ the *modulus of uniform continuity* of f, that is, a function from $\mathbb{R}_+$ to itself where $\delta(\varepsilon)$ is the largest value such that

$$\forall x, y \in X,\ d_X(x, y) < \delta(\varepsilon) \implies d_Y(f(x), f(y)) < \varepsilon.$$

For instance, if f is a Lipschitz function, then $\delta(\varepsilon) \leq L^{-1}\varepsilon$, where L is a Lipschitz constant for f.

THEOREM 1.4.1. *Let $X = (X, d, \mu)$ be an mm-space, and let f be a uniformly continuous real-valued function on X with the modulus of uniform continuity δ.*

Denote by $M = M_f$ a median value for f. Then for every $\varepsilon > 0$

$$\mu\{|f(x) - M| > \varepsilon\} \leq 2\alpha_X(\delta(\varepsilon)). \tag{1.6}$$

PROOF. If we denote

$$A = \{x \in X \mid f(x) \leq M\}$$

and

$$B = \{x \in X \mid f(x) - M \leq \varepsilon\},$$

then clearly

$$A_{\delta(\varepsilon)} \subseteq B$$

and consequently

$$\mu(B) \geq 1 - \alpha(\delta(\varepsilon)).$$

A similar argument shows that

$$\mu\{x \mid M - f(x) \leq \varepsilon\} \geq 1 - \alpha(\delta(\varepsilon)).$$

The two halves come together to imply Eq. (1.6). □

EXERCISE 1.4.2. Refine the proof of Theorem 1.2.5, using Theorem 1.4.1 and the fact that the spheres $\mathbb{S}^n$ form a *normal* Lévy family, so as to prove the following result.

THEOREM 1.4.3. *Let $f\colon \mathbb{S}^N \to \mathbb{R}$ be a 1-Lipschitz function (with regard to the Euclidean distance say), let $\varepsilon > 0$, and let $F \subset \mathbb{S}^N$ be a finite set of cardinality*

$$k \leq \frac{2}{\sqrt{\pi}} e^{\varepsilon^2 N/2}.$$

Then there is a unitary operator $u \in SU(N)$ such that

$$\|f|_{uF} - M(f)\|_{sup} < \varepsilon, \tag{1.7}$$

where $M(f)$ denotes the median value of f on $\mathbb{S}^N$. □

REMARK 1.4.4. More generally, if f is a Lipschitz function with Lipschitz constant $\mathrm{Lip}(f) \leq L$, then under the same assumptions one obtains the conclusion

$$\|f|_{uF} - M(f)\|_{sup} < L\varepsilon.$$

Equivalently (after a change of variables $L\varepsilon \mapsto \varepsilon$), one can obtain the conclusion in Eq. (1.7) on the assumption that

$$k \leq \frac{2}{\sqrt{\pi}} e^{\varepsilon^2 N/2L^2}.$$

(2) Instead of the conclusion $\mathrm{osc}\,(f|uF) < \varepsilon$, that is, a small oscillation on a finite subset, what we really need, is a small oscillation on the unit sphere of a subspace of lower dimension k. This is achieved by approximating the sphere $\mathbb{S}^k$ with a finite ε-net, using a very simple, but surprisingly efficient, argument.

EXERCISE 1.4.5. Let $k \in \mathbb{N}$ and $\varepsilon > 0$. Show that the Euclidean sphere $\mathbb{S}^k$ admits an ε-net of cardinality at most

$$\left(1+\frac{2}{\varepsilon}\right)^k \leq e^{2k/\varepsilon}.$$

Hint: let $\mathscr{N}$ be a maximal $\varepsilon/2$-discrete subset of $\mathbb{S}^k$, that is, for all $x, y \in \mathscr{N}$, $x \neq y$, one has $\|x-y\| \geq \varepsilon/2$. The $\varepsilon/2$-balls with centres at elements of $\mathscr{N}$ are pairwise disjoint and all are contained in a single ball of radius $1+\varepsilon/2$. Now compare the volumes.

Since the values of a 1-Lipschitz function on $\mathbb{S}^k$ are controlled by the values at elements of a sufficiently fine ε-net, Theorem 1.4.3, Remark 1.4.4 and Exercise 1.4.5 (all applied with $\varepsilon/2$ instead of ε) lead to the following.

LEMMA 1.4.6. *Let $\varepsilon > 0$, $n \in \mathbb{N}$, $L > 0$, and let a natural number k satisfy*

$$e^{4k/\varepsilon} \leq \frac{2}{\sqrt{\pi}} e^{\varepsilon^2 n/8L^2}. \tag{1.8}$$

Let $f\colon \mathbb{S}^N \to \mathbb{R}$ is a Lipschitz function with $\mathrm{Lip}(f) = L$. *Then there exists a Euclidean subspace V of $\mathbb{R}^{N+1}$ of dimension k such that*

$$\|f|_{\mathbb{S}^N \cap V} - M(f)\|_{sup} < \varepsilon,$$

*where $M(f)$ is the median value of f on the sphere $\mathbb{S}^N$ (*not on $\mathbb{S}^N \cap V$*).* □

EXERCISE 1.4.7. In the proof of Dvoretzky theorem, we will use the above result with $L = \sqrt{n}/\sqrt{\log n}$. Verify that in this case Eq. (1.8) is satisfied whenever $k \geq c\varepsilon^2 n/|\log \varepsilon|$, where $c > 0$ is a suitable universal constant.

(3) Given a finite-dimensional Banach space $(E, \|\cdot\|)$, choose an equivalent Euclidean norm $|\cdot|$ on E. Suppose that $|\cdot| \geq \|\cdot\|$, so that the norm $\|\cdot\|$, considered as a function on the unit Euclidean sphere, is 1-Lipschitz. By force of finite oscillation stability, the values of $\|\cdot\|$ on the Euclidean unit sphere of a certain subspace V of E concentrate near the median value. It follows that the restrictions of norms $|\cdot|$ and $\|\cdot\|$ to V are nearly multiples of each other, and the Dvoretzky theorem follows.... unless the median value in question is not too close to zero, in which case the result gives no useful information!

To assure that this is not the case, one needs additional tools.

By an *ellipsoid* in a finite-dimensional vector space we mean the unit ball with regard to some Euclidean distance, or in other words, the image of the standard Euclidean ball, $\mathbb{B}$, under a linear transformation. Clearly, among all ellipsoids contained in a given bounded convex body with a non-empty interior in $\mathbb{R}^n$ there exists one having the largest volume (with regard to the standard Lebesgue measure in $\mathbb{R}^n$).

THEOREM 1.4.8 (Dvoretzky–Rogers lemma). *Let $E = (\mathbb{R}^n, \|\cdot\|)$ be an n-dimensional normed space with the property that the Euclidean unit ball forms an ellipsoid of maximal volume contained in the unit ball of E. Then there exists an orthonormal basis $(x_i)_{i=1}^n$ in $\mathbb{R}^n$ with the property*

$$\|x_i\| \geq 2^{-n/(n-i+1)}$$

for all i. In particular, for $i = 1, 2, \ldots, \lfloor n/2 \rfloor$ one has

$$\|x_i\| \geq \frac{1}{2}.$$

Proof. Choose elements x_i, $i = 1, 2, \ldots, n$ by recursion so that each x_i maximizes the value of $\|x_i\|_E$ on the compact set $\{x_j : j < i\}^\perp \cap \mathbb{S}_E$.

Note in particular that, since the Euclidean ball must touch the unit sphere $\mathbb{S}_E$, one necessarily has $\|x_1\|_E = 1$. Also, the sequence $\|x_i\|$ is (non-strictly) decreasing.

Denote $a = \|x_i\|$. The set

$$\mathscr{E} = \left\{ x = \sum_{j=1}^{n} t_j x_j : 2\sum_{j=1}^{i-1} t_j^2 + 2a^2 \sum_{j=i}^{n} t_j^2 \leq 1 \right\}$$

is an ellipsoid. The determinant of a linear mapping transforming the Euclidean unit ball $\mathbb{B}$ into $\mathscr{E}$ is $(\sqrt{2})^{-n} a^{n-j+1}$, therefore the volume of the ellipsoid $\mathscr{E}$ equals $(\sqrt{2})^{-n} a^{n-j+1} \mathrm{vol}\,\mathbb{B}$. It only remains to show that $\mathscr{E}$ is contained in the unit ball $\mathbb{B}_E$, because in such a case, by Theorem's assumptions, $\mathrm{vol}\,\mathscr{E} \leq \mathrm{vol}\,\mathbb{B}$, which inequality gives a required estimate on the value of $a = \|x_i\|$.

Let us prove that $\mathscr{E} \subseteq \mathbb{B}_E$. The norm of an $x \in \mathscr{E}$ is bounded from above by

$$\left\| 2\sum_{j=1}^{i-1} t_j \right\|_E + \left\| 2a^2 \sum_{j=i}^{n} t_j \right\|_E,$$

where the first sum is easily seen not to exceed $2|\sum_{j=1}^{i-1} t_j x_i|$, and the second $2a|\sum_{j=1}^{i-1} t_j x_i|$. Since the elements given by two sums are orthogonal to each other, the sum of the two latter norms is in its turn bounded by

$$\sqrt{2} \left(\sum_{j=1}^{i-1} t_j + a^2 \sum_{j=1}^{i-1} t_j \right)^{1/2} \leq 1.$$

□

In some applications of Theorem 1.4.1 it is convenient to replace the median value of a Lipschitz function f on an mm-space (X, d, μ) with the average value, or expectation, of f:

$$E(f) = \int_X f(x)\, d\mu(x).$$

The following series of exercises shows that such a substitution is painless, as the difference between the two values is sufficiently small.

Exercise 1.4.9. Let $X = (X, d, \mu)$ be an mm-space with concentration function α_X. Show that for every real-valued 1-Lipschitz function f

$$|M(f) - E(f)| \leq \int_0^\infty \alpha_X(\varepsilon)\, d\varepsilon.$$

Exercise 1.4.10. Deduce that, if $X_n = (X_n, d_n, \mu_n)$, $n \in \mathbb{N}_+$ is a Lévy family of mm-spaces and $f_n : X_n \to \mathbb{R}$ are 1-Lipschitz functions, then

$$|M(f_n) - E(f_n)| \to 0 \text{ as } n \to \infty.$$

EXERCISE 1.4.11. Now let $\mathscr{X} = (X_n, d_n, \mu_n)_{n=1}^{\infty}$ be a normal Lévy family. Show that there exists a constant $C_{\mathscr{X}} > 0$ such that for every n and each 1-Lipschitz function $f\colon X_n \to \mathbb{R}$,

$$|M(f_n) - E(f_n)| \le \frac{C_{\mathscr{X}}}{\sqrt{n}}.$$

EXERCISE 1.4.12. By considering the example of spheres $\mathbb{S}^n$, show that the above asymptotic estimate cannot be improved.

Now we will reduce estimating the average value of a norm to estimating the average value of the ℓ^∞-norm on the Euclidean sphere in $\ell^2(N)$.

LEMMA 1.4.13. *Let E be an N-dimensional normed space. There exists a linear embedding $\ell^2(n) \hookrightarrow E$ with $n \ge \lfloor N/2 \rfloor$ such that for each $x \in \ell^2(n)$ one has $\|x\|_E \le |x|$ and*

$$E(\|\cdot\|_E) \ge \frac{1}{2} E(\|\cdot\|_\infty),$$

where the expectation is taken over the sphere $\mathbb{S}^n$.

PROOF. Embed $\ell^2(n)$ with $n = \lfloor N/2 \rfloor$ by sending the standard basic vectors into the elements x_i chosen as in the Dvoretzky–Rogers lemma 1.4.8. One has $\|x_i\| \ge 1/2$ for $i = 1, 2, \dots, n$.

For every $\varepsilon \in \{1, -1\}^n$ and $x \in \ell^2(n)$, denote $\varepsilon x = (\varepsilon_1 x_1, \dots, \varepsilon_n x_n)$. Because of the second property of the norm, it is easily seen that the average value $\mathrm{ave}_{\varepsilon \in \{\pm\}^n} \|\varepsilon x\|_E$, $x = (x_1, x_2, \dots, x_n)$, is bounded from below by

$$\max_{i=1}^{n} |x_i| \, \|e_i\|_E \,,$$

and consequently, by $(1/2)\,\|x\|_\infty$.

Every $\varepsilon \in \{1, -1\}^n$ determines a generalized octant of the sphere $\mathbb{S}^n$, given by the condition $\operatorname{sign} x_i = \varepsilon_i$, $i = 1, 2, \dots, n$. Denote this octant by $\mathbb{S}^n_\varepsilon$. The integral of a function f over $\mathbb{S}^n$ is the sum of integrals over all $\mathbb{S}^n_\varepsilon$. Equivalently, one can first sum up the values of f at every point x over all vertices of the hypercube $\{\varepsilon x \colon \varepsilon \in \{1, -1\}^n\}$ and then integrate the result over $\mathbb{S}^n_{1,1,\dots,1} = \mathbb{S}^n_+$, or, which is the same, to average f over the hypercube and then integrate the result over $\mathbb{S}^n$.

Applying this to $f = \|\cdot\|_E$, one obtains

$$\begin{aligned} E(\|\cdot\|_E) &= \int_{\mathbb{S}^{n-1}} \mathrm{ave}_{\varepsilon \in \{1,-1\}^n} (f(\varepsilon x))\, d\nu_{n-1} \\ &\ge \frac{1}{2} \int_{\mathbb{S}^{n-1}} \|x\| \, d\nu_{n-1} \\ &= \frac{1}{2} E(\|\cdot\|_E). \end{aligned}$$

□

All that remains to show, is that the average value of the ℓ_∞ norm of a point of the Euclidean unit sphere is not as hopelessly small as one might fear. Indeed, since the mass of the unit hypercube (the unit ball of $\ell^\infty(n)$) is concentrated near the corners, whose scalar multiples on $\mathbb{S}^{n-1}$ only have $\ell^\infty(n)$-norm $1/\sqrt{n}$, one might rush to a conclusion that the expectation of $E(\|\cdot\|_\infty)$ is of the order $O(1/\sqrt{n})$ and therefore unusable: the median value differs from the expectation by the same quantity, and cannot be separated away from zero! In reality, this estimate is way too rough and the actual expected value differs from it by a multiple of $\sqrt{\log n}$, which makes all the difference.

The following is, essentially, a well-known lemma in probability theory in a slight disguise.

LEMMA 1.4.14. *For some absolute constant $c > 0$ and all n, one has*

$$\int_{\mathbb{S}^{n-1}} \|x\|_\infty \, d\nu_{n-1}(x) \geq \frac{c\sqrt{\log n}}{\sqrt{n}}.$$

PROOF. The key observation is that the Haar measure ν_{n-1} can be replaced with any rotation-invariant probability measure on $\mathbb{R}^n$, for instance, a Gaussian measure γ_n with density $e^{-\pi\|x\|^2}$. The integrals of any positive homogeneous function with regard to the two measures will only differ by a constant factor, because both measures are rotation-invariant. At the same time, the Gaussian measure is much better suited to integrate over parallelepipeds.

To estimate the constant of proportionality,

$$\int_{\mathbb{R}^n} f(x)\, d\gamma_n(x) = \lambda_n \int_{\mathbb{S}^{n-1}} f(x)\, d\nu_{n-1},$$

consider the case $f(x) = \|x\|$. The integral on the r.h.s. is one, while the integral on the l.h.s. is bounded above by $\left(\int_{\mathbb{R}^n} \|x\|^2 \, d\gamma_n(x)\right)^{1/2}$, where the integral is just the product of one-dimensional integrals of Gaussians and can be calculated explicitely, giving the bound $\lambda_n = O(\sqrt{n})$, that is, $\lambda_n \leq C\sqrt{n}$ for some $C > 0$.

It remains to show that asymptotically $\int_{\mathbb{R}^n} f(x)\, d\gamma_n(x) = \Omega(\sqrt{\log n})$. With this purpose, let us estimate the γ_n-volume of the n-dimensional cube with side $\alpha > 0$:

$$\begin{aligned} \int_{\|x\|_\infty \leq \alpha} 1 \cdot d\gamma_n(x) &= \int_{-\alpha}^{\alpha}\int_{-\alpha}^{\alpha} \cdots \int_{-\alpha}^{\alpha} \varepsilon^{-\pi \sum_{i=1}^n t_i^2}\, d_1 \ldots dt_n \\ &= \left(\int_{-\alpha}^{\alpha} e^{-\pi t^2} dt\right)^n \\ &= \left(1 - 2\int_{\alpha}^{\infty} e^{-\pi t^2} dt\right)^n \\ &\leq \left(1 - c \cdot e^{-\pi\alpha^2}\right)^n, \end{aligned} \tag{1.9}$$

where the last inequality for sufficiently large values of α can be verified by integrating the inequality $e^{-\pi t^2} \geq 2te^{-\pi t^2}$ from α to ∞, while for smaller values of α the constant $c > 0$ fills the gap.

Substituting $\alpha = \delta\sqrt{\log n}$ into the last expression in Eq. (1.9), one gets a quantity on the order of $\Theta(e^{-c/\pi delta^2})$, which can be made asymptotically less than than $1/2$ by selecting a suitably small $\delta > 0$. We conclude: the γ_n-volume of the cube of radius $\delta\sqrt{\log n}$ around zero with regard to the ℓ^∞-norm is less than half. Consequently, the measure of the complement to this cube is at least half, and the average value of the norm $\|\cdot\|_\infty$ with regard to the Gaussian measure is, asymptotially, of the order $\Omega(\sqrt{\log n})$, as required. $\square$

PROOF OF DVORETZKY'S THEOREM 0.0.1. Lemma 1.4.14 and Exercise 1.4.11 imply a lower bound of $\Omega(\sqrt{\log n/n})$ for the *median* value of the ℓ^∞ norm on the unit sphere of $\ell^2(n)$. Given an n-dimensional normed space E, Lemma 1.4.13 allows, without loss in generality and cutting the dimension by only a half, apply the above observation to get the same bound on the median value of $\|\cdot\|_E$ on the unit sphere of an equivalent Euclidean norm $|\cdot| \geq \|\cdot\|_E$.

Now replace the norm $\|\cdot\|_E$ with its multiple $\|\cdot\|' = c(\sqrt{n}/\sqrt{\log n})\,\|\cdot\|_E$ for a suitably chosen $c > 0$ so that the median value M of $\|\cdot\|'$ on the Euclidean unit sphere satisfies $M \geq 1$.

The Lipschitz constant of the function $\|\cdot\|'$ on the Euclidean sphere does not exceed $c(\sqrt{n}/\sqrt{\log n})$. By Lemma 1.4.6 and Exercise 1.4.7, there is a further Euclidean subspace V of dimension $k \geq c\varepsilon^2 n/|\log\varepsilon|$ on whose unit sphere the norm $\|\cdot\|'$ differs from M by less than ε. Since $M \geq 1$, the multiplicative Banach-Mazur distance between the normed space E and $\ell^2(k)$ does not exceed ε. □

Our presentation essentially follows Milman's proof of the Dvoretzky theorem [**147**] in a form it is presented in [**60**]. We were using [**153**], sections 3–5, and [**20**], section 12.1.

Among things to note, is a remarkable equilibrium between concentration and "complexity" (metric entropy of the sphere), cf. Exercise 1.4.7 and results leading up to it. This effect is often stressed by Vitali Milman: the two numbers are on the same order of magnitude and keep each other in check. One would expect that sooner or later this effect will be put to use in the dynamical theory forming the subject of the present lectures, but to the date it did not happen.

1.5. Second example: finite Ramsey theorem

Denote by S_∞ the infinite symmetric group, that is, the group of all self-bijections of a countably infinite set ω. For every $n \in \mathbb{N}_+$, let $[\omega]^n$ denote the set of all n-subsets of ω, that is, of all subsets $A \subset \omega$ of cardinality exactly n. The group S_∞ acts on $[\omega]^n$ by permutations in a natural fashion: for each $\tau \in S_\infty$ and $A \in [\omega]^n$, $\tau(A) = \{\tau(a) \colon a \in A\}$.

Notice that the action of S_∞ on $\omega = [\omega]^1$ is *ultratransitive,* that is, every bijection between two finite subsets $A, B \subset \omega$ of the same cardinality extends to a global bijection $\tau \in S_\infty$.

Equip the set $[\omega]^n$ with a discrete uniform structure, that is, a uniform structure $\mathcal{U}_\omega$ containing the diagonal as an element. This uniform structure is generated by a discrete metric:

$$d(x,y) = \begin{cases} 1, & \text{if } x = y, \\ 0, & \text{otherwise.} \end{cases}$$

The pair $([\omega]^n, S_\infty)$ forms a uniform G-space.

THEOREM 1.5.1. *For every $n \in \mathbb{N}_+$, the pair $([\omega]^n, S_\infty)$ is finitely oscillation stable.*

The proof is based on, and is in fact equivalent to, the following classical result. Recall that a *colouring* of a set A with k colours is a map $c \colon A \to [k]$, where $[k] = \{1, 2, \dots, k\}$ is a k-element set.

THEOREM 1.5.2 (Finite Ramsey Theorem). *Let n, m and k be natural numbers. There exists a natural number $N = N(n,m,k)$ such that for every set A of* *cardinality N and each colouring c of $[A]^n$ with k colours there is a subset $B \subseteq A$ of cardinality m such that $[B]^n$ is monochromatic: $c|_{[B]^n} = const$.*

Finite Ramsey Theorem (1.5.2) $\implies$ Theorem 1.5.1. We will use finite oscillation stability in the form of the equivalent condition (9) of Theorem 1.1.18. Since the uniform structure on $[\omega]^n$ is discrete, it is enough to set $V = \Delta$. Let $k \in \mathbb{N}$, and let $F \subseteq [\omega]^n$ be finite. Denote $\operatorname{supp} F = \cup_{x \in F} x \subset \omega$ and let $m = |\operatorname{supp} F|$. Choose $N = N(n, m, k)$ as in Finite Ramsey Theorem. Let $A \subseteq \omega$ be any subset of cardinality N, and let $K = [A]^n$. Now consider an arbitrary colouring c of K with k colours. According to Theorem 1.5.2, there is a subset $B \subseteq A$ of cardinality m and such that $[B]^n$ is monochromatic. Since the action of S_∞ on ω is ultratransivite, there is a $\tau \in S_\infty$ such that $\tau(\operatorname{supp} F) = B$, and consequently $\tau(F) \subseteq [B]^n$ is monochromatic. In particular, $\tau(F)$ is monochromatic to within Δ, verifying condition (9). □

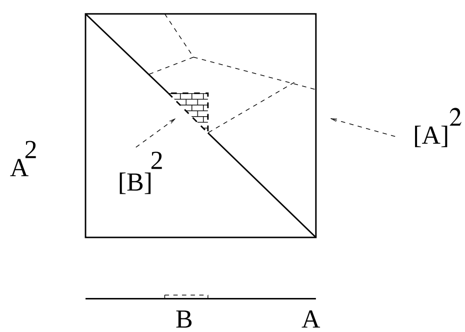

Figure 1.7. To the Finite Ramsey Theorem ($n = 2$, $k = 4$).

Remark 1.5.3. The only property of the symmetric group S_∞ that we have needed in the above proof is the transitivity of the action of S_∞ on $[\omega]^m$ for an arbitrarily large m. Therefore, Theorem 1.5.1 can be extended as follows.

Theorem 1.5.4. *Let G be a subgroup of the symmetric group S_∞ acting ultratransitively on ω. Then the pair $([\omega]^n, G)$ is finitely oscillation stable for every $n \in \mathbb{N}$.* □

Example 1.5.5. Here is one important example of such an acting group. Fix on ω a linear order of type η, that is, a dense linear order without the first and the last elements. Equipped with this order, ω is order-isomorphic to the rational numbers $\mathbb{Q}$ equipped with the usual order. Denote by $\operatorname{Aut}(\mathbb{Q}, \leq)$ the group of all order-preserving bijections of the rationals. Clearly, the action of this group on the rationals is ultratransitive.

Corollary 1.5.6. *For every $n \in \mathbb{N}_+$, the pair $([\mathbb{Q}]^n, \operatorname{Aut}(\mathbb{Q}, \leq))$ is finitely oscillation stable.* □

In the remaining part of this Section we will show that oscillation stability of the pair $([\omega]^n, S_\infty)$ indeed implies the Ramsey theorem, as well as sketch a proof of the latter.

THEOREM 1.5.1 $\implies$ FINITE RAMSEY THEOREM. Assume that $([\omega]^n, S_\infty)$ is finitely oscillation stable. Let m, k be natural numbers, and let $C \subset \omega$ be any subset with m elements. Apply the condition (9) of Theorem 1.1.18 to k as above, $X = [\omega]^n$, $F = [C]^n$, and $V = \Delta$ being the diagonal of $[\omega]^n$. There exists a finite subset $K \subseteq [\omega]^n$ such that for every colouring of K with k colours there is a $\tau \in S_\infty$ with the properties that $\tau(F) \subseteq K$ and $\tau(F)$ is "monochromatic to within Δ," that is, monochromatic. We will show that $N = |\mathrm{supp}\, K|$ has the required property.

Let $A \subset \omega$ be any subset of cardinality N, and let $c\colon A \to [k]$ be a colouring. Because of ultratransitivity of the action of S_∞ on ω, there is a $\sigma \in S_\infty$ with the property $\sigma(\mathrm{supp}\, K) = A$. Denote by the same letter σ the induced mapping $K \to [A]^n$. By the choice of K, there is a $\tau \in S_\infty$ with the properties that $\tau([C]^n) \subseteq K$ and $\tau(F)$ is monochromatic with respect to the colouring $c \circ \sigma$ of K. Equivalently, if one denotes $B = \sigma \circ \tau(C)$, one has $B \subseteq A$, $|B| = m$, and $[B]^n$ is c-monochromatic, as required. □

Now we will outline, following the book [**80**] (which is a nice introduction to Ramsey theory), a proof of Ramsey Theorem, in its infinite version.

THEOREM 1.5.7 (Infinite Ramsey Theorem). *Let k be a natural number. For every finite colouring of $[\omega]^k$ there exists an infinite subset $A \subseteq X$ such that the set $[A]^k$ is monochromatic.*

(One says that such a set A is *homogeneous*.)

EXAMPLE 1.5.8. For $k = 1$ the statement is simply the familiar pigeonhole principle: if one partitions an infinite set into finitely many subsets, then one of them ought to be infinite.

EXAMPLE 1.5.9. Here is a popular reformulation of the result in the case $k = 2$. Among infinitely many people, either there exists an infinite subset such that everyone in this subset knows each other, or there exists an infinite subset such that no two persons in this subset know each other.

REMARK 1.5.10. One deduces the finite version of the Ramsey theorem 1.5.2 from the infinite version via the same routine ultrafilter construction used by us in the proof of Theorem 1.1.18 (implication (8) $\implies$ (10)). To avoid repeating the same construction all over again, just notice that the Infinite Ramsey Theorem implies the oscillation stability of the pair $([\omega]^n, S_\infty)$ (and thus Theorem 1.5.1) through an argument exactly similar to our deduction 1.5.2 $\implies$ 1.5.1 above, but using condition (7) of Theorem 1.1.18, rather than (9).

PROOF OF THE INFINITE RAMSEY THEOREM 1.5.7. We will give a proof in the case $k = 2$, letting the reader to figure out how to handle further induction over k.

Let χ be an r-colouring of $[X]^2$. We will think of the latter set as the set of all edges of the complete graph $K_{\mathbb{N}}$ on natural numbers as vertices.

Infinitely many edges incident to 0 have the same colour, say c_1 (an application of the pigeonhole principle); denote

$$X = \{n \in \mathbb{N}\colon n > 0, \chi(\{0, n\}) = c_1\} = \{x_1, x_2, \dots\},$$

where $x_1 < x_2 < \dots$.

Infinitely many of the edges of the form $\{x_1, x_k\}$, where $k > 1$, have the same colour, say c_2. Denote

$$Y = \{x \in X : x > x_1, \chi(\{x_1, x\}) = c_2\} = \{y_1, y_2, \dots\},$$

where $y_1 < y_2 < \dots$.

Now, infinitely many edges incident to y_1 and a vertex from Y have the same colour, say c_3, and we let

$$Z = \{y \in Y : y > y_1, \chi(\{y_1, y\}) = c_3\} = \{z_1, z_2, \dots\},$$

where $z_1 < z_2 < \dots$.

Denote now $T = \{0, x_0, y_0, z_0, \dots\}$. This is an infinite set. Notice that for a pair of elements $a, b \in T$, the colour of the edge $\{a, b\}$ only depends on $\min\{a, b\}$. Indeed, by the choice of our sets $X, Y, Z, \dots$, the colour of each edge joining a minimal element of the form $0, x_0, y_0, \dots$ with any other element in T is completely determined by the minimal element: e.g., for any $t \in T$, $\chi(\{0, t\}) = c_1$, $\chi(\{x_1, t\}) = c_2$, and so forth.

Thus, one can define a new colouring of T as follows: $\chi^*(s)$ is the colour of any edge of the form $\{s, t\}$, where $s < t$. This colour is independent of t. By the pigeonhole principle, there is an infinite χ^*-monochromatic subset $A \subseteq T$. Now it is clear that for every $a, b \in A$, the colour of the edge $\{a, b\}$ is the same. □

Note that a straightforward extension of the Infinite Ramsey Theorem to uncountable sets fails.

Theorem 1.5.11 (W. Sierpiński [**198**]). *Let X be a set of cardinality $2^{\aleph_0}$. There exists a colouring $c\colon [X]^2 \to [2]$ without uncountable homogeneous subsets $A \subseteq X$ (that is, such A that the set $[A]^2$ is monochromatic).*

If one is allowed to use countably infinitely many colours, the proof is immedaite.

Exercise 1.5.12. Prove a weaker version of Sierpiński's theorem (a colouring with countably many colours) by identifying X with the set of all 0-1 sequences, and for each pair of distinct sequences considering the first position where they differ.

Sierpiński's argument, also surprisingly simple, provides however a deeper insight into the nature of things. The following construction is known as *Sierpiński's partition argument.* We will use it again in Chapter 8.

Proof of Theorem 1.5.11. Let $<$ and $\prec$ be two total orders on X. Denote their characteristic functions by $\chi_<$ and $\chi_\prec$, respectively. The function

$$\chi_< + \chi_\prec \mod 2$$

on X^2 is invariant with regard to the flip $(x, y) \mapsto (y, x)$, and so gives rise to a function $c\colon [X]^2 \to \{0, 1\}$; this is a colouring.

Let $A \subseteq X$ be a homogeneous subset. Then the restrictions $< |_A$ and $\prec |_A$ are either isomorphic or anti-isomorphic between themselves. It is, therefore, enough to find two orders on X whose restrictions to every uncountable subset are neither isomorphic nor anti-isomorphic, or, as they say, *incompatible* orders.

In our case, an obvious choice is the usual order on the real line $X = \mathbb{R}$ and a minimal well-ordering. □

1.6. Counterexample: ordered pairs

Here is the third and last major example illustrating the concept of finite oscillation stability. This time, we will show a G-space that is not finitely oscillation stable.

The space X is $\omega^2 \setminus \Delta_\omega$, that is, the set of all ordered pairs of distinct elements of ω:

$$X = \{(x, y) \colon x, y \in \omega,\ x \neq y\}.$$

We equip X with a discrete uniformity. The group $G = S_\infty$ is acting upon X by double permutations:

$$\tau(x, y) = (\tau x, \tau y).$$

Let $<$ denote the usual linear order on ω identified with the set of natural numbers, and let $\chi_<$ be the characteristic function of the corresponding order relation:

$$\chi_<(x, y) = \begin{cases} 1, & \text{if } x < y, \\ 0, & \text{otherwise.} \end{cases}$$

The function $\chi_<$ is bounded and uniformly continuous as a matter of course. At the same time, $\chi_<$ is not finitely oscillation stable. To see this, define a finite subset of $X = \omega^2 \setminus \Delta_\omega$ by

$$F = \{(0, 1), (1, 0)\}.$$

For every $\tau \in S_\infty$, the image $\tau(F)$ is of the form

$$\tau(F) = \{(a, b), (b, a)\},\ a \neq b,$$

and therefore

$$\operatorname{osc}(f|\tau(F)) = 1.$$