Sums, Products and Projections of Discretized Sets

Weikun He

NNT : 2017SACLS335

THÈSE DE DOCTORAT

de

L'UNIVERSITÉ PARIS-SACLAY

École doctorale de mathématiques Hadamard (EDMH, ED 574)

Établissement d'inscription : Université Paris-Sud

Laboratoire d'accueil : Laboratoire de mathématiques d'Orsay, UMR 8628 CNRS

Spécialité de doctorat : Mathématiques fondamentales

Weikun HE

Sommes, produits et projections des ensembles discrétisés

Date de soutenance : 22 septembre 2017

Après avis des rapporteurs :

JULIEN BARRAL (Université Paris 13)
ELON LINDENSTRAUSS (Université Hébraïque de Jérusalem)

Jury de soutenance :

JULIEN BARRAL	(Université Paris 13) Rapporteur
EMMANUEL BREUILLARD	(Université de Münster) Directeur de thèse
BERNARD HOST	(Université Paris-Est) Examinateur
FRÉDÉRIC PAULIN	(Université Paris-Sud) Examinateur
STÉPHANE SEURET	(Université Paris-Est) Président du jury
PÉTER VARJÚ	(Université de Cambridge) Directeur de thèse

Sums, products and projections of discretized sets

Weikun He

Thesis prepared under the supervision of
Emmanuel Breuillard and Péter Varjú

Abstract

In the discretized setting, the size of a set is measured by its covering number by δ-balls (a.k.a. metric entropy), where δ is the scale. In this document, we investigate combinatorial properties of discretized sets under addition, multiplication and orthogonal projection. There are three parts. First, we prove sum-product estimates in matrix algebras, generalizing Bourgain's sum-product theorem in the ring of real numbers and improving higher dimensional sum-product estimates previously obtained by Bourgain-Gamburd. Then, we study orthogonal projections of subsets in the Euclidean space, generalizing Bourgain's discretized projection theorem to higher rank situations. Finally, in a joint work with Nicolas de Saxcé, we prove a product theorem for perfect Lie groups, generalizing previous results of Bourgain-Gamburd and Saxcé.

Résumé

Dans le cadre discrétisé, la taille d'un ensemble à l'échelle δ est évaluée par son nombre de recouvrement par δ-boules (également connu sous le nom de l'entropie métrique). Dans cette thèse, nous étudions les propriétés combinatoires des ensembles discrétisés sous l'addition, la multiplication et les projections orthogonales. Il y a trois parties principales. Premièrement, nous démontrons un théorème somme-produit dans les algèbres de matrices, qui généralise un théorème somme-produit de Bourgain concernant l'anneau des réels. On améliore aussi des estimées somme-produit en dimension supérieure obtenues précédemment par Bougain et Gamburd. Deuxièmement, on étudie les projections orthogonales des sous-ensembles de l'espace euclidien et étend ainsi le théorème de projection discrétisé de Bourgain aux projections de rang supérieur. Enfin, dans un travail en commun avec Nicolas de Saxcé, nous démontrons un théorème produit dans les groupes de Lie parfaits. Ce dernier résultat généralise les travaux antérieurs de Bourgain-Gamburd et de Saxcé.

Remerciement

Je voudrais remercier tout d'abord mon directeur Emmanuel Breuillard pour sa disponibilité et sa patience durant toutes ces années, pour m'avoir proposé un sujet de recherche très passionnant, pour ses remarques éclairantes et pour m'avoir fait découvrir de nouveaux horizons mathématiques.

I would like to thank my second advisor Péter Varjú for his help (despite being his "half student", I enjoyed from him the full attention of an advisor), for guiding me through my research, for sharing his insights and for correcting carefully my manuscripts filled with errors.

It is an honor to have Julien Barral and Elon Lindenstrauss as the referees of my thesis. I am grateful for their careful reading. The corrections they suggested improved greatly the quality of this memoir.

Je voudrais également remercier très chaleureusement Julien Barral, Bernard Host, Frédéric Paulin et Stéphane Seuret pour avoir accepté de faire partie de mon jury de thèse.

Cela a été un plaisir de collaborer avec Nicolas de Saxcé. Je le remercie pour de nombreuses discussions très enrichissantes, pour son enthousiasme et son encouragement.

Je remercie aussi Arindam, Çagrı, Mikołaj, Lison, Kajal, Matthew, Richard, Jonas, Shu[1], Yeping[2], Bingxiao[3], Wei Guo, Zicheng[4], Camille, Davi, Gabriel et Maxence pour des discussions inspirantes autour des groupes, la géométrie et la dynamique.

Cette thèse a été préparée à l'Université Paris-Sud, je suis reconnaissant envers mesdames Blandin-Lavigne, Jacquemin, Rey ainsi que mesdames Mignier, Rigal, Roussas et Toro pour leur aide dans les démarches administratives.

Enfin, je remercie mes ami-e-s doctorant-e-s (les noms de certains d'entre eux sont déjà apparus un peu plus haut) Xiaodong[5], Yi[6], Eddie, Benjamin, Guillaume, Émilien, Vincent, Linxiao[7], Salim, Sasha, Thibault, Ruoci[8], Cong[9], Lucile, Tiago, Yang[10], Joseph, Robert, Maxime, Mor, Anthony, Jeanne, Luc, Thomas, Yi[11], ... pour avoir partagé le bureau, les repas au CESFO, les cafés/thés autour des énigmes mathématiques, les sorties sportives et surtout la passion pour les mathématiques.

Mes derniers remerciements vont à ma famille : 最后，感谢我的家人，爸爸、妈妈和伟鹏，对我一直以来的支持与鼓励。

[1] 申述 [2] 张野平 [3] 刘冰萧 [4] 钱子诚 [5] 王晓东 [6] 黄益 [7] 陈林晓 [8] 孙若词 [9] 薛聪 [10] 曹阳 [11] 潘亿

Contents

Chapitre 0

Introduction en français

Pour un ensemble A, notons $|A|$ son cardinal. On utilise les notations de Landau $O(f)$ et les notations de Vinogradov $f \ll g$ tout au long de ce mémoire. De plus, nous écrivons $f \asymp g$ pour dire $f \ll g$ et $g \ll f$.

0.1 La notion d'ensembles discrétisés

Depuis les travaux de Katz-Tao [38], la notion d'ensembles discrétisés est connue pour être un cadre général pour étudier les propriétés fractales des ensembles continus tout en utilisant les outils combinatoires comme la combinatoire arithmétique. Dans cette section, nous allons introduire les notions de base.

0.1.1 Le nombre de δ-recouvrement

Soit (E, d) un espace métrique. Pour $\rho > 0$ et $x \in E$ notons $\mathbf{B}(x, \rho)$ la boule fermée de centre x et de rayon ρ. On écrit $\mathbf{B}_E(x, \rho)$ quand on veut préciser dans quel espace ambiant se trouve la boule. Pour un sous-ensemble $A \subset E$, notons $A^{(\rho)}$ le ρ-voisinage de A :

$$A^{(\rho)} = \{x \in E \mid d(x, A) \leq \rho\}.$$

Un sous-ensemble A est dit *ρ-séparé* si pour tout $a \in A$, $A \cap \mathbf{B}(a, \rho) = \{a\}$. Tout au long de ce mémoire, la variable δ désigne un réel strictement positif que l'on appelle *l'échelle*.

Définition. Soit A un sous-ensemble relativement compact de (E, d). Son nombre de δ-recouvrement, également connu sous le nom de l'entropie métrique, est défini par

$$\mathcal{N}_\delta(A) = \min\Big\{N \geq 0 \mid \exists x_1, \dots, x_N \in E,\ A \subset \bigcup_{i=1}^{N} \mathbf{B}(x_i, \delta)\Big\}.$$

Exemple. Soient p un nombre premier et $(\mathbb{Z}_p, |\cdot|_p)$ l'anneau des entiers p-adiques muni de sa valeur absolue usuelle. Si A est un sous-ensemble de $\mathbb{Z}_p$, alors pour tout $k \in \mathbb{N}$,

$$\mathcal{N}_{p^{-k}}(A) = |\pi_k(A)|$$

où $\pi_k \colon \mathbb{Z}_p \to \mathbb{Z}_p/p^k\mathbb{Z}_p \simeq \mathbb{Z}/p^k\mathbb{Z}$ désigne la réduction modulo p^k.

Si nous permettons à δ d'être 0, alors $\mathcal{N}_0(A)$ est exactement le cardinal de A. Il est plus intéressant d'étudier le comportement asymptotique de $\mathcal{N}_\delta(A)$ quand δ tend vers 0.

Exemple. Les dimensions de Minkowski inférieure et supérieure de $A \subset (E, d)$ peuvent être définies par les formules suivantes, voir [45, §5.3].

$$\underline{\dim}_{\mathrm{M}}(A) = \liminf_{\delta \to 0} \frac{\log \mathcal{N}_\delta(A)}{-\log \delta} \quad \text{et} \quad \overline{\dim}_{\mathrm{M}}(A) = \limsup_{\delta \to 0} \frac{\log \mathcal{N}_\delta(A)}{-\log \delta}.$$

En particulier, si $\mathcal{N}_\delta(A) \asymp_{\delta \to 0} \delta^{-\alpha}$ alors elles coïncident et sont égales à α.

Maintenant supposons que E est l'espace euclidien $\mathbb{R}^n$, $n \geq 1$. Notons λ la measure de Lebesgue sur $\mathbb{R}^n$. Outre le nombre de δ-recouvrement, nous avons d'autre moyen de mesurer la taille d'un ensemble à l'échelle δ.

Lemma 2.1. *Soient $\delta > 0$ et A un sous-ensemble borné de $\mathbb{R}^n$. Soit $\tilde{A}$ un sous-ensemble 2δ-séparé maximal de A. Alors*

$$\mathcal{N}_{2\delta}(A) \leq |\tilde{A}| \leq \mathcal{N}_\delta(A) \leq \mathcal{N}_1(\mathbf{B}(0,2))\, \mathcal{N}_{2\delta}(A),$$

et

$$\delta^{-n} \lambda(A^{(\delta)}) \asymp_n \mathcal{N}_\delta(A).$$

En particulier, le fait de changer l'échelle par un facteur constant ne change le nombre de δ-recouvrement que par un facteur constant. De plus, A et son δ-voisinage $A^{(\delta)}$ ont à peu près la même taille à l'échelle δ :

$$\mathcal{N}_\delta(A^{(\delta)}) \asymp_n \mathcal{N}_\delta(A).$$

Et c'est pour cette raison-là que dans [38], un ensemble δ-discrétisé est défini comme une réunion de boules de rayon δ. Dans les démonstrations de ce mémoire, nous pouvons presque toujours remplacer A par son δ-voisinage.

Exemple. Soit μ une mesure de probabilité sur $\mathbb{R}^n$ à support compact. Sa transformée de Fourier–Stieltjes $\hat{\mu}(\xi) = \int e^{-i\langle \xi, x \rangle} \, \mathrm{d}\mu(x)$, $\xi \in \mathbb{R}^n$, est K-lipschitzienne pour un certain $K > 0$ dépendant seulement de $\max_{x \in \mathrm{Supp}(\mu)} \|x\|$. Par conséquent, si l'on pose, pour $t > 0$,

$$A_t = \{\xi \in \mathbb{R}^n \mid |\hat{\mu}(\xi)| \geq t\},$$

alors $A_{2t}^{(\frac{t}{K})} \subset A_t$. Cela montre qu'il est naturel de regarder l'ensemble A_t à l'échelle t.

0.1.2 Non-concentration

Si un ensemble borné $A \subset \mathbb{R}^n$ a la taille $\mathcal{N}_\delta(A) \asymp \delta^{-\alpha}$, alors on a envie de penser α comme étant la dimension de A. Mais, une boule de rayon $\delta^{1-\frac{\alpha}{n}}$ vérifie aussi $\mathcal{N}_\delta(\mathbf{B}(0, \delta^{1-\frac{\alpha}{n}})) \asymp \delta^{-\alpha}$. Évidemment, une boule a des propriétés très différentes de celles d'une fractale. Pour éviter de telles situations dégénérées, nous imposons souvent une propriété de non-concentration sur l'ensemble A. Soient $\kappa > 0$ et $\epsilon > 0$ des paramètres. Par exemple, on pourrait demander

$$\forall \rho \geq \delta, \quad \mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}. \tag{0.1}$$

Parfois, nous avons besoin d'une condition plus forte comme suit.

$$\forall \rho \geq \delta,\ \forall x \in E, \quad \mathcal{N}_\delta(A \cap \mathbf{B}(x,\rho)) \leq \delta^{-\epsilon}\rho^\kappa \mathcal{N}_\delta(A). \tag{0.2}$$

Dans ces conditions, plus κ est grande, plus la non-concentration est forte. La notion de $(\delta,\alpha)_n$-ensemble dans [38] correspond à la condition (0.2) avec κ maximal, i.e. $\kappa = \alpha$. Ici, ϵ est un petit paramètre pour relâcher légèrement ces conditions. Ajouter le terme δ^ϵ est de dire que la non-concentration est exigée de l'échelle δ jusqu'à l'échelle δ^ϵ. Clairement, la condition (0.2) implique (0.1). La réciproque n'est pas vraie. Néanmoins, si l'on se permet de passer dans des sous-ensembles relativement grandes, on peut extraire un sous-ensemble ayant la propriété (0.2) d'un ensemble ayant la propriété (0.1).

Exemple. Soit A un sous-ensemble borélien borné de $\mathbb{R}^n$ de dimension de Hausdorff $\dim_{\mathrm{H}} A > \alpha$. Soit μ la mesure donnée par le lemme de Frostman.

Théorème 0.1 (Lemme de Frostman [45, Theorem 8.8])**.** *Soit A un sous-ensemble borélien de $\mathbb{R}^n$. Si $\dim_{\mathrm{H}}(A) > \alpha$ alors il existe une mesure borélienne finie non-nulle μ de support compact telle que $\mathrm{Supp}(\mu) \subset A$ et*

$$\forall \rho > 0,\ \forall x \in \mathbb{R}^n, \quad \mu(\mathbf{B}(x,\rho)) \leq \rho^\alpha.$$

Il est alors immédiat que pour tout $\rho > 0$,

$$\mathcal{N}_\rho(A) \geq \mu(A)\rho^{-\alpha}.$$

C'est-à-dire, nous avons la non-concentration (1.1) avec $\kappa = \alpha$.

En guise de conclusion, dans ce mémoire, un ensemble δ-discrétisé veut dire (parfois informellement) un ensemble dont la taille est mesurée par le nombre de δ-recouvrement et qui vérifie en plus une condition de non-concentration du type (0.1) ou (0.2). Le sujet de cette thèse est d'étudier le comportement (surtout la croissance) des ensembles discrétisés sous des opérations telles que l'addition, la multiplication, l'action d'un groupe ou les projections orthogonales. Dans les sections suivantes, nous discuterons selon l'opération qui nous intéresse.

0.2 La théorie additive

Dans cette section, supposons que l'espace ambiant E est muni d'une structure de groupe abélien dont la loi est notée additivement. Soient $A, B \subset E$ des sous-ensembles. Notons

$$-A = \{-a \mid a \in A\},$$
$$A + B = \{a + b \mid a \in A, b \in B\}$$

et

$$A - B = \{a - b \mid a \in A, b \in B\}.$$

Soit s un entier naturel non-nul. Le notation sA désigne la somme de A avec lui-même s fois, $A + \cdots + A$. Par la théorie additive, on entend l'étude de la croissance de sA ou bien celle de $sA - s'A$, $s, s' \in \mathbb{N}$.

Considérons $E = \mathbb{R}^n$ dans le cadre discrétisé. En approchant $\mathbb{R}^n$ par le réseau $\delta \cdot \mathbb{Z}^n$, on peut transférer la théorie additive du cadre discret vers le cadre δ-discrétisé. Par exemple, on a l'inégalité de Plünnecke-Ruzsa.

Lemma 2.4 (Inégalité de Plünnecke-Ruzsa). *Pour tout paramètre $K \geq 1$, si $\mathcal{N}_\delta(A+B) \leq K\mathcal{N}_\delta(B)$ alors pour tous les entiers $k \geq 1$ et $l \geq 0$,*

$$\mathcal{N}_\delta(kA - lA) \ll_n K^{k+l}\mathcal{N}_\delta(B).$$

En particulier, pour tout $s \geq 2$, $\mathcal{N}_\delta(sA)$ est considérablement plus grand que $\mathcal{N}_\delta(A)$ si et seulement si $\mathcal{N}_\delta(A+A)$ est déjà considérablement plus grand que $\mathcal{N}_\delta(A)$. C'est pour cela qu'en combinatoire additive, une question centrale est de classer les ensembles avec petit doublement (i.e. ensembles A tels que $|A+A| \leq K|A|$). Le théorème de Freiman affirme qu'un tel ensemble est forcément contenu dans une progression arithmétique généralisée de taille comparable (voir [59, Chapter 5]). Transféré dans le cadre discrétisé, ce théorème devient

Théorème 0.2 (Version discrétisée du théorème de Freiman, [57, Proposition 7.3]). *Soit $K \geq 2$ un paramètre. Soit A un sous-ensemble borné de $\mathbb{R}^n$. Si $\mathcal{N}_\delta(A+A) \leq K\mathcal{N}_\delta(A)$ alors il existe un ensemble P qui est une somme de $O_{K,n}(1)$ progressions arithmétiques dans $\mathbb{R}^n$ et tel que $A \subset P + \mathbf{B}(0,\delta)$ et $|P| \asymp_{K,n} \mathcal{N}_\delta(A)$.*

L'un des problèmes fondamentaux en combinatoire additive est d'obtenir une bonne dépendance explicite en K des constantes implicites dans le théorème précédent.

Dans le sens inverse, tout résultat démontré dans le cadre discrétisé entraînera sa contrepartie dans le cadre discret. Dans ce mémoire, nous nous contentons des résultats existants et ne démontrons rien de nouveau pour la théorie additive.

0.3 Estimées somme-produit

Maintenant, supposons que notre espace ambiant E est en plus muni d'une structure d'anneau. Soient $A, B \subset E$ des sous-ensembles. Notons

$$A \cdot B = \{ab \mid a \in A, b \in B\}.$$

Soit $s \in \mathbb{N}^*$. On écrit A^s pour le produit itéré de A avec lui-même s fois, $A \cdots A$. De plus, on définit récursivement $\langle A \rangle_1 = A \cup (-A)$ et $\forall s \in \mathbb{N}^*$,

$$\langle A \rangle_s = \langle A \rangle_{s-1} \cup \bigcup_{k=1}^{s-1} \big(\langle A \rangle_k + \langle A \rangle_{s-k}\big) \cup \bigcup_{k=1}^{s-1} \big(\langle A \rangle_k \cdot \langle A \rangle_{s-k}\big).$$

Autrement dit, $\langle A \rangle_s$ est l'ensemble des éléments qui peuvent être obtenus en additionnant et multipliant au plus s éléments de $A \cup (-A)$.

Grosso modo, le problème somme-produit demande, étant donné un sous-ensemble de E, s'il croît vite sous l'addition, la soustraction et la multiplication et si ce n'est pas le cas, quels sont les obstructions. Ainsi, une estimée somme-produit est une minoration pour la taille de $\langle A \rangle_s$ avec $s \geq 2$. Le premier résultat de ce genre est dû à Erdős et Szemerédi [28] pour l'anneau des réels $\mathbb{R}$.

Théorème 0.3 (Erdős-Szemerédi [28]). *Il existe une constante absolue $c > 0$ telle que pour tout sous-ensemble fini non-vide A de $\mathbb{R}$, on ait*

$$\max(|A+A|, |A \cdot A|) \geq c|A|^{1+c}.$$

Cela est le point de départ de beaucoup de travaux qui ou bien établissent des estimées similaires pour des classes plus larges d'anneaux, ou bien améliorent des bornes existantes. Voir [58] pour une approche élégante et plus d'histoire.

Bien que le théorème de Erdős-Szemerédi concerne les réels[1], il ne s'agit pas d'un résultat dans le cadre discrétisé. À la différence de la théorie additive, l'étude du phénomène somme-produit dans le cadre discrétisé diffère de son analogue discret en plusieurs aspects. Par exemple, la multiplication par δ préserve le cardinal alors qu'elle réduit le nombre de δ-recouvrement de tout ensemble borné à une constante. De plus quand x varie de δ à 1, son comportement varie continûment entre celui d'un diviseur de zéro et celui d'un élément inversible. Malgré cela, on espère qu'un ensemble discrétisé typique croît très vite sous l'addition et la multiplication. Pour l'anneau $\mathbb{R}$, cela a été conjecturé par Katz et Tao dans [38] et résolu par Bourgain [6, 7].

Théorème 0.4 (Théorème somme-produit discrétisé de Bourgain [7])**.** *Étant donnés $\kappa > 0$ et $\sigma < 1$, il existe une constante $\epsilon > 0$ telle que la proposition suivante soit vraie pour $\delta > 0$ suffisamment petit. Soit A un sous-ensemble de $\mathbb{R}$ tel que*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$.

Alors $\mathcal{N}_\delta(A+A) + \mathcal{N}_\delta(A \cdot A) \geq \delta^{-\epsilon}\mathcal{N}_\delta(A)$.

La démonstration de ce théorème utilise une analyse multi-échelle. On étudie d'abord la structure additive à chaque niveau. Cela peut être fait à l'aide du théorème de Freiman rappelé plus haut. Notamment, la première démonstration de Bourgain [6] fait appel à un théorème de Freiman quantitatif dû à Chang [20]. Ensuite, on utilise la multiplication qui interagit entre les différents niveaux pour déduire une croissance.

L'importance du théorème somme-produit de Bourgain peut être justifiée par ses nombreuse applications. Lorsque le cadre discrétisé a été introduit dans [38], il était conçu comme une stratégie générale pour démontrer les résultats dans le régime continu tout en utilisant les idées et les résultats venant du régime discret où l'on dispose de théories bien développées telles que la combinatoire arithmétique. Ainsi, l'une des motivations de Bourgain était d'établir la conjecture de Erdős-Volkmann [29] qui affirme qu'aucun sous-anneau borélien de $\mathbb{R}$ n'est de dimension de Hausdorff strictement entre 0 et 1. La conjecture a aussi été résolue par Edgar et Miller [27] avec d'autres idées. Cependant, leur méthode ne sont pas quantitative, comparée à la méthode de Bourgain. De plus, comme prévu par Katz et Tao [38], le théorème 0.4 fait également progrès sur le problème de distance de Falconer et sur une conjecture de Furstenberg en géométrie fractale. Parmi d'autres applications directes, citons aussi la construction explicite d'expanseurs monotones due à Bourgain et Yehudayoff [15]. Nous discuterons plus d'applications plus tard lorsque nous parlerons des théorèmes de projection et des théorèmes produit.

[1] Dans leur article original, le théorème est énoncé pour l'anneau $\mathbb{Z}$, mais leur démonstration est valable pour tout anneau intègre totalement ordonné.

Bourgain et Gamburd ont obtenu des estimations similaires pour l'anneau $\mathbb{C}$ dans [9] et plus tard pour l'anneau $\mathbb{C}^n$, le produit direct de $\mathbb{C}$ avec lui-même n fois, $n \geq 2$, dans [11]. Les énoncés sont un peu différentes puisque $\mathbb{R}$ n'a pas de sous-algèbre réel propre alors que $\mathbb{C}$ et $\mathbb{C}^n$ en ont. De plus, l'anneau $\mathbb{C}^n$ a des idéaux propres non-triviaux , i.e. il n'est pas simple. Dans le chapitre 3, nous démontrons une estimation somme-produit pour les algèbres simples de dimension finie sur $\mathbb{R}$. Notons que par le théorème de Wedderburn et le théorème de Frobenius, une telle algèbre est isomorphe à $\mathcal{M}_n(\mathbb{R})$, $\mathcal{M}_n(\mathbb{C})$ ou $\mathcal{M}_n(\mathbb{H})$, l'algèbre des matrices de taille $n \times n$ à coefficients réels, complexes ou dans les quaternions, avec $n \geq 1$. Avant d'énoncer le résultat, mentionnons que dans le contexte discret, Chang [21] a étudié le problème somme-produit pour les matrices réelles et Tao [58] a obtenu un théorème somme-produit concernant une algèbre quelconque.

Théorème 3.1. *Soit E une algèbre réelle normée*[2] *simple de dimension finie. Étant donnés $\kappa > 0$ et $\sigma < \dim(E)$, il existe une constante $\epsilon > 0$ dépendant seulement de E, κ et σ telle que la proposition suivante soit vraie pour $\delta > 0$ suffisamment petit. Soit A un sous-ensemble de E satisfaisant*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$,

(iv) pour tout sous-algèbre propre W de E, il existe $a \in A$ tel que $d(a, W) \geq \delta^\epsilon$.

Alors,

$$\mathcal{N}_\delta(A + A) + \mathcal{N}_\delta(A + A \cdot A) \geq \delta^{-\epsilon} \mathcal{N}_\delta(A).$$

La stratégie de la démonstration consiste à produire un tore riche afin d'utiliser l'estimée somme-produit de Bourgain-Gamburd dans $\mathbb{C}^n$. Ici, un tore riche veut dire un sous-ensemble de taille relativement grande constitué d'éléments simultanément diagonalisables. L'idée du tore riche remonte jusqu'aux travaux de Helfgott [36] et s'est montrée très utile pour démontrer des théorèmes produit dans beaucoup de situations, par exemple, le théorème produit discrétisé de Saxcé (théorem 0.7) dont nous parlerons un peu plus tard. En fait, notre méthode donne une stratégie pour obtenir des estimées somme-produit dans les algèbres simples à partir d'une estimée somme-produit dans le corps de base.

Nous obtenons également des estimations concernant la croissance d'un sous-ensemble discrétisé d'un espace euclidien sous l'action linéaire. Soient X un sous-ensemble borné de $\mathbb{R}^n$ et A une collection d'endomorphismes de $\mathbb{R}^n$. On peut se demander si X croît vite sous l'addition et les transformations par les éléments de A, pourvu que A soit assez riche.

Théorème 3.2. *Soit $n \in \mathbb{N}^*$. Étant donnés $\kappa > 0$ et $\sigma < n$, il existe une constante $\epsilon > 0$ telle que la proposition suivante soit vraie pour $\delta > 0$ suffisamment petit. Soient A un sous-ensemble de* $\mathrm{End}(\mathbb{R}^n)$ *et X un sous-ensemble de $\mathbb{R}^n$. Supposons que*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

[2] Par « normée » on entend une norme qui fait de l'espace vectoriel sous-jacent un espace vectoriel normé. Et comme toutes les normes sur un espace vectoriel de dimension finie sont équivalentes, on peut la supposer sous-multiplicative.

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) *pour tout sous-espace propre non-trivial* W *de* $\mathbb{R}^n$, *il existe* $a \in A$ *et* $w \in \mathbf{B}_W(0,1)$ *tel que* $d(aw, W) \geq \delta^\epsilon$,

(iv) $X \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(v) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(X) \geq \delta^\epsilon \rho^{-\kappa}$,

(vi) $\mathcal{N}_\delta(X) \leq \delta^{-\sigma-\epsilon}$.

Alors,

$$\mathcal{N}_\delta(X+X) + \max_{a \in A} \mathcal{N}_\delta(X + aX) \geq \delta^{-\epsilon} \mathcal{N}_\delta(X),$$

où $aX = \{ax \mid x \in X\}$.

Cela améliore un résultat de Bourgain-Gamburd [11, Proposition 1] où on exige une constante à la place de δ^ϵ dans la condition (iii). Cette condition (iii) est une condition d'irréductibilité. Son affaiblissement est le défi technique principal dans la démonstration du théorème 3.2. Le facteur δ^ϵ signifie que la condition doit être vérifiée à l'échelle δ^ϵ. Ainsi, le théorème 3.2 affirme que la croissance a lieu même si l'on ne connaît rien à des échelles entre δ^ϵ et 1. Cette amélioration est importante pour beaucoup d'applications parce que ce genre d'estimations sont souvent utilisées avec le théorème de Balog-Szemerédi-Gowers qui nécessite de restreindre les ensembles avec lesquels on travaille à des sous-ensembles de taille δ^ϵ fois la taille initiale. Cette procédure détruit généralement toutes les informations au-dessus de l'échelle δ^ϵ.

Revenons aux phénomènes somme-produit dans les algèbres. Au lieu de penser la somme et le produit comme étant les lois dans un anneau, on peut penser la somme comme étant l'addition dans un espace vectoriel et le produit comme étant les endomorphismes de cet espace vectoriel obtenus en considérant les multiplications à gauche et à droite. Avec ce point de vue, on retrouve facilement le théorème 3.1 à partir du théorème 3.2. Donc en fait le phénomène « somme-produit » ne concerne pas seulement les anneaux. Le « produit » peut venir de l'extérieur, d'une action. On verra dans le chapitre 5 une estimation de type somme-produit pour les représentations de groupe de Lie.

On peut aussi déduire du théorème 3.2 une estimation « somme-crochet » dans les algèbres de Lie simples. Si A est un sous-ensemble d'une algèbre de Lie $\mathfrak{g}$, on écrit $[A, A] = \{[a, b] \mid a, b \in A\}$.

Corollaire 3.3. *Soit* $\mathfrak{g}$ *une algèbre de Lie réelle de dimension finie qui est en plus munie d'une norme. Étant donnés* $\kappa > 0$ *et* $\sigma < \dim(\mathfrak{g})$, *il existe une constante* $\epsilon > 0$ *telle que la proposition suivante soit vraie pour* $\delta > 0$ *suffisamment petit. Soit* A *un sous-ensemble de* $\mathfrak{g}$ *satisfaisant*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$,

(iv) *pour tout sous-algèbre de Lie propre* W *de* $\mathfrak{g}$, *il existe* $a \in A$ *tel que* $d(a, W) \geq \delta^\epsilon$.

Alors,

$$\mathcal{N}_\delta(A+A) + \mathcal{N}_\delta(A + [A, A]) \geq \delta^{-\epsilon} \mathcal{N}_\delta(A).$$

0.4 Orthogonal projections

Soient $0 < m < n$ des entiers positifs. Soit $\mathrm{Gr}(\mathbb{R}^n, m)$ la grassmannienne des sous-espaces vectoriels de dimension m dans l'espace euclidien $\mathbb{R}^n$. Pour $V \in \mathrm{Gr}(\mathbb{R}^n, m)$, notons $\pi_V \colon \mathbb{R}^n \to V$ la projection orthogonale sur V.

Les propriétés fractales des projections orthogonales ont été beaucoup étudiées en géométrie fractale, en commençant par le théorème de projection de Marstrand [43]. Voir par exemple [31] pour plus d'histoire. De manière informelle, un théorème de projection est une minoration pour la taille des projections d'un sous-ensemble sur de différentes directions. Comme en générale on n'espère pas que la projection soit grande pour toutes les directions, le théorème s'exprime souvent par une majoration de la taille de l'ensemble des directions exceptionnelles, c'est-à-dire, des directions sur lesquelles la projection est petite. Par exemple, rappelons le résultat suivant obtenu dans une série de travaux dûs à Mattila et Falconer.

Théorème 0.5 (Mattila [44], Falconer [32], voir aussi [46, §5.3]). *Soit $A \subset \mathbb{R}^n$ un sous-ensemble borélien de dimension de Hausdorff $\dim_{\mathrm{H}}(A) = \alpha$. Soit $s > 0$ un réel.*

(i) Si $0 < s < \alpha \leq m$, alors

$$\dim_{\mathrm{H}}\{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \dim_{\mathrm{H}}(\pi_V(A)) < s\} \leq m(n-m) - (m-s).$$

(ii) Si $0 < s \leq m \leq \alpha$, alors

$$\dim_{\mathrm{H}}\{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \dim_{\mathrm{H}}(\pi_V(A)) < s\} \leq m(n-m) - (\alpha - s).$$

Notons que $m(n-m)$ est la dimension de $\mathrm{Gr}(\mathbb{R}^n, m)$ et que la notion de la dimension de Hausdorff sur $\mathrm{Gr}(\mathbb{R}^n, m)$ ne dépend pas du choix spécifique de la métrique car toutes les distances sur un compact sont équivalentes tant qu'elles définissent la même topologie.

À l'aide de son théorème somme-produit discrétisé, Bourgain a établi un théorème de projection discrétisé pour les projections de rang 1. Avant d'énoncer le théorème, introduisons quelques notations afin de formuler des conditions de non-concentration dans les espaces projectifs ou dans les grassmanniennes. Pour $V \in \mathrm{Gr}(\mathbb{R}^n, m)$ et $W \in \mathrm{Gr}(\mathbb{R}^n, n-m)$, on définit

$$\mathrm{d}_{\measuredangle}(V, W) = |\det(v_1, \ldots, v_m, w_1, \ldots, w_{n-m})|$$

où $(v_1, \ldots, v_m)$ est une base orthonormée de V et $(w_1, \ldots, w_{n-m})$ une base orthonormée de W. Par exemple, $\mathrm{d}_{\measuredangle}(V, W) = 0$ si et seulement si V et W ont une intersection non-triviale. Pour $\rho \geq 0$, on pose

$$\mathcal{V}_{\measuredangle}(W, \rho) = \{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \mathrm{d}_{\measuredangle}(V, W) \leq \rho\}.$$

Théorème 0.6 (Bourgain [7, Theorem 5]). *Soit $n \geq 2$ un entier. Étant donnés $0 < \alpha < n$ et $\kappa > 0$, il existe $\epsilon > 0$ tel que la proposition suivante soit vraie pour $\delta > 0$ suffisamment petit. Soit $A \subset \mathbf{B}_{\mathbb{R}^n}(0, 1)$. Soit μ une mesure de probabilité sur l'espace projectif $\mathrm{Gr}(\mathbb{R}^n, 1)$. Supposons que*

$$\mathcal{N}_\delta(A) \geq \delta^{-\alpha+\epsilon};$$

$$\forall \rho \geq \delta,\ \forall x \in \mathbb{R}^n, \quad \mathcal{N}_\delta(A \cap \mathbf{B}(x,\rho)) \leq \delta^{-\epsilon}\rho^\kappa \mathcal{N}_\delta(A);$$
$$\forall \rho \geq \delta,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, n-1), \quad \mu(\mathcal{V}_\measuredangle(W,\rho)) \leq \delta^{-\epsilon}\rho^\kappa.$$

Alors il existe un sous-ensemble $A' \subset A$ et un ensemble de direction $\mathcal{D} \subset \mathrm{Gr}(\mathbb{R}^n, 1)$ tels que $\mathcal{N}_\delta(A') \geq \delta^\epsilon \mathcal{N}_\delta(A)$, $\mu(\mathcal{D}) \geq 1 - \delta^\epsilon$ et

$$\mathcal{N}_\delta(\pi_\theta(A'')) \geq \delta^{-\frac{\alpha}{n}-\epsilon}$$

dès que $\theta \in \mathcal{D}$ et que $A'' \subset A'$ vérifie $\mathcal{N}_\delta(A'') \geq \delta^{2\epsilon}\mathcal{N}_\delta(A)$.

Plus tard, dans [48], pour les sous-ensembles de $\mathbb{R}^2$, Orponen a obtenu une borne plus forte sous une codition de non-concentration plus forte. L'approche qui y est utilisée est différente mais fait également appel à la théorie autour du phénomène somme-produit.

Il convient aussi de remarquer qu'il y a un ensemble de problèmes en parallèle concernant les projections des ensembles ou mesures auto-similaires. Par exemple, récemment, une conjecture de Furstenberg concernant transversalité entre les ensembles $\times 2$-invariants et $\times 3$-invariants a été résolue dans deux travaux indépendants, Shmerkin [56] et Wu [61]. L'approche de Shmerkin utilise aussi les outils venant de la combinatoire additive et quelques idées dans la démonstration du théorème somme-produit de Bourgain. Voir aussi [55] pour plus d'informations.

L'objectif du chapitre 4 est de généraliser le théorème 0.6 aux projections de rang supérieur.

Théorème 4.1. *Soient $1 \leq m < n$ des entiers. Étant donnés $0 < \alpha < n$ et $\kappa > 0$, il existe $\epsilon > 0$ tel que la proposition suivante soit vraie pour $\delta > 0$ suffisamment petit. Soient $A \subset \mathbf{B}_{\mathbb{R}^n}(0,1)$ et μ une mesure de probabilité sur $\mathrm{Gr}(\mathbb{R}^n, m)$. Supposons que*

$$\mathcal{N}_\delta(A) \geq \delta^{-\alpha+\epsilon};$$
$$\forall \rho \geq \delta,\ \forall x \in \mathbb{R}^n, \quad \mathcal{N}_\delta(A \cap \mathbf{B}(x,\rho)) \leq \delta^{-\epsilon}\rho^\kappa \mathcal{N}_\delta(A);$$
$$\forall \rho \geq \delta,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, n-m), \quad \mu(\mathcal{V}_\measuredangle(W,\rho)) \leq \delta^{-\epsilon}\rho^\kappa.$$

Alors il existe $\mathcal{D} \subset \mathrm{Gr}(\mathbb{R}^n, m)$ tel que $\mu(\mathcal{D}) \geq 1 - \delta^\epsilon$ et

$$\mathcal{N}_\delta(\pi_V(A')) \geq \delta^{-\frac{m}{n}\alpha-\epsilon}$$

dès que $V \in \mathcal{D}$ et que $A' \subset A$ vérifie $\mathcal{N}_\delta(A') \geq \delta^\epsilon \mathcal{N}_\delta(A)$.

Ce résultat est nouveau pour $m > 1$. En plus, comparée au théorème 0.6, notre borne inférieure est établie pour tout sous-ensemble assez grand de A au lieu de seulement les sous-ensembles assez grand d'une certaine partie de A. La démonstration du théorème 4.1 se fait par une récurrence sur le couple (n, m). L'approche dans la démonstration du théorème 0.6 combinée avec l'estimation somme-produit en dimension supérieure (théorème 3.2) montre le cas où m est un diviseur de n. Ensuite, de nouvelles idées sont utilisées pour ramener à ce cas spécial.

0.4.1 Motivation ergodique

Le théorème de projection discrétisé de Bourgain constitue un ingrédient important dans le théorème de Bourgain-Furman-Lindenstrauss-Mozes [8] sur l'équidistribution des orbites sur le tore $\mathbb{R}^d/\mathbb{Z}^d$ sous les actions des sous-semi-groupes de $\mathrm{SL}_d(\mathbb{Z})$. Ce résultat ergodique est d'autant plus intéressant qu'il est quantitatif. Par exemple, il a permis à Bourgain et Varjú [14] d'établir l'existence de trous spectraux uniformes dans $\mathrm{SL}_d(\mathbb{Z}/q\mathbb{Z})$, avec q entier arbitraire.

Le théorème de Bourgain-Furman-Lindenstrauss-Mozes donne des énoncés qualitatifs sur les mesures stationnaires. Les quatre auteurs ont notamment démontré une propriété de raideur[3]. Dans [4], par les méthodes plus ergodiques, Benoist et Quint ont généralisé ce résultat de « stiffness » à une classe beaucoup plus large de systèmes dynamiques. En particulier, pour l'action linéaire de $\mathrm{SL}_d(\mathbb{Z})$ sur le tore $\mathbb{R}^d/\mathbb{Z}^d$, ils n'ont pas besoin de l'hypothèse de proximalité dans [8]. Cependant, les résultats dans [4] ne sont pas quantitatifs.

Donc si l'on veut des résultats quantitatifs, il faudrait reprendre la méthode de [8]. L'approche dans [8] est Fourier-analytique. Lorsqu'un sous-groupe $\Gamma \subset \mathrm{SL}_d(\mathbb{Z})$ agit sur le tore, son transposé ${}^t\Gamma$ agit sur les coefficients de Fourier. Une grande partie de la démonstration se concentre sur l'étude des grands coefficients de Fourier sous cette action. Par la théorie des produits aléatoires de matrices, si Γ est proximal, alors les grands produits aléatoires dans Γ se comporte comme des projections de rang 1 composées avec des rotations, s'ils sont observés à la bonne échelle. C'est ainsi que le théorème 0.4 intervient. Maintenant, si Γ n'est pas proximal, les projections de rang supérieur vont jouer un rôle. Et nous espérons que le théorème 4.1 va être utile dans cette situation.

0.4.2 Conséquence sur les projections des fractales

Tout comme le théorème de projection discrétisé de Bourgain, le théorème 4.1 peut être utilisé pour déduire un théorème de projection en terme de dimension de Hausdorff.

Corollaire 4.2. *Soient $1 \leq m < n$ des entiers. Étant donnés $0 < \alpha < n$ et $\kappa > 0$, il existe $\epsilon > 0$ tel que la proposition suivante soit vraie. Soit $A \subset \mathbb{R}^n$ un sous-ensemble borélien de dimension $\dim_{\mathrm{H}}(A) = \alpha$. Alors l'ensemble des directions exceptionnelles*

$$\big\{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \dim_{\mathrm{H}}(\pi_V(A)) \leq \frac{m}{n}\alpha + \epsilon\big\}$$

ne peut pas supporter de mesure non-nulle μ sur $\mathrm{Gr}(\mathbb{R}^n, m)$ ayant la propriété de non-concentration suivante,

$$\forall \rho > 0,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, n-m), \quad \mu(\mathcal{V}_\angle(W, \rho)) \leq \rho^\kappa.$$

Appliqué à une mesure de Frostman supportée par les directions exceptionnelles, cela donne

Corollaire 4.3. *Soient $1 \leq m < n$ des entiers. Étant donnés $0 < \alpha < n$ et $\kappa > 0$, il existe $\epsilon > 0$ tel que la proposition suivante soit vraie. Soit $A \subset \mathbb{R}^n$ un sous-ensemble borélien de dimension $\dim_{\mathrm{H}}(A) = \alpha$. Alors*

$$\dim_{\mathrm{H}}\big\{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \dim_{\mathrm{H}}(\pi_V(A)) \leq \frac{m}{n}\alpha + \epsilon\big\} \leq m(n-m) - 1 + \kappa.$$

[3] Une action $\Gamma \curvearrowright X$ est dite μ-raide pour une mesure μ sur Γ si toute mesure μ-stationnaire est $\langle \mathrm{Supp}(\mu) \rangle$-invariante. La terminologie est due à Furstenberg [33].

Par conséquent, on peut poser $\kappa = \epsilon = 0$ dans l'estimation précédente. Comparée à ce qui est connu (le théorème 0.5), la constante 1 dans notre estimation est très faible pour la plupart des valeurs de m et de α. Néanmoins, notre résultat n'est pas entièrement recouvert par le théorème 0.5, notamment dans les deux situations suivantes

(i) (Projection sur des droites) $m = 1$ et $\alpha \in]0, 1 + \frac{1}{n-1}[$,

(ii) (Projection sur des hyperplans) $m = n - 1$ et $\alpha \in]n - 1 - \frac{1}{n-1}, n[$.

Par exemple, pour $n = 2$ et $m = 1$ (cas traité dans Bourgain [7]), cela donne,

$$\dim_{\mathrm{H}}\big\{\theta \in \mathrm{Gr}(\mathbb{R}^2, 1) \mid \dim_{\mathrm{H}}(\pi_\theta(A)) \leq \frac{1}{2}\dim_H(A)\big\} = 0,$$

pour tout ensemble borélien A avec $0 < \dim_{\mathrm{H}}(A) < 2$.

0.5 Estimées produit

Maintenant supposons que notre espace ambiant est un groupe multiplicatif G. Comme plus haut, pour des sous-ensembles A et B de G, leur ensemble produit est noté par

$$AB = \{ab \mid a \in A,\ b \in B\}.$$

Nous écrivons aussi $A^{-1} = \{a^{-1} \mid a \in A\}$ pour l'ensemble inverse et $A^s = A \cdots A$ pour l'ensemble produit itéré s-fois. Bien que l'inégalité de Plünnecke-Ruzsa soit fausse dans les groupes non-commutatifs en général, nous avons toutefois (voir [59, Proposition 2.40])

$$|(A \cup \{1\} \cup A^{-1})^s| \ll_s \Big(\frac{|A^3|}{|A|}\Big)^{O_s(1)}|A|. \tag{0.3}$$

En vue de cette inégalité, quand on étudie la croissance de A^s, on s'intéresse particulièrement à la constante de triplement $\frac{|A^3|}{|A|}$.

Une petite précision est nécessaire pour éviter de la confusion. Quand on parle de croissance dans les groupes en géométrie des groupes, on s'intéresse au comportement asymptotique de $|A^s|$ avec A une partie génératrice. Citons le théorème de Gromov qui caractérise les groupes de croissance polynomiale[4]. Alors qu'en combinatoire arithmétique, le cadre dans lequel nous nous plaçons, nous analysons de manière plus fine la quantité $|A^s|$ pour les valeurs petites de s. En particulier, les questions dans ce cadre ont du sens et sont intéressantes même si G est un groupe fini.

Dans le cadre discret, nous avons une très bonne compréhension des ensembles à petit triplement. La théorie commence par les travaux de Helfgott [36, 37] et mène à une généralisation aux groupes finis de type Lie par Breuillard-Green-Tao [18] et indépendamment par Pyber-Szabó [51] et aboutit aussi à une classification des ensembles à petit triplement due à Breuillard-Green-Tao [19]. Le dernier résultat établit une conjecture de Helfgott-Lindenstrauss et peut être vu comme une généralisation du théorème de Freiman dans le cadre non-commutatif.

[4] Un groupe est dit de croissance polynomiale si pour une/toute partie génératrice A, $|A^s|$ est majoré par un polynôme en s quand s tend vers $+\infty$.

Contrairement à ce qui se passe dans le cas commutatif, il y a peu d'espoir de transférer directement les estimées produit depuis le cadre discret au cadre discrétisé.

Néanmoins, l'analogue de l'inégalité (0.3) reste valable, comme montré par Tao [57, Theorem 6.8]. Et c'est pour cela qu'on se concentre sur le triplement $\frac{\mathcal{N}_\delta(A^3)}{\mathcal{N}_\delta(A)}$ dans le cadre discrétisé aussi. En utilisant le théorème somme-produit de Bourgain, Bourgain et Gamburd [9, 11] ont les premiers résultats dans ce cadre, à savoir, un théorème produit pour les groupes de Lie $\mathrm{SU}(d)$, $d \geq 2$. Cela a été ensuite généralisé par Saxcé [24] aux groupes de Lie simples.

Théorème 0.7 (Théorème produit pour les groupes de Lie simples, Saxcé [24])**.** *Soit G un groupe de Lie simple. Il existe un voisinage U de l'élément neutre de G tel que la proposition suivante soit vraie. Étant donnés $\sigma < \dim(G)$ et $\kappa > 0$, il existe $\epsilon > 0$ tel que pour tout $\delta > 0$ suffisamment petit, si $A \subset U$ satisfait*

(i) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$;

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$;

(iii) pour tout sous-groupe propre fermé connexe $H < G$, il existe $a \in A$ avec $d(a, H) \geq \delta^\epsilon$;

alors $\mathcal{N}_\delta(A^3) \geq \delta^{-\epsilon}\mathcal{N}_\delta(A)$.

Ce résultat peut être comparé à celui de Breuillard-Green-Tao [18] dans le cadre discret. Tous les deux affirment qu'un sous-ensemble typique dans un groupe simple a un grand triplement sauf s'il est coincé dans un sous-groupe.

Les théorèmes produit peuvent être appliqués pour obtenir des résultats de trou spectral. Cela est fait pour les groupes $\mathrm{SU}(d)$, $d \geq 2$ par Bourgain-Gamburd [9, 11] et plus généralement pour les groupes compacts simples par Benoist-Saxcé [3]. Ensuite, Boutonnet, Ioana et Golsefidy [16] ont introduit la notion de trou spectral local et ont étendu davantage ces résultats aux cas non-compacts. De l'autre côté, un théorème produit discrétisé peut avoir des conséquences concernant la dimension de Hausdorff des sous-groupes boréliens. Ainsi, en combinant les techniques d'analyse de Fourier [23], Lindenstrauss-Saxcé [41] et Saxcé [25] ont montré qu'il n'y a pas de sous-groupe borélien dense de dimension de Hausdorff intermédiaire dans un groupe de Lie connexe simple.

Si le groupe G est nilpotent, alors il n'est pas difficile de construire des sous-ensembles qui satisfont les hypothèses du théorème 0.7 mais qui ont un petit triplement. Par exemple, on peut prendre une somme de progressions arithmétiques si G est abélien. Et pour un groupe nilpotent quelconque, il y a une notion de nilprogression qui généralise les progressions arithmétiques, voir par exemple [19]. Ce contraste entre les groupes simples et les groupes nilpotents apparaît aussi avec l'existence de sous-groupe borélien dense de dimension intermédiaire. Erdős et Volkmann [29] ont construit des sous-groupes de $\mathbb{R}$ de dimension de Hausdorff arbitraire entre 0 et 1. Saxcé [22] a étendu cette construction aux groupes de Lie nilpotents et à une large classe de groupes résolubles.

Dans un travail en commun avec Nicolas de Saxcé, nous étendons le théorème 0.7 aux groupes de Lie parfaits. Ceci constitue le résultat principal du chapitre 5. Rappelons qu'un groupe de Lie G est dit parfait si son algèbre de

Lie $\mathfrak{g}$ satisfait la condition $[\mathfrak{g},\mathfrak{g}] = \mathfrak{g}$. En particulier, les groupes de Lie semi-simples sont parfaits. Rappelons que le quotient d'un groupe de Lie G par son radical R est semi-simple. Si G est simplement connexe alors G/R l'est aussi et donc un produit direct de groupes de Lie simples et simplement connexes. Les facteurs dans ce produit direct sont appelés les facteurs simples de G.

Théorème 5.1 (Théorème produit pour les groupes de Lie parfaits)**.** *Soit G un groupe de Lie parfait simplement connexe. Il existe un voisinage U de l'élément neutre de G tel que la proposition suivante soit vraie. Étant donnés $\sigma < \dim(G)$ et $\kappa > 0$, il existe $\epsilon > 0$ tel que pour tout $\delta > 0$ suffisamment petit, si $A \subset U$ satisfait*

(i) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$ *;*

(ii) pour tout facteur simple S de G, en notant $\pi_S \colon G \to S$ la projection canonique, on a
$$\forall \rho \geq \delta, \quad \mathcal{N}_\rho(\pi_S(A)) \geq \delta^\epsilon \rho^{-\kappa};$$

(iii) pour tout sous-groupe propre fermé connexe $H < G$, il existe $a \in A$ avec $d(a,H) \geq \delta^\epsilon$;

alors $\mathcal{N}_\delta(A^3) \geq \delta^{-\epsilon}\mathcal{N}_\delta(A)$.

Ici c'est pour des raisons pratiques que nous demandons à ce que G soit simplement connexe. Le résultat concerne en fait les groupes de Lie locaux, car nous restreignons à un voisinage de l'élément neutre.

L'idée de la démonstration est de réduire d'abord au cas où G est un produit semi-direct d'un groupe semi-simple par un groupe abelien. Et puis, dans ce cas spécial, on considère sa représentation adjointe et y montre une croissance. Enfin, cette croissance dans l'algèbre de Lie est transférée au groupe de Lie grâce à l'utilisation de la formule de Campbell-Hausdorff.

Ainsi, en étape intermédiaire, nous obtenons une estimation de type somme-produit pour les représentations de groupe de Lie. C'est un résultat intéressant à part entière. Soit V une représentation linéaire de dimension finie sur $\mathbb{R}$ d'un groupe de Lie G. Soient X un sous-ensemble de V et A un sous-ensemble de G. Nous nous demandons si X croît vite sous l'addition et l'action de A. Pour $s \geq 1$, on écrit $\langle A, X\rangle_s$ pour désigner l'ensemble des éléments de V qui peuvent être exprimés comme sommes, différences et produits d'au plus s éléments de A et de X.

Théorème 5.2 (Estimée somme-produit dans les représentations)**.** *Soient G un groupe de Lie et V un G-module de dimension finie. Il existe un voisinage U de l'élément neutre de G tel que la proposition suivante soit vraie. Étant donnés ϵ_0 et $\kappa > 0$, il existe $s \geq 1$ et $\epsilon > 0$ tels que pour tout $\delta > 0$ suffisamment petit, si $A \subset U$ et $X \subset \mathbf{B}_V(0,1)$ satisfont*

(i) il existe une suite Jordan-Hölder[5] *$0 = V_0 < \ldots < V_l = V$ telle que pour $i = 1, \ldots, l$,*
$$\forall \rho \geq \delta, \quad \mathcal{N}_\rho(\pi_{V_i/V_{i-1}}(A)) \geq \delta^\epsilon \rho^{-\kappa}.$$
où $\pi_{V_i/V_{i-1}} \colon G \to \mathrm{GL}(V_i/V_{i-1})$ désigne la représentation de G sur V_i/V_{i-1} ;

[5] Par « Jordan-Hölder », nous entendons que $0 = V_0 < \ldots < V_l = V$ sont des sous-modules tels que les quotients V_i/V_{i-1} sont tous des G-modules simples.

(ii) pour tout sous-groupe propre fermé connexe $H < G$, il existe $a \in A$ avec $d(a, H) \geq \delta^{\epsilon}$;

(iii) pour tout sous-module propre $W < V$, il existe $x \in X$ avec $d(x, W) \geq \delta^{\epsilon}$;

Alors,

$$\mathbf{B}_V(0, \delta^{\epsilon_0}) \subset \langle A, X \rangle_s + \mathbf{B}_V(0, \delta).$$

Remarquons que pour les représentations irréductibles, ce théorème est une variante du théorème 3.2. C'est dans ce cas spécial qu'on va utiliser les résultats du chapitre 3. Le reste de la démonstration du théorème 5.2 consiste à ramener à ce cas spécial par une récurrence sur la longueur de la décomposition de Jordan-Hölder de V.

En considérant l'action de $\mathbb{R}^*$ sur $\mathbb{R}$, nous retrouvons le théorème somme-produit discrétisé de Bourgain (théorème 1.4). De même, nous pouvons retrouver les estimations somme-produit discrétisées pour $\mathbb{C}$ ou $\mathbb{H}$. En plus, la méthode dans la démonstration du théorème 5.2 peut être utilisée pour obtenir des estimations somme-produit dans les algèbres semi-simples (i.e. sommes directes d'algèbres simples).

Chapter 1

Introduction

For any set A, we denote by $|A|$ or $\#A$ its cardinality. Landau notations $O(f)$ and Vinogradov notations $f \ll g$ are used throughout this document. We also write $f \asymp g$ to say $f \ll g$ and $g \ll f$. In this chapter we introduce the main results by providing backgrounds and motivations. For impatient readers, the main new results in this document are : (all in the discretized a.k.a metric entropy setting)

- A sum-product theorem for matrix algebras (Theorem 3.1).
- A sum-product estimate for sets in Euclidean space under linear transformation (Theorem 3.2).
- A projection theorem for orthogonal projections of rank ≥ 2 (Theorem 4.1).
- (Joint work with N. de Saxcé) A product theorem in perfect Lie groups (Theorem 5.1).
- (Joint work with N. de Saxcé) A sum-product estimate for representations of Lie groups (Theorem 5.2).

1.1 The notion of discretized sets

The notion of discretized sets was highlighted by Katz and Tao in [38] as a framework for studying fractal properties of continuous sets by using combinatorial tools such as arithmetic combinatorics. In this section we introduce the setup.

1.1.1 The δ-covering number

Let (E, d) be a metric space. For $\rho > 0$ and $x \in E$, we denote by $\mathbf{B}(x, \rho)$ or by $x^{(\rho)}$ the closed ball centered at x and of radius ρ. And we write $\mathbf{B}_E(x, \rho)$ to specify the ambient space when it is not clear. For $A \subset E$, we denote by $A^{(\rho)}$ the closed ρ-neighborhood of A :

$$A^{(\rho)} = \{x \in E \mid d(x, A) \leq \rho\}.$$

A subset A is said to be *ρ-separated* if for any $a \in A$, a is the only element in the intersection $\mathbf{B}(a,\rho) \cap A$. Throughout this document, the variable δ stands for a positive real number which we refer to as the scale.

Definition. Let A be a relatively compact subset of (E,d). The *δ-covering number* of A, also known as *the metric entropy* of A, is defined as

$$\mathcal{N}_\delta(A) = \min\Big\{N \geq 0 \mid \exists x_1,\dots,x_N \in E,\ A \subset \bigcup_{i=1}^{N} \mathbf{B}(x_i,\delta)\Big\}.$$

Example. Let p be a prime number and let $(\mathbb{Z}_p, |\cdot|_p)$ be the ring of p-adic integers equipped with its usual p-adic absolute value. If A is a subset of $\mathbb{Z}_p$, then for all $k \in \mathbb{N}$,

$$\mathcal{N}_{p^{-k}}(A) = |\pi_k(A)|$$

where $\pi_k \colon \mathbb{Z}_p \to \mathbb{Z}_p/p^k\mathbb{Z}_p \simeq \mathbb{Z}/p^k\mathbb{Z}$ denotes the reduction modulo p^k.

If we allow the scale δ to be zero, then $\mathcal{N}_0(A)$ is exactly the cardinality of A. It is more interesting to study the asymptotic behavior of $\mathcal{N}_\delta(A)$ when δ goes to 0 from above.

Example. The lower and upper Minkowski dimensions of $A \subset (E,d)$ can be defined as (see [45, §5.3])

$$\underline{\dim}_{\mathrm{M}}(A) = \liminf_{\delta\to 0} \frac{\log \mathcal{N}_\delta(A)}{-\log\delta} \quad \text{and} \quad \overline{\dim}_{\mathrm{M}}(A) = \limsup_{\delta\to 0} \frac{\log \mathcal{N}_\delta(A)}{-\log\delta}.$$

In particular, if $\mathcal{N}_\delta(A) \asymp_{\delta\to 0} \delta^{-\alpha}$ then they agree and are equal to α.

Now let E be the Euclidean space $\mathbb{R}^n$, $n \geq 1$. Denote by λ the Lebesgue measure on $\mathbb{R}^n$. Besides the δ-covering number, we have other means to measure the size of a set at scale δ.

Lemma 2.1. *Let $\delta > 0$ and let A be a bounded subset of $\mathbb{R}^n$. Let $\tilde{A}$ be a maximal 2δ-separated subset of A. Then*

$$\mathcal{N}_{2\delta}(A) \leq |\tilde{A}| \leq \mathcal{N}_\delta(A) \leq \mathcal{N}_1(\mathbf{B}(0,2))\,\mathcal{N}_{2\delta}(A),$$

and

$$\delta^{-n}\lambda(A^{(\delta)}) \asymp_n \mathcal{N}_\delta(A).$$

In particular, changing the scale by a constant factor only affects the covering number by a constant factor. Moreover, A and its δ-neighborhood $A^{(\delta)}$ have the same size when viewed at scale δ :

$$\mathcal{N}_\delta(A^{(\delta)}) \asymp_n \mathcal{N}_\delta(A).$$

That is why in [38], a δ-discretized set is defined to be a union of balls of radius δ. And in the proofs in later chapters, we can almost always replace A by its δ-neighborhood.

Example. Let μ be a compactly supported probability measure on $\mathbb{R}^n$. Its Fourier–Stieltjes transform (see [39, Chapter VI, §2]) $\hat{\mu}(\xi) = \int e^{-i\langle \xi, x\rangle}\,\mathrm{d}\mu(x)$, $\xi \in \mathbb{R}^n$, is K-Lipschitz for some $K > 0$ depending only on $\max_{x\in \mathrm{Supp}(\mu)}\|x\|$. Therefore, if we set, for $t > 0$,

$$A_t = \{\xi \in \mathbb{R}^n \mid |\hat{\mu}(\xi)| \geq t\},$$

then $A_{2t}^{(\frac{t}{K})} \subset A_t$. Hence, it is natural to look at the set A_t at scale t.

1.1.2 Non-concentration

If a bounded subset $A \subset \mathbb{R}^n$ has size $\mathcal{N}_\delta(A) \asymp \delta^{-\alpha}$, we would like to think α as its dimension. But a ball of radius $\delta^{1-\frac{\alpha}{n}}$ also satisfies $\mathcal{N}_\delta(\mathbf{B}(0, \delta^{1-\frac{\alpha}{n}})) \asymp \delta^{-\alpha}$. Clearly, a ball has very different properties compared to a fractal. To avoid this degenerate situation, we will often require some non-concentration properties on the set A. Let $\kappa > 0$ and $\epsilon > 0$ be parameters. For example we often ask that

$$\forall \rho \geq \delta, \quad \mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}. \tag{1.1}$$

And sometimes we need a stronger condition as follows :

$$\forall \rho \geq \delta, \ \forall x \in E, \quad \mathcal{N}_\delta(A \cap x^{(\rho)}) \leq \delta^{-\epsilon} \rho^\kappa \mathcal{N}_\delta(A). \tag{1.2}$$

In these conditions the larger κ is, the stronger the non-concentration becomes. The notion of $(\delta, \alpha)_n$-sets in [38] corresponds to condition (1.2) with κ maximal, i.e. $\kappa = \alpha$. Here ϵ is a small parameter to relax these conditions. We may understand it as saying that the non-concentration is required from scale δ to scale δ^ϵ. Obviously, condition (1.2) implies (1.1). The converse is not true. Nevertheless, if we allow passing to a large subset, we can extract a subset having property (1.2) from a set having property (1.1).

Example. Let A be a bounded Borel set in $\mathbb{R}^n$ of Hausdorff dimension $\dim_{\mathrm{H}} A > \alpha$. Let μ be the measure given by Frostman's lemma.

Theorem 1.1 (Frostman's lemma [45, Theorem 8.8])**.** *Let A be a Borel set of $\mathbb{R}^n$. If $\dim_{\mathrm{H}}(A) > \alpha$ then there exists a finite nonzero compactly supported Borel measure μ with $\mathrm{Supp}(\mu) \subset A$ such that*

$$\forall \rho > 0, \ \forall x \in \mathbb{R}^n, \quad \mu(\mathbf{B}(x, \rho)) \leq \rho^\alpha.$$

Then it is immediate that for any $\rho > 0$, we have

$$\mathcal{N}_\rho(A) \geq \mu(A) \rho^{-\alpha}.$$

In other words, we have the non-concentration (1.1) with $\kappa = \alpha$.

1.2 Additive theory

In this section, suppose that the ambient space E is moreover equipped with a structure of abelian group. For subsets $A, B \subset E$, we write

$$-A = \{-a \mid a \in A\},$$

$$A + B = \{a + b \mid a \in A, b \in B\}$$

and

$$A - B = \{a - b \mid a \in A, b \in B\}.$$

Let s be a positive integer, we write sA for the s-fold sum-set $A + \cdots + A$. By additive theory, we mean the study of the growth of sA or that of $sA - s'A$, $s, s' \geq 1$.

Consider the discretized setting in $E = \mathbb{R}^n$. It turns out that by approximating $\mathbb{R}^n$ by the lattice $\delta \cdot \mathbb{Z}^d$, we can transfer any result concerning additive properties in the discrete setting to the δ-discretized setting. For example, we have the Plünnecke-Ruzsa inequality.

Lemma 2.4 (Plünnecke-Ruzsa inequality). *For all $K \geq 1$, if $\mathcal{N}_\delta(A+B) \leq K\mathcal{N}_\delta(B)$ then for all integers $k \geq 1$ and $l \geq 0$,*

$$\mathcal{N}_\delta(kA - lA) \ll_n K^{k+l}\mathcal{N}_\delta(B).$$

In particular, for any $s \geq 2$, $\mathcal{N}_\delta(sA)$ is significantly larger than $\mathcal{N}_\delta(A)$ if and only if $\mathcal{N}_\delta(A+A)$ is already significantly larger than $\mathcal{N}_\delta(A)$. That is why an important problem in additive combinatorics is the classification of sets with small doubling (i.e. sets A such that $|A+A| \leq K|A|$). Freiman's theorem says such sets are contained in generalized arithmetic progressions of comparable size (see [59, Chapter 5]). Transferred into the discretized setting, it becomes

Theorem 1.2 (Discretized version of Freiman's Theorem, [57, Proposition 7.3]). *Let $K \geq 2$ be a parameter. Let A be a bounded subset of $\mathbb{R}^n$. If $\mathcal{N}_\delta(A+A) \leq K\mathcal{N}_\delta(A)$ then there exists a set P which is the sum of $O_{K,n}(1)$ arithmetic progressions in $\mathbb{R}^n$ such that $A \subset P + \mathbf{B}(0,\delta)$ and $|P| \asymp_{K,n} \mathcal{N}_\delta(A)$.*

Having a good dependence of the implied constants on K is one of the central problems in additive combinatorics.

In the other way around, anything proved in the δ-discretized setting will imply its counterpart in the discrete setting. So in this document, we do not prove anything new for the additive theory of discretized sets.

1.3 Sum-product estimates

Now we suppose that our ambient space E has moreover a ring structure. For A and B subsets of E, we write

$$A \cdot B = \{ab \mid a \in A, b \in B\}.$$

Let s be a positive integer, we write A^s for the s-fold product-set $A \cdots A$. Moreover, we define recursively $\langle A \rangle_1 = A \cup (-A)$ and for all positive integers $s \geq 2$,

$$\langle A \rangle_s = \langle A \rangle_{s-1} \cup \bigcup_{k=1}^{s-1} \big(\langle A \rangle_k + \langle A \rangle_{s-k}\big) \cup \bigcup_{k=1}^{s-1} \big(\langle A \rangle_k \cdot \langle A \rangle_{s-k}\big).$$

In other words, $\langle A \rangle_s$ is the set of elements obtained from at most s elements of $A \cup (-A)$ by adding and multiplying them.

Roughly speaking, the sum-product problem asks, given a set A, whether A grows fast under addition, multiplication and subtraction and if not, what are the obstructions. Thus, a sum-product estimate is a lower bound for the size of $\langle A \rangle_s$ with $s \geq 2$. The first result of this type is due to Erdős and Szemerédi [28] for the ring of real numbers $\mathbb{R}$.

Theorem 1.3 (Erdős-Szemerédi [28]). *There is an absolute constant $c > 0$ such that for any nonempty finite subset A of $\mathbb{R}$, we have*

$$\max(|A+A|, |A \cdot A|) \geq c|A|^{1+c}.$$

This was the starting point of numerous works. They either establish similar estimates for broader classes of rings or improve existing bounds. See [58] for an elegant treatment and more history.

Although the Erdős-Szemerédi theorem concerns the ring of real numbers[1], it is not about discretized sets. Unlike the additive theory, the study of the sum-product phenomenon in the discretized setting differs in many ways from its discrete analogue. For example, the multiplication by δ preserves the cardinality while reduces the δ-covering number of any bounded set to a constant. And when x changes from δ to 1, its behavior continuously changes form that of a zero-divisor to that of an invertible element. Despite this, we expect typical discretized sets to grow quickly under addition and multiplication. For the ring $\mathbb{R}$, this was conjectured by Katz and Tao in [38] and settled by Bourgain [6, 7].

Theorem 1.4 (Bourgain [7]). *Given $\kappa > 0$ and $\sigma < 1$, there is $\epsilon > 0$ such that the following holds for $\delta > 0$ sufficiently small. Let A be a subset of $\mathbb{R}$, assume that*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$.

Then,

$$\mathcal{N}_\delta(A+A) + \mathcal{N}_\delta(A \cdot A) \geq \delta^{-\epsilon} \mathcal{N}_\delta(A). \tag{1.3}$$

The proof of this theorem uses multi-scale analysis. Additive structure is analyzed at each level (using a quantitative Freiman's theorem due to Chang [20] in Bourgain's first proof in [6]) and then multiplicative information is used to show growth in size.

Bourgain's sum-product theorem is very influential. It is used directly or indirectly in almost all results about discretized sets mentioned in this document. When the discretized setting was introduced in [38], it was conceived as a general strategy for proving results in the continuous regime (where sets are measured by its Hausdorff dimension) while using ideas and results in the discrete regime where we have well-developed theory such as arithmetic combinatorics. Thus, one of the original motivations of Bourgain's theorem was the Erdős-Volkmann ring conjecture [29] which asserts that no Borel subring of $\mathbb{R}$ has Hausdorff dimension between 0 and 1. This conjecture was also settled by Edgar and Miller [27] using different ideas. However, their proof is not quantitative. Moreover, as anticipated in [38], Bourgain's theorem also makes progress on the Falconer distance problem and the Furstenberg conjecture in fractal geometry. Among other direct applications, let us mention the construction of explicit monotone expanders by Bourgain and Yehudayoff [15]. We will discuss more applications later when we talk about projection and product estimates.

Bourgain and Gamburd have obtained similar estimates for the ring $\mathbb{C}$ in [9] and later for the ring $\mathbb{C}^n$, the n-fold direct product of $\mathbb{C}$ with itself, in [11]. The statements are slightly different since $\mathbb{R}$ has no proper subalgebra while $\mathbb{C}$ and $\mathbb{C}^n$ have. Moreover, $\mathbb{C}^n$ has nontrivial proper ideals, i.e. it is not simple. In Chapter 3, we prove a sum-product estimate for finite-dimensional simple algebras over $\mathbb{R}$. Note that by the Wedderburn structure theorem and the Frobenius theorem, such algebras are isomorphic to $\mathcal{M}_n(\mathbb{R})$, $\mathcal{M}_n(\mathbb{C})$ or $\mathcal{M}_n(\mathbb{H})$,

[1] In their original paper, the theorem was stated for $\mathbb{Z}$ but their proof works in any totally ordered ring without zero-divisor.

the algebra of $n \times n$ matrices over the real numbers, the complex numbers, or the quaternions, with $n \geq 1$. Before stating the result, let us mention that in the discrete context, Chang [21] investigated the sum-product problem for real matrices and Tao [58] obtained a sum-product theorem concerning general algebras.

Theorem 3.1. *Let E be a normed*[2] *simple real algebra of finite dimension. Given $\kappa > 0$ and $\sigma < \dim(E)$, there is $\epsilon > 0$ depending on E, κ and σ such that the following holds for $\delta > 0$ sufficiently small. Let A be a subset of E, assume that*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$,

(iv) for every proper subalgebra $W \subset E$, there is $a \in A$ such that $d(a, W) \geq \delta^\epsilon$.

Then,

$$\mathcal{N}_\delta(A + A) + \mathcal{N}_\delta(A + A \cdot A) \geq \delta^{-\epsilon} \mathcal{N}_\delta(A).$$

The strategy is to produce a rich torus, i.e. a relatively large subset consisting of simultaneously diagonalizable elements, and then use the Bourgain-Gamburd sum-product estimate in $\mathbb{C}^n$. The idea of rich torus originated from the work of Helfgott [36] and since then has been used to show product theorems in a lot of situations. For instance, the proof of Saxcé's product theorem (Theorem 1.7) consists also of producing a rich torus and applying the Bourgain-Gamburd sum-product estimate in $\mathbb{C}^n$. Our method provides a general strategy for obtaining sum-product estimates in simple algebras from sum-product estimates in the base field.

We also obtain related estimates concerning linear actions on Euclidean spaces. Let X be a bounded subset of the Euclidean space $\mathbb{R}^n$. Let $A \subset \mathrm{End}(\mathbb{R}^n)$ be a collection of linear endomorphisms. We can ask whether X grows under addition and transformation by elements of A, provided that A is sufficiently rich.

Theorem 3.2. *Let n be a positive integer. Given $\kappa > 0$ and $\sigma < n$, there is $\epsilon > 0$ such that the following holds for $\delta > 0$ sufficiently small. Let A be a subset of $\mathrm{End}(\mathbb{R}^n)$ and X a subset of $\mathbb{R}^n$, assume that*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) for every nonzero proper linear subspace $W \subset \mathbb{R}^n$, there is $a \in A$ and $w \in \mathbf{B}_W(0, 1)$ such that $d(aw, W) \geq \delta^\epsilon$,

(iv) $X \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(v) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(X) \geq \delta^\epsilon \rho^{-\kappa}$,

[2] By "normed" we mean there is a norm which makes the underlying linear space E a normed vector space. Since all norms on E are equivalent, we can always assume that the norm is submultiplicative if we want.

(vi) $\mathcal{N}_\delta(X) \leq \delta^{-\sigma-\epsilon}$.

Then,

$$\mathcal{N}_\delta(X+X) + \max_{a\in A} \mathcal{N}_\delta(X+aX) \geq \delta^{-\epsilon}\mathcal{N}_\delta(X),$$

where $aX = \{ax \mid x \in X\}$.

This improves a previous result of Bourgain and Gamburd [11, Proposition 1] where a constant is required instead of δ^ϵ in the irreducibility condition (iii). The proof of Bourgain and Gamburd relies on this irreducibility hypothesis at all scales in a crucial way. It seems to us not easy to modify if we relax their assumption. Relaxing this hypothesis is the most important technical challenge in the proof of Theorem 3.2. A reason for which this improvement is important is that this kind of estimates are often used together with the Balog-Szemerédi-Gowers theorem, which requires restricting the sets we work with to subsets of size δ^ϵ times the original size. This procedure usually destroys all information above scale δ^ϵ.

Theorem 3.2 is actually a better formulation for the sum-product phenomenon since we can easily recover Theorem 3.1 form it by considering the left and right multiplications by elements of A as the collection of endomorphisms of E. Similarly, we can also obtain a "sum-bracket" estimate in simple Lie algebras. If A is a subset of a Lie algebra $\mathfrak{g}$, write $[A,A] = \{[a,b] \mid a,b \in A\}$.

Corollary 3.3. *Let* $\mathfrak{g}$ *be a normed*[3] *simple Lie algebra of finite dimension. Given* $\kappa > 0$ *and* $\sigma < \dim(\mathfrak{g})$*, there is* $\epsilon > 0$ *such that the following holds for* $\delta > 0$ *sufficiently small. Let* A *be a subset of* $\mathfrak{g}$*, assume that*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$,

(iv) for every proper Lie subalgebra W *of* $\mathfrak{g}$*, there is* $a \in A$ *such that* $d(a,W) \geq \delta^\epsilon$.

Then,

$$\mathcal{N}_\delta(A+A) + \mathcal{N}_\delta(A+[A,A]) \geq \delta^{-\epsilon}\mathcal{N}_\delta(A).$$

1.4 Orthogonal projections

Let $0 < m < n$ be positive integers. We write $\mathrm{Gr}(\mathbb{R}^n, m)$ do denote the Grassmannian of m-dimensional subspaces in $\mathbb{R}^n$. For $V \in \mathrm{Gr}(\mathbb{R}^n, m)$, denote by $\pi_V : \mathbb{R}^n \to V$ the orthogonal projection onto V.

Fractal properties of orthogonal projections of subsets the Euclidean space have been intensively studied in fractal geometry starting from the renowned Marstrand projection theorem [43]. See the survey [31] for more history. Roughly speaking the problem asks for lower bounds for the size of the projections of the set to different directions. Since in general, the projection can not be large for every direction. We ask more precisely for a bound on the size of exceptional directions where a exceptional direction means a subspace onto which the

[3] We mean a norm which makes the underlying linear structure a normed vector space.

projection is small. Recall the following higher-dimensional projection theorem obtained in a series of works of Mattila and Falconer.

Theorem 1.5 (Mattila [44], Falconer [32], see also [46, §5.3]). *Let $A \subset \mathbb{R}^n$ be a Borel set of Hausdorff dimension $\dim_{\mathrm{H}}(A) = \alpha$. For all $s > 0$,*

(i) if $0 < s < \alpha \leq m$ then

$$\dim_{\mathrm{H}}\{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \dim_{\mathrm{H}}(\pi_V(A)) < s\} \leq m(n-m) - (m-s);$$

(ii) if $0 < s \leq m \leq \alpha$ then

$$\dim_{\mathrm{H}}\{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \dim_{\mathrm{H}}(\pi_V(A)) < s\} \leq m(n-m) - (\alpha - s).$$

Note that $m(n-m)$ is the dimension of $\mathrm{Gr}(\mathbb{R}^n, m)$ and the notion of Hausdorff dimension on $\mathrm{Gr}(\mathbb{R}^n, m)$ does not depend on the specific choice of the metric since all distances on a compact set are equivalent as long as they define the same topology.

Using his discretized sum-product theorem, Bourgain established a discretized projection theorem concerning rank one projections. Before stating the theorem, let us introduce some notations in order to formulate non-concentration conditions in projective spaces or Grassmannians. For $V \in \mathrm{Gr}(\mathbb{R}^n, m)$ and $W \in \mathrm{Gr}(\mathbb{R}^n, n-m)$, we define

$$\mathrm{d}_{\measuredangle}(V, W) = |\det(v_1, \ldots, v_m, w_1, \ldots, w_{n-m})|$$

where $(v_1, \ldots, v_m)$ is an orthonormal basis of V and $(w_1, \ldots, w_{n-m})$ an orthonormal basis of W. For example $\mathrm{d}_{\measuredangle}(V, W) = 0$ if and only if V and W have nontrivial intersection. For $\rho \geq 0$, we define

$$\mathcal{V}_{\measuredangle}(W, \rho) = \{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \mathrm{d}_{\measuredangle}(V, W) \leq \rho\}.$$

Theorem 1.6 (Bourgain [7, Theorem 5]). *Let $n \geq 2$ be an integer. Given $0 < \alpha < n$ and $\kappa > 0$, there exists $\epsilon > 0$ such that the following holds for sufficiently small $\delta > 0$. Let A be a subset of $\mathbb{R}^n$ contained in the unit ball $\mathbf{B}(0,1)$. Let μ be a probability measure on the projective space $\mathrm{Gr}(\mathbb{R}^n, 1)$. Assume that*

$$\mathcal{N}_\delta(A) \geq \delta^{-\alpha+\epsilon};$$

$$\forall \rho \geq \delta,\ \forall x \in \mathbb{R}^n, \quad \mathcal{N}_\delta(A \cap \mathbf{B}(x, \rho)) \leq \delta^{-\epsilon}\rho^\kappa \mathcal{N}_\delta(A);$$

$$\forall \rho \geq \delta,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, n-1), \quad \mu(\mathcal{V}_{\measuredangle}(W, \rho)) \leq \delta^{-\epsilon}\rho^\kappa.$$

Then there is a subset $A' \subset A$ and a set of directions $\mathcal{D} \subset \mathrm{Gr}(\mathbb{R}^n, 1)$ such that $\mathcal{N}_\delta(A') \geq \delta^\epsilon \mathcal{N}_\delta(A)$, $\mu(\mathcal{D}) \geq 1 - \delta^\epsilon$ and

$$\mathcal{N}_\delta(\pi_\theta(A'')) \geq \delta^{-\frac{\alpha}{n}-\epsilon}$$

whenever $\theta \in \mathcal{D}$ and $A'' \subset A'$ is a subset such that $\mathcal{N}_\delta(A'') \geq \delta^{2\epsilon}\mathcal{N}_\delta(A)$.

Later in [48], for sets in the plane $\mathbb{R}^2$, Orponen obtained shaper bound under stronger non-concentration condition for μ. The approach is different but nevertheless exploits the sum-product phenomenon.

It is also worth noting that there is a parallel set of problems which concerns projections of self-similar measures and self-similar sets. For instance, recently, a conjecture of Furstenberg concerning transversality between $\times 2$ and $\times 3$ invariant sets was settled in two independent works, Shmerkin [56] and Wu [61]. Shmerkin's approach also uses tools form additive combinatorics and ideas from Bourgain's discretized sum-product theorem. See also the survey paper [55] for further references.

The main result of Chapter 4 is a generalization of Theorem 1.6 to higher rank projections.

Theorem 4.1. *Let $m < n$ be positive integers. Given $0 < \alpha < n$ and $\kappa > 0$, there exists $\epsilon > 0$ such that the following holds for sufficiently small $\delta > 0$. Let A be a subset of $\mathbb{R}^n$ contained in the unit ball $\mathbf{B}(0,1)$. Let μ be a probability measure on $\mathrm{Gr}(\mathbb{R}^n, m)$. Assume that*

$$\mathcal{N}_\delta(A) \geq \delta^{-\alpha+\epsilon};$$

$$\forall \rho \geq \delta,\ \forall x \in \mathbb{R}^n, \quad \mathcal{N}_\delta(A \cap \mathbf{B}(x,\rho)) \leq \delta^{-\epsilon}\rho^\kappa \mathcal{N}_\delta(A);$$

$$\forall \rho \geq \delta,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, n-m), \quad \mu(\mathcal{V}_\measuredangle(W,\rho)) \leq \delta^{-\epsilon}\rho^\kappa.$$

Then there is a set $\mathcal{D} \subset \mathrm{Gr}(\mathbb{R}^n, m)$ such that $\mu(\mathcal{D}) \geq 1 - \delta^\epsilon$ and

$$\mathcal{N}_\delta(\pi_V(A')) \geq \delta^{-\frac{m}{n}\alpha-\epsilon}$$

whenever $V \in \mathcal{D}$ and $A' \subset A$ is a subset such that $\mathcal{N}_\delta(A') \geq \delta^\epsilon \mathcal{N}_\delta(A)$.

For $m > 1$, our result is new. Moreover compared to Theorem 1.6, we establish the lower bound for all relatively large subsets of A. The proof goes by induction on the dimension pair (n, m). The approach in the proof of Theorem 1.6 combined with the higher dimensional sum-product estimate (Theorem 3.2) proves the cases where m divides n. New ideas are used to reduce other cases to these cases.

1.4.1 Ergodic theoretic motivation

Bourgain's discretized projection theorem is one of the main ingredients in the Bourgain-Furman-Lindenstrauss-Mozes theorem [8] on equidistribution for orbits of subsemigroups of $\mathrm{SL}_d(\mathbb{Z})$ on the torus. This last result can be used to study stationary measures. Namely, they obtained a stiffness result. Note also that Bourgain and Varjú [14] used Bourgain-Furman-Lindenstrauss-Mozes theorem in combination with the result in [60] to show expansion in $\mathrm{SL}_d(\mathbb{Z}/q\mathbb{Z})$, q arbitrary integer.

In [4], Benoist and Quint have generalized the stiffness result of Bourgain-Furman-Lindenstrauss-Mozes to a much broader class of dynamical systems. In particular, for linear actions on tori, they do not need the proximality assumption in [8]. However, the results in [4] are not quantitative.

The approach in [8] is Fourier-analytic. While a subgroup $\Gamma \subset \mathrm{SL}_d(\mathbb{Z})$ acts on the torus, its transpose ${}^t\Gamma$ acts on Fourier coefficients. A large part of the proof focuses on the study of large Fourier coefficients under this action. By the theory of random matrix products, if Γ is proximal, then large random products in Γ behave like rank one projections composed with rotations, if viewed at an appropriate scale. That is how Theorem 1.6 comes into play. If Γ is not proximal, higher rank projections will be involved. We hope that Theorem 4.1 can be used in this situation.

1.4.2 Consequence on fractal projections

Just like Bourgain's discretized projection theorem can be used to derive a projection theorem in terms of Hausdorff dimension [7, Theorem 4], Theorem 4.1 has the following consequence.

Corollary 4.2. *Let $m < n$ be positive integers. Given $0 < \alpha < n$ and $\kappa > 0$, there is $\epsilon > 0$ such that the following is true. Let $A \subset \mathbb{R}^n$ is a Borel subset of dimension $\dim_{\mathrm{H}}(A) = \alpha$. Then the set of exceptional directions*

$$\big\{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \dim_{\mathrm{H}}(\pi_V(A)) \leq \frac{m}{n}\alpha + \epsilon\big\}$$

does not support any nonzero measure μ on $\mathrm{Gr}(\mathbb{R}^n, m)$ with the following non-concentration property,

$$\forall \rho > 0,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, n-m), \quad \mu(\mathcal{V}_\measuredangle(W, \rho)) \leq \rho^\kappa.$$

Applied to a Frostman measure supported on the set of exceptional directions, this gives

Corollary 4.3. *Let $m < n$ be positive integers. Given $0 < \alpha < n$ and $\kappa > 0$, there is $\epsilon > 0$ such that the following holds. Let $A \subset \mathbb{R}^n$ be a Borel set of dimension $\dim_{\mathrm{H}}(A) = \alpha$. Then*

$$\dim_{\mathrm{H}}\big\{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \dim_{\mathrm{H}}(\pi_V(A)) \leq \frac{m}{n}\alpha + \epsilon\big\} \leq m(n-m) - 1 + \kappa.$$

In particular, as $\kappa \to 0$, we get

$$\dim_{\mathrm{H}}\big\{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \dim_{\mathrm{H}}(\pi_V(A)) \leq \frac{m}{n}\alpha\big\} \leq m(n-m) - 1.$$

This should be compared to the estimates already known in Theorem 1.5. Our result is not covered by Theorem 1.5 in the following two situations:

(i) (Projection to lines) $m = 1$ and $\alpha \in]0, 1 + \frac{1}{n-1}[$,

(ii) (Projection to hyperplanes) $m = n - 1$ and $\alpha \in]n - 1 - \frac{1}{n-1}, n[$.

For example, as known in Bourgain [7], for $n = 2$ and $m = 1$, we have

$$\dim_{\mathrm{H}}\big\{\theta \in \mathrm{Gr}(\mathbb{R}^2, 1) \mid \dim_{\mathrm{H}}(\pi_\theta(A)) \leq \frac{1}{2}\dim_H(A)\big\} = 0,$$

for all Borel set A such that $0 < \dim_{\mathrm{H}}(A) < 2$.

1.5 Product estimates

Now let the ambient space be a multiplicative group G. As before, for subsets A and B of G, we write

$$AB = \{ab \mid a \in A,\ b \in B\}$$

to denote the product set. We write also $A^{-1} = \{a^{-1} \mid a \in A\}$ for the inverse set and $A^s = A \cdots A$ for the s-fold product set. This time we are interested in

the growth of $(A \cup \{1\} \cup A^{-1})^s$. Although the Plünnecke-Ruzsa inequality is false for general groups, we still have (see [59, Proposition 2.40])

$$|(A \cup \{1\} \cup A^{-1})^s| \ll_s \Big(\frac{|A^3|}{|A|}\Big)^{O_s(1)} |A|. \tag{1.4}$$

That is why we are interested in sets with small tripling (i.e. $|A^3| \leq K|A|$).

In the discrete setting, we have very good understanding of such sets, starting with the works of Helfgott [36, 37], generalized by Breuillard-Green-Tao [18] and independently by Pyber-Szabó [51] to finite groups of Lie types on the one hand and leading to a classification of small tripling set due to Breuillard-Green-Tao [19] on the other. The last result solves a conjecture of Helfgott-Lindenstrauss and can be viewed as a generalization of Freiman's theorem.

Unlike the commutative case, there is little hope to transfer directly product estimates from the discrete setting to the discretized setting.

Nevertheless, the analogue of (1.4) in the discretized setting holds, as shown by Tao [57, Theorem 6.8]. That is why in the discretized setting, we also focus on the tripling. Using Bourgain's sum-product theorem, Bourgain and Gamburd [9, 11] obtained product theorem for $\mathrm{SU}(d)$, $d \geq 2$. This is then generalized by Saxcé [24] to simple Lie groups.

Theorem 1.7 (Product theorem for simple Lie groups, Saxcé [24])**.** *Let G be a simple Lie group. There is a neighborhood U of the identity in G such that the following holds. Given $\sigma < \dim(G)$ and $\kappa > 0$, there exists $\epsilon > 0$ such that for all $\delta > 0$ sufficiently small, if $A \subset U$ satisfies*

(i) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$;

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$;

(iii) for any proper closed connected subgroup $H < G$, there exists $a \in A$ with $d(a, H) \geq \delta^\epsilon$;

then $\mathcal{N}_\delta(A^3) \geq \delta^{-\epsilon} \mathcal{N}_\delta(A)$.

This result can be compared to Breuillard-Green-Tao [18] in the discrete setting. It says a typical discretized set in a simple group has large tripling unless it is trapped in a subgroup.

Product theorems can be applied to get spectral gap results. This is done for the groups $\mathrm{SU}(d)$, $d \geq 2$ by Bourgain-Gamburd [9, 11] and for compact simple Lie groups by Benoist-Saxcé [3]. Then Boutonnet-Ioana-Golsefidy [16] introduced the notion of local spectral gap and further extended the expansion result to noncompact settings. Discretized product theorems can also be applied to study the Hausdorff dimension of measurable subgroups. Thus, combined with Fourier techniques [23], Lindenstrauss-Saxcé [41] and Saxcé [25] showed that there is no measurable dense subgroup of intermediate Hausdorff dimension in connected simple Lie groups.

If the group G is nilpotent, then it is easy to construct a subset satisfying the assumptions of Theorem 1.7 but having small tripling. For example, we can take a sum of arithmetic progressions if G is abelian. And for a general nilpotent group, there is a notion of nilprogression which generalizes arithmetic progressions, see e.g. [19]. This contrast between simple groups and nilpotent groups

also appears with the existence of measurable dense subgroups of intermediate dimension. Erdős and Volkmann [29] constructed measurable subgroups of arbitrary Hausdorff dimension in $\mathbb{R}$. Saxcé [22] extended this construction to nilpotent Lie groups and a large class of solvable groups.

In a joint work with Nicolas de Saxcé, we extend Theorem 1.7 to perfect Lie groups. This is the main result of Chapter 5. Recall that a Lie group G is said to be perfect if its Lie algebra $\mathfrak{g}$ satisfies the condition $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$. It is a weaker condition than semisimplicity. The quotient of G by its radical R is semisimple. If G is simply-connected, then G/R is simply-connected, hence a direct product of simply-connected simple Lie groups. Factors appearing in this direct product are called simple factors of the group G.

Theorem 5.1 (Product theorem for perfect Lie groups)**.** *Let G be a simply-connected perfect Lie group. There is a neighborhood U of the identity in G such that the following holds. Given $\sigma < \dim(G)$ and $\kappa > 0$, there exists $\epsilon > 0$ such that for all $\delta > 0$ sufficiently small, if $A \subset U$ satisfies*

(i) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$*;*

(ii) for any simple factor S of G, denote by $\pi_S \colon G \to S$ the canonical projection,
$$\forall \rho \geq \delta, \quad \mathcal{N}_\rho(\pi_S(A)) \geq \delta^\epsilon \rho^{-\kappa};$$

(iii) for any proper closed connected subgroup $H < G$, there exists $a \in A$ with $d(a, H) \geq \delta^\epsilon$;

then $\mathcal{N}_\delta(A^3) \geq \delta^{-\epsilon} \mathcal{N}_\delta(A)$.

We require G to be simply-connected only for notational convenience. The theorem is in fact about local Lie groups (a.k.a. analytic group chunk, see [54]) because we restrict to a neighborhood of the identity. If G is a connected perfect Lie group, we can restrict to a neighborhood U' such that U'^3 lifted to $\tilde{G}$, the universal covering of G, is contained in the neighborhood given by Theorem 5.1 applied to $\tilde{G}$.

The idea of the proof is to first reduce to the case where G is the semidirect product of a semisimple group by an abelian group. Then in this special case we consider its adjoint representation and show a growth statement in the adjoint representation. Finally, this growth in the Lie algebra can be transferred to the group by using the Campbell-Hausdorff formula.

Thus, as an intermediate result, we obtain a sum-product type estimate for representations of Lie groups. It is an interesting result on its own right. The setting is the following. Let V be a finite-dimensional linear representation over $\mathbb{R}$ of a Lie group G. Let X be a bounded subset of V and A a bounded subset of G. We can ask whether X grows fast under addition and the action of A. For $s \geq 1$, we define $\langle A, X \rangle_s$ to be the set of elements in V that can be obtained as sums, differences and products of at most s elements of A and X.

Theorem 5.2 (Sum-product estimates in representations)**.** *Let G be a Lie group and V finite-dimensional linear representation of G over $\mathbb{R}$. There exists a neighborhood U of the identity in G depending only on V such that the following holds. Given $\epsilon_0, \kappa > 0$, there exists $s \geq 1$ and $\epsilon > 0$ such that for all $\delta > 0$ sufficiently small, if $A \subset U$ and $X \subset \mathbf{B}_V(0, 1)$ satisfy the following :*

(i) There is a Jordan-Hölder[4] *sequence* $0 = V_0 < \ldots < V_l = V$ *such that for every* $i = 1, \ldots, l$,

$$\forall \rho \geq \delta, \quad \mathcal{N}_\rho(\pi_{V_i/V_{i-1}}(A)) \geq \delta^\epsilon \rho^{-\kappa}.$$

where $\pi_{V_i/V_{i-1}} \colon G \to \mathrm{GL}(V_i/V_{i-1})$ *denotes the representation of* G *on* V_i/V_{i-1}.

(ii) For any proper closed connected subgroup $H < G$, *there exists* $a \in A$ *with* $d(a, H) \geq \delta^\epsilon$.

(iii) For any proper submodule $W < V$, *there exists* $x \in X$ *with* $d(x, W) \geq \delta^\epsilon$.

Then,

$$\mathbf{B}_V(0, \delta^{\epsilon_0}) \subset \langle A, X \rangle_s + \mathbf{B}_V(0, \delta).$$

Note that for irreducible representations, this is a variant of Theorem 3.2. It is in this special case that we use results from Chapter 3. The rest of the proof of Theorem 5.2 consists of reducing to this special case by an induction on the length of the Jordan-Hölder decomposition of V.

Considering $\mathbb{R}^*$ acting on $\mathbb{R}$, we can recover from Theorem 5.2 Bourgain's discretized sum-product theorem (Theorem 1.4). Similarly, we can recover the discretized sum-product estimates for $\mathbb{C}$ or $\mathbb{H}$. Moreover, the method in the proof of Theorem 5.2 can be used to obtain sum-product estimates in semisimple algebras (i.e. sums of simple algebras).

[4] We mean $0 = V_0 < \ldots < V_l = V$ are submodules such that V_i/V_{i-1} are all irreducible.

Chapter 2

Generalities on discretized sets

In this chapter we collect some elementary and well-known results, especially arithmetic combinatorial tools, in the discretized setting. We provide proofs that are important or are not easily accessible in the literature. Some of the results are not used in later chapters.

2.1 Basics

We will first work in the Euclidean space $\mathbb{R}^n$. Anything we prove in $\mathbb{R}^n$ can be easily transferred to an arbitrary finite-dimensional normed vector space (at the price of losing a constant) if the proof is not already valid in general vector spaces.

Recall that λ denotes the Lebesgue measure.

Lemma 2.1. *Let $\delta > 0$ and let A be a bounded subset of $\mathbb{R}^n$. Let $\tilde{A}$ be a maximal 2δ-separated subset of A. Then*

$$\mathcal{N}_{2\delta}(A) \leq |\tilde{A}| \leq \mathcal{N}_\delta(A) \leq \mathcal{N}_1(\mathbf{B}(0,2))\,\mathcal{N}_{2\delta}(A), \tag{2.1}$$

and

$$|\tilde{A}| \leq \frac{\lambda(A^{(\delta)})}{\lambda(\mathbf{B}(0,\delta))} \leq 2^n \mathcal{N}_\delta(A).$$

As a consequence, $\mathcal{N}_\delta(A^{(\delta)}) \ll_n \mathcal{N}_\delta(A)$.

Proof. Let A and $\tilde{A}$ be as in the statement of the lemma. Let $\mathbf{B}(x_1,\delta),\dots,\mathbf{B}(x_N,\delta)$ be a cover of A by $N = \mathcal{N}_\delta(A)$ balls of radius δ.

If the balls $\mathbf{B}(a,2\delta)$, $a \in \tilde{A}$ do not cover A then we could find $b \in A \setminus \tilde{A}$ such that $\tilde{A} \cup \{b\}$ is still 2δ-separated, contradicting the maximality of $\tilde{A}$. Therefore, $\mathcal{N}_{2\delta}(A) \leq |\tilde{A}|$.

Consider the map $\tilde{A} \to \{1,\dots,N\}$ which maps each $a \in \tilde{A}$ to an index i such that $a \in \mathbf{B}(x_i,\delta)$. Since $\tilde{A}$ is 2δ-separated, this map is injective. Hence $|\tilde{A}| \leq \mathcal{N}_\delta(A)$.

The last inequality in (2.1) follows from a more general estimate

$$\forall \rho \geq \delta, \quad \mathcal{N}_\delta(A) \leq \mathcal{N}_\rho(A) \max_{x\in\mathbb{R}^n} \mathcal{N}_\delta(A \cap \mathbf{B}(x,\rho)) \tag{2.2}$$

and the fact that $\mathcal{N}_\delta(\mathbf{B}(x,2\delta)) = \mathcal{N}_1(\mathbf{B}(0,2))$ for any $x \in \mathbb{R}^n$. To see inequality (2.2), we first cover A by ρ-balls and then cover each of these ρ-balls by δ-balls.

Since $\tilde{A}$ is 2δ-separeted, the balls $\mathbf{B}(a,\delta)$, $a \in \tilde{A}$ are pairwise disjoint and included in $A^{(\delta)}$. Hence $|\tilde{A}|\lambda(\mathbf{B}(0,\delta)) \le \lambda(A^{(\delta)})$.

From $A \subset \bigcup_{i=1}^N \mathbf{B}(x_i,\delta)$ follows that $A^{(\delta)} \subset \bigcup_{i=1}^N \mathbf{B}(x_i,2\delta)$. Hence $\lambda(A^{(\delta)}) \le \mathcal{N}_\delta(A)\lambda(\mathbf{B}(0,2\delta))$. □

It is sometimes useful to change scales. Clearly, $\mathcal{N}_\delta(A)$ is nonincreasing in δ. Conversely, for all $\delta' \ge \delta$, we have

$$\mathcal{N}_\delta(A) \ll_n \Big(\frac{\delta'}{\delta}\Big)^n \mathcal{N}_{\delta'}(A). \tag{2.3}$$

It is a consequence of (2.2) and the fact $\mathcal{N}_\delta(\mathbf{B}(0,\delta')) \ll_n (\delta'\delta^{-1})^n$ which can be proved using Lemma 2.1.

If $f\colon \mathbb{R}^{n'} \to \mathbb{R}^n$ is a linear map with $\|f\| \le K$ where $K \ge 1$, or more generally if $f : A \to \mathbb{R}^n$ is K-Lipschitz, we have

$$\mathcal{N}_\delta(fA) \ll_n K^n \mathcal{N}_\delta(A). \tag{2.4}$$

For a family of bounded sets (A_i) we have

$$\mathcal{N}_\delta\big(\bigcup A_i\big) \le \sum \mathcal{N}_\delta(A_i).$$

This need not be an equality even if (A_i) is a disjoint family. However, we have an equality if $A_i^{(\delta)}$ are pairwise disjoint.

When we want to intersect two discretized sets $A, B \subset \mathbb{R}^n$, we shall take the δ-neighborhood of at least one of the sets before intersecting. Note that $\mathcal{N}_\delta(A^{(\delta)} \cap B^{(\delta)})$ can be large while at the same time $A \cap B$ is empty. The same is true for $A^{(2\delta)} \cap B^{(2\delta)}$ and $A^{(\delta)} \cap B^{(\delta)}$. However, we have

$$\mathcal{N}_\delta(A^{(2\delta)} \cap B) \ll_n \mathcal{N}_\delta(A^{(\delta)} \cap B^{(\delta)}) \ll_n \mathcal{N}_\delta(A \cap B^{(2\delta)}). \tag{2.5}$$

2.2 Ruzsa calculus

Let $\delta > 0$ be the scale. Let A, B, C be bounded subsets of $\mathbb{R}^n$. All implied constants in Landau and Vinogradov notations depend on n. Again results in this section can be transferred to an arbitrary finite-dimensional normed vector space with a loss of a constant factor.

Let $K \ge 2$ be a parameter. We denote by $\mathbf{O}(K)$ an unspecified finite set of cardinality $O(K)$. Moreover, we write $\mathbf{O}(K;A)$ to indicate that it is a subset of A. For example, we have

$$\mathbf{O}(K;A) + \mathbf{O}(L;B) \subset \mathbf{O}(KL, A+B).$$

Lemma 2.2 (Ruzsa's covering lemma)**.** *Let $K \ge 2$ be a parameter. If $\mathcal{N}_\delta(A+B) \le K\mathcal{N}_\delta(A)$, then*

$$B \subset A - A + \mathbf{O}(K;B) + \mathbf{B}(0,\delta).$$

Proof. Let B_0 be a maximal subset of B such that the translates $\left(b + A + \mathbf{B}(0, \frac{\delta}{2})\right)_{b \in B_0}$ are disjoint. We have, on the one hand, for any $b \in B$, the translate $b + A + \mathbf{B}(0, \frac{\delta}{2})$ is not disjoint from $b' + A + \mathbf{B}(0, \frac{\delta}{2})$ for some $b' \in B_0$ which means $b \in A - A + B_0 + \mathbf{B}(0, \delta)$. On the other hand, by the disjointness,

$$\mathcal{N}_\delta(A + B) \gg \mathcal{N}_{\frac{\delta}{2}}(A + B_0) \geq |B_0| \mathcal{N}_{\frac{\delta}{2}}(A).$$

Hence $|B_0| \ll K$. □

We can approximate $\mathbb{R}^n$ by the lattice $\delta \cdot \mathbb{Z}^n$. More precisely for a subset $A \subset \mathbb{R}^n$, we define

$$\tilde{A} = \left\{a \in \delta \cdot \mathbb{Z}^n \mid A \cap a^{(n\delta)} \neq \varnothing\right\}.$$

Then $A \subset \tilde{A}^{(n\delta)}$ and $\tilde{A} \subset A^{(n\delta)}$. Consequently $\mathcal{N}_\delta(A) \asymp |\tilde{A}|$. Moreover these inclusions behave nicely under addition and subtraction. Namely, for all $k, l \geq 0$,

$$kA - lA \subset k\tilde{A} - l\tilde{A} + \mathbf{B}(0, n(k + l)\delta)$$

and conversely

$$k\tilde{A} - l\tilde{A} \subset kA - lA + \mathbf{B}(0, n(k + l)\delta).$$

Using this approximation and the scale change estimate (2.1) we can translate additive combinatorial results in the discrete setting to the discretized setting. For the discrete version see the book [59] or [49] for the Plünnecke-Ruzsa inequality.

Lemma 2.3 (Ruzsa triangular inequality)**.** *We have*

$$\mathcal{N}_\delta(B)\mathcal{N}_\delta(A - C) \ll \mathcal{N}_\delta(A - B)\mathcal{N}_\delta(B - C).$$

Lemma 2.4 (Plünnecke-Ruzsa inequality)**.** *For all $K \geq 1$, if $\mathcal{N}_\delta(A + B) \leq K\mathcal{N}_\delta(B)$ then for all natural numbers k and l,*

$$\mathcal{N}_\delta(kA - lA) \leq O(K)^{k+l}\mathcal{N}_\delta(B).$$

2.3 Energy and Balog-Szemerédi-Gowers theorem

Before stating the Balog-Szemerédi-Gowers theorem in the discretized setting let us recall some basic facts about energy in the discrete setting. Let $\varphi\colon X \to Y$ be a map between discrete sets and A a finite subset of X, define the φ-energy of A to be

$$\omega(\varphi, A) = \sum_{y \in Y} |A \cap \varphi^{-1}(y)|^2.$$

In other words, it is the square of the l^2-norm of the push-forward of the counting measure on A under φ or the number of collisions of the map $\varphi_{|A}$:

$$\omega(\varphi, A) = \|\varphi_* \mathbf{1}_A\|_2^2 = \#\left\{(a_1, a_2) \in A \times A\colon \varphi(a_1) = \varphi(a_2)\right\}.$$

For example, the usual additive energy between two subsets A and B of an abelian group G is $\omega(+, A \times B)$ where $+$ is the group law of G.

When nothing is known about φ, $\omega(\varphi, A)$ can be as small as $|A|$ (when φ is injective) and as large as $|A|^2$ (when φ is constant on A). If the image of A by φ is small then the energy is large by the Cauchy-Schwarz inequality :

$$\omega(\varphi, A) \geq \frac{|A|^2}{|\varphi(A)|}. \tag{2.6}$$

The converse is not true. Nevertheless, we have a partial converse.

Lemma 2.5. *Suppose there are $K, M > 0$ such that $\omega(\varphi, A) \geq \frac{M}{K}|A|$ and for all $y \in Y$, $|A \cap \varphi^{-1}(y)| \leq M$. Then there exists $A' \subset A$ such that $|A'| \geq \frac{1}{2K}|A|$ and $|\varphi(A')| \leq \frac{2K}{M}|A|$.*

Proof. The idea is to trim off small fibers. We consider

$$Y' = \big\{y \in Y \mid |A \cap \varphi^{-1}(y)| \geq \frac{M}{2K}\big\}$$

and let $A' = \varphi^{-1}(Y')$. By the definition Y', we have

$$|A| \geq \sum_{y \in Y'} |A \cap \varphi^{-1}(y)| \geq \frac{M}{2K}|Y'|.$$

Hence $|\varphi(A')| \leq \frac{2K}{M}|A|$.

From the definition of the energy and the assumptions of the lemma,

$$\begin{aligned}\omega(\varphi, A) &\leq \frac{M}{2K}\sum_{y \notin Y}|A \cap \varphi^{-1}(y)| + M\sum_{y \in Y'}|A \cap \varphi^{-1}(y)| \\ &\leq \frac{M}{2K}|A| + M|A'|.\end{aligned}$$

It follows that $|A'| \geq \frac{1}{2K}|A|$. □

What Balog-Szemerédi-Gowers theorem roughly says is that if φ is a group law (or has some injectivity property similar to a group law) and A is a Cartesian product then the set A' in the conclusion of Lemma 2.5 can be chosen to be a Cartesian product.

For discretized sets we have an analogous notion of energy.

Definition. Let $\varphi\colon X \to Y$ be a Lipschitz map between metric spaces and A a bounded subset of X. We define the *φ-energy of A at scale δ* as

$$\omega_\delta(\varphi, A) = \mathcal{N}_\delta\big(\{(a, a') \in A \times A \mid d(\varphi(a), \varphi(a')) \leq \delta\}\big).$$

Here we adhere to the convention that the distance on a Cartesian product $X \times Y$ of metric spaces is such that

$$d\big((x, y), (x', y')\big)^2 = d(x, x')^2 + d(y, y')^2$$

for all pairs (x, y), $(x', y') \in X \times Y$.

Lemma 2.6. *Let n and m be positive integers. Let $A \subset \mathbb{R}^n$ be a bounded subset and $\varphi\colon \mathbb{R}^n \to \mathbb{R}^m$ a K-Lipschitz map.*

(i) The analogue of inequality (2.6) *is true :*

$$\omega_\delta(\varphi, A) \gg_{n,m} \frac{\mathcal{N}_\delta(A)^2}{\mathcal{N}_\delta(\varphi(A))}. \tag{2.7}$$

(ii) If $\psi\colon A \to \mathbb{R}^n$ *is another* K*-Lipschitz map, then*

$$\omega_\delta(\varphi, \psi A) \ll_n K^{2n}\, \omega_\delta(\varphi \circ \psi, A). \tag{2.8}$$

(iii) Let $\tilde{A}$ *be a maximal* δ*-separated subset of* A*. Then*

$$\omega_\delta(\varphi, A) \ll_n \#\left\{(a, a') \in \tilde{A} \times \tilde{A} \mid \|\varphi(a) - \varphi(a')\| \le (1+2K)\delta\right\}.$$

Proof. Let $\tilde{A}$ be a maximal δ-separated subset of A.

(i) Let Y be a finite subset of $\mathbb{R}^m$ such that $\varphi(A)$ is covered by the balls of radius $\frac{\delta}{2}$ centered at points in Y. Then

$$|\tilde{A}| \le \sum_{y \in Y} \big|\tilde{A} \cap \varphi^{-1}\big(\mathbf{B}(y, \frac{\delta}{2})\big)\big|,$$

and by the definition of ω_δ,

$$\sum_{y \in Y} \big|\tilde{A} \cap \varphi^{-1}\big(\mathbf{B}(y, \frac{\delta}{2})\big)\big|^2 \ll_n \omega_\delta(\varphi, A).$$

It follows from the Cauchy-Schwarz inequality that

$$|\tilde{A}|^2 \ll_n |Y|\, \omega_\delta(\varphi, A).$$

(ii) It is an easy consequence of (2.4).

(iii) For each $a \in A$, choose $\tilde{a} \in \tilde{A}$ such that $\|a - \tilde{a}\| \le \delta$. Let Ω be a maximal 8δ-separated subset of $\{(a, a') \in A \times A \mid \|\varphi(a) - \varphi(a')\| \le \delta\}$. Then the map $(a, a') \mapsto (\tilde{a}, \tilde{a}')$ is injective from Ω to the set on the right-hand side of the desired inequality. □

We will need the following additive version of Balog-Szemerédi-Gowers Theorem which gives a criterion for the additive energy between two sets to be large. See [57, Theorem 6.10] where it is proved in a much broader context.

Theorem 2.7 (Balog-Szemerédi-Gowers Theorem)**.** *Let* A*,* B *be bounded subsets of* $\mathbb{R}^n$*. If*

$$\omega_\delta(+, A \times B) \ge \frac{1}{K} \mathcal{N}_\delta(A)^{\frac{3}{2}} \mathcal{N}_\delta(B)^{\frac{3}{2}},$$

then there exists $A' \subset A$ *and* $B' \subset B$ *such that* $\mathcal{N}_\delta(A') \gg_n K^{-O(1)} \mathcal{N}_\delta(A)$*,* $\mathcal{N}_\delta(B') \gg_n K^{-O(1)} \mathcal{N}_\delta(Y)$ *and*

$$\mathcal{N}_\delta(A' + B') \ll_n K^{O(1)} \mathcal{N}_\delta(A)^{\frac{1}{2}} \mathcal{N}_\delta(B)^{\frac{1}{2}}.$$

2.4 Basic sum-product estimates

First consider a situation where $A \subset \mathrm{End}(\mathbb{R}^n)$ is a set of endomorphisms acting on a bounded subset $X \subset \mathbb{R}^n$ in a Euclidean space. Let $K \geq 2$ be a parameter. Similarly to the consideration of "good elements" in [13, Proposition 3.3] and the basic construction in [58, Proposition 3.1], we define at scale $\delta > 0$,

$$S_\delta(X;K) = \left\{ f \in \mathbf{B}_{\mathrm{End}(\mathbb{R}^n)}(0,K) \mid \mathcal{N}_\delta(X+fX) \leq K\mathcal{N}_\delta(X) \right\}.$$

Here are some basic properties of $S_\delta(X;K)$. The main idea is that it behaves like an approximate ring.

Lemma 2.8. *Let $X \subset \mathbf{B}(0,K)$ be a subset of $\mathbb{R}^n$, we have*

(i) If $a \in S_\delta(X;K)$ and $b \in \mathrm{End}(\mathbb{R}^n)$ such that $\|a-b\| \leq K\delta$, then $b \in S_\delta(X;K^{O(1)})$

(ii) If $\mathrm{Id}, a, b \in S_\delta(X;K)$, then $a+b$, $a-b$ and ab all belong to $S_\delta(X;K^{O(1)})$.

(iii) Suppose that a invertible and $\|a^{-1}\| \leq K$. If $a \in S_\delta(X;K)$, then $a^{-1} \in S_\delta(X;K^{O(1)})$.

(iv) If $\mathrm{Id}, a_1, \dots, a_s \in S_\delta(X;K)$, then

$$\mathcal{N}_\delta(X + a_1X + \cdots + a_sX) \leq K^{O(1)}\mathcal{N}_\delta(X).$$

(v) If $\mathrm{Id}, a \in S_\delta(X;K)$, then for all $\rho \geq \delta$, we have

$$\mathcal{N}_\rho(X+aX) \leq K^{O(1)}\mathcal{N}_\rho(X).$$

In other words, $a \in S_\rho(X;K^{O(1)})$.

Proof. (i) If $\mathcal{N}_\delta(X+aX) \leq K\mathcal{N}_\delta(X)$ and $\|a-b\| \leq K\delta$, then

$$X + bX \subset X + aX + \mathbf{B}(0,K^2\delta).$$

Hence $\mathcal{N}_\delta(X+bX) \leq K^{O(1)}\mathcal{N}_\delta(X)$.

(ii) Let $a, b \in S_\delta(X;K)$. By Ruzsa's covering lemma (Lemma 2.2),

$$aX \subset X - X + \mathbf{O}(K) + \mathbf{B}(0,\delta),$$

and

$$bX \subset X - X + \mathbf{O}(K) + \mathbf{B}(0,\delta).$$

Hence,

$$X + (a+b)X \subset 3X - 2X + \mathbf{O}(K^2) + \mathbf{B}(0,2\delta),$$
$$X + (a-b)X \subset 3X - 2X + \mathbf{O}(K^2) + \mathbf{B}(0,2\delta).$$

Finally by the Plünnecke-Ruzsa inequality (Lemma 2.4),

$$\mathcal{N}_\delta(X+(a+b)X),\ \mathcal{N}_\delta(X+(a-b)X) \leq K^{O(1)}\mathcal{N}_\delta(X).$$

Moreover, since $\|a\| \leq K$, we have

$$\begin{aligned} X + abX &\subset X + a(X - X + \mathbf{O}(K) + \mathbf{B}(0,\delta)) \\ &\subset X + aX - aX + \mathbf{O}(K) + \mathbf{B}(0,K\delta) \\ &\subset 3X - 2X + \mathbf{O}(K^3) + \mathbf{B}(0,3K\delta) \end{aligned}$$

Hence, $\mathcal{N}_\delta(X+abX) \leq K^{O(1)}\mathcal{N}_\delta(X)$.

(iii) If $a \in \mathrm{GL}_n(\mathbb{R})$ and $\|a^{-1}\| \leq K$, then $X + a^{-1}X = a^{-1}(X + aX)$ and hence $\mathcal{N}_\delta(X + a^{-1}X) \leq K^{O(1)}\mathcal{N}_\delta(X + aX)$.

(iv) The argument is similar to that of (ii).

(v) For all $\rho \geq \delta$ we have

$$\mathcal{N}_\delta(X) \leq \max_{x \in \mathbb{R}^n} \mathcal{N}_\delta(X \cap \mathbf{B}(x, \rho))\mathcal{N}_\rho(X)$$

and for all $x \in \mathbb{R}^n$,

$$\mathcal{N}_\delta(X + X + aX) \gg \mathcal{N}_\delta(X \cap \mathbf{B}(x, \rho))\mathcal{N}_\rho(X + aX).$$

If $\mathrm{Id}, a \in S_\delta(X; K)$, then by (iv),

$$\mathcal{N}_\delta(X + X + aX) \leq K^{O(1)}\mathcal{N}_\delta(X).$$

We obtain the desired estimate by combining the three inequalities above. □

Now let A be a bounded subset of a finite-dimensional normed algebra E. We can view left multiplications by $a \in A$ as elements of $\mathrm{End}(E)$. Lemma 2.8 tells us that, if for some $K \geq 2$,

$$\mathcal{N}_\delta(A + a \cdot A) \leq K\mathcal{N}_\delta(A)$$

for all $a \in A$, then for any positive integer s there is $s' = s'(s) \geq 1$ such that

$$\mathcal{N}_\delta(A + a \cdot A) \leq K^{s'}\mathcal{N}_\delta(A)$$

for all $a \in \langle A \rangle_s$.

Lemma 2.9. *Let $K \geq 2$ be a parameter. If $A \subset \mathbf{B}_E(0, K)$ is a subset of a finite-dimensional normed algebra E satisfying*

$$\mathcal{N}_\delta(A + A) + \mathcal{N}_\delta(A + A \cdot A) \leq K\mathcal{N}_\delta(A),$$

then for any positive integer s,

$$\mathcal{N}_\delta(\langle A \rangle_s) \leq K^{O_s(1)}\mathcal{N}_\delta(A).$$

Moreover, for any $\rho \geq \delta$,

$$\mathcal{N}_\rho(\langle A \rangle_s) \leq K^{O_s(1)}\mathcal{N}_\rho(A).$$

This lemma is a sum-product analogue of property (1.4) in groups . This discretized version can be proved by mimicking the proof of its discrete counterpart in [17, Lemma 5.5].

Proof. Without loss of generality, we can assume that the norm on E is submultiplicative and $0 \in A$. We prove by induction on s that there exists $k_s \geq 1$ such that

$$\langle A \rangle_s \subset k_s A - k_s A + \mathbf{O}(K^{k_s}; \langle A \rangle_{k_s}) + \mathbf{B}(0, K^{k_s}\delta). \tag{2.9}$$

This combined with the Plünnecke-Ruzsa inequality (Lemma 2.4) will finish the proof. For $s = 2$, the inclusion (2.9) is immediate from Ruzsa's covering lemma (Lemma 2.2). In particular, we have

$$A \cdot A \subset A - A + \mathbf{O}(K; A \cdot A) + \mathbf{B}(0, \delta). \tag{2.10}$$

For the induction step, we show that if (2.9) is true then $\langle A \rangle_s \pm A$ and $\langle A \rangle \cdot A$ also satisfy an inclusion of the form (2.9). For $\langle A \rangle_s \pm A$, this is clear. Multiplying A on the right to both sides of (2.9) yields

$$\langle A \rangle_s \cdot A \subset k_s(A \cdot A) - k_s(A \cdot A) + \mathbf{O}(K^{k_s}; \langle A \rangle_{k_s}) \cdot A + \mathbf{B}(0, K^{k_s+1}\delta).$$

By (2.10),

$$k_s(A \cdot A) \subset k_s A - k_s A + \mathbf{O}(K^{k_s^2}; k_s A \cdot A) + \mathbf{B}(0, k_s \delta).$$

For any $x \in \langle A \rangle_{k_s}$, by the discussion after the proof of Lemma 2.8, $\mathcal{N}_\delta(A + x \cdot A) \leq K^{O_s(1)} \mathcal{N}_\delta(A)$. Hence by Ruzsa's covering lemma (Lemma 2.2) again,

$$x \cdot A \subset A - A + \mathbf{O}(K^{O_s(1)}; \langle A \rangle_{k_s+1}) + \mathbf{B}(0, \delta).$$

Taking union of $O(K^{k_s})$ of such sets, we obtain

$$\mathbf{O}(K^{k_s}; \langle A \rangle_{k_s}) \cdot A \subset A - A + \mathbf{O}(K^{O_s(1)}; \langle A \rangle_{k_s+1}) + \mathbf{B}(0, \delta).$$

Putting these inclusions together gives

$$\langle A \rangle_s \cdot A \subset (2k_s + 1)A - (2k_s + 1)A + \mathbf{O}(K^{O_s(1)}; \langle A \rangle_{O_s(1)}) + \mathbf{B}(0, K^{O_s(1)}\delta),$$

which finishes the proof of the induction step.

The "moreover" part is obtained by the same argument as in the proof of Lemma 2.8(v). □

As a consequence of Lemma 2.9, in order to show that $\langle A \rangle_3$ is substantially larger than A, we can show that $\langle A \rangle_s$ is much larger than A. That is why the strategy of the proof of Theorem 3.1 is to show that for any $\epsilon_0 > 0$ there is $s \geq 1$ such that $\langle A \rangle_s$ contains a ball of radius δ^{ϵ_0}.

Conversely, if a growth statement such as Theorem 3.1 is true then within a bounded number of steps, the set always grows to become δ-dense in a large ball. This kind of phenomenon is often proved using Fourier analysis, as shown by Bourgain [7, Theorem 6]. Below we extend this argument to higher dimensional situation. It is also worth noting that these arguments are analogue of arguments using the high dimension of irreducible representations in groups, such as those used in Sarnak-Xue [52], Gowers [34], Bourgain-Gamburd [10] in the discrete setting and Bourgain-Gamburd [9, 11], Saxcé [23] in the discretized setting. In the following paragraphs, $\mathrm{End}(\mathbb{R}^n)$ is endowed with the usual operator norm.

Proposition 2.10. *Let $\epsilon > 0$. Given $\alpha > 0$ and $\beta > 0$ such that $\alpha + \beta > n$, there exists $s = s(\alpha + \beta - n, n) \geq 1$ such that the following is true for all $\delta > 0$ sufficiently small. Let μ be a probability measure on $\mathbf{B}_{\mathrm{End}(\mathbb{R}^n)}(0, 1)$ and ν a probability measure on $\mathbf{B}_{\mathbb{R}^n}(0, 1)$. Assume that*

(i) for all $\rho \geq \delta$ and all $\xi, \eta \in \mathbb{R}^n$ with $\|\xi\| = 1$,

$$\mu(\{a \in \mathrm{End}(\mathbb{R}^n) \mid {}^t a\xi \in \mathbf{B}(\eta, \rho)\}) \leq \delta^{-\epsilon} \rho^\alpha;$$

(ii) for all $\rho \geq \delta$ and all $x \in \mathbb{R}^n$, $\nu(\mathbf{B}(x,\rho)) \leq \delta^{-\epsilon}\rho^{\beta}$.

Then there exists $a_1, \ldots, a_s \in \mathrm{Supp}(\mu)$ such that

$$\mathbf{B}(0,\delta^{s\epsilon}) \subset a_1 X + \cdots + a_s X - a_1 X - \cdots - a_s X + \mathbf{B}(0,\delta)$$

where $X = \mathrm{Supp}(\nu)$.

Lemma 2.11. *Let ϵ, α, β, μ and ν be as in Proposition 2.10. Then for all $\xi \in \mathbb{R}^n$ with $1 \leq \|\xi\| \leq \delta^{-1}$,*

$$\int |\hat{\nu}({}^t a\xi)| \,\mathrm{d}\mu(a) \ll_n \delta^{-2\epsilon}\|\xi\|^{-\frac{\alpha+\beta-n}{n+3}}. \tag{2.11}$$

This lemma can be viewed as a higher dimensional extension of [7, Theorem 7].

Proof. First we prove that for any $1 \leq R \leq 2\delta^{-1}$,

$$\int_{\mathbf{B}(0,R)} |\hat{\nu}(\eta)|^2 \,\mathrm{d}\eta \ll_n \delta^{-\epsilon} R^{n-\beta}. \tag{2.12}$$

For any function φ, denote by φ^- the function $x \mapsto \varphi(-x)$. Similarly define ν^- for any measure on $\mathbb{R}^n$. Choose a smooth function $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ supported on $\mathbf{B}(0,1)$ and having $\int_{\mathbb{R}^n} \varphi = 1$. Replace it with $\varphi * \varphi^-$ if necessary, we can assume that $\hat{\varphi}(\xi) \in \mathbb{R}_+$ for all $\xi \in \mathbb{R}^n$. From the continuity of $\hat{\varphi}$, there is a constant $c > 0$ such that

$$\hat{\varphi}(\xi) \geq \frac{1}{2} \quad \text{for all } \xi \in \mathbf{B}(0,c).$$

For $R > 0$, set $\varphi_R(x) = c^{-n}R^n\varphi(c^{-1}Rx)$, $x \in \mathbb{R}^n$ so that $\|\varphi_R\|_\infty \ll_n R^n$ and $\mathrm{Supp}(\varphi_R) \subset \mathbf{B}(0, cR^{-1})$ and moreover

$$\forall \eta \in \mathbf{B}(0,R), \quad \hat{\varphi_R}(\eta) = \hat{\varphi}\big(\frac{c\eta}{R}\big) \geq \frac{1}{2}.$$

Thus for all $1 \leq R \leq 2\delta^{-1}$,

$$\int_{\mathbf{B}(0,R)} |\hat{\nu}(\eta)|^2 \,\mathrm{d}\eta \ll \int_{\mathbb{R}^n} |\hat{\nu}(\eta)|^2 \hat{\varphi_R}(\eta) \,\mathrm{d}\eta.$$

By the Parseval identity, the right-hand side is equal to

$$(2\pi)^n \int_{\mathbb{R}^n} \varphi_R \,\mathrm{d}\nu * \nu^- \ll_n \|\varphi_R\|_\infty (\nu * \nu^-)(\mathbf{B}(0, cR^{-1})).$$

Using assumption (ii), the right-hand side can be bounded by $O_n(\delta^{-\epsilon}R^{n-\beta})$. This finishes the proof of (2.12).

Now, consider $H_{R,t} = \{\eta \in \mathbf{B}_{\mathbb{R}^n}(0,R) \mid |\hat{\nu}(\eta)| \geq t.\}$ for $1 \leq R \leq \delta^{-1}$ and $0 < t \leq 1$. Since $\hat{\nu}$ is K-Lipschitz for some constant K depending only on n, we have

$$H_{R,t} + \mathbf{B}(0, \frac{t}{2K}) \subset H_{R+1,\frac{t}{2}}.$$

Thus $\lambda(H_{R,t}^{(t/2K)})t^2 \ll \int_{\mathbf{B}(0,R+1)}|\hat{\nu}|^2 \ll_n \delta^{-\epsilon}R^{n-\beta}$ and hence

$$\mathcal{N}_t(H_{R,t}) \ll_n \delta^{-\epsilon}t^{-n-2}R^{n-\beta}.$$

Let $\xi \in \mathbb{R}^n$ be a vector of length $1 \leq \|\xi\| = R \leq \delta^{-1}$, then

$$\int |\hat{\nu}({}^t a\xi)|\,\mathrm{d}\mu(a) \leq t + \mu(H_{R,t}^{\xi}).$$

where $H_{R,t}^{\xi} = \{a \in \mathrm{End}(\mathbb{R}^n) \mid {}^t a\xi \in H_{R,t}\}$. From a covering of $H_{R,t}$ by t-Balls we can lift to a covering of $H_{R,t}^{\xi}$ by sets of the form $\{a \in \mathrm{End}(\mathbb{R}^n) \mid {}^t a\xi \in \mathbf{B}(\eta,t)\}$. Using $t \leq 1$, $R \leq \delta^{-1}$ and assumption (i), we obtain, for all $\eta \in \mathbb{R}^n$,

$$\begin{aligned}\mu(\{a \in \mathrm{End}(\mathbb{R}^n) \mid {}^t a\xi \in \mathbf{B}(\eta,t)\}) &\leq \mu(\{a \in \mathrm{End}(\mathbb{R}^n) \mid {}^t a\frac{\xi}{R} \in \mathbf{B}(\frac{\eta}{R},\frac{1}{R})\}) \\ &\leq \delta^{-\epsilon}R^{-\alpha}.\end{aligned}$$

Therefore,

$$\mu(H_{R,t}^{\xi}) \leq \delta^{-\epsilon}R^{-\alpha}\mathcal{N}_t(H_{R,t}).$$

We obtain the desired estimate by setting $t = R^{-\frac{\alpha+\beta-n}{n+3}}$. □

Proof of Proposition 2.10. Write $\tau = \frac{\alpha+\beta-n}{n+3}$ and set $R = \min\{\delta^{-\frac{4\epsilon}{\tau}}, \delta^{-1}\}$. Let s be a positive integer. Taking the s-power of (2.11) and integrating over the annulus $\mathbf{B}(0,\delta^{-1}) \setminus \mathbf{B}(0,R)$, we obtain

$$\int_{R\leq\|\xi\|\leq\delta^{-1}} \Big(\int |\hat{\nu}({}^t a\xi)|\,\mathrm{d}\mu(a)\Big)^s \,\mathrm{d}\xi \leq \delta^{-3s\epsilon}R^{n-\tau s}$$

for $\delta > 0$ sufficiently small. Developing the left-hand side and using Fubini's theorem gives existence of $a_1, \ldots, a_s \in \mathrm{Supp}(\mu)$ such that

$$\int_{R\leq\|\xi\|\leq\delta^{-1}} |\hat{\nu}({}^t a_1\xi)|\cdots|\hat{\nu}({}^t a_s\xi)|\,\mathrm{d}\xi \leq \delta^{-3s\epsilon}R^{n-\tau s}.$$

Denote by ν' the image measure of $\nu^{\otimes s}$ by the map $(x_1,\cdots,x_s) \mapsto a_1x_1 + \cdots + a_sx_s$. The above inequality and the choice of R yield

$$\int_{R\leq\|\xi\|\leq\delta^{-1}} |\hat{\nu'}(\xi)|\,\mathrm{d}\xi \leq \delta^{s\epsilon-\frac{4n}{\tau}\epsilon}.$$

Let $\psi\colon \mathbb{R}^n \to \mathbb{R}$ be a smooth function supported on $\mathbf{B}(0,1)$ and with $\int_{\mathbb{R}^n}\psi = 1$ and moreover $\hat{\psi}(\xi) \in \mathbb{R}_+$ for all $\xi \in \mathbb{R}^n$. Set $\psi_\delta(x) = \delta^{-n+n\epsilon}\psi(\delta^{-1+\epsilon}x)$ so that it is supported on $\mathbf{B}(0,\delta^{1-\epsilon})$ and $\hat{\psi}_\delta(\xi) = \hat{\psi}(\delta^{1-\epsilon}\xi)$. The convolution $\nu' * \psi_\delta$ is a smooth function supported on $\mathbf{B}(0,s+1)$, hence there is $x_0 \in \mathrm{Supp}(\nu' * \psi_\delta)$ such that

$$\nu' * \psi_\delta(x_0) \gg_n s^{-n}. \tag{2.13}$$

By the Parseval identity, for all $x \in \mathbb{R}^n$

$$\nu' * \psi_\delta(x) = \int_{\mathbb{R}^n} e^{-i\langle\xi,x\rangle}\hat{\nu'}(\xi)\hat{\varphi}_\delta(\xi)\,\mathrm{d}\xi.$$

Hence for all $x \in \mathbf{B}(x_0, \delta^{s\epsilon})$

$$\begin{aligned}
&|\nu' * \psi_\delta(x_0) - \nu' * \psi_\delta(x)| \\
&\leq \int_{\mathbf{B}(0,R)} |1 - e^{-i\langle \xi, x - x_0\rangle}| \, \mathrm{d}\xi + 2\int_{R \leq \|\xi\| \leq \delta^{-1}} |\hat{\nu'}(\xi)| \, \mathrm{d}\xi + 2\int_{\|\xi\| \geq \delta^{-1}} |\hat{\psi_\delta}(\xi)| \, \mathrm{d}\xi \\
&\ll_n \delta^{s\epsilon} R^{n+1} + 2\delta^{s\epsilon - \frac{4n}{\tau}\epsilon} + \int_{\|\xi\| \geq \delta^{-1}} |\hat{\psi}(\delta^{1-\epsilon}\xi)| \, \mathrm{d}\xi
\end{aligned}$$

Pick $s = \frac{4(n+1)}{\tau} + 1$ so that the first two terms are $O(\delta^\epsilon)$. To bound the last integral, we note that since ψ is smooth, we have $|\hat{\psi}(\xi)| \ll_m \|\xi\|^{-m}$ for any $m > n$. Hence the last integral is

$$\int_{\|\xi\| \geq \delta^{-1}} |\hat{\psi}(\delta^{1-\epsilon}\xi)| \, \mathrm{d}\xi \ll_m \delta^{-m+m\epsilon} \int_{\|\xi\| \geq \delta^{-1}} \|\xi\|^{-m} \, \mathrm{d}\xi \ll_{n,m} \delta^{-n+m\epsilon}.$$

Now pick m large enough so that $m\epsilon - n \geq \epsilon$ and we obtain

$$|\nu' * \psi_\delta(x_0) - \nu' * \psi_\delta(x)| \ll_{n,\epsilon} \delta^\epsilon.$$

Combined with (2.13) we have $B(0, \delta^{s\epsilon}) \subset \mathrm{Supp}(\nu') - \mathrm{Supp}(\nu') + \mathbf{B}(0, 2\delta^{1-\epsilon})$ when δ is small enough. This is almost what we want except here we need a $2\delta^{1-\epsilon}$-neighborhood instead of δ-neighborhood. To finish the proof it suffices to see that we can work at the scale $\delta_1 = (\frac{\delta}{2})^{\frac{1}{1-\epsilon}}$ from the beginning. For example, the hypothesis (ii) implies that for all $\rho \geq \delta_1$ and all $x \in \mathbb{R}^n$, $\nu(\mathbf{B}(x, \rho)) \leq \delta_1^{-O_n(\epsilon)} \rho^\beta$. □

Recall that all algebras in this document are associative and unital.

Proposition 2.12. *Let E be a finite-dimensional normed algebra over $\mathbb{R}$. There is an integer s depending on E such that the following holds for all $\epsilon > 0$ and for $\delta > 0$ sufficiently small*[1]*. Let $A \subset \mathbf{B}_E(0,1)$. Assume that $\mathcal{N}_\delta(A) \geq \delta^{-\dim(E)+\epsilon}$. Then there exist elements $a_1, \ldots, a_s \in A$ such that*

$$\mathbf{B}(0, \delta^{O_E(\epsilon)}) \subset a_1 A + \cdots + a_s A - a_1 A - \cdots - a_s A + \mathbf{B}(0, \delta).$$

This proposition can be viewed as a discretized analogue of [17, Lemma 6.4]. Note also that the statement with right multiplication instead of left one is equally true.

Proof. This is a direct consequence of Proposition 2.10. Without loss of generality, we can suppose that the norm on E is Euclidean. Denote by n the dimension of E. Let ν be the normalized restriction of the Lebesgue measure to $A^{(\delta)}$. The assumption (ii) of Proposition 2.10 for $\beta = n$ follows immediately from Lemma 2.1.

Let μ be the push forward of ν under the map $E \to \mathrm{End}(E)$, $a \mapsto l_a$ where l_a denote the left multiplication by a. We claim that μ satisfies the assumption (i) of Proposition 2.10 with $\alpha = 1$.

For any $\xi \in E$, denote by m_ξ the linear endomorphism $a \mapsto {}^t l_a \xi$. The linear application $E \to \mathrm{End}(E)$, $\xi \mapsto m_\xi$ is injective since E is unital. Thus, there exists $c > 0$ such that $\|m_\xi\| \geq c\|\xi\|$. Let $\xi, \eta \in E$ with $\|\xi\| = 1$ and let $\rho \geq \delta$.

[1] smaller than a constant depending on E and ϵ.

We want be bound $\nu(m_\xi^{-1}(\mathbf{B}(\eta,\rho)))$. We know that $\|m_\xi\| \geq c$. Consequently, from the Cartan decomposition of m_ξ we see that

$$m_\xi^{-1}(\mathbf{B}(0,\rho)) \subset W_\xi + \mathbf{B}(0, c^{-1}\rho)$$

where W_ξ is a $n-1$-dimensional subspace. Hence either $m_\xi^{-1}(\mathbf{B}(\eta,\rho))$ is empty or there is $a_0 \in E$ such that

$$m_\xi^{-1}(\mathbf{B}(\eta,\rho)) \subset a_0 + W_\xi + \mathbf{B}(0, 2c^{-1}\rho).$$

We conclude that

$$\mu\big(\{f \in \mathrm{End}(E) \mid {}^t f\xi \in \mathbf{B}(\eta,\rho)\big) = \nu(m_\xi^{-1}(\mathbf{B}(\eta,\rho))) \ll_E \delta^{-\epsilon}\rho,$$

which proves the claim and hence the proposition. □

2.5 Noncommutative analogues

Let G be connected real Lie group. Let μ be a left-invariant Haar measure on G and d a left-invariant smooth metric on G.

Due to non-zero curvature, the precise estimates in Lemma 2.1 valid in $\mathbb{R}^n$ does not hold in general. Nevertheless, the same type of estimates hold if we restrict ourself to a compact set and allow constants to depend on this compact. Actually, (G,d) is a *locally reasonable metric group* following the definition of Tao [57, Definition 6.3]. Namely,

(i) The topology on G is compatible with the metric d. Closed balls for d are compact.

(ii) For any $R \geq 1$, we have

$$d(xg, yg),\ d(gx, gy),\ d(x^{-1}, y^{-1}) \asymp_{G,R} d(x,y)$$

for all $x, y, g \in \mathbf{B}(1,R)$.

(iii) For any $R \geq 1$, we have

$$\mu(\mathbf{B}(1,2r)) \ll_{G,R} \mu(\mathbf{B}(1,r))$$

for all $0 < r < R$.

Form these properties, estimates similar to Lemma 2.1 follows. Let us recall some of them, for more details see [57, §6]. Let $R \geq 1$. Let $X \subset \mathbf{B}(1,R)$. For all $0 < \delta \leq 1$, we have

$$X\,\mathbf{B}(1,\delta) \subset \mathbf{B}(1, O_{G,R}(\delta))X,$$

$$\mathbf{B}(1,\delta)X \subset X\,\mathbf{B}(1, O_{G,R}(\delta)),$$

$$\frac{\mu(X^{(\delta)})}{\mu(\mathbf{B}(1,\delta))} \asymp_{G,R} \mathcal{N}_\delta(X), \tag{2.14}$$

and

$$\mathcal{N}_{2\delta}(X) \asymp_{G,R} \mathcal{N}_\delta(X) \asymp_{G,R} \mathcal{N}_\delta(X^{(\delta)}).$$

There is also a useful property specific to Lie groups. For all $0 < \delta \leq 1$ and all closed connected subgroups $H < G$,

$$\mathcal{N}_\delta(\mathbf{B}_H(1,1)) \asymp_H \delta^{-\dim(H)}.$$

The Plünnecke-Ruzsa inequality does not hold for noncommutative groups. However the Ruzsa triangular inequality still holds in the noncommutative setting. From it we can deduce a weak but noncommutative version of Plünnecke-Ruzsa inequality. In the discretized setting, its statement is the following.

Lemma 2.13 (Tao [57, Theorem 6.8])**.** *Let $R \geq 1$ and $K \geq 2$ be parameters. If $A \subset \mathbf{B}_G(1,R)$ satisfies $\mathcal{N}_\delta(A^3) \leq K\mathcal{N}_\delta(A)$ then for all $s \geq 1$,*

$$\mathcal{N}_\delta\big((A \cup \{1\} \cup A^{-1})^s\big) \ll_{G,R} K^{O_s(1)}\mathcal{N}_\delta(A).$$

It is worth noting that the Balog-Szemerédi-Gowers Theorem also holds in the noncommutative and δ-discretized setting, see [57, Theorem 6.10].

2.5.1 Metric entropy version of the Petridis lemma

In this subsection, we prove an discretized version of a result known as the Petridis lemma [49]. It implies the Plünnecke-Ruzsa inequality in the commutative setting and thus can be viewed as a substitute for Plünnecke-Ruzsa inequality in the noncommutative setting.

Lemma 2.14 (Metric entropy version of the Petridis lemma)**.** *Let (G,d) be a locally reasonable metric group. Let $R \geq 1$ and $K \geq 2$ be parameters and let $0 < \delta < 1$ be a scale. If A and B are subsets of $\mathbf{B}_G(1,R)$ such that $\mathcal{N}_\delta(AB) \leq K\mathcal{N}_\delta(A)$, then there exists $A_0 \subset A$ such that for any subset $Y \subset \mathbf{B}_G(1,R)$,*

$$\mathcal{N}_\delta(YA_0B) \ll_{G,R} K\mathcal{N}_\delta(YA_0).$$

Like in [57], the strategy is to first prove a continuous version. Then the lemma follows from properties of locally reasonable metric groups such as estimate (2.14). Till the end of this section, G denote a locally compact Lie group and μ an left-invariant Haar-measure.

The main idea in Petridis' proof is to consider a subset $A_0 \subset A$ which minimizes the ratio $\frac{|A_0B|}{|A_0|}$. So in the continuous setting we want a subset A_0 minimizing the ratio $\frac{\mu(A_0B)}{\mu(A_0)}$.

Let $\mathcal{B}$ denote the Borel σ-algebra of a locally compact group G. Write $\mathcal{B}_{>0}$ for the set of all Borel subsets of positive measure. Let $B \subset G$ be a measurable relatively compact subset with positive measure and define $f\colon \mathcal{B} \to \mathbb{R}_+ \cup \{+\infty\}$ as

$$f(A) = \mu(AB), \quad \forall A \in \mathcal{B}.$$

Properties of the function f can be summarized as follows.

(i) (Monotonicity) for all $A, A' \in \mathcal{B}$, if $A \subset A'$, then $f(A) \leq f(A')$.

(ii) (Continuity from below) if (A_n) is sequence of measurable sets with $A_n \subset A_{n+1}$ for all n, then

$$f(\bigcup_{n\geq 1} A_n) \leq \sup_n f(A_n).$$

(iii) (Submodularity) for all $A, A' \in \mathcal{B}$,

$$f(A \cup A') + f(A \cap A') \leq f(A) + f(A').$$

(iv) (Bounded away from 0) there exists $c > 0$ such that $f(A) \geq c$ for all $A \in \mathcal{B}_{>0}$.

(v) (Left-invariance) for all $A \subset \mathcal{B}$ and all $g \in G$, $f(gA) = f(A)$.

From now on we can forget the definition of f and only use the above properties. The only difficulty for transferring Petridis' proof to the continuous setting is to prove the existence of sets minimizing the ratio $\frac{f(A')}{\mu(A')}$ where A' ranges over all subsets of A with positive measure.

Lemma 2.15. *Let $A \subset G$ be a measurable subset of finite positive measure. Let $\mathcal{A}_{>0}$ denote the set $\{A' \in \mathcal{B}_{>0} \mid A' \subset A\}$. Assume that $f \colon \mathcal{A} \to \mathbb{R}_+$ is a function satisfying properties (i)-(iv) above. Then the infimum*

$$K = \inf_{A' \in \mathcal{A}_{>0}} \frac{f(A')}{\mu(A')}$$

is reached. Moreover, countable unions of minimizing sets are minimizing. Consequently, up to a null set, there is a unique minimizing set of maximal measure. Any minimizing set is a subset of this set up to a null set.

Proof. We first show that finite unions of nearly minimizing sets are nearly minimizing. More precisely we claim that if $A', A'' \in \mathcal{A}_{>0}$ are such that

$$f(A') \leq (K + \epsilon')\mu(A') \text{ and } f(A'') \leq (K + \epsilon'')\mu(A'') \tag{2.15}$$

for some $\epsilon', \epsilon'' \geq 0$, then

$$f(A' \cup A'') \leq (K + \epsilon' + \epsilon'')\mu(A' \cup A'').$$

To prove the claim, we note that $f(A' \cap A'') \geq K\mu(A' \cap A'')$ by the definition of K. Substituting this and (2.15) into the submodularity inequality yields

$$f(A' \cup A'') \leq (K + \epsilon')\mu(A') + (K + \epsilon'')\mu(A'') - K\mu(A' \cap A'').$$

Then the monotonicity and modularity of μ gives the claim.

Now choose a sequence $(A_n)_{n \geq 1}$ of measurable sets with positive measure such that $f(A_n) \leq (K + 2^{-n})$. Note that by the condition (iv), $\mu(A_n) \geq \frac{c}{K+1}$ for any $n \geq 1$. Hence

$$\mu(\limsup_{n \to +\infty} A_n) = \lim_{N \to +\infty} \mu(\bigcup_{n \geq N} A_n) > 0.$$

By the claim, for all $k \geq 1$

$$f(\bigcup_{n=N}^{N+k} A_n) \leq (K + 2^{1-N})\mu(\bigcup_{n=N}^{N+k} A_n).$$

Taking the limit for $k \to +\infty$ and using condition (ii), we obtain

$$f(\bigcup_{n \geq N} A_n) \leq (K + 2^{1-N})\mu(\bigcup_{n \geq N} A_n).$$

Finally using condition (i) and taking the limit for $N \to +\infty$, we obtain

$$f(\limsup_{n\to+\infty} A_n) \leq K\mu(\limsup_{n\to+\infty} A_n).$$

This proves the existence of minimizing sets.

The claim above already shows that finite unions of minimizing sets are minimizing. The fact that this is also true for countable unions follows from the continuity from below of f and μ. The union of a sequence (A_n) of minimizing sets with $\mu(A_n)$ tending to the supremum of the measure of all minimizing sets is a minimizing set of maximal measure. The uniqueness follows immediately. □

Lemma 2.16 (Continuous version of the Petridis lemma)**.** *Let $f\colon \mathcal{B} \to \mathbb{R}_+ \cup \{+\infty\}$ be a function satisfying properties (i)-(v) on page 26. Let A be a measurable subset with $0 < \mu(A) < +\infty$. Then there exists a Borel set $A_0 \subset A$ such that for any countable subset $Y \subset G$,*

$$f(YA_0) \leq \frac{f(A)}{\mu(A)}\mu(YA_0).$$

Proof. The previous lemma shows the existence of a Borel subset $A_0 \subset A$ such that

$$\frac{f(A_0)}{\mu(A_0)} = K = \inf_{\substack{A'\subset A\\ \mu(A')>0}} \frac{f(A')}{\mu(A')} \leq \frac{f(A)}{\mu(A)}.$$

We consider the collection of subsets $Y \subset G$ such that $f(YA_0) \leq K\mu(YA_0) < +\infty$. This collection contains $\{1\}$. It is closed under G-translation on the left by the left-invariance of f and μ. We claim that it is also closed under adding the identity element. Indeed, by the submodularity,

$$f(YA_0 \cup A_0) + f(YA_0 \cap A_0) \leq f(YA_0) + f(A_0).$$

By the definition of K, $f(YA_0 \cap A_0) \geq K\mu(YA_0 \cap A_0)$. Hence

$$f((Y \cup \{1\})A_0) \leq K(\mu(YA_0) + \mu(A_0) - \mu(YA_0 \cap A_0)) = K\mu((Y \cup \{1\}) \cup A_0).$$

Thus, the conclusion of the lemma holds for any finite subset Y. Too see that it is also true if Y is countable it suffices to use the continuity from below. □

Now we deduce Lemma 2.14 from Lemma 2.16.

Proof of Lemma 2.14. Let G, R, A, B and K be as in the statement of Lemma 2.14. From $\mathcal{N}_\delta(AB) \leq K\mathcal{N}_\delta(A)$ and the basic properties of locally reasonable metric groups, we deduce that

$$\mu(A^{(\delta)}B^{(\delta)}) \ll_{G,R} \mathcal{N}_\delta(AB)\mu(\mathbf{B}(1,\delta)) \ll_{G,R} K\mathcal{N}_\delta(A)\mu(\mathbf{B}(1,\delta)) \ll_{G,R} K\mu(A^{(\delta)})$$

Now we apply Lemma 2.16 to the set of positive measure $A^{(\delta)}$ and the function $f\colon A' \mapsto \mu(A'B^{(\delta)})$. We get $A_0 \subset A^{(\delta)}$ such that for any countable set $Y \subset G$,

$$\mu(YA_0B^{(\delta)}) \ll_{G,R} K\mu(YA_0).$$

Set $A_1 = A_0^{(\delta)} \cap A$ so that we have both $A_1 \subset A_0^{(\delta)}$ and $A_0 \subset A_1^{(\delta)}$. Let $Y \subset \mathbf{B}_G(1,R)$. We claim that

$$\mathcal{N}_\delta(YA_1B) \ll_{G,R} K\mathcal{N}_\delta(YA_1),$$

which finishes the proof of Lemma 2.14.

To prove the claim, consider $\tilde{Y} \subset Y$ a finite subset such that $Y \subset \tilde{Y}^{(\delta)}$. We know

$$\mu(\tilde{Y}A_0B^{(\delta)}) \ll_{G,R} K\mu(\tilde{Y}A_0).$$

On the one hand, we can bound from above the right hand side by

$$\mu(\tilde{Y}A_0) \leq \mu(YA_1^{(\delta)}) \leq \mu((YA_1)^{(O_{G,R}(\delta))}) \ll_{G,R} \mathcal{N}_\delta(YA_1)\mu(\mathbf{B}(1,\delta)).$$

On the other hand, we can bound from below the left hand side : there exists $C \ll_{G,R} 1$ such that

$$(\tilde{Y}A_0B)^{(\frac{\delta}{C})} \subset \tilde{Y}A_0B^{(\delta)} \quad \text{and} \quad YA_1B \subset \tilde{Y}^{(\delta)}A_0^{(\delta)}B \subset (\tilde{Y}A_0B)^{(C\delta)}.$$

Hence

$$\mu(\tilde{Y}A_0B^{(\delta)}) \geq \mu((\tilde{Y}A_0B)^{(\frac{\delta}{C})}) \gg_{G,R} \mathcal{N}_{\frac{\delta}{C}}(\tilde{Y}A_0B)\mu(\mathbf{B}(1,\frac{\delta}{C})).$$

Moreover, we have $\mu(\mathbf{B}(1,\frac{\delta}{C})) \gg_{G,R} \mu(\mathbf{B}(1,\delta))$ and

$$\mathcal{N}_{\frac{\delta}{C}}(\tilde{Y}A_0B) \gg_{G,R} \mathcal{N}_{C\delta}((\tilde{Y}A_0B)^{(C\delta)}) \gg_{G,R} \mathcal{N}_{C\delta}(YA_1B) \gg_{G,R} \mathcal{N}_\delta(YA_1B).$$

Putting all these inequalities together, we obtain the claim. □

Chapter 3

Sum-product estimates in matrix algebras

In the chapter we generalize Bourgain's discretized sum-product theorem to matrix algebras. By normed algebra we mean an associative unital algebra over $\mathbb{R}$ endowed with a norm that makes the underlying linear space a normed vector space. Let E be a normed algebra. We want to understand, given a bounded subset $A \subset E$ satisfying similar properties as in Theorem 1.4, how $\mathcal{N}_\delta(\langle A \rangle_s)$ grows and whether (1.3) or similar estimates hold.

If we ask these questions for general real algebras, they can be as hard as the Freiman problem[1] as illustrated by the following example. Let A_0 be a bounded subset of $\mathbb{R}$ containing 0. Consider A the set of matrices of the form $\left(\begin{smallmatrix} 1 & a \\ 0 & 1 \end{smallmatrix}\right)$ with $a \in A_0$. Then for any positive integer s, every element in $\langle A \rangle_s$ is of the form $\left(\begin{smallmatrix} k & a \\ 0 & k \end{smallmatrix}\right)$ with $a \in s'A_0$ and $k \in \{-s', \dots, s'\}$ where s' is an integer depending on s. Conversely, for every $a \in sA_0$, we have $\left(\begin{smallmatrix} s & a \\ 0 & s \end{smallmatrix}\right) \in \langle A \rangle_s$. Hence

$$\mathcal{N}_\delta(sA_0) \leq \mathcal{N}_\delta(\langle A \rangle_s) \ll_s \mathcal{N}_\delta(s'A_0).$$

This means the growth of A under addition and multiplication is somehow equivalent to the growth of A_0 under only addition.

That is why we will restrict our attention to simple algebras. By the Wedderburn structure theorem and the Frobenius theorem, we know that a simple real algebra of finite dimension is isomorphic to $\mathcal{M}_n(\mathbb{R})$, $\mathcal{M}_n(\mathbb{C})$ or $\mathcal{M}_n(\mathbb{H})$, the algebra of $n \times n$ matrices over the real numbers, the complex numbers, or the quaternions, for some $n \geq 1$.

Theorem 3.1. *Let E be a normed simple real algebra of finite dimension. Given $\kappa > 0$ and $\sigma < \dim(E)$, there is $\epsilon > 0$ depending on E, κ and σ such that the following holds for $\delta > 0$ sufficiently small. Let A be a subset of E, assume that*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$,

[1] Freiman problem asks, in a given abelian group, which subsets grow slowly under addition. Freiman's Theorem asserts that they are "close" to generalized arithmetic progressions. Obtaining a polynomial bound for this theorem is one of the fundamental open problems in additive combinatorics. See [59, Chapter 5].

(iii) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$,

(iv) for every proper subalgebra $W \subset E$, *there is* $a \in A$ *such that* $d(a, W) \geq \delta^\epsilon$.

Then,

$$\mathcal{N}_\delta(A+A) + \mathcal{N}_\delta(A+A\cdot A) \geq \delta^{-\epsilon}\mathcal{N}_\delta(A). \tag{3.1}$$

The case $E = \mathbb{C}$ is due to Bourgain and Gamburd [9]. So our result is new for $\dim(E) > 2$.

The assumption (ii) is a non-concentration condition. It is to avoid the situation where A is a union of a bounded number of small balls. A subset $A \subset E$ satisfying the condition (iv) will be said to be δ^ϵ-away from subalgebras. This is the additional condition compared to the one-dimensional case. Without it A can be trapped in a small neighborhood of a proper subalgebra. Note the conclusion (3.1) is slightly weaker than that of Theorem 1.4. Here, instead of $A \cdot A$, we need $A + A \cdot A$ to see the growth. Actually, the estimate (1.3) fails under the same assumptions as soon as $\dim(E)$ is greater than 1. Indeed, if A is a union of a segment of unit length and an orthonormal basis of E, then the set A satisfies the assumptions[2] of Theorem 3.1 but $A + A$ and $A \cdot A$ are both unions of a bounded number of unit segments. Thus (1.3) fails for such A.

Using the fact that all norms on a finite-dimensional linear space are equivalent, it is easy to see that the constant $\epsilon > 0$ can be made independent of the choice of the norm while scale δ need to be smaller than a constant δ_0 depending on E, its norm, σ and κ. Moreover, this constant δ_0 can be made uniform for norms ranging in a compact set (for the topology of pointwise convergence). However, the dependence on the norm can not be removed completely as illustrated by the following example. Let $E = \mathbb{C}$ and $A = [0,1] \cdot i$. For any $r > 0$, consider the norm defined by $\forall x, y \in \mathbb{R}$, $\|x + iy\|_r = r|x| + |y|$. Then the assumptions of Theorem 3.1 are satisfied but $\mathcal{N}_\delta(A+A)+\mathcal{N}_\delta(A+A\cdot A) \ll \mathcal{N}_\delta(A)$ if $r < \delta$.

Our second result concerns linear actions on Euclidean spaces. Let X be a bounded subset of the Euclidean space $\mathbb{R}^n$. Let $A \subset \mathrm{End}(\mathbb{R}^n)$ be a collection of linear endomorphisms. We can ask whether X grows under addition and transformation by elements of A, provided that A is sufficiently rich.

Theorem 3.2. *Let* n *be a positive integer. Given* $\kappa > 0$ *and* $\sigma < n$, *there is* $\epsilon > 0$ *such that the following holds for* $\delta > 0$ *sufficiently small. Let* A *be a subset of* $\mathrm{End}(\mathbb{R}^n)$ *and* X *a subset of* $\mathbb{R}^n$, *assume that*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) for every nonzero proper linear subspace $W \subset \mathbb{R}^n$, *there is* $a \in A$ *and* $w \in \mathbf{B}_W(0,1)$ *such that* $d(aw, W) \geq \delta^\epsilon$.

(iv) $X \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(v) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(X) \geq \delta^\epsilon \rho^{-\kappa}$,

(vi) $\mathcal{N}_\delta(X) \leq \delta^{-\sigma-\epsilon}$.

[2] The condition $\dim(E) > 1$ is needed to have assumption (iv).

Then,

$$\mathcal{N}_\delta(X+X) + \max_{a\in A} \mathcal{N}_\delta(X+aX) \geq \delta^{-\epsilon}\mathcal{N}_\delta(X), \tag{3.2}$$

where $aX = \{ax \mid x \in X\}$.

This improves a previous result of Bourgain and Gamburd [11, Proposition 1].

As a simple corollary, we can obtain a "sum-bracket" estimate in simple Lie algebras. If A is a subset of a Lie algebra $\mathfrak{g}$, write $[A,A] = \{[a,b] \mid a,b \in A\}$.

Corollary 3.3. *Let* $\mathfrak{g}$ *be a normed*[3] *simple Lie algebra of finite dimension. Given* $\kappa > 0$ *and* $\sigma < \dim(\mathfrak{g})$*, there is* $\epsilon > 0$ *such that the following holds for* $\delta > 0$ *sufficiently small. Let* A *be a subset of* $\mathfrak{g}$*, assume that*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$,

(iv) for every proper Lie subalgebra W *of* $\mathfrak{g}$*, there is* $a \in A$ *such that* $d(a,W) \geq \delta^\epsilon$.

Then,

$$\mathcal{N}_\delta(A+A) + \mathcal{N}_\delta(A+[A,A]) \geq \delta^{-\epsilon}\mathcal{N}_\delta(A).$$

3.0.1 Outline of the proofs

Both Theorem 3.1 and Theorem 3.2 are deduced from the following theorem.

Theorem 3.4. *Let* E *be a normed simple real algebra of finite dimension. Given* $\kappa > 0$ *and* $\epsilon_0 > 0$*, there is* $\epsilon > 0$ *and an integer* $s \geq 1$ *such that the following holds for* $\delta > 0$ *sufficiently small. Let* A *be a subset of* E*, assume that*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) A *is* δ^ϵ*-away from subalgebras.*

Then,

$$\mathbf{B}(0, \delta^{\epsilon_0}) \subset \langle A \rangle_s + \mathbf{B}(0, \delta).$$

Note that each of the conditions (ii) and (iii) rules out one obvious obstruction for $\langle A \rangle_s$ to grow. Indeed, firstly, if A is covered by a bounded number of balls of radius ρ with $\rho < \delta^{\epsilon_0}$, then $\langle A \rangle_s$ is covered by $O_s(1)$ balls of radius ρ. Secondly, if A is contained in the unit ball[4] and in the ρ-neighborhood of a proper subalgebra with $\rho < \delta^{\epsilon_0}$, then $\langle A \rangle_s$ is contained in the $O_s(\rho)$-neighborhood of the same proper subalgebra.

The main ingredient in the proof of Theorem 3.4 is a sum-product theorem [11, Corollary 8] due to Bourgain-Gamburd concerning the ring $\mathbb{C}^n$, the n-fold direct product of $\mathbb{C}$ with itself. Let n be a positive integer. We denote by Δ the set of diagonal matrices in $\mathcal{M}_n(\mathbb{C})$.

[3] We mean a norm which makes the underlying linear structure a normed vector space.
[4] In this example, the norm on E is submultiplicative, i.e. $\forall x,y \in E$, $\|xy\| \leq \|x\|\|y\|$. This assumption is not restrictive since every norm on E is equivalent to a submultiplicative one.

Theorem 3.5 (Bourgain-Gamburd [11]). *Given $\kappa > 0$ and n a positive integer, there is a positive integer $s \geq 1$ such that, for $\delta > 0$ sufficiently small, the following holds. Let A be a subset of $\mathcal{M}_n(\mathbb{C})$. Assume that*

(i) $A \subset \mathbf{B}(0,1)$,

(ii) $\mathcal{N}_\delta(A) \geq \delta^{-\kappa}$,

(iii) $A \subset \Delta + \mathbf{B}(0,\delta)$.

Then there is $\eta \in \Delta$ with $\|\eta\| = 1$ such that

$$[0,\delta^\alpha]\eta \subset \langle A\rangle_s + \mathbf{B}(0,\delta^{\alpha+\beta}),$$

with some $0 \leq \alpha < C(n,\kappa)$ and some $\beta > c(n,\kappa) > 0$.

Let us sketch the proof of Theorem 3.4. In the following paragraphs, each s stands for some unspecified integer that can be bounded in terms of E, κ. In order to use the Bourgain-Gamburd theorem above, we need first to embed the algebra E in $\mathcal{M}_n(\mathbb{C})$ and then produce a lot of nearly simultaneously diagonalizable elements in $\langle A\rangle_s$. The standard way (since [36]) to produce such elements is to use the fact that the centralizer of a matrix with n distinct eigenvalues is simultaneously diagonalizable. Since the set of matrices with at least one multiple eigenvalue is an algebraic subvariety of $\mathcal{M}_n(\mathbb{C})$, to find an element $a \in \langle A\rangle_s$ with n distinct eigenvalues we use the technique of "escape from subvarieties", first developed in [30]. For our discretized setting, a quantitative version of this technique is required since distance matters. For Lie groups, this is established in [24]. Here we adapt the argument in the sum-product setting.

Once we have such an element a, we consider the map $\varphi\colon x \mapsto ax - xa$. We distinguish two cases.

(a) If $\varphi(A)$ is large ($\mathcal{N}_\delta(\varphi(A)) \geq \delta^{\kappa'}\mathcal{N}_\delta(A)$ with $\kappa' = \frac{\kappa}{3\dim(E)}$), then we will prove $\mathcal{N}_\delta(\langle A\rangle_s) \geq \delta^{-\kappa'}\mathcal{N}_\delta(A)$ in this case. We remark that all element in $\varphi(A)$ have zero trace. Hence if B is a set of matrices with a lot of different traces, then $\varphi(A) + B$ contains a lot of disjoint translates of $\varphi(A)$. In particular, $\mathcal{N}_\delta(\varphi(A) + B) \gg \mathcal{N}_\delta(\varphi(A))\mathcal{N}_\delta(\mathrm{tr}(B))$. Thus, it suffices to establish a lower bound on the size of the set of traces of $\langle A\rangle_s$. Indeed, we can prove $\mathrm{tr}(\langle A\rangle_s) \geq \delta^{-2\kappa'}$ using the fact that the bilinear form $(x,y) \mapsto \mathrm{tr}(xy)$ is non-degenerate.

(b) Otherwise the set A must have a large intersection with a fiber of φ, i.e. there is $y \in \mathcal{M}_n(\mathbb{C})$ such that $\mathcal{N}_\delta(A \cap \varphi^{-1}(\mathbf{B}(y,\delta))) \geq \delta^{-\kappa'}$. The difference set of the above intersection consists of nearly simultaneously diagonalizable matrices. Then we can apply Theorem 3.5 to get a small segment at a smaller scale (a segment of length δ^α is inside the $\delta^{\alpha+\beta}$-neighborhood of $\langle A\rangle_s$, where α, β and s are the constants given by Theorem 3.5).

What we do is to repeat the same argument to $\langle A\rangle_s$ if case (a) happens. After a bounded number of times, case (a) won't be possible because $\langle A\rangle_s \subset \mathbf{B}(0, O_s(\delta^{-O_s(\epsilon)}))$. Hence eventually, case (b) is true, i.e. inside $\langle A\rangle_s$, there is a segment of direction ξ and length δ^α at scale $\delta^{\alpha+\beta}$. Then, using the fact that the two-sided ideal generated by ξ is the whole algebra E, we can prove that

the small segment will grow into a small ball under left and right multiplication by elements of A.

This almost finishes the proof. The only problem is that the ball obtained is not large enough and it is at a different scale than δ. As in [24], this issue can be solved by applying the above argument at various scales ranging from $\delta^{\frac{1}{\alpha+\beta}}$ to $\delta^{\frac{\varepsilon_0}{\alpha}}$.

That is how the proof of Theorem 3.4 goes. To deduce Theorem 3.1 from it, we argue by contradiction and use Lemma 2.9 : if both $A+A$ and $A+A\cdot A$ are small (i.e. (3.1) fails), then for every s, $\langle A\rangle_s$ is small, and thus cannot grow into a large ball as Theorem 3.4 asserts.

To prove Theorem 3.2, a little more work is needed. First, in the special case where the collection of endomorphisms is so large that for every $x \in X$, $Ax = \{ax \mid a \in A\}$ contains a ball of radius $\|x\|$, a Fubini-type argument yields (3.2). Then, using Lemma 2.8, we know that if (3.2) fails, then we have an upper bound for $\mathcal{N}_\delta(X + aX)$ for every $a \in \langle A\rangle_s$, $s \geq 1$. Therefore, the idea of the proof is to apply Theorem 3.4 to make A grow into a fat ball in some subalgebra $E \subset \mathcal{M}_n(\mathbb{R})$ so that we can use the special case. Here the subalgebra E can be understood as the subalgebra approximately generated by the set A. It inherits the irreducibility property (assumption (iii) in Theorem 3.2) from A. In particular, $\mathbb{R}^n$ is an irreducible representation of E. Hence, by the Wedderburn structure theorem, E is isomorphic to $\mathcal{M}_n(\mathbb{R})$ or $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})$ or $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})$. Here, a technical issue appears : in Theorem 3.4, the result depends on the norm on E. In the present situation, the norm on E is induced from that on $\mathcal{M}_n(\mathbb{R})$. To have a control on it, we need a quantitative version of the Wedderburn theorem. Indeed, we show that under the quantitative irreducibility condition (iii) of Theorem 3.2, the normed algebra E is isomorphic to one of the three matrix algebras endowed with standard operator norm via a bi-Lipschitz map with Lipschitz constant controlled independently of A.

3.0.2 Organization of Chapter 3

In Section 3.1 we introduce some definitions and notation and then recall the Łojasiewicz inequality from the theory of semianalytic sets. Sections 3.2–3.4 prepare for the proof of the main results. More precisely, Section 3.2 is dedicated to the "escape from subvariety" technique. Section 3.3 deals with a lower bound on the size the set of traces. Then in Section 3.4 we establish an effective version of the Wedderburn structure theorem. We complete the proof of Theorem 3.4 and deduce Theorem 3.1 in Section 3.5. Finally Theorem 3.2 is proved in Section 3.6 and Corollary 3.3 is deduced in Section 3.7.

3.1 Preliminaries

We first set up notation and terminology and then recall the Łojasiewicz inequality.

3.1.1 Notations and definitions

Throughout this chapter, n denotes a positive integer. Most of our estimates are about objects in some ambient space (a normed vector space or a normed

algebra) and we write $f \ll_V g$ and $f = O_V(g)$ to indicate that the implied constant depends not only on the dimension of V but also on the norm of V. We will omit the subscript when it depends only on the dimension n.

We endow the space $\mathbb{R}^n$ with its usual Euclidean norm $\|\cdot\|$ and $\mathbb{C}^n$ and $\mathbb{H}^n$ with their respective l^2-norm. All algebras are over $\mathbb{R}$ and unital. In an algebra E, 1_E denotes the multiplicative identity. All subalgebras of E contain 1_E. When K is a division algebra over $\mathbb{R}$, denote by $\mathcal{M}_n(K)$ the algebra of n by n matrices with coefficients in K. For a real linear space V, denote by $\mathrm{End}(V)$ the algebra of real endomorphisms of V. We identify $\mathrm{End}(\mathbb{R}^n)$ with $\mathcal{M}_n(\mathbb{R})$ in the usual way.

Since all our spaces are normed, we will need a notion of good bases : those which are well spaced. When V is $\mathbb{R}^n$ or $\mathbb{C}^n$ endowed with l^2-norm, recall that its norm induces an l^2-norm on each of its exterior powers. In this case the best bases are clearly orthonormal ones. Note that a basis $(a_1, \dots, a_n)$ is orthonormal if and only if $\forall k, \|a_k\| \leq 1$ and $\|a_1 \wedge \cdots \wedge a_n\| \geq 1$. If we loosen this condition, we get a notion of good bases. However, the norm on an exterior power of V is properly defined only when V is equipped with an l^2-norm and we will deal with other norms such as the operator norm on $\mathrm{End}(\mathbb{R}^n)$. Thus, we need an equivalent formulation.

Lemma 3.6. *Let $(a_1, \dots, a_n)$ be a basis of a normed vector space V over $\mathbb{R}$ or $\mathbb{C}$, then the following conditions are equivalent in the sense that if the i-th condition holds for some $0 < \rho_i \leq 1$ then the j-th condition holds for some $\rho_j \gg_V \rho_i^{O(1)}$.*

(i) For all $k = 1, \dots, n$, $\|a_k\| \leq \rho_1^{-1}$ and $d(a_k, \mathrm{Span}(a_1, \dots, a_{k-1})) \geq \rho_1$.

(ii) For all $k = 1, \dots, n$, $\|a_k\| \leq \rho_2^{-1}$ and all $x \in V$, its coordinates $(x_k)_k$ in the basis $(a_k)_k$ satisfy, $\forall k$, $|x_k| \leq \rho_2^{-1}\|x\|$.

Moreover, if the norm on V is an l^2-norm, then they are also equivalent to the following conditions.

(iii) For all $k = 1, \dots, n$, $\|a_k\| \leq \rho_3^{-1}$ and $\|a_1 \wedge \cdots \wedge a_n\| \geq \rho_3$.

(iv) Any endomorphism that maps an orthonormal basis to $(a_1, \dots, a_n)$ is ρ_4^{-1}-bi-Lipschitz.

In condition (i), we adhere to the convention that $\mathrm{Span}(\varnothing)$ means the zero subspace. This lemma is already known in [30, Lemma 7.5] and [24, Lemma 2.16]. We give an alternative proof.

Proof. Every norm on a finite-dimensional linear space is equivalent to an l^2-norm. Hence it suffices to prove the equivalences in the case where $V = \mathbb{R}^n$ or $\mathbb{C}^n$ endowed with the standard norm. First, (i) implies (iii) since we have

$$\|a_1 \wedge \cdots \wedge a_n\| = \prod_{k=1}^{n} d(a_k, \mathrm{Span}(a_1, \dots, a_{k-1})).$$

To see that (iii) implies (ii), let $x \in E$, then $x = x_1 a_1 + \cdots + x_n a_n$ with $(x_i)_i$ the coordinates of x in $(a_i)_i$. On the one hand,

$$\|x \wedge a_2 \wedge \cdots \wedge a_n\| \leq \|x\|\|a_2\| \cdots \|a_n\| \leq \rho_3^{-(n-1)}\|x\|.$$

On the other hand, $x \wedge a_2 \wedge \cdots \wedge a_n = x_1 a_1 \wedge \cdots \wedge a_n$ so

$$\|x \wedge a_2 \wedge \cdots \wedge a_n\| = |x_1| \, \|a_1 \wedge \cdots \wedge a_n\| \geq \rho_3 |x_1|.$$

Hence $|x_1| \leq \rho_3^{-n} \|x\|$ and the proof is similar for the other coordinates.

Equivalence between (ii) and (iv) is clear.

Finally, (iv) implies (i) because the inequality in (i) holds for an orthonormal basis with $\rho_1 = 1$ and a ρ_4^{-1}-bi-Lipschitz map will only introduce a factor ρ_4^{-1} or ρ_4 to these inequalities. □

Remark. From the proof we see that the implied constant in the notation $\gg_V$ in the lemma can be 1 if V is endowed with an l^2-norm. Also, if V_0 is a fixed normed vector space, then this implied constant is uniform for all subspaces V of V_0.

Lemma 3.6 suggests the following definition.

Definition. Let $0 < \rho \leq 1$ be a parameter. We say a basis $(a_1, \ldots, a_n)$ of a normed vector space V is *ρ-almost orthonormal* if it satisfies the condition (i) in Lemma 3.6 with $\rho_1 = \rho$.

Definition. Let $0 < \rho \leq 1$ be a parameter. Let V be a normed vector space. We say that a subset $A \subset V$ is *ρ-away from linear subspaces* if for every proper linear subspace $W \subset V$, there is $a \in A$ such that $d(a, W) \geq \rho$.

Let E be a normed algebra. We say that a subset $A \subset E$ is *ρ-away from subalgebras* if for every proper subalgebra $W \subset E$, there is $a \in A$ such that $d(a, W) \geq \rho$.

In a similar way, we define the notion of being *ρ-away from Lie subalgebras*.

We have the following observation.

Lemma 3.7. *Let $0 < \rho \leq \frac{1}{2}$ be a parameter. In a normed vector space V of finite dimension, if a subset $A \subset \mathbf{B}(0, \rho^{-1})$ is ρ-away from linear subspaces, then A contains a ρ-almost orthonormal basis. Conversely, if A contains a ρ-almost orthonormal basis, then A is $\rho^{O_V(1)}$-away from subspaces.*

Proof. Assume that $A \subset \mathbf{B}(0, \rho^{-1})$ is ρ-away from linear subspaces. We can construct a ρ-basis from the set A by induction. For $k = 1, \ldots, \dim(V)$, suppose that $a_1, \ldots, a_{k-1}$ are constructed, then $\mathrm{Span}(a_1, \ldots, a_{k-1})$ is a proper subspace of V. Hence there is $a_k \in A$ such that $d(a_k, \mathrm{Span}(a_1, \ldots, a_{k-1})) \geq \rho$.

Conversely, assume that A contains a ρ-almost orthonormal basis (a_i). For any proper linear subspace $W \subset V$, there is $x \in V$ such that $\|x\| = d(x, W) = 1$. By Lemma 3.6, we can write $x = \sum_i x_i a_i$, with $|x_i| \leq \rho^{-O_V(1)}$, for all i. Consequently,

$$d(x, W) \leq \sum_i |x_i| d(a_i, W) \leq \rho^{-O_V(1)} \sum_i d(a_i, W).$$

Hence there is i such that $d(a_i, W) \geq \rho^{O_V(1)}$. □

Definition. Let $0 < \rho \leq 1$ be a parameter. Let A be subset of $\mathrm{End}(\mathbb{R}^n)$. We say that A *acts ρ-irreducibly on* $\mathbb{R}^n$ if for every nonzero proper linear subspace $W \subset \mathbb{R}^n$, there is $a \in A$ and $w \in \mathbf{B}_W(0, 1)$ such that $d(aw, W) \geq \rho$.

We say that a subalgebra $E \subset \mathrm{End}(\mathbb{R}^n)$ *acts ρ-irreducibly on* $\mathbb{R}^n$ if the set $\mathbf{B}_E(0, 1)$ acts ρ-irreducibly on $\mathbb{R}^n$.

3.1.2 Łojasiewicz inequality

The Łojasiewicz inequality [42, Théorème 2, page 62] is a powerful tool which allows us to extract quantitative estimates from algebraic facts. Let us recall it here.

Theorem 3.8 (Łojasiewicz inequality)**.** *Let M be a real analytic manifold endowed with a riemannian distance d and let $f\colon M \to \mathbb{R}$ be a real analytic map. If K is a compact subset of M, then there is $C > 0$ depending on K and f such that for all $x \in K$,*

$$|f(x)| \geq \frac{1}{C} \min(1, d(x, Z))^C$$

where $Z = \{x \in M \mid f(x) = 0\}$.

Note that we take the minimum between 1 and $d(x, Z)$ to make the inequality true even if Z is empty.

3.2 Escaping from subvarieties

In this section we show that if a subset A of a simple algebra is not trapped in any subalgebra then we can escape from any subvariety within a bouned number of steps using addition and multiplication. The number of necessary steps depends only on the ambient algebra E. This is achieved in two steps. First, using only multiplication we can escape from linear subspaces (Proposition 3.11). Then, once the set is away from linear subspaces, we can escape from subvarieties using only addition (Lemma 3.13). Note that everything is quantitative. By escaping a subvariety we mean getting outside a neighborhood of that subvariety.

3.2.1 Escaping from linear subspaces

Let A be subset of a normed algebra E of finite dimension. Obviously, if A is away from linear subspaces, then it is away from subalgebras. We will see in this subsection that the converse is true if we are allowed to replace A with its product set A^s.

The following is the subalgebra (and simpler) version of the Lemma 2.5 in [24]. The proof is essentially the same.

Lemma 3.9. *Let $0 < \rho \leq \frac{1}{2}$ be a parameter. Let A be a subset of a normed algebra E of finite dimension. If $A \subset \mathbf{B}(0, \rho^{-1})$ and A is ρ-away from subalgebras, then A contains a subset of cardinality at most $\dim(E)$ which is $\rho^{O_E(1)}$-away from subalgebras.*

Proof. Let $C \geq 1$ be a large constant. Suppose that $a_1, \ldots, a_{k-1}$ are constructed. If $\{a_1, \ldots, a_{k-1}\}$ is ρ^C-away from subalgebras, then we are done. Otherwise there is a proper subalgebra W such that for all $i = 1, \ldots, k-1$, $d(a_i, W) < \rho^C$. Then choose from A an element a_k such that $d(a_k, W) \geq \rho$.

We prove by induction that at each step $(a_1, \ldots, a_k)$ is a $\rho^{O_E(1)}$-almost orthonormal basis of its linear span $V_k = \mathrm{Span}(a_1, \ldots, a_k)$. This is obvious for $k = 1$. For $k \geq 2$, suppose that $(a_1, \ldots, a_{k-1})$ is a $\rho^{O_E(1)}$-almost orthonormal basis of V_{k-1}. We can write $a_k = x_1 a_1 + \cdots + x_{k-1} a_{k-1} + a_k'$ with $x_i \in \mathbb{R}$ and $\|a_k'\| = d(a_k, V_{k-1})$. Consequently, $x_1 a_1 + \cdots + x_{k-1} a_{k-1}$ is the decomposition of

the vector $a_k - a'_k$ in the $\rho^{O_E(1)}$-almost orthonormal basis $(a_1, \dots, a_{k-1})$. Hence for every $i = 1, \dots k-1$, $|x_i| \leq 2\rho^{-O_E(1)}\|a_k\| \leq \rho^{-O_E(1)}$. Then

$$\begin{aligned}\rho \leq d(a_k, W) &\leq \sum_{i=1}^{k-1} |x_i| d(a_i, W) + \|a'_k\| \\ &\leq \rho^{-O_E(1)} \max_{1 \leq i \leq k-1} d(a_i, W) + \|a'_k\| \\ &\leq \rho^{C - O_E(1)} + \|a'_k\|\end{aligned}$$

When C is large enough, this implies $d(a_k, V_{k-1}) = \|a'_k\| \geq \rho^2$. Hence, $(a_1, \dots, a_k)$ is a $\rho^{O_E(1)}$-almost orthonormal basis of V_k.

We conclude that the construction must stop after at most $\dim(E)$ steps. □

Then we have the analogue of Proposition 2.7 in [24].

Lemma 3.10. *Let $0 < \rho \leq \frac{1}{2}$ be a parameter. Let A be a subset of a normed simple algebra E of finite dimension. If $A \subset \mathbf{B}(0, \rho^{-1})$ and A is ρ-away from subalgebras, then for every nonzero proper linear subspace W of E, there is $w \in \mathbf{B}_W(0,1)$ and $a \in A$ such that*

$$d(aw, W) \geq \rho^{O_E(1)} \quad \text{or} \quad d(wa, W) \geq \rho^{O_E(1)}.$$

Proof. In view of Lemma 3.9, we can assume that A has exactly $n = \dim(E)$ elements $a_1, \dots, a_n$. We can further assume that $A \subset \mathbf{B}(0,1)$ for we can replace A with its contraction $\rho \cdot A$. We will treat the case where the norm on E is Euclidean. The general case follows easily since every norm on E is equivalent to an Euclidean one. Suppose the lemma were false. Then there would be a linear subspace W_0 of dimension $0 < k < n$ such that for all $w \in \mathbf{B}_{W_0}(0,1)$ and all $i = 1, \dots, n$, $d(a_i w, W_0) < \rho^C$ and $d(w a_i, W_0) < \rho^C$ for some large C. The actual value of this constant will be determined by the Łojasiewicz inequality used below.

Consider the map $f\colon \mathrm{Gr}(E,k) \times E^n \to \mathbb{R}$ defined by

$$f(W; x_1, \dots, x_n) = \sum_{i=1}^{n} \int_{\mathbf{B}_W(0,1)} d(x_i w, W)^2 + d(w x_i, W)^2 \, \mathrm{d}w$$

where the integration $\mathrm{d}w$ is with respect to the k-dimensional Lebesgue measure on $\mathbf{B}_W(0,1)$. This map is well-defined and real analytic. This can be seen by observing that the tautological bundle $\pi\colon T \to \mathrm{Gr}(E,k)$ of the Grassmannian has around every point an analytic local trivialization $\varphi\colon \mathcal{U} \times \mathbb{R}^k \to \pi^{-1}(\mathcal{U})$ such that $\forall W \in \mathcal{U}$, $\varphi(W, \cdot)\colon \mathbb{R}^k \to \pi^{-1}(\{W\}) \simeq W$ is an isometry of Euclidean spaces.

From the choice of W_0 it follows that $f(W_0; a_1, \dots, a_n) \ll \rho^C$. Hence by the Łojasiewicz inequality (Theorem 3.8) applied to the compact set $\mathrm{Gr}(E,k) \times \mathbf{B}(0,1)^n$, there is $W_1 \in \mathrm{Gr}(E,k)$ and $b_1, \dots, b_n \in E$ such that $f(W_1; b_1, \dots, b_n) = 0$ and $\forall i$, $\|a_i - b_i\| < \rho$ when the constant C is chosen large enough. The map f vanishing on $(W_1; b_1, \dots, b_n)$ is equivalent to every b_i being in the subalgebra

$$E_{W_1} = \{x \in E \mid xW_1 \subset W_1 \text{ and } W_1 x \subset W_1\}.$$

Now our set A is not ρ-away from the subalgebra E_{W_1}. Therefore E_{W_1} must be the whole algebra E, which in turn implies that W_1 is a two-sided ideal in E. This contradicts the assumption that E is simple. □

Proposition 3.11. *Let $0 < \rho < \frac{1}{2}$ be a parameter. Let A be a subset of a normed simple algebra E of dimension n. Assume that $A \subset \mathbf{B}(0, \rho^{-1})$ and A is ρ-away from subalgebras. Write $A_1 = \{1_E\} \cup A$. Then for any $\eta \in E$ with $\|\eta\| \geq \rho$, the set $A_1^n \eta A_1^n$ contains a $\rho^{O_E(1)}$-almost orthonormal basis of E. Equivalently, $A_1^n \eta A_1^n$ is $\rho^{O_E(1)}$-away from linear subspaces in E.*

Proof. We construct the basis inductively. First, let $\eta_1 = \eta$. Then for $k = 1, \dots, n-1$, suppose that after k steps, we have constructed $\eta_1, \dots, \eta_k \in A^k \eta A^k$ such that $(\eta_1, \dots, \eta_k)$ is a $\rho^{O_E(1)}$-almost orthonormal basis of $W_k = \mathrm{Span}(\eta_1, \dots, \eta_k)$. By Lemma 3.10, there is $w \in \mathbf{B}_{W_k}(0,1)$ and $a \in A$ such that either $d(aw, W) \geq \rho^{O_E(1)}$ or $d(wa, W) \geq \rho^{O_E(1)}$. Let us deal with the former case, the latter case being similar. We can write w in the $\rho^{O_E(1)}$-almost orthonormal basis $(\eta_1, \dots, \eta_k)$,

$$w = w_1 \eta_1 + \cdots + w_k \eta_k.$$

By Lemma 3.6, the coefficients $|w_i| \leq \rho^{-O_E(1)}$, $\forall i$. Hence

$$d(aw, W_k) \leq \sum_{i=1}^{k} |w_i| d(a\eta_i, W_k) \leq \rho^{-O_E(1)} \max_{i=1,\dots,k} d(a\eta_i, W_k).$$

Hence it is possible to pick $i_* \in \{1, \dots, k\}$ so that putting $\eta_{k+1} = a\eta_{i_*}$ we have $d(\eta_{k+1}, W_k) \geq \rho^{O_E(1)}$. Then $(\eta_1, \dots, \eta_{k+1})$ is a $\rho^{O_E(1)}$-almost orthonormal basis of its linear span. □

3.2.2 Escaping from subvarieties, $\mathbb{R}^n$ case

Lemma 3.12. *Let $d \geq 1$ and $0 < \rho \leq 1$. Let $(a_1, \dots, a_n)$ be a ρ-almost orthonormal basis of $\mathbb{R}^n$. Define*

$$\Gamma_d = \big\{ x_1 a_1 + \cdots + x_n a_n \mid \forall i,\ x_i \in \{0, \dots, d\} \big\}.$$

For any polynomial function $P \colon \mathbb{R}^n \to \mathbb{R}$ of degree at most d, we have

$$\sup_{x \in \mathbf{B}(0,1)} |P(x)| \ll_{n,d} \rho^{-O_{n,d}(1)} \max_{\gamma \in \Gamma_d} |P(\gamma)|.$$

Proof. We write $\mathbb{R}_d[X_i]_{1 \leq i \leq n}$ for the space of polynomials on $\mathbb{R}^n$ of degree at most d. First consider the case where $(a_1, \dots, a_n)$ is the canonical basis. Let

$$\Lambda_d = \big\{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid \forall i,\ x_i \in \{0, \dots, d\} \big\}.$$

By a simple induction on the number of variables n, we see that if $P \in \mathbb{R}_d[X_i]_{1 \leq i \leq n}$ and vanishes on Λ_d then it must be the zero polynomial. Therefore the linear map

$$\begin{array}{ccc} \mathbb{R}_d[X_i]_{1 \leq i \leq n} & \to & \mathbb{R}^{\Lambda_d} \\ P & \mapsto & (P(\lambda))_{\lambda \in \Lambda_d}, \end{array}$$

is injective. Hence the coefficients of P are controlled by $\max_{\lambda \in \Lambda_d} |P(\lambda)|$ and consequently for all $r \geq 0$,

$$\sup_{x \in \mathbf{B}(0,r)} |P(x)| \ll_{n,d} r^d \max_{\lambda \in \Lambda_d} |P(\lambda)|. \tag{3.3}$$

For the general case consider $\psi\colon \mathbb{R}^n \to \mathbb{R}^n$ the unique linear map which sends the canonical basis to $(a_1, \ldots, a_n)$. Thus $\psi(\Lambda_d) = \Gamma_d$ and, by Lemma 3.6, $\psi(\mathbf{B}(0, \rho^{-O_n(1)})) \supset \mathbf{B}(0,1)$. We obtain the desired estimate by applying inequality (3.3) to the polynomial $P \circ \psi$ and $r = \rho^{-O_n(1)}$. □

Combining this lemma with Lemma 3.7, we have the following immediate consequence.

Lemma 3.13. *Let V be a normed vector space and $P\colon V \to \mathbb{R}$ a nonzero polynomial map. Then there is an positive integer s depending only on the degree of P and the dimension of V such that the following holds for any parameters $0 < \rho \leq \frac{1}{2}$. Let A be a subset of V with $A \subset \mathbf{B}(0, \rho^{-1})$. If A is ρ-away from linear subspaces then there is $a \in s(A \cup \{0\})$ such that*

$$|P(a)| \gg_{V,P} \rho^{O_{V,P}(1)}.$$

3.3 Trace set estimates

To each real algebra E we can associate a trace function tr_E and a bilinear form τ_E. Let x be an element of E, we define $\mathrm{tr}_E(x)$ to be the trace of the left multiplication by x as an endomorphism of E. For example, if $E = \mathcal{M}_n(\mathbb{R})$, then tr_E is n times the usual trace for matrices.

Let $x, y \in E$. We define $\tau_E(x,y) = \mathrm{tr}_E(xy)$. Thus $\tau_E\colon E \times E \to \mathbb{R}$ is a symmetric bilinear form. Observe that its kernel $\ker \tau_E$ is a two-sided ideal of E. It follows that τ_E is non-degenerate if E is semisimple, i.e. direct sum of simple algebras. Note that the converse is true, but we won't need this fact here.

Let A be a bounded subset of a semisimple algebra E. In this section, we are interested in the size of the trace set of A^2 :

$$\mathrm{tr}_E(A^2) = \{\mathrm{tr}_E(ab) \mid a, b \in A\}.$$

The aim is to establish a lower bound under appropriate conditions. Here is the result.

Lemma 3.14. *Given $\epsilon > 0$. The following is true for sufficiently small $\delta > 0$. Let E be a normed semisimple algebra of dimension $n < +\infty$. Let A be a subset of E such that*

- $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,
- *A is δ^{ϵ}-away from linear subspaces in E.*

Then $\mathcal{N}_\delta(\mathrm{tr}_E(A^2)) \gg_E \delta^{O(\epsilon)} \mathcal{N}_\delta(A)^{\frac{1}{n}}$, where $n = \dim(E)$.

We begin with a lemma.

Lemma 3.15. *Let $0 < \rho \leq \frac{1}{2}$ be a parameter. Let V be a normed vector space of dimension n and $\tau\colon V \times V \to \mathbb{R}$ a non-degenerate bilinear form. If (a_i) is a ρ-almost orthonormal basis of V, then for all $x \in V$,*

$$\|x\| \ll_{V,\tau} \rho^{-O(1)} \max_i |\tau(a_i, x)|.$$

The V and τ in the subscript indicate that the implied constant depends not only on V but also on the bilinear form.

Proof. It suffices to deal with the special case where $V = \mathbb{R}^n$ endowed with the standard Euclidean norm. Let g be the endomorphism which sends the canonical basis to the ρ-almost orthonormal basis (a_i). By Lemma 3.6, $\|g^{-1}\| \leq \rho^{-O(1)}$. Let s be the matrix of τ in the canonical basis. Then a straightforward computation tells us that for all $x \in V$, ${}^t gsx$ is the column vector $(\tau(a_i, x))_i$. Hence

$$\|x\| \leq \|s^{-1}\|\|{}^t g^{-1}\|\|{}^t gsx\| \ll_{n,\tau} \rho^{-O(1)} \max_i |\tau(a_i, x)|. \qquad \square$$

Proof of Lemma 3.14. By Lemma 3.7, A contains a δ^ϵ-almost orthonormal basis $(a_i)_{1 \leq i \leq n}$ of E. Thus, by Lemma 3.15 applied to the non-degenerate bilinear form τ_E, for all $x \in E$, $\|x\| < \delta^{1-O(\epsilon)}$ whenever $|\mathrm{tr}_E(a_i x)| < \delta$ for all $i = 1, \ldots, n$.

Consider the map $\Theta\colon A \to \mathbb{R}^n$ defined by $x \mapsto \big(\mathrm{tr}_E(a_i x)\big)_i$. On the one hand, this map is "almost injective" with the δ-blurred vision : for all $x, y \in A$, if $\|\Theta(x) - \Theta(y)\| < \delta$ then $\|x - y\| < \delta^{1-O(\epsilon)}$. It follows that

$$\mathcal{N}_\delta(\Theta(A)) \gg_n \mathcal{N}_{\delta^{1-O(\epsilon)}}(A) \gg_E \delta^{O(\epsilon)} \mathcal{N}_\delta(A).$$

On the other hand, the set $\Theta(A)$ is contained in the n-fold Cartesian product $\mathrm{tr}(A^2) \times \cdots \times \mathrm{tr}(A^2)$, hence

$$\mathcal{N}_\delta(\Theta(A)) \ll_n \mathcal{N}_\delta(\mathrm{tr}(A^2))^n.$$

We obtain the desired estimate by combining the two inequalities above. $\square$

3.4 Effective Wedderburn theorem

The Wedderburn theorem (see [40, Chapter XVII, §3] or [53]) states that if $\mathbb{R}^n$ is an irreducible representation of an finite-dimensional algebra E, then E is isomorphic to either $\mathcal{M}_n(\mathbb{R})$ or $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})$ or $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})$. In this section we prove a quantified version of this algebraic fact. From now on, we endow these matrix algebras with operator norms[5]. Let $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})_{\mathbb{R}}$ denote a fixed embedding of $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})$ in $\mathcal{M}_n(\mathbb{R})$ and $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})_{\mathbb{R}}$ a fixed embedding of $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})$ in $\mathcal{M}_n(\mathbb{R})$.

Proposition 3.16. *Let $0 < \rho \leq \frac{1}{2}$ be a parameter. Let E be a subalgebra of $\mathcal{M}_n(\mathbb{R})$ acting ρ-irreducibly on $\mathbb{R}^n$. Then either $E = \mathcal{M}_n(\mathbb{R})$ or E is conjugate to $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})_{\mathbb{R}}$ or to $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})_{\mathbb{R}}$ by a change of basis matrix $g \in \mathrm{GL}(\mathbb{R}^n)$ satisfying $\|g\| + \|g^{-1}\| \leq \rho^{-O(1)}$.*

In particular, E is isomorphic to $\mathcal{M}_n(\mathbb{R})$ or $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})$ or $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})$ by an isomorphism which is $\rho^{-O(1)}$-bi-Lipschitz.

3.4.1 Effective diagonalization

The following lemma is implicit in [30, proof of Proposition 7.4]. We include its proof for the sake of completeness.

[5] Each matrix is seen as an endomorphism of the Euclidean space $\mathbb{R}^n$, $\mathbb{C}^n \simeq \mathbb{R}^{2n}$ or $\mathbb{H}^n \simeq \mathbb{R}^{4n}$.

Lemma 3.17. *Let $0 < \rho \leq \frac{1}{2}$ be a parameter. Let $a \in \mathcal{M}_n(\mathbb{C})$ be a diagonalizable matrix. Assume that $\|a\| < \rho^{-1}$ and its spectrum is a ρ-separated set (it may contain multiple eigenvalues). Then a is diagonal in a $\rho^{O(1)}$-almost orthonormal basis of $\mathbb{C}^n$.*

Proof. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of a (they appear with the corresponding multiplicity). Let $v_1, \ldots, v_n$ be the corresponding eigenvectors. Each v_i can be chosen to be of unit length and we can further assume that $(v_i, v_j) = 0$ whenever $i \neq j$ but $\lambda_i = \lambda_j$. For $k = 1, \ldots, n$, write $V_k = \mathrm{Span}(v_1, \ldots, v_k)$. We will prove by induction on k that $(v_1, \ldots, v_k)$ is a $\rho^{O_k(1)}$-almost orthonormal basis of V_k. This is clear when $k = 1$.

Let k be an integer between 1 and $n-1$. Suppose that $(v_1, \ldots, v_k)$ is a $\rho^{O_k(1)}$-almost orthonormal basis of V_k. Let us show that $d(v_{k+1}, V_k) > \rho^{O_k(1)}$ and thus $(v_1, \ldots, v_k, v_{k+1})$ is a $\rho^{O_k(1)}$-almost orthonormal basis of V_{k+1}. Without loss of generality we can assume that among the eigenvalues $\lambda_1, \ldots, \lambda_k$, only $\lambda_{l+1}, \ldots, \lambda_k$ are equal to λ_{k+1} for some $l \leq k$. Considering the orthogonal projection of v_{k+1} onto V_k, we can decompose the vector v_{k+1} into

$$v_{k+1} = \sum_{i=1}^{k} x_i v_i + v'_{k+1}, \tag{3.4}$$

where $x_1, \ldots, x_k \in \mathbb{C}$ and $v'_{k+1} \in V_k^{\perp}$. In particular, $d(v_{k+1}, V_k) = \|v'_{k+1}\|$. We can then express av_{k+1} in two different ways :

$$av_{k+1} = \lambda_{k+1} v_{k+1} = \sum_{i=1}^{k} \lambda_{k+1} x_i v_i + \lambda_{k+1} v'_{k+1},$$

and

$$av_{k+1} = \sum_{i=1}^{k} \lambda_i x_i v_i + a v'_{k+1}.$$

It follows that

$$\sum_{i=1}^{l} (\lambda_{k+1} - \lambda_i) x_i v_l = \lambda_{k+1} v'_{k+1} - a v'_{k+1}.$$

Denote by w_{k+1} the vector on the right-hand side. Its norm can be bounded : $\|w_{k+1}\| \leq 2\|a\|\|v'_{k+1}\| \leq \rho^{-2}\|v'_{k+1}\|$. The left-hand side gives the coordinates of w_{k+1} in the basis $(v_1, \ldots, v_k)$ which is $\rho^{O_k(1)}$-almost orthonormal by the induction hypothesis. Hence, by Lemma 3.6,

$$\forall i = 1, \ldots, l, \quad |\lambda_{k+1} - \lambda_i||x_i| \leq \rho^{-O_k(1)}\|w_{k+1}\| \leq \rho^{-O_k(1)}\|v'_{k+1}\|.$$

Hence for all $i = 1, \ldots, l$, $|x_i| \leq \rho^{-O_k(1)}\|v'_{k+1}\|$ thanks to the assumption $|\lambda_{k+1} - \lambda_i| \geq \rho$.

In order to bound $|x_j|$ for $j \geq l+1$, we take the scalar product with v_j on both sides of (3.4). We obtain

$$0 = \sum_{i=1}^{l} x_i (v_j, v_i) + x_j.$$

Hence for all $j = l+1, \ldots k$, $|x_j| \leq \rho^{-O_k(1)}\|v'_{k+1}\|$. Using (3.4) again we obtain $1 = \|v_{k+1}\| \ll \rho^{-O_k(1)}\|v'_{k+1}\|$ and then $\|v'_{k+1}\| \geq \rho^{O_k(1)}$. This finishes the proof of the inductive step. □

3.4.2 Effective Wedderburn theorem in $\mathcal{M}_n(\mathbb{R})$

For $x \in \mathcal{M}_n(\mathbb{R})$ and $A \subset \mathcal{M}_n(\mathbb{R})$ denote by $\mathcal{C}(x)$ and $\mathcal{C}(A)$ their respective centralizers, i.e.

$$\mathcal{C}(x) = \{y \in \mathcal{M}_n(\mathbb{R}) \mid xy = yx\},$$
$$\mathcal{C}(A) = \{y \in \mathcal{M}_n(\mathbb{R}) \mid \forall x \in A, xy = yx\}.$$

Lemma 3.18. *Let $0 < \rho \leq 1$ be a parameter. For any $g \in \mathrm{SL}_n(\mathbb{R})$, if $\|g\| > \rho^{-n}$, then the subalgebra $\mathcal{C}(g)$ does* not *act ρ-irreducibly on $\mathbb{R}^n$.*

Proof. Let $g = kal$ be the Cartan decomposition of g, with $k, l \in \mathrm{SO}(n)$ and $a = \mathrm{diag}(a_1, \ldots, a_n)$ where its singular values $a_1, \ldots, a_n$ are arranged so that $a_1 \geq a_2 \geq \cdots \geq a_n > 0$. Assume that $a_1 = \|g\| > \rho^{-n}$. Since $a_1 \cdots a_n = \det(g) = 1$, there is $p \in \{1, \ldots, n-1\}$ such that

$$\frac{a_{p+1}}{a_p} < \rho.$$

Put $W = k\,\mathrm{Span}(e_1, \ldots, e_p)$. It is a nonzero proper linear subspace of $\mathbb{R}^n$. We claim that for all $x \in \mathcal{C}(g) \cap \mathbf{B}(0,1)$ and all $w \in \mathbf{B}_W(0,1)$, $d(xw, W) < \rho$.

Indeed, decomposing every vector $v \in \mathbb{R}^n$ as $v = v' + v''$ with $v' \in l^{-1}\,\mathrm{Span}(e_1, \ldots, e_p)$ and $v'' \in l^{-1}\,\mathrm{Span}(e_{p+1}, \ldots e_n)$, we see that

$$d(gv, W) = \|av''\| \leq a_{p+1}\|v\|.$$

Moreover, for all $w \in W$,

$$\|g^{-1}w\| \leq a_p^{-1}\|w\|.$$

Consequently, for all $x \in \mathcal{C}(g) \cap \mathbf{B}(0,1)$ and all $w \in \mathbf{B}_W(0,1)$, we have $xw = gxg^{-1}w$, and thus

$$d(xw, W) \leq a_{p+1}\|xg^{-1}w\| \leq a_{p+1}\|g^{-1}w\| \leq \frac{a_{p+1}}{a_p} < \rho.$$

□

Lemma 3.19. *Let $0 < \rho \leq \frac{1}{2}$ be a parameter. If two real matrices x and $y \in \mathcal{M}_n(\mathbb{R})$ are conjugate by a complex matrix $g \in \mathrm{GL}_n(\mathbb{C})$ with $\|g\| + \|g^{-1}\| < \rho^{-1}$ then there is a real matrix $h \in \mathrm{GL}_n(\mathbb{R})$ which also conjugates them and moreover $\|h\| + \|h^{-1}\| \leq \rho^{-O(1)}$.*

Proof. Assume that $gxg^{-1} = y$ with $x, y \in \mathcal{M}_n(\mathbb{R})$ and $g \in \mathrm{GL}_n(\mathbb{C})$. We write $g = g_\Re + ig_\Im$ with $g_\Re, g_\Im \in \mathcal{M}_n(\mathbb{R})$ its real and imaginary part. From $gx = yg$ we see that $g_\Re x = yg_\Re$ and $g_\Im x = yg_\Im$. For $\lambda \in \mathbb{C}$, consider $h_\lambda = g_\Re + \lambda g_\Im$. For all $\lambda \in \mathbb{C}$, we have $h_\lambda x = y h_\lambda$. Hence whenever $\det(h_\lambda) \neq 0$, h_λ conjugates x and y. What remains to do is to find appropriate $\lambda \in \mathbb{R}$ such that h_λ and h_λ^{-1} have bounded norms.

Define $P(\lambda) = \det(h_\lambda)$. It is a polynomial with real coefficients and its degree is at most n. We know that

$$|P(i)| = |\det(g)| = |\det(g^{-1})|^{-1} \gg \|g^{-1}\|^{O(1)} \gg \rho^{O(1)}.$$

It is easy to see that the coefficients of P are controlled by $\max_{\lambda \in [0,1]} |P(\lambda)|$. Hence

$$|P(i)| \ll \max_{\lambda \in [0,1]} |P(\lambda)|.$$

So there is $\lambda_0 \in [0,1]$ such that $|P(\lambda_0)| \gg \rho^{O(1)}$. Take $h = h_{\lambda_0}$. we have

$$\|h\| \leq \|g_\Re\| + |\lambda_0|\|g_\Im\| \leq 2\|g\| \ll \rho^{-1}.$$

and $|\det(h)| \gg \rho^{O(1)}$, which implies

$$\|h^{-1}\| \leq \|h\|^{n-1}|\det(h)|^{-1} \ll \rho^{-O(1)}.$$

Here the first inequality can be seen from the Cartan decomposition of h. □

Proof of Proposition 3.16. Consider $K = \mathcal{C}(E)$ the centralizer of E in $\mathcal{M}_n(\mathbb{R})$. From (the proof of) Wedderburn's theorem, K is a division algebra over $\mathbb{R}$ and E is equal to $\mathcal{C}(K)$, the centralizer of K. By the Frobenius theorem, the real division algebra K is isomorphic to $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$. The action of K on $\mathbb{R}^n$ makes $\mathbb{R}^n$ a K linear space. Hence n is even if $K \simeq \mathbb{C}$ and a multiple of 4 if $K \simeq \mathbb{H}$.

If $K \simeq \mathbb{R}$ then $E = \mathcal{M}_n(\mathbb{R})$, we are done. If $K \simeq \mathbb{C}$ or respectively if $K \simeq \mathbb{H}$, then E is isomorphic to $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})$ or respectively to $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})$. The Skolem-Noether theorem (see [50, §12.6]) tells us all embeddings of $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})$ or $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})$ in $\mathcal{M}_n(\mathbb{R})$ are conjugate (by a matrix in $\mathrm{GL}(\mathbb{R}^n)$). We are going to show that under our quantitative irreducibility assumption, the change of basis matrix can be nicely chosen.

From the discussion above, we see that an embedding of $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})$ is uniquely determined by an embedding of $\mathbb{C}$ in $\mathcal{M}_n(\mathbb{R})$. Moreover, if K is conjugate to a fixed embedding of $\mathbb{C}$, then E is conjugate to a fixed embedding of $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})$ by the same change of basis matrix. The same applies to the quaternion case. That is why we are going to study embeddings of $\mathbb{C}$ and of $\mathbb{H}$. Let $\varphi_0 \colon \mathbb{C} \to \mathcal{M}_n(\mathbb{R})$ denote the embedding of $\mathbb{C}$ such that $\mathcal{C}(\varphi_0(\mathbb{C})) = \mathcal{M}_{\frac{n}{2}}(\mathbb{C})_\mathbb{R}$ and $\psi_0 \colon \mathbb{H} \to \mathcal{M}_n(\mathbb{R})$ the embedding of $\mathbb{H}$ such that $\mathcal{C}(\psi_0(\mathbb{H})) = \mathcal{M}_{\frac{n}{4}}(\mathbb{H})_\mathbb{R}$.

When K is isomorphic to $\mathbb{C}$, denote by $\varphi \colon \mathbb{C} \to \mathcal{M}_n(\mathbb{R})$ the embedding of $\mathbb{C}$ whose image is K. The homomorphism φ is uniquely determined by $\varphi(i)$ and we have $E = \mathcal{C}(\varphi(i))$. The matrix $\varphi(i)$ satisfies $\varphi(i)^2 = -\mathrm{Id}_n$, which implies that $\varphi(i)$ is diagonalizable over $\mathbb{C}$ and its eigenvalues can only be i or $-i$. Its trace is real. Therefore the multiplicities of the two eigenvalues must be equal. We conclude that $\varphi(i)$ is conjugate (over $\mathbb{C}$) to the diagonal matrix $\mathrm{diag}(i\,\mathrm{Id}_{\frac{n}{2}}, -i\,\mathrm{Id}_{\frac{n}{2}})$. From Lemma 3.18 and the irreducibility assumptions on E, we have $\|\varphi(i)\| \ll \rho^{-O(1)}$. Then by Lemma 3.17, $\varphi(i)$ is conjugate to $\mathrm{diag}(i\,\mathrm{Id}_{\frac{n}{2}}, -i\,\mathrm{Id}_{\frac{n}{2}})$ by a change of basis matrix $g \in \mathrm{GL}_n(\mathbb{C})$ satisfying $\|g\| + \|g^{-1}\| \ll \rho^{-O(1)}$. The same is true for $\varphi_0(i)$. We conclude that $\varphi(i)$ is conjugate to $\varphi_0(i)$ by a change of basis matrix $g' \in \mathrm{GL}_n(\mathbb{C})$ with $\|g'\| + \|g'^{-1}\| \ll \rho^{-O(1)}$. Finally, thanks to Lemma 3.19, g' can be chosen to be real. This finishes the proof for the case where $K \simeq \mathbb{C}$.

When K is isomorphic to $\mathbb{H}$, denote by $\psi \colon \mathbb{H} \to \mathcal{M}_n(\mathbb{R})$ the embedding of $\mathbb{H}$ whose image is K. The homomorphism $\psi \colon \mathbb{H} \to \mathcal{M}_n(\mathbb{R})$ is uniquely determined by $\psi(i)$ and $\psi(j)$ and $E = \mathcal{C}(\psi(i)) \cap \mathcal{C}(\psi(j))$.

The two matrices $\psi(i)$ and $\psi(j)$ satisfy

$$\psi(i)^2 = \psi(j)^2 = -\mathrm{Id}_n \text{ and } \psi(i)\psi(j) = -\psi(j)\psi(i). \tag{3.5}$$

Repeating the argument in the complex case, Lemma 3.18 gives the estimates $\|\psi(i)\|, \|\psi(j)\| \ll \rho^{-O(1)}$. Moreover, $\psi(i)$ is conjugate to $\mathrm{diag}(i\,\mathrm{Id}_{\frac{n}{2}}, -i\,\mathrm{Id}_{\frac{n}{2}})$ by

a change of basis matrix satisfying the desired norm estimate. Write $\psi(i)$ and $\psi(j)$ in this new basis,

$$\psi(i) = \left[\begin{array}{c|c} i\,\mathrm{Id}_{\frac{n}{2}} & 0 \\ \hline 0 & -i\,\mathrm{Id}_{\frac{n}{2}} \end{array}\right] \quad \text{and} \quad \psi(j) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

The condition (3.5) is then equivalent to $A = D = 0$ and $BC = CB = -\mathrm{Id}_{\frac{n}{2}}$. It is easy to check that the conjugation by the block diagonal matrix $g' = \mathrm{diag}(C, \mathrm{Id}_{\frac{n}{2}})$ preserve $\psi(i)$ and conjugates $\psi(j)$ to

$$\left[\begin{array}{c|c} 0 & -\mathrm{Id}_{\frac{n}{2}} \\ \hline \mathrm{Id}_{\frac{n}{2}} & 0 \end{array}\right].$$

Note that $\|g'\| + \|g'^{-1}\| = \|C\| + \|B\| \ll \|\psi(j)\| \ll \rho^{-O(1)}$. To conclude the proof we use Lemma 3.19 to make sure the change of basis matrix is real. □

3.5 Sum-product estimate in simple algebras

We prove Theorem 3.4 and Theorem 3.1 in this section.

3.5.1 Sum-product theorem in $\mathbb{C}^n$

We will use Bourgain-Gamburd's sum-product theorem in $\mathbb{C}^n$ (Theorem 3.5) in the a slightly stronger form.

Corollary 3.20. *Given $\kappa > 0$ and $n \geq 1$, there is $\alpha \geq 0$, $\beta > 0$ and a positive integer $s \geq 1$ such that, for $\epsilon > 0$ sufficiently small and $\delta > 0$ sufficiently small, the following holds. Let A be a subset of $\mathcal{M}_n(\mathbb{C})$. Assume that*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\mathcal{N}_\delta(A) \geq \delta^{-\kappa}$,

(iii) $A \subset \Delta + \mathbf{B}(0, \delta^{1-\epsilon})$.

Then there exists η in the algebra generated by A such that $\|\eta\| = 1$ and

$$[0, \delta^\alpha]\eta \subset \langle A \rangle_s + \mathbf{B}(0, \delta^{\alpha+\beta}). \tag{3.6}$$

The corollary mainly says that the segment length δ^α and the new scale $\delta^{\alpha+\beta}$ in Theorem 3.5 can be chosen independently of A. This will be very useful when we use this diagonal case. Recall that Δ denotes the set of diagonal matrices in $\mathcal{M}_n(\mathbb{C})$.

Proof. In this proof, let $\pi_i(a)$ denote the i-th diagonal entry of a for any $i = 1, \ldots, n$ and any $a \in \mathcal{M}_n(\mathbb{C})$. Partitioning A into at most $\delta^{-O_n(\epsilon)}$ parts of diameter 1 and choosing the part with the largest δ-covering number, we see that $\mathcal{N}_\delta(\mathbf{B}(0,1) \cap (A - A)) \geq \delta^{-\kappa + O_n(\epsilon)}$. Thus we can assume that $A \subset \mathbf{B}(0,1)$. By working at scale $\delta^{1-\epsilon}$, we can further assume $A \subset \Delta + \mathbf{B}(0, \delta)$. Theorem 3.5 says that (3.6) is true for some α and β which may depend on A. Nevertheless, they can be bounded by constants depending only on n and κ. What we need to show is that they can actually be chosen independently of A. This is evident

for β since (3.6) gets only weaker when β becomes smaller. Hence there exist $0 \le \alpha_0 < C = C(n,\kappa)$, $s_0 = s_0(n,\kappa) \ge 1$ and $\eta \in \Delta$ such that $\|\eta\| = 1$ and

$$[0, \delta^{\alpha_0}]\eta \subset \langle A\rangle_{s_0} + \mathbf{B}(0, \delta^{\alpha_0+\beta}). \tag{3.7}$$

By replacing α_0 with $\alpha_0 + \frac{1}{2}\beta$ and β with $\frac{1}{2}\beta$ we can assume that $\alpha_0 \ge \beta$. Fix an index i such that $|\pi_i(\eta)| \gg_n 1$. Consider for any $s \ge 1$, the set

$$\Omega_s = \big\{\omega \in \mathbb{R}_+ \mid \exists \xi \in \langle A\rangle_s,\ \|\xi\| \le s\delta^{\omega},\ d(\xi,\Delta) \le s\delta^{\omega+\frac{1}{2}\beta},\\ |\pi_i(\xi)| \ge \frac{1}{s}\delta^{\omega} \text{ and } [0,1]\xi \subset \langle A\rangle_s + \mathbf{B}(0, s\delta^{\omega+\beta})\big\}.$$

We need to prove existence of $\alpha > 0$ and $s \ge 1$ depending only on n and κ such that $\alpha \in \Omega_s$.

It follows from (3.7) that $\alpha_0 \in \Omega_{s_0}$. Now we show that if $\omega \in \Omega_s$ for some $s \ge 1$ then there is $s' = s'(s,n,\kappa) \ge 1$ such that $[\omega + \alpha_0, \omega + \alpha_0 + \frac{1}{2}\beta] \subset \Omega_{s'}$. Indeed, for any $\gamma \in [\alpha_0, \alpha_0 + \frac{1}{2}\beta]$, by (3.7), there exists $a \in \langle A\rangle_{s_0}$ such that $\delta^{\gamma}\eta \in a + \mathbf{B}(0, \delta^{\alpha_0+\beta})$. Thus $d(a,\Delta) \le \delta^{\alpha_0+\beta} \le \delta^{\gamma+\frac{1}{2}\beta}$ and

$$\delta^{\gamma} \ll_n \pi_i(a) \le \|a\| \ll \delta^{\gamma}.$$

By multiplying a to the relation $[0,1]\xi \subset \langle A\rangle_s + \mathbf{B}(0, s\delta^{\omega+\beta})$ we obtain

$$[0,1]\xi a \subset \langle A\rangle_{s+s_0} + \mathbf{B}(0, s\|a\|\delta^{\omega+\beta}) \subset \langle A\rangle_{s+s_0} + \mathbf{B}(0, O(s)\delta^{\omega+\gamma+\beta}).$$

Moreover, $\|\xi a\| \ll_n \|\xi\|\|a\| \ll s\delta^{\omega+\gamma}$ and

$$d(\xi a, \Delta) \ll_n d(\xi,\Delta)\|a\| + \|\xi\| d(a,\Delta) \ll s\delta^{\omega+\gamma+\frac{1}{2}\beta},$$

and for $\delta > 0$ sufficiently small,

$$|\pi_i(\xi a)| \ge |\pi_i(\xi)||\pi_i(a)| - O_n(d(\xi,\Delta)d(a,\Delta)) \gg_n \frac{1}{s}\delta^{\omega+\gamma} - O_n(\delta^{\omega+\gamma+\beta}) \gg_n \frac{1}{s}\delta^{\omega+\gamma}$$

Hence $\omega + \gamma \in \Omega_{s'}$ for some $s' = s'(s,n,\kappa)$.

A simple induction yields that there exists a sequence $(s_k)_{k\ge 0}$ depending only on n and κ such that for all $k \ge 0$,

$$[(k+1)\alpha_0, (k+1)\alpha_0 + \frac{k}{2}\beta] \subset \Omega_{s_k}.$$

Recall that α_0 depends on A but it is bounded by $\beta \le \alpha_0 \le C$ where β and C are constants given by Theorem 3.5 and depend only on n and κ. Put $\alpha = \big(\left\lceil \frac{2C}{\beta} \right\rceil + 1\big)C$ and $K = \left\lceil \frac{2\alpha}{\beta} \right\rceil$. For any choice of $\alpha_0 \in [\beta, C]$, the equation

$$(k+1)\alpha_0 \le \alpha \le (k+1)\alpha_0 + \frac{k}{2}\beta$$

has a solution k satisfying $k \le K$. It follows that $\alpha \in \Omega_{s_k} \subset \Omega_{s_K}$. This concludes the proof since α and s_K depend only on n and κ. □

3.5.2 Small segment

The key step in the proof of Theorem 3.4 is to produce a small segment in $\langle A\rangle_s$.

Proposition 3.21. *Given a finite-dimensional normed simple algebra E and $\kappa > 0$, there is $s \geq 1$ and $\epsilon > 0$, $\alpha, \beta > 0$ such that the following is true for $\delta > 0$ sufficiently small. Let A be subset of E. Assume that*

(i) $A \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\mathcal{N}_\delta(A) \geq \delta^{-\kappa}$,

(iii) *A is δ^ϵ-away from subalgebras.*

Then there is $\eta \in E$, $\|\eta\| = 1$ such that

$$[0, \delta^\alpha]\eta \subset \langle A\rangle_s + \mathbf{B}(0, \delta^{\alpha+\beta}). \tag{3.8}$$

Proof of Proposition 3.21. First observe that we can assume without loss of generality that E is a real subalgebra of $\mathcal{M}_n(\mathbb{C})$ for some positive integer n and it contains a least one element with n distinct eigenvalues. Indeed, this is evident if E is isomorphic to $\mathcal{M}_n(\mathbb{C})$ or $\mathcal{M}_n(\mathbb{R})$ since for the latter case we can embed naturally $\mathcal{M}_n(\mathbb{R})$ in $\mathcal{M}_n(\mathbb{C})$. We don't need to worry about the norm because all linear isomorphisms are bi-Lipschitz and bi-Lipschitz maps only change the constants in the assumption and conclusion of the proposition. If E is isomorphic to $\mathcal{M}_n(\mathbb{H})$, then we can embed $\mathcal{M}_n(\mathbb{H})$ in $\mathcal{M}_{2n}(\mathbb{C})$ by sending each entry $x + iy + jz + kw \in \mathbb{H}$ to a 2×2 block $\left(\begin{smallmatrix} x+iy & z+iw \\ -z+iw & x-iy \end{smallmatrix}\right)$. It is easy to check that this embedding of E contains a diagonal matrix with distinct diagonal entries.

In this proof, s stands for an unspecified positive integer depending on n and κ that may increase from one line to another. Since A is δ^ϵ-away from subalgebras, $\langle A\rangle_s$ is $\delta^{O(\epsilon)}$-away from linear subspaces by Proposition 3.11 applied to any $\eta \in A$ with $\|\eta\| \geq \delta^\epsilon$. Therefore, without loss of generality, we can assume that A is δ^ϵ-away from linear subspaces in E.

Consider $P\colon \mathcal{M}_n(\mathbb{C}) \to \mathbb{C}$ defined by

$$P(x) = \prod_{i<j} (\lambda_i - \lambda_j)^2 \tag{3.9}$$

where $\lambda_1, \ldots, \lambda_n$ are eigenvalues of $x \in \mathcal{M}_n(\mathbb{C})$ (the n roots of the characteristic polynomial of x). The right-hand side of (3.9) is symmetric in (λ_i) and thus polynomial in the coefficients of the characteristic polynomial of x. Hence $x \mapsto |P(x)|^2$ is a real polynomial on $\mathcal{M}_n(\mathbb{C})$. Apply Lemma 3.13 to $|P|^2$ restricted to E. Since E contains an element with n distinct eigenvalues, we obtain an element $a \in \langle A\rangle_s$ such that $|P(a)| > \delta^{O(\epsilon)}$ and consequently the eigenvalues $\lambda_1, \ldots, \lambda_n$ of a satisfy

$$\forall i \neq j, \quad |\lambda_i - \lambda_j| > \delta^{O(\epsilon)}.$$

Now consider the map

$$\begin{array}{rccc} \varphi\colon & \mathcal{M}_n(\mathbb{C}) & \to & \mathcal{M}_n(\mathbb{C}) \\ & x & \mapsto & ax - xa. \end{array}$$

Let $\kappa' = \frac{\kappa}{9n^2}$. We distinguish two cases according to the size of the image $\varphi(A)$. First consider the case where

$$\mathcal{N}_\delta\big(\varphi(A)\big) \geq \delta^{2\kappa'}\mathcal{N}_\delta(A).$$

In this case we can show a growth estimate. By Lemma 3.14, there is a subset $A' \subset A^2$ such that $\mathrm{tr}_E(A')$ is δ-separated and of size $|\mathrm{tr}_E(A')| \geq \delta^{-4\kappa'}$. Here, we used the fact that $\dim(E) \leq 2n^2$. Observe that $\mathrm{tr}_E(\varphi(x)) = 0$ for all $x \in E$, hence

$$\mathcal{N}_\delta\big(\varphi(A) + A'\big) \gg \mathcal{N}_\delta(\varphi(A))|\mathrm{tr}_E(A')| \geq \delta^{-2\kappa'}\mathcal{N}_\delta(A).$$

Consequently,

$$\mathcal{N}_\delta\big(\langle A\rangle_s\big) \geq \delta^{-\kappa'}\mathcal{N}_\delta(A). \tag{3.10}$$

Otherwise, we have

$$\mathcal{N}_\delta\big(\varphi(A)\big) < \delta^{2\kappa'}\mathcal{N}_\delta(A),$$

then by cutting A into "radius δ" fibers, we see that

$$\mathcal{N}_\delta(A) \ll \mathcal{N}_\delta\big(\varphi(A)\big)\max_y \mathcal{N}_\delta\big(\varphi^{-1}\big(\mathbf{B}(y,\delta)\big) \cap A\big).$$

Hence there is $y_* \in \mathcal{M}_n(\mathbb{C})$ such that $\mathcal{N}_\delta\big(\varphi^{-1}\big(\mathbf{B}(y_*,\delta)\big) \cap A\big) \geq \delta^{-\kappa'}$. Put $A'' = \varphi^{-1}\big(\mathbf{B}(y_*,\delta)\big)\cap A - \varphi^{-1}\big(\mathbf{B}(y_*,\delta)\big)\cap A$ so that $\mathcal{N}_\delta(A'') \geq \delta^{-\kappa'}$ and $\varphi(A'') \subset \mathbf{B}(0,2\delta)$.

Recall that $|\lambda_i - \lambda_j| > \delta^{O(\epsilon)}$ for all $i \neq j$. Hence we can apply Lemma 3.17 to the matrix a. We obtain a change of basis matrix $g \in \mathrm{GL}_n(\mathbb{C})$ such that $\|g\| + \|g^{-1}\| < \delta^{-O(\epsilon)}$ and gag^{-1} is diagonal. Conjugation by g will change any estimate only by a factor of $\delta^{-O(\epsilon)}$ or $\delta^{O(\epsilon)}$. Hence without loss of generality we can assume that $a = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Then an explicit computation gives an expression for φ in the standard basis : if $x = (x_{ij})_{i,j} \in \mathcal{M}_n(\mathbb{C})$, then

$$\varphi(x) = \big((\lambda_i - \lambda_j)x_{ij}\big)_{i,j}.$$

Therefore for any $x \in \mathcal{M}_n(\mathbb{C})$, if $\|\varphi(x)\| \ll \delta$ then $d(x,\Delta) < \delta^{1-O(\epsilon)}$. Consequently $A'' \subset \Delta + \mathbf{B}(0,\delta^{1-O(\epsilon)})$. Then Corollary 3.20 gives constants $\alpha, \beta > 0$ and integer $s \geq 1$ depending only on n and κ' and a unit vector $\eta \in \mathcal{M}_n(\mathbb{C})$ such that (3.8) holds.

What we have proved is that either the proposition holds or we have (3.10). If we are in the latter case, we can iterate the same argument to $\langle A\rangle_s$. After at most $O(\frac{n^2}{\kappa'})$ iterations, (3.10) cannot be possible anymore, hence the proposition must be true. □

3.5.3 Proof of Theorem 3.4.

Once $\langle A\rangle_s$ contains a segment, it takes only a few more steps to produce a small ball.

Proposition 3.22. *Under the assumptions of Proposition 3.21, we have*

$$\mathbf{B}(0,\delta^\alpha) \subset \langle A\rangle_s + \mathbf{B}(0,\delta^{\alpha+\beta}). \tag{3.11}$$

Proof. Suppose that $\eta \in E$ is a unit vector satisfying (3.8). Since A is δ^ϵ-away from subalgebras, by Proposition 3.11, there is a $\delta^{O(\epsilon)}$-almost orthonormal basis of E of the form $(a_i \eta b_i)_i$ with $a_i, b_i \in \langle A \rangle_{\dim(E)}$. Then (3.8) implies that for all $i = 1, \dots, \dim(E)$,

$$[0, \delta^\alpha] a_i \eta b_i \subset \langle A \rangle_{s+2\dim(E)} + \mathbf{B}(0, \delta^{\alpha+\beta-O(\epsilon)}).$$

Moreover, Lemma 3.6 yields

$$\mathbf{B}(0, \delta^{\alpha+O(\epsilon)}) \subset \sum_i [-\delta^\alpha, \delta^\alpha] a_i \eta b_i.$$

Hence (3.11) holds for sufficiently small ϵ and slightly worse α, β and s. □

Proof of Theorem 3.4. The idea is to apply Proposition 3.22 at various scales ranging from δ to δ^ϵ. Let $\epsilon_1, \alpha, \beta$ and s be the constants given by Proposition 3.22 applied to $\frac{\kappa}{2}$ in place of κ. Let $r = \left\lceil \frac{\ln(\epsilon_0)}{\ln(\alpha) - \ln(\alpha+\beta)} \right\rceil$ and for $k = 0, \dots, r$, define $\delta_k = \delta^{\frac{1}{\alpha}(\frac{\alpha}{\alpha+\beta})^k}$ so that $\delta_0^\alpha = \delta$, $\delta^{\epsilon_0} \le \delta_r^\alpha$ and $\delta_k^{\alpha+\beta} = \delta_{k-1}^\alpha$ for all $k = 1, \dots, r$.

For all $k = 1, \dots, r$, assumptions of Proposition 3.21 at scale δ_k are satisfied provided that $\epsilon < \frac{1}{\alpha}(\frac{\alpha}{\alpha+\beta})^r \min\{\epsilon_1, \frac{\kappa}{2}\}$. Thus

$$\mathbf{B}(0, \delta_k^\alpha) \subset \langle A \rangle_s + \mathbf{B}(0, \delta_{k-1}^\alpha).$$

Hence,

$$\begin{aligned} \mathbf{B}(0, \delta^{\epsilon_0}) &\subset \mathbf{B}(0, \delta_r^\alpha) \\ &\subset \langle A \rangle_s + \mathbf{B}(0, \delta_{r-1}^\alpha) \\ &\subset \langle A \rangle_s + \langle A \rangle_s + \mathbf{B}(0, \delta_{r-2}^\alpha) \\ &\dots \\ &\subset \langle A \rangle_s + \dots + \langle A \rangle_s + \mathbf{B}(0, \delta_0^\alpha) \end{aligned}$$

Hence, $\mathbf{B}(0, \delta^{\epsilon_0}) \subset \langle A \rangle_{rs} + \mathbf{B}(0, \delta)$. □

3.5.4 Proof of Theorem 3.1

We deduce Theorem 3.1 from Theorem 3.4 and Lemma 2.9.

Proof of Theorem 3.1. Suppose for a contradiction that for arbitrarily small $\epsilon > 0$, there exists $A \subset E$ satisfying the assumptions of Theorem 3.1 but

$$\mathcal{N}_\delta(A + A) + \mathcal{N}_\delta(A + A \cdot A) \le \delta^{-\epsilon} \mathcal{N}_\delta(A).$$

We will show a contradiction when ϵ is smaller than a constant depending only on E, κ and σ.

On the one hand, applying Theorem 3.4 with $\epsilon_0 = \frac{\dim(E)-\sigma}{2\dim(E)}$, we obtain an integer $s \ge 1$ depending only on E, κ and σ and such that

$$\mathbf{B}(0, \delta^{\epsilon_0}) \subset \langle A \rangle_s + \mathbf{B}(0, \delta).$$

Hence

$$\delta^{-\frac{1}{2}(\dim(E)+\sigma)} = \delta^{-(1-\epsilon_0)\dim(E)} \ll_E \mathcal{N}_\delta(\langle A \rangle_s). \tag{3.12}$$

On the other hand, by Lemma 2.9,

$$\mathcal{N}_\delta(\langle A\rangle_s) \ll_E \delta^{-O_s(\epsilon)}\mathcal{N}_\delta(A) \ll \delta^{-\sigma-O_s(\epsilon)}. \tag{3.13}$$

The inequalities (3.12) and (3.13) lead to a contradiction when ϵ is sufficiently small. □

Remark. Conversely, a growth statement like Theorem 3.1 always implies a statement like Theorem 3.4. The idea is to use the growth statement repeatedly until the set is nearly "full-dimensional" and then Proposition 2.10 shows that within a few more steps, it grows to "full dimension".

3.6 Growth under linear action

We prove Theorem 3.2 in this section.

3.6.1 Acting on $\mathbb{R}^n$, probabilistic method

We endow $\mathrm{End}(\mathbb{R}^n)$ with its usual operator norm. Recall that we denote by $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})_\mathbb{R}$ the standard embedding of $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})$ in $\mathrm{End}(\mathbb{R}^n)$ and by $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})_\mathbb{R}$ that of $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})$. First we study the special case where the collection of endomorphisms A is the unit ball in $\mathrm{End}(\mathbb{R}^n)$ or one of these two subalgebras. Actually, it is a direct consequence of [11, Proposition 1]. Here, we present an elementary proof of this easier fact.

Lemma 3.23. *Given $\kappa > 0$ and $\sigma < n$, there is $\epsilon > 0$ such that the following holds for $\delta > 0$ sufficiently small. Let E be $\mathrm{End}(\mathbb{R}^n)$ or $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})_\mathbb{R}$ or $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})_\mathbb{R}$. Let X be a subset of $\mathbb{R}^n$. Assume that*

(i) $X \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(X) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) $\mathcal{N}_\delta(X) \leq \delta^{-\sigma-\epsilon}$,

then

$$\max_{f\in\mathbf{B}_E(0,1)} \mathcal{N}_\delta(X + fX) \geq \delta^{-\epsilon}\mathcal{N}_\delta(X).$$

Proof. Let μ be the normalized Lebesgue measure on $\mathbf{B}_E(0,1)$. It is easy to verify that μ satisfies the assumptions of the following proposition with $\tau = n$. Note that $\mathbf{B}_E(0,1)$ contains the identity Id. □

Proposition 3.24. *Given $\kappa > 0$ and $0 < \sigma < \tau$, there is $\epsilon > 0$ such that the following holds for $\delta > 0$ sufficiently small. Let X be a subset of $\mathbb{R}^n$ and μ a probability measure on $\mathrm{End}(\mathbb{R}^n)$. If*

(i) $X \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(ii) $\forall \rho \geq \delta$, $\mathcal{N}_\rho(X) \geq \delta^\epsilon \rho^{-\kappa}$,

(iii) $\mathcal{N}_\delta(X) \leq \delta^{-\sigma-\epsilon}$,

(iv) The support of μ, $\mathrm{Supp}(\mu) \subset \mathbf{B}(0, \delta^{-\epsilon})$,

(v) For all $\rho \geq \delta$ and all $v, w \in \mathbb{R}^n$ with $\|v\| = 1$,

$$\mu\big(\{f \in \mathrm{End}(\mathbb{R}^n) \mid fv \in w + \mathbf{B}(0,\rho)\}\big) \leq \delta^{-\epsilon}\rho^{\tau},$$

then

$$\mathcal{N}_\delta(X+X) + \max_{f \in \mathrm{Supp}(\mu)} \mathcal{N}_\delta(X + fX) \geq \delta^{-\epsilon}\mathcal{N}_\delta(X).$$

Proof. Let X and μ be as in the statement. Assume that $\mathcal{N}_\delta(X+X) \leq \delta^{-\epsilon}\mathcal{N}_\delta(X)$. For all $\rho \geq \delta$ we have

$$\mathcal{N}_\delta(X+X) \gg \mathcal{N}_\rho(X) \max_{w \in \mathbb{R}^n} \mathcal{N}_\delta(X \cap \mathbf{B}(w,\rho)).$$

Therefore,

$$\max_w \mathcal{N}_\delta(X \cap \mathbf{B}(w,\rho)) \leq \delta^{-O(\epsilon)}\rho^{\kappa}\mathcal{N}_\delta(X). \tag{3.14}$$

Let f be a random variable following the law μ. Define $\varphi_f \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ by

$$\varphi_f(x,y) = x + fy.$$

This map is $\delta^{-\epsilon}$-Lipschitz by assumption (iv). Consider the φ_f-energy of $X \times X$. By Lemma 2.6(i),

$$\mathcal{N}_\delta(X + fX) \gg \frac{\mathcal{N}_\delta(X)^4}{\omega_\delta(\varphi_f, X \times X)}.$$

Hence by Jensen's inequality,

$$\mathbb{E}\big[\mathcal{N}_\delta(X + fX)\big] \gg \frac{\mathcal{N}_\delta(X)^4}{\mathbb{E}\big[\omega_\delta(\varphi_f, X \times X)\big]}.$$

The rest of the proof consists of bounding the expectation $\mathbb{E}\big[\omega_\delta(\varphi_f, X \times X)\big]$ from above. Fix $\tilde{X}$ a maximal δ-separated subset of X. By Lemma 2.6(iii),

$$\mathbb{E}\big[\omega_\delta(\varphi_f, X \times X)\big] \ll \sum_{x,x',y,y' \in \tilde{X}} \mathbb{P}\big[f(y-y') \in x' - x + B(0,\delta^{1-2\epsilon})\big].$$

Let $\rho > 0$ be a constant to be chosen later. We distinguish two cases according to whether $\|y - y'\| \geq \rho$. If it is the case then the assumption (v) yields for all $x, x' \in \tilde{X}$,

$$\mathbb{P}\big[f(y-y') \in x' - x + B(0,\delta^{1-2\epsilon})\big] \leq \delta^{-O(\epsilon)}\rho^{-\tau}\delta^{\tau}.$$

Otherwise, the number of pairs (y, y') such that $\|y - y'\| \leq \rho$ can be bounded using (3.14).

$$\#\{(y,y') \in \tilde{X} \times \tilde{X} \mid \|y-y'\| \leq \rho\} \leq \delta^{-O(\epsilon)}\rho^{\kappa}\mathcal{N}_\delta(X)^2.$$

Moreover, we have for all $x, y, y' \in \tilde{X}$,

$$\sum_{x' \in \tilde{X}} \mathbb{P}\big[f(y-y') \in x' - x + B(0,\delta^{1-2\epsilon})\big] \leq \delta^{-O(\epsilon)}$$

since the events on the left-hand side can occur simultaneously for at most $\delta^{-O(\epsilon)}$ different $x' \in \tilde{X}$.

By combining these inequalities and assumption (iii) and taking $\rho = \delta^{\frac{\tau-\sigma}{\tau+\kappa}}$, we obtain

$$\begin{aligned}\mathbb{E}\big[\omega_\delta(\varphi_f, X \times X)\big] &\leq \delta^{-O(\epsilon)}\big(\rho^{-\tau}\delta^\tau|\tilde{X}|^4 + \rho^\kappa|\tilde{X}|^3\big)\\ &\leq \delta^{-O(\epsilon)}(\delta^{\tau-\sigma}\rho^{-\tau} + \rho^\kappa)\mathcal{N}_\delta(X)^3\\ &\leq \delta^{\frac{\kappa(\tau-\sigma)}{\tau+\kappa}-O(\epsilon)}\mathcal{N}_\delta(X)^3.\end{aligned}$$

It follows that when ϵ is small enough, $\mathbb{E}\big[\mathcal{N}_\delta(X + fX)\big] \geq \delta^{-\epsilon}\mathcal{N}_\delta(X)$. □

3.6.2 Almost-generating a subalgebra

In view of Lemma 2.8 we know that in order to establish (3.2), it suffices to prove it with $\langle A\rangle_s + \mathbf{B}(0,\delta)$ in the place of A for some $s \geq 1$. That's why we can focus on growth of A as a set of matrices. We cannot use Theorem 3.4 yet since we do not know if A is away from subalgebras. In this subsection, we show that we can find a subalgebra E_0 of $\mathrm{End}(\mathbb{R}^n)$ such that A is effectively away from subalgebras in E_0 at some scale. This subalgebra E_0 can be viewed as approximately generated by A. Moreover, under the quantitative irreducibility condition, E_0 shall be described by Proposition 3.16.

Proposition 3.25. *For all $\epsilon_1 > 0$ there is $c > 0$ such that for all $0 < \epsilon < c$, the following holds for all $\delta > 0$ sufficiently small. Let A be subset of $\mathrm{End}(\mathbb{R}^n)$. If*

- *$A \subset \mathbf{B}(0, \delta^{-\epsilon})$,*
- *A acts δ^ϵ-irreducibly on $\mathbb{R}^n$;*

then there exists $\delta_1 \in [\delta, \delta^c]$ and $g \in \mathrm{GL}_n(\mathbb{R})$ with $\|g\| + \|g^{-1}\| \leq \delta^{-O(\epsilon)}$ such that for $E = \mathrm{End}(\mathbb{R}^n)$, $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})_\mathbb{R}$ or $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})_\mathbb{R}$,

$$gAg^{-1} \subset E + \mathbf{B}(0, \delta_1)$$

and for all proper subalgebras F of E,

$$\exists a \in A, d(gag^{-1}, F) > \delta_1^{\epsilon_1}.$$

Proof. Let l_0 be the largest among all integer $l \in \mathbb{N}$ such that there exists a subalgebra E of $\mathrm{End}(\mathbb{R}^n)$ of codimension l and such that $A \subset E + \mathbf{B}(0, \delta^{(\frac{\epsilon_1}{4})^l})$. We know l_0 exists since 0 is clearly such an l. Set $\delta_1 = \delta^{\frac{1}{2}(\frac{\epsilon_1}{4})^{l_0}}$. Thus $\delta \leq \delta_1 \leq \delta^{c_1}$ with $c_1 = \frac{1}{2}(\frac{\epsilon_1}{4})^{n^2}$. By the definition of l_0 there is a subalgebra E_0 of $\mathrm{End}(\mathbb{R}^n)$ such that

$$A \subset E_0 + \mathbf{B}(0, \delta_1^2) \tag{3.15}$$

and for any proper subalgebra F of E_0, there is $a \in A$ such that

$$d(a, F) > \delta_1^{\frac{\epsilon_1}{2}} \tag{3.16}$$

We shall apply Proposition 3.16 to this subalgebra E_0 in order to conjugate it into one of the three "model" subalgebras. For any nonzero linear subspace

W of $\mathbb{R}^n$, since A acts δ^ϵ-irreducibly , there is $w \in \mathbf{B}_W(0,1)$ and $a \in A$ such that $d(aw, W) > \delta^\epsilon$. Then there is $a' \in E_0$ such that $\|a - a'\| < \delta_1$. Hence $\|a'\| < \delta^{-\epsilon} + \delta_1 < \delta^{-2\epsilon}$ and

$$d(\frac{a'}{\|a'\|}w, W) > \delta^{2\epsilon}(\delta^\epsilon - \delta_1) \geq \delta^{2\epsilon}(\delta^\epsilon - \delta^{c_1}) > \delta^{4\epsilon}$$

when $\epsilon \leq \frac{c_1}{2}$. Thus E_0 acts $\delta^{4\epsilon}$-irreducibly on $\mathbb{R}^n$. We conclude that there is $g \in \mathrm{GL}_n(\mathbb{R})$ with $\|g\| + \|g^{-1}\| < \delta^{-O(\epsilon)}$ and $E = gE_0g^{-1}$ is one of these three subalgebras : $\mathrm{End}(\mathbb{R}^n)$, $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})_\mathbb{R}$, $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})_\mathbb{R}$.

The map $x \mapsto gxg^{-1}$ is $\delta^{-O(\epsilon)}$-bi-Lipschitz. Hence by (3.15) and (3.16),

$$gAg^{-1} \subset E + \mathbf{B}(0, \delta^{-O(\epsilon)}\delta_1^2)$$

and gAg^{-1} is $\delta^{O(\epsilon)}\delta_1^{\frac{\epsilon_1}{2}}$-away from proper subalgebras of E. When $\epsilon \leq \frac{c_1\epsilon_1}{C'}$, we have $\delta_1^{\epsilon_1} < \delta^{C'\epsilon}$ and hence both

$$\delta^{-O(\epsilon)}\delta_1^2 \leq \delta_1 \quad \text{and} \quad \delta^{O(\epsilon)}\delta_1^{\frac{\epsilon_1}{2}} \geq \delta_1^{\epsilon_1}.$$

This completes the proof. □

Remark. From the proof we see that the new scale δ_1 can be chosen such that δ is an integer power of δ_1.

3.6.3 Proof of Theorem 3.2.

The proof consists of putting together what precedes. The only technical difficulty is due to change of working scale required by Proposition 3.25.

Proof of Theorem 3.2. Assume for a contradiction that for $\epsilon > 0$ arbitrarily small, there exists $A \subset \mathrm{End}(\mathbb{R}^n)$ and $X \subset \mathbb{R}^n$ such that the assumptions in Theorem 3.2 are met but the conclusion fails, i.e. $A \cup \{\mathrm{Id}\} \subset S_\delta(X; \delta^{-\epsilon})$ using the notation introduced in §2.4.

Let $\epsilon_1 > 0$ be a small constant to be chosen later. By applying Proposition 3.25 to the set A, we get a constant $c > 0$ depending on ϵ_1, a new scale $\delta_1 \in [\delta, \delta^c]$, an element $g \in \mathrm{GL}_n(\mathbb{R})$ and a subalgebra $E = \mathrm{End}(\mathbb{R}^n)$ or $\mathcal{M}_{\frac{n}{2}}(\mathbb{C})_\mathbb{R}$ or $\mathcal{M}_{\frac{n}{4}}(\mathbb{H})_\mathbb{R}$ such that $\|g\| + \|g^{-1}\| \leq \delta^{-O(\epsilon)}$ and

$$gAg^{-1} \subset E + \mathbf{B}(0, \delta_1)$$

and the projection of gAg^{-1} on E, which we will denote by A', is $\delta_1^{\epsilon_1}$-away from subalgebras in E. Moreover, we have

$$A' \subset gAg^{-1} + \mathbf{B}(0, \delta_1) \tag{3.17}$$

Thus

$$A' \subset \mathbf{B}(0, \delta^{-O(\epsilon)} + \delta_1) \subset \mathbf{B}(0, \delta_1^{-O(\frac{\epsilon}{c})}).$$

Morever, for all $\rho \geq \delta_1$, since $gAg^{-1} \subset A' + \mathbf{B}(0, \rho)$, we have

$$\mathcal{N}_\rho(A') \gg \mathcal{N}_\rho(gAg^{-1}) \gg \delta^{O(\epsilon)}\mathcal{N}_\rho(A) \geq \delta^{O(\epsilon)}\rho^{-\kappa} \geq \delta_1^{O(\frac{\epsilon}{c})}\rho^{-\kappa}.$$

So when ϵ is sufficiently small depending on ϵ_1 and c, we have both $A' \subset \mathbf{B}(0, \delta_1^{-\epsilon_1})$ and $\forall \rho \geq \delta_1, \mathcal{N}_\rho(A') \geq \delta_1^{\epsilon_1} \rho^{-\kappa}$.

Let $\epsilon_0 > 0$ be a small constant to be chosen later in the proof. Applying Theorem 3.4 to the set A' at scale δ_1 inside the algebra E, we obtain an integer $s \geq 1$ depending only on n, κ and ϵ_0 such that

$$\mathbf{B}_E(0, \delta_1^{\epsilon_0}) \subset \langle A' \rangle_s + \mathbf{B}(0, \delta_1), \tag{3.18}$$

if ϵ_1 is chosen small enough depending on n, κ and ϵ_0.

From $A \subset S_\delta(X; \delta^{-\epsilon})$, we get for all $a \in A$,

$$\mathcal{N}_\delta(gX + gag^{-1}gX) = \mathcal{N}_\delta\big(g(X + aX)\big) \leq \delta^{-O(\epsilon)} \mathcal{N}_\delta(X) \leq \delta^{-O(\epsilon)} \mathcal{N}_\delta(gX).$$

That is to say $gAg^{-1} \subset S_\delta(gX, \delta^{-O(\epsilon)})$. By Lemma 2.8(v), $gAg^{-1} \subset S_{\delta_1}(gX; \delta^{-O(\epsilon)})$. Then by Lemma 2.8(i),

$$A' \subset S_{\delta_1}(gX; \delta_1^{-O(\frac{\epsilon}{c})}) \subset S_{\delta_1}(gX; \delta_1^{-\epsilon_1}),$$

provided that ϵ is sufficiently small compared to $c\epsilon_1$. Repeated use of Lemma 2.8(ii) yields

$$\langle A' \rangle_s \subset S_{\delta_1}(gX; \delta_1^{-O_s(\epsilon_1)}).$$

Then by (3.18) and Lemma 2.8(i),

$$\mathbf{B}_E(0, \delta_1^{\epsilon_0}) \subset S_{\delta_1}(gX; \delta_1^{-O_s(\epsilon_1)}).$$

In particular, $\delta_1^{\epsilon_0} \operatorname{Id} \in S_{\delta_1}(gX; \delta_1^{-O_s(\epsilon_1)})$. Hence, by Lemma 2.8(iii),

$$\delta_1^{-\epsilon_0} \operatorname{Id} \in S_{\delta_1}(gX; \delta_1^{-O_s(\epsilon_1)-O(\epsilon_0)}).$$

Then again by Lemma 2.8(ii), $\mathbf{B}_E(0, 1) \subset S_{\delta_1}(gX; \delta_1^{-O_s(\epsilon_1)-O(\epsilon_0)})$.

Hence for any given $\epsilon_2 > 0$, we can choose sufficiently small $\epsilon_0 > 0$ and $\epsilon_1 > 0$ accordingly so that

$$\mathbf{B}_E(0, 1) \subset S_{\delta_1}(gX; \delta_1^{-\epsilon_2}). \tag{3.19}$$

Take ϵ_2 to be the constant given by Lemma 3.23 depending on σ and κ. We would like to apply Lemma 3.23 to the set gX at scale δ_1. It's easy to see that when ϵ is small enough,

$$gX \subset \mathbf{B}(0, \delta^{-O(\epsilon)}) \subset \mathbf{B}(0, \delta_1^{-\epsilon_2}),$$

and for all $\rho \geq \delta_1$,

$$\mathcal{N}_\rho(gX) \geq \delta^{O(\epsilon)} \mathcal{N}_\delta(X) \geq \delta^{O(\epsilon)} \rho^{-\kappa} \geq \delta_1^{\epsilon_2} \rho^{-\kappa}.$$

So the first two assumptions in Lemma 3.23 are satisfied but the conclusion fails by (3.19). This means the assumption (iii) must fail, namely, $\mathcal{N}_{\delta_1}(gX) > \delta_1^{-\sigma-\epsilon_2}$. Therefore,

$$\mathcal{N}_{\delta_1}(X) \geq \delta^{O(\epsilon)} \delta_1^{-\sigma-\epsilon_2}. \tag{3.20}$$

In other words, at scale δ_1, the set X is almost full in $\mathbf{B}(0, \delta^{-\epsilon})$. The idea of the rest of the proof is to use multiplication of elements of A to propagate

estimate (3.20) to smaller scales until we reach the original scale δ where we have the assumption $\mathcal{N}_\delta(X) \leq \delta^{-\sigma-\epsilon}$.

We have by (3.18) and (3.17),

$$\delta_1^{\frac{1}{2}} \operatorname{Id} \in \langle A' \rangle_s + \mathbf{B}(0, \delta_1) \subset g\langle A \rangle_s g^{-1} + \mathbf{B}(0, \delta^{-O_s(\epsilon)}\delta_1).$$

Hence there exists $a \in \langle A \rangle_s$ such that $a \in \delta_1^{\frac{1}{2}} \operatorname{Id} + \mathbf{B}(0, \delta^{-O_s(\epsilon)}\delta_1)$. Taking the square, we obtain $a^2 \in \delta_1 \operatorname{Id} + \mathbf{B}(0, \delta^{-O_s(\epsilon)}\delta_1^{\frac{3}{2}})$. Thus, when ϵ is sufficiently small,

$$a^2 \in \delta_1 \operatorname{Id} + \mathbf{B}(0, \frac{\delta_1}{2}).$$

For all $\rho > 0$, the multiplication by a^2 will transform a ρ-separated set in $\mathbb{R}^n$ into a $\frac{\delta_1\rho}{2}$-separated set. Hence,

$$\mathcal{N}_{\delta_1\rho}(a^2 X) \gg \mathcal{N}_\rho(X).$$

Moreover $a^2 X \subset \mathbf{B}(0, \delta^{-2\epsilon}\delta_1)$. Hence

$$\mathcal{N}_{\delta_1\rho}(X + a^2 X) \geq \delta^{O(\epsilon)} \mathcal{N}_{\delta^{-2\epsilon}\delta_1}(X) \mathcal{N}_{\delta_1\rho}(a^2 X) \geq \delta^{O(\epsilon)} \mathcal{N}_{\delta_1}(X) \mathcal{N}_\rho(X).$$

Lemma 2.8 tells us $a^2 \in S_\delta(X, \delta^{-O_s(\epsilon)})$ and for all $\rho \geq \delta_1^{-1}\delta$,

$$\mathcal{N}_{\delta_1\rho}(X + a^2 X) \leq \delta^{-O_s(\epsilon)} \mathcal{N}_{\delta_1\rho}(X).$$

Combining the two estimates above, we obtain

$$\mathcal{N}_{\delta_1\rho}(X) \geq \delta^{O_s(\epsilon)} \mathcal{N}_{\delta_1}(X) \mathcal{N}_\rho(X) \tag{3.21}$$

whenever $\delta_1\rho \geq \delta$.

According to the remark after the proof of Proposition 3.25, we can assume that there is an integer k with $1 \leq k \leq \frac{1}{c}$ such that $\delta = \delta_1^k$. Applying (3.21) to $\rho = \delta_1^i$, $i = 1, \ldots, k-1$, we obtain

$$\mathcal{N}_{\delta_1^k}(X) \geq \delta^{O_s(k\epsilon)} \mathcal{N}_{\delta_1}(X)^k.$$

Hence, by (3.20) and the assumption on X,

$$\delta^{-\sigma-\epsilon} \geq \mathcal{N}_\delta(X) \geq \delta^{-\sigma-\epsilon_2+O_s(k\epsilon)},$$

which is clearly impossible when ϵ is small enough and this finishes the proof. □

Remark. Theorem 3.2 together with Lemma 3.10 yields another proof of Theorem 3.1. It suffices to consider the action of left and right multiplications of elements of A on the set A. Actually, this proves something slightly stronger. Namely, the conclusion of Theorem 3.1 can be improved to

$$\mathcal{N}_\delta(A + A) + \max_{a \in A} \mathcal{N}_\delta(A + a \cdot A) + \max_{a \in A} \mathcal{N}_\delta(A + A \cdot a) \geq \delta^{-\epsilon} \mathcal{N}_\delta(A).$$

3.7 A sum-product estimate in simple Lie algebras

In this last section, we prove Corollary 3.3. Let $\mathfrak{g}$ be a normed real simple Lie algebra. We consider the adjoint representation $\mathrm{ad}\colon \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$, i.e. $\mathrm{ad}(a)x = [a,x]$ for all $a, x \in \mathfrak{g}$.

Proof of Corollary 3.3. It suffices to apply Theorem 3.2 to the set $A \subset \mathfrak{g}$ and the set of endomorphisms $\mathrm{ad}(A) \subset \mathrm{End}(\mathfrak{g})$. Note that the kernel of ad is the center of $\mathfrak{g}$ which is trivial for $\mathfrak{g}$ is simple. Therefore, ad is a bi-Lipschitz map from $\mathfrak{g}$ to its image. This gives the non-concentration condition on $\mathrm{ad}(A)$. Moreover, the quantitative irreducibility condition of Theorem 3.2 is guaranteed by the following lemma. □

Lemma 3.26. *Let $0 < \rho \leq \frac{1}{2}$ be a parameter. Let A be a subset in a normed simple Lie algebra $\mathfrak{g}$ of finite dimension. Assume that $A \subset \mathbf{B}(0, \rho^{-1})$ and that A is ρ-away from Lie subalgebras in $\mathfrak{g}$. Then $\mathrm{ad}(A)$ acts $\rho^{O_{\mathfrak{g}}(1)}$-irreducibly on $\mathfrak{g}$.*

Proof. Without loss of generality, we can assume the norm on $\mathfrak{g}$ to be Euclidean. The statement and proof of Lemma 3.9 remains valid when the word "subalgebra" is replaced by "Lie subalgebra". Actually this fact is implicit in [24, Lemma 2.5]. Therefore, as in the proof of Lemma 3.10, we can assume that A is finite of cardinality at most n and contained in $\mathbf{B}(0,1)$. Write $A = \{a_1, \dots, a_n\}$.

Suppose the conclusion were false, which means there is W_0 linear subspace of dimension $0 < k < n$ such that for all $w \in \mathbf{B}_{W_0}(0,1)$ and all $i = 1, \dots, n$, $d(\mathrm{ad}(a_i)w, W_0) < \rho^C$ where C is a large constant to be determined by the use of the Łojasiewicz inequality below. Consider the following real analytic map :

$$f\colon \begin{array}{ccc} \mathrm{Gr}(\mathfrak{g},k) \times \mathfrak{g}^n & \to & \mathbb{R} \\ (W; x_1, \dots, x_n) & \mapsto & \displaystyle\sum_{i=1}^{n} \int_{\mathbf{B}_W(0,1)} d(\mathrm{ad}(x_i)w, W)^2 \, \mathrm{d}w \end{array}$$

From the above, $f(W_0; a_1, \dots, a_n) \ll \rho^C$. Application of the Łojasiewicz inequality (Theorem 3.8) to the compact set $\mathrm{Gr}(\mathfrak{g},k) \times \mathbf{B}(0,1)^n$ gives $W_1 \in \mathrm{Gr}(\mathfrak{g},k)$ and $b_1, \dots, b_n \in \mathfrak{g}$ such that $f(W_1; b_1, \dots, b_n) = 0$ and $\forall i$, $\|a_i - b_i\| < \rho$ when the constant C is chosen large enough. The fact $f(W_1; b_1, \dots, b_n) = 0$ is equivalent to every b_i being in the Lie subalgebra

$$\mathfrak{g}_{W_1} = \{x \in \mathfrak{g} \mid \mathrm{ad}(x)W_1 \subset W_1\}.$$

Now our set A is not ρ-away from the Lie subalgebra $\mathfrak{g}_{W_1}$. Hence $\mathfrak{g}_{W_1}$ must be $\mathfrak{g}$, which in turn implies that W_1 is an ideal in $\mathfrak{g}$. This contradicts the simplicity of $\mathfrak{g}$. □

Chapter 4

Orthogonal projections of discretized sets

In this chapter we generalize Bourgain's discretized projection theorem (Theorem 1.6) to higher rank projections. Let $0 < m < n$ be positive integers. We denote by $\mathrm{Gr}(\mathbb{R}^n, m)$ the Grassmannian of m-dimensional subspaces in $\mathbb{R}^n$. For $V \in \mathrm{Gr}(\mathbb{R}^n, m)$, $\pi_V \colon \mathbb{R}^n \to V$ stands for the orthogonal projection onto V. If $W \in \mathrm{Gr}(\mathbb{R}^n, n-m)$, we define

$$\mathrm{d}_{\measuredangle}(V, W) = \|v_1 \wedge \cdots \wedge v_m \wedge w_1 \wedge \cdots \wedge w_{n-m}\|$$

where $(v_1, \ldots, v_m)$ is an orthonormal basis of V and $(w_1, \ldots, w_{n-m})$ an orthonormal basis of W. For example $\mathrm{d}_{\measuredangle}(V, W) = 0$ if and only if V and W have nontrivial intersection. For $\rho \geq 0$, we denote by $\mathcal{V}_{\measuredangle}(W, \rho)$ the set of all $V \in \mathrm{Gr}(\mathbb{R}^n, m)$ such that $\mathrm{d}_{\measuredangle}(V, W) \leq \rho$. Recall that $\mathcal{V}_{\measuredangle}(W, 0)$ is a submanifold of codimension 1 in $\mathrm{Gr}(\mathbb{R}^n, m)$ and belongs to the class of algebraic subvarieties known as Schubert cycles (see for example [35, Chapter 1, §5]).

Our main result is the following.

Theorem 4.1. *Let $m < n$ be positive integers. Given $0 < \alpha < n$ and $\kappa > 0$, there exists $\epsilon > 0$ such that the following holds for sufficiently small $\delta > 0$. Let A be a subset of $\mathbb{R}^n$ contained in the unit ball $\mathbf{B}(0, 1)$. Let μ be probability measure on $\mathrm{Gr}(\mathbb{R}^n, m)$. Assume that*

$$\mathcal{N}_\delta(A) \geq \delta^{-\alpha+\epsilon}; \tag{4.1}$$

$$\forall \rho \geq \delta,\ \forall x \in \mathbb{R}^n, \quad \mathcal{N}_\delta(A \cap \mathbf{B}(x, \rho)) \leq \delta^{-\epsilon} \rho^{\kappa} \mathcal{N}_\delta(A); \tag{4.2}$$

$$\forall \rho \geq \delta,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, n-m), \quad \mu(\mathcal{V}_{\measuredangle}(W, \rho)) \leq \delta^{-\epsilon} \rho^{\kappa}. \tag{4.3}$$

Then there is a set $\mathcal{D} \subset \mathrm{Gr}(\mathbb{R}^n, m)$ such that $\mu(\mathcal{D}) \geq 1 - \delta^{\epsilon}$ and

$$\mathcal{N}_\delta(\pi_V(A')) \geq \delta^{-\frac{m}{n}\alpha - \epsilon}$$

whenever $V \in \mathcal{D}$ and $A' \subset A$ is a subset such that $\mathcal{N}_\delta(A') \geq \delta^{\epsilon} \mathcal{N}_\delta(A)$.

For $1 < m < n$, our result is new. Hypothesis (4.2) is a Frostman type[1] non-concentration condition on A. Without it we can have example like $A = \mathbf{B}(0, \delta^{1-\frac{\alpha}{n}})$, a ball of radius $\delta^{1-\frac{\alpha}{n}}$, whose size is $\mathcal{N}_\delta(A) \asymp \delta^{-\alpha}$ but whose projection to any $V \in \mathrm{Gr}(\mathbb{R}^n, m)$ is of size

$$\mathcal{N}_\delta(\pi_V(A)) \asymp \delta^{-\frac{m}{n}\alpha}.$$

Hypothesis (4.3) is a non-concentration condition on the distribution of the subspace V. The set $\mathcal{V}_\measuredangle(W, \rho)$ can be thought of as a ρ-neighborhood of the Schubert cell $\mathcal{V}_\measuredangle(W, 0)$. For example if $m = 1$, V lives in the projective space and (4.3) is asking μ to be not concentrated around any projective subspace. Note that the factor $\delta^{-\epsilon}$ in both (4.2) and (4.3) means that the non-concentration property needs to be satisfied up to scale δ^ϵ. So the parameter κ is about how good the assumptions are and ϵ is about how much the assumptions can be relaxed and how good the conclusion is.

Just like Bourgain's discretized projection theorem can be used to derive a projection theorem in terms of Hausdorff dimension [7, Theorem 4], Theorem 4.1 has the following consequence.

Corollary 4.2. *Let $m < n$ be positive integers. Given $0 < \alpha < n$ and $\kappa > 0$, there is $\epsilon > 0$ such that the following is true. Let $A \subset \mathbb{R}^n$ is a Borel set of dimension $\dim_{\mathrm{H}}(A) = \alpha$. Then the set of exceptional directions*

$$\big\{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \dim_{\mathrm{H}}(\pi_V(A)) \le \frac{m}{n}\alpha + \epsilon\big\}$$

does not support any finite Borel measure μ on $\mathrm{Gr}(\mathbb{R}^n, m)$ with the following non-concentration property,

$$\forall \rho > 0,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, n-m), \quad \mu(\mathcal{V}_\measuredangle(W, \rho)) \le \rho^\kappa.$$

Applied to a Frostman measure supported on the set of exceptional directions, we get

Corollary 4.3. *Let $m < n$ be positive integers. Given $0 < \alpha < n$ and $\kappa > 0$, there is $\epsilon > 0$ such that the following holds. Let $A \subset \mathbb{R}^n$ be a Borel set of dimension $\dim_{\mathrm{H}}(A) = \alpha$. Then*

$$\dim_{\mathrm{H}}\big\{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \dim_{\mathrm{H}}(\pi_V(A)) \le \frac{m}{n}\alpha + \epsilon\big\} \le m(n-m) - 1 + \kappa.$$

Compared to what is already known (Theorem 1.5), the number 1 in our estimate is very weak. However, our result does provide something new for specific choice of n, m and α, namely when $\frac{n}{m}(m-1) < \alpha \le m$ or when $m < \alpha < \frac{n}{n-m}$.

4.0.1 Strategy of the proof

Fix n, m and α. For $\epsilon > 0$ and bounded subset $A \subset \mathbb{R}^n$ we define the set of exceptional directions to be

$$\mathcal{E}(A) = \{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \exists A' \subset A,\ \mathcal{N}_\delta(A') \ge \delta^\epsilon \mathcal{N}_\delta(A) \text{ and } \mathcal{N}_\delta(\pi_V(A')) < \delta^{-\frac{m}{n}\alpha - \epsilon}\}.$$

[1] Cf. Frostman's lemma (Theorem 1.1).

When we want to specify ϵ, we write $\mathcal{E}(A, \epsilon)$ instead. Our task is to bound $\mu(\mathcal{E}(A))$. In order to prove Theorem 4.1 which says $\mu(\mathcal{E}(A)) \leq \delta^{\epsilon}$ under the assumptions of the theorem, we prove instead that $\mu(\mathcal{E}(A')) \leq \delta^{\epsilon}$ for some subset A' of A.

Theorem 4.4. *Let $m < n$ be positive integers. Given $0 < \alpha < n$ and $\kappa > 0$, there exists $\epsilon > 0$ such that the following holds for sufficiently small $\delta > 0$. Let A be a subset of $\mathbb{R}^n$ contained in the unit ball $\mathbf{B}(0,1)$. Let μ be a probability measure on $\mathrm{Gr}(\mathbb{R}^n, m)$. Assume (4.1), (4.2) and (4.3), then there exists $A' \subset A$ such that*

$$\mu(\mathcal{E}(A')) \leq \delta^{\epsilon}.$$

This statement is seemingly weaker, but there is actually a rather formal argument which allows to deduce Theorem 4.1 from Theorem 4.4. We will show this implication in Proposition 4.17.

The proof of Theorem 4.4 starts with the special case where $n = 2m$.

Proposition 4.5. *Theorem 4.4 is true if $n = 2m$.*

As in the $m = 1$ case in [7], this special case is proved using a sum-product theorem. For $m > 1$, we need the higher dimensional sum-product estimate Theorem 3.2.

To see why this sum-product estimate can be helpful we remark that the space of m by m matrices, $\mathrm{End}(\mathbb{R}^m)$, is diffeomorphic to the complement of a codimension 1 Schubert cell in $\mathrm{Gr}(\mathbb{R}^{2m}, m)$. More precisely, if $\mathbb{R}^{2m}$ is the direct sum of two subspaces V_1 and V_2 both of dimension m, then the map $\mathrm{Gr}(\mathbb{R}^{2m}, m) \setminus \mathcal{V}_{\measuredangle}(V_1^{\perp}, 0) \to \mathcal{L}(V_2, V_1)$ defined by $V \mapsto \pi_{V|V_1}^{-1} \circ \pi_{V|V_2}$ is a diffeomorphism. Here $\mathcal{L}(V_2, V_1)$ stands for the space of linear maps from V_1 to V_2.

Once we have Proposition 4.5 we would like reduce other cases to this special case. If m divides n, this is done easily by considering large slices of dimension $n - m$ in A.

Proposition 4.6. *Let $q \geq 3$ be an integer. If Theorem 4.4 is true for $n' = (q-1)m$ and m then it is also true for $n = qm$ and m.*

If m does not divide n and $m < \frac{n}{2}$, the idea is the following. Write $n = qm + r$ with $0 < r < m$. Let $V_1, \ldots, V_q$ be linear subspaces of dimension m in "generic" position. If the projection of A to the sum subspace $V_1 + \cdots + V_q$ is large then its projection to one of the V_i must be large as well.

Proposition 4.7. *Let $0 < m < n$ be such that $qm < n$ where $q \geq 1$. If Theorem 4.4 is true for n and $m' = qm$ then it is also true for n and m.*

If m does not divide n and $m > \frac{n}{2}$, we are in a dual situation to the previous one. So we consider intersections instead of sums of subspaces. Write $n = q(n-m) + r$ with $0 < r \leq n - m$. Necessarily $q \geq 2$. Let $V_1, \ldots, V_q$ be linear subspaces of dimension m in "generic" position. The intersection $V_1 \cap \cdots \cap V_q$ has dimension r. If the projection of A to $V_1 \cap \cdots \cap V_q$ is large then we would like to conclude that its projection to one of the subspace V_i must be large as well. However, this is not true unless we assume that A does not have any large slice orthogonal to V (see Proposition 4.26). If A does have a large slice of dimension $n - r$, we can produce large projections using this slice.

Proposition 4.8. *Let $0 < m < n$ be such that $n = q(n-m) + r$ where $q \geq 1$ and $0 < r \leq n - m$. If Theorem 4.4 is true for n and $m' = r$ then it is also true for n and m.*

Let us see how we prove Theorem 4.4 by putting these propositions together.

Proof of Theorem 4.4. Propositions 4.5 and 4.6 imply the theorem for all pairs (n, m) such that m divides n. Consider the following order on pairs of positive integers. We say $(n, m) \prec (n', m')$ if $(n, \min(m, n-m), m)$ is smaller than $(n', \min(m', n'-m'), m')$ for the lexicographical order.

If the theorem were false then let (n, m) be a $\prec$-minimal pair for which the theorem fails. We know that m does not divide n. If $m < \frac{n}{2}$ then write $n = qm + r$ with $0 < r < m$. We have $(n, qm) \prec (n, m)$. Hence Proposition 4.7 contradicts the minimality of (n, m). Otherwise $m > \frac{n}{2}$, then write $n = q(n - m) + r$ with $0 < r \leq n - m$. We have $(n, r) \prec (n, m)$ and then Proposition 4.8 contradicts the minimality of (n, m). □

4.1 Preliminaries

In this section we collect some elementary estimates about the Grassmannian and establish two useful lemmata about intersections.

4.1.1 Distance on the Grassmannian

For linear subspaces V, W of $\mathbb{R}^n$, we define

$$\mathrm{d}_\measuredangle(V, W) = \|v_1 \wedge \cdots \wedge v_r \wedge w_1 \wedge \cdots \wedge w_s\|$$

where $(v_1, \ldots, v_r)$ is an orthonormal basis of V and $(w_1, \ldots, w_s)$ an orthonormal basis of W. It is a distance when restricted to the projective space $\mathrm{Gr}(\mathbb{R}^n, 1)$ but only in this case. For example, $\mathrm{d}_\measuredangle(V, W) = 0$ if and only if V and W have nontrivial intersection and $\mathrm{d}_\measuredangle(V, W) = 1$ if and only if they are orthogonal to each other. For other cases, $\mathrm{d}_\measuredangle(V, W)$ falls between 0 and 1.

If $v_1, \ldots, v_r$ are vectors and $\mathbf{w} = w_1 \wedge \cdots \wedge w_s$ the wedge product of an orthonormal basis of W, then

$$\|v_1 \wedge \cdots \wedge v_r \wedge \mathbf{w}\| = \|\pi_{W^\perp}(v_1) \wedge \cdots \wedge \pi_{W^\perp}(v_r)\|. \tag{4.4}$$

In particular, if $(v_1, \ldots, v_r)$ is an orthonormal basis of V, then

$$\mathrm{d}_\measuredangle(V, W) = \|\pi_{W^\perp}(v_1) \wedge \cdots \wedge \pi_{W^\perp}(v_r)\|. \tag{4.5}$$

If $f \colon V \to W$ is a linear map between euclidean spaces of same dimension, then the determinant of its matrix expressed in orthonormal bases up to a sign does not depend on the choice of the bases. Moreover, we have

$$|\det(f)| = \|f(v_1) \wedge \cdots \wedge f(v_r)\|$$

where $(v_1, \ldots, v_r)$ is an orthonormal basis of V. Together with (4.5) this gives yet another definition of $\mathrm{d}_\measuredangle(V, W)$ if $\dim(V) + \dim(W) = n$,

$$\mathrm{d}_\measuredangle(V, W) = |\det(\pi_{W^\perp | V})|, \tag{4.6}$$

where $\pi_{W^\perp|V}\colon V \to W^\perp$ denotes the restriction of π_W to V.

The natural action of the orthogonal group $\mathrm{O}(n)$ on the Grassmannian preserves $\mathrm{d}_\measuredangle$, i.e.

$$\forall g \in \mathrm{O}(n), \quad \mathrm{d}_\measuredangle(gV, gW) = \mathrm{d}_\measuredangle(V, W).$$

Consequently if $\dim V + \dim W = n$ then

$$\mathrm{d}_\measuredangle(V^\perp, W^\perp) = \mathrm{d}_\measuredangle(V, W), \tag{4.7}$$

because in this case we can always send V to $W^\perp$ (hence W to $V^\perp$) by an element of $\mathrm{O}(n)$.

Moreover, when we have several subspaces, $V_1, V_2, \ldots, V_q$ of $\mathbb{R}^n$, we define

$$\mathrm{d}_\measuredangle(V_1, \ldots, V_q) = \|\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_q\|$$

where for each $i = 1, \ldots, q$, $\mathbf{v}_i$ is the wedge product of the elements of an orthonormal basis of V_i. For example, if $x_1, \ldots, x_n \in \mathbb{R}^n$ are unit vectors, then

$$\mathrm{d}_\measuredangle(\mathbb{R}x_1, \ldots, \mathbb{R}x_n) = |\det(x_1, \ldots x_n)|.$$

Obviously, $\mathrm{d}_\measuredangle(V_1, \ldots, V_q)$ is symmetric in the variables $V_1, \ldots, V_q$. Below are some other elementary properties of $\mathrm{d}_\measuredangle$.

Lemma 4.9. *If U, V, W are linear subspaces of $\mathbb{R}^n$, then*

$$\mathrm{d}_\measuredangle(U, V, W) = \mathrm{d}_\measuredangle(U + V, W)\, \mathrm{d}_\measuredangle(U, V). \tag{4.8}$$

Consequently, if $V_1, \ldots, V_q$ are also linear subspaces, then

$$\mathrm{d}_\measuredangle(V_1, \ldots, V_q) = \mathrm{d}_\measuredangle(V_2, V_1)\, \mathrm{d}_\measuredangle(V_3, V_1 + V_2) \cdots \mathrm{d}_\measuredangle(V_q, V_1 + \cdots + V_{q-1}); \tag{4.9}$$

$$\mathrm{d}_\measuredangle(V_1 + \cdots + V_q, W) \geq \mathrm{d}_\measuredangle(V_1, W)\, \mathrm{d}_\measuredangle(V_2, V_1 + W) \cdots \mathrm{d}_\measuredangle(V_q, V_1 + \cdots + V_{q-1} + W). \tag{4.10}$$

Proof. If the sum $U + V$ is not a direct sum, then $\mathrm{d}_\measuredangle(U, V, W) = 0$ and $\mathrm{d}_\measuredangle(U, V) = 0$. Otherwise, let $\mathbf{u}$ and $\mathbf{v}$ be wedge products of orthonormal bases of U and V respectively. Then $\mathbf{u} \wedge \mathbf{v}/\|\mathbf{u} \wedge \mathbf{v}\|$ is the wedge product of an orthonormal basis of $U + V$. Then (4.8) follows immediately from the definition.

The estimates (4.9) can be obtained by a simple induction. The inequality (4.10) follows from (4.9) since, by (4.9), the right hand side of (4.10) is equal to $\mathrm{d}_\measuredangle(V_1, \ldots, V_q, W)$ which, by (4.9) again, is equal to $\mathrm{d}_\measuredangle(V_1, \ldots, V_q)\, \mathrm{d}_\measuredangle(V_1 + \cdots + V_q, W)$. □

Lemma 4.10. *Let $q \geq 2$. Let $V_1, \ldots, V_q$ be linear subspaces of $\mathbb{R}^n$. If $z \in V_1 + \cdots + V_q$ then*

$$\|z\|\, \mathrm{d}_\measuredangle(V_1, \ldots, V_q) \leq \|\pi_{V_1}(z)\| + \|\pi_{V_2}(z)\| + \cdots + \|\pi_{V_q}(z)\| \tag{4.11}$$

Proof. We will proceed by induction. Let $q = 2$. Obviously, there is nothing to prove if $V_1 + V_2$ is not a direct sum. Moreover, without loss of generality, we can assume that $\mathbb{R}^n = V_1 + V_2$. Then also $\mathbb{R}^n = V_1^\perp + V_2^\perp$. Write $z = z_1 + z_2$ with $z_1 \in V_1^\perp$ and $z_2 \in V_2^\perp$. Then

$$\|\pi_{V_1}(z)\| = \|\pi_{V_1}(z_2)\| = \|z_2\|\, \mathrm{d}_\measuredangle(V_1^\perp, \mathbb{R}z_2) \geq \|z_2\|\, \mathrm{d}_\measuredangle(V_1^\perp, V_2^\perp) = \|z_2\|\, \mathrm{d}_\measuredangle(V_1, V_2).$$

Similarly, $\|\pi_{V_2}(z)\| \geq \|z_1\| \, \mathrm{d}_\angle(V_1, V_2)$. We get the lemma for $q = 2$ using the triangular inequality.

Now, suppose the lemma is true for some $q \geq 2$. Let us show the lemma for $q+1$. Let $V_q' = V_q + V_{q+1}$ and $z' = \pi_{V_q'}(z)$. The induction hypothesis applied to z and $(V_1, \ldots, V_{q-1}, V_q')$ gives

$$\|z\| \, \mathrm{d}_\angle(V_1, \ldots, V_{q-1}, V_q + V_{q+1}) \leq \|\pi_{V_1}(z)\| + \cdots + \|\pi_{V_{q-1}}(z)\| + \|z'\|.$$

The lemma applied to z' and (V_q, V_{q+1}) gives

$$\|z'\| \, \mathrm{d}_\angle(V_q, V_{q+1}) \leq \|\pi_{V_q}(z')\| + \|\pi_{V_{q+1}}(z')\| = \|\pi_{V_q}(z)\| + \|\pi_{V_{q+1}}(z)\|.$$

Recall that $\mathrm{d}_\angle(V_1, \ldots, V_{q+1}) = \mathrm{d}_\angle(V_1, \ldots, V_{q-1}, V_q + V_{q+1}) \, \mathrm{d}_\angle(V_q, V_{q+1})$. We obtain the desired estimate by multiplying the first inequality by $\mathrm{d}_\angle(V_q, V_{q+1})$ and combining it with the second. □

Lemma 4.11. *If $\mathbb{R}^n$ is a direct sum of $V_1, \ldots, V_q$ then for any bounded subset $A \subset \mathbb{R}^n$,*

$$\mathcal{N}_\delta(A) \ll_n \mathrm{d}_\angle(V_1, \ldots, V_q)^{-n} \prod_{i=1}^{q} \mathcal{N}_\delta(\pi_{V_i}(A)). \tag{4.12}$$

Proof. Suppose for each $i \in \{1, \ldots, q\}$, $\pi_{V_i}(A)$ is covered by the balls $x_i^{(\delta)}$, $x_i \in X_i \subset V_i$. For each $(x_i)_i \in X_1 \times \cdots \times X_q$, there is a unique $x \in \mathbb{R}^n$ such that $\forall i$, $\pi_{V_i}(x) = x_i$. By Lemma 4.10, we have

$$\pi_{V_1}^{-1}(x_1^{(\delta)}) \cap \cdots \cap \pi_{V_q}^{-1}(x_q^{(\delta)}) \subset x^{(\delta')},$$

where $\delta' = \mathrm{d}_\angle(V_1, \ldots, V_q)^{-1} q \delta$. So A is covered by the balls centered at such x. Hence $\mathcal{N}_{\delta'}(A) \leq |X_1| \cdots |X_q|$. We then conclude by using the scale change estimate (2.3). □

Lemma 4.12. *Let V, W be linear subspaces of $\mathbb{R}^n$. If $V' = \pi_W(V)$, then for all $x \in W$,*

$$\mathrm{d}_\angle(V, W^\perp) \|\pi_{V'}(x)\| \leq \|\pi_V(x)\| \leq \|\pi_{V'}(x)\|. \tag{4.13}$$

In particular, if moreover $\dim V = \dim W$, then for all $x \in W$,

$$\mathrm{d}_\angle(V, W^\perp) \|x\| \leq \|\pi_V(x)\| \leq \|x\|.$$

Proof. Since $V' = \pi_W(V)$, we have $V'^\perp \cap W \subset V^\perp$. Hence we can write $x = y + z$ with $y = \pi_{V'}(x) \in V'$ and $z \in V'^\perp \cap W \subset V^\perp$. Then $\pi_V(x) = \pi_V(y)$. This gives the second inequality in (4.13).

It is clear that V and V' have different dimensions if and only if V and $W^\perp$ have nontrivial intersection, which is equivalent to $\mathrm{d}_\angle(V, W^\perp) = 0$. In this case, the first inequality in the lemma holds.

Let us assume $\dim V = \dim V'$. In this case there is $g \in \mathrm{O}(n)$ which exchanges V with V'. We have

$$\|\pi_V(y)\| = \mathrm{d}_\angle(\mathbb{R}y, V^\perp \cap (V + V')) \|y\| = \mathrm{d}_\angle(\mathbb{R}gy, V'^\perp \cap (V + V')) \|y\|.$$

We have $gy \in V$ and since $W = V' + V'^{\perp} \cap W \subset V' + V'^{\perp} \cap V^{\perp}$,

$$V'^{\perp} \cap (V + V') = (V' + V'^{\perp} \cap V^{\perp})^{\perp} \subset W^{\perp}.$$

Hence $\mathrm{d}_{\measuredangle}(\mathbb{R}gy, V'^{\perp} \cap (V + V')) \geq \mathrm{d}_{\measuredangle}(V, W^{\perp})$. This proves the first inequality in (4.13). □

Lemma 4.13. *Let V, W be linear subspaces of $\mathbb{R}^n$ with $\dim V \leq \dim W$ and $\mathrm{d}_{\measuredangle}(V, W^{\perp}) > 0$. Write $V' = \pi_W(V)$. For any bounded subset $A \subset W$,*

$$\mathcal{N}_{\delta}(\pi_{V'}(A)) \ll_n \mathrm{d}_{\measuredangle}(V, W^{\perp})^{-n} \mathcal{N}_{\delta}(\pi_V(A)). \tag{4.14}$$

In particular, if moreover $\dim V = \dim W$, then for any bounded subset $A \subset W$,

$$\mathcal{N}_{\delta}(A) \ll_n \mathrm{d}_{\measuredangle}(V, W^{\perp})^{-n} \mathcal{N}_{\delta}(\pi_V(A)). \tag{4.15}$$

Proof. If $\mathrm{d}_{\measuredangle}(V, W^{\perp}) > 0$ then π_V restricted to W is surjective. Hence we can cover $\pi_V(A)$ by the balls $\pi_V(b)^{(\delta)}$, $b \in \tilde{A} \subset W$ with $|\tilde{A}| = \mathcal{N}_{\delta}(\pi_V(A))$. Then $\pi_{V'}(A)$ is covered by the balls $\pi_{V'}(b)^{(\delta')}, b \in \tilde{A}$ with $\delta' = \mathrm{d}_{\measuredangle}(V, W^{\perp})^{-1}\delta$. Indeed, $\forall a \in A$, there is $b \in \tilde{A}$ such that $\|\pi_V(a - b)\| \leq \delta$. Hence, by (4.13), $\|\pi_{V'}(a - b)\| \leq \delta'$. We then conclude by using (2.3). □

Lemma 4.14. *Let V, W, U be linear subspaces of $\mathbb{R}^n$, with $U \subset W$. We have*

$$\mathrm{d}_{\measuredangle}(V, U + W^{\perp}) = \mathrm{d}_{\measuredangle}(V, W^{\perp})\, \mathrm{d}_{\measuredangle}(\pi_W(V), U). \tag{4.16}$$

Proof. Both sides of (4.16) vanish if the dimension of $V' = \pi_W(V)$ is smaller than V. So we can assume that $\dim V' = \dim V = r$. Let $(v_1, \ldots, v_r)$ be an orthonormal basis of V. Then $(\pi_W(v_1), \ldots, \pi_W(v_r))$ is a basis of V'. Moreover, by (4.4), we have

$$\|\pi_W(v_1) \wedge \cdots \wedge \pi_W(v_r)\| = \mathrm{d}_{\measuredangle}(V, W^{\perp})$$

and

$$\|\pi_W(v_1) \wedge \cdots \wedge \pi_W(v_r) \wedge \mathbf{u}\| = \mathrm{d}_{\measuredangle}(V, U, W^{\perp}),$$

where $\mathbf{u}$ is the wedge product an orthonormal basis of U. The desired equality (4.16) follows from the fact

$$\mathrm{d}_{\measuredangle}(V', U) = \frac{\|\pi_W(v_1) \wedge \cdots \wedge \pi_W(v_r) \wedge \mathbf{u}\|}{\|\pi_W(v_1) \wedge \cdots \wedge \pi_W(v_r)\|}$$

and Lemma 4.9 applied to $V, U, W^{\perp}$:

$$\mathrm{d}_{\measuredangle}(V, U, W^{\perp}) = \mathrm{d}_{\measuredangle}(U, W^{\perp})\, \mathrm{d}_{\measuredangle}(V, U + W^{\perp}) = \mathrm{d}_{\measuredangle}(V, U + W^{\perp}).$$ □

4.1.2 Intersections

Here we collect two useful lemmata about intersections and unions of intersections.

The first one is about intersections of large subsets. Let A be a Borel set in $\mathbb{R}^n$. Let Θ be an index set equipped with a probability measure μ and for each $\theta \in \Theta$, we have a Borel subset A_θ of A. We need an appropriate measurability, namely, the map $(x, \theta) \mapsto \mathbf{1}_{A_\theta}(x)$ is required to be measurable.

Lemma 4.15. *In the situation described above, if there is $K \geq 1$ such that $\forall \theta \in \Theta$, $\lambda(A_\theta) \geq \lambda(A)/K$, then for any positive integer $q > 0$,*

$$\mu^{\otimes q}(\{(\theta_1, \ldots, \theta_q) \mid \lambda(A_{\theta_1} \cap \cdots \cap A_{\theta_q}) \geq \frac{\lambda(A)}{2K^q}\}) \geq \frac{1}{2K^q}.$$

Proof. By Fubini's theorem and then Jensen's inequality,

$$\begin{aligned}
&\int \lambda(A_{\theta_1} \cap \cdots \cap A_{\theta_q}) \, \mathrm{d}\mu^{\otimes q}(\theta_1, \ldots, \theta_q) \\
=&\int_A \int \mathbf{1}_{A_{\theta_1}}(x) \cdots \mathbf{1}_{A_{\theta_q}}(x) \, \mathrm{d}\mu^{\otimes q}(\theta_1, \ldots, \theta_q) \, \mathrm{d}\lambda(x) \\
=&\lambda(A) \int_A \big(\int \mathbf{1}_{A_\theta}(x) \, \mathrm{d}\mu(\theta) \big)^q \frac{\mathrm{d}\lambda(x)}{\lambda(A)} \\
\geq&\lambda(A) \big(\int_A \int \mathbf{1}_{A_\theta}(x) \, \mathrm{d}\mu(\theta) \frac{\mathrm{d}\lambda(x)}{\lambda(A)} \big)^q \\
=&\lambda(A) \big(\int \frac{\lambda(A_\theta)}{\lambda(A)} \, \mathrm{d}\mu(\theta) \big)^q \\
\geq&\frac{\lambda(A)}{K^q}
\end{aligned}$$

The lemma follows. □

The next lemma is about small probability events happening simultaneously. Let (E, μ) be a probability space. Suppose we have a collection of subsets $(E_i)_{i \in \{1, \ldots, N\}}$ of E. We will think E_i as events with small probability and we want to estimate the probability such that a lot of them happen together. Here "a lot" is relatively to weights we give to the events. Let $(a_i)_{i \in \{1, \ldots, N\}}$ be non-negative real numbers such that $\sum_{i=1}^N a_i = 1$. For $I \subset \{1, \ldots, N\}$, write $a_I = \sum_{i \in I} a_i$. The following lemma is an easy consequence of Markov's inequality.

Lemma 4.16. *With the notations above, we have, for any $a > 0$,*

$$\mu\big(\bigcup_{I | a_I \geq a} \big(\bigcap_{i \in I} E_i \big) \big) \leq a^{-1} \max_{i \in \{1, \ldots, N\}} \mu(E_i).$$

Proof. Consider the Bernoulli random variables $X_i = \mathbf{1}_{E_i}$ for $i = 1, \ldots, N$ so that $\mu(E_i) = \mathbb{E}\big[X_i\big]$ and

$$\mu\big(\bigcup_{I | a_I \geq a} \bigcap_{i \in I} E_i \big) = \mathbb{P}\big[\sum_{i=1}^N a_i X_i \geq a \big].$$

Then it follows from Markov's inequality that

$$\mathbb{P}\big[\sum_{i=1}^N a_i X_i \geq a \big] \leq a^{-1} \mathbb{E}\big[\sum_{i=1}^N a_i X_i \big] \leq a^{-1} \max_{i \in \{1, \ldots, N\}} \mathbb{E}\big[X_i\big].$$

This finishes the proof. □

4.2 Technical lemmata

In this section, we show the deduction of Theorem 4.1 from Theorem 4.4 and collect several other lemmata which are needed in the next section. Since they are mostly about technical details, it is advisable to skip their proofs for a first reading. In this section, implied constants in Landau notations $O(f)$ and Vinogradov notations $f \ll g$ may depend on the dimension n and the parameter κ. Every statement is true only for $\delta > 0$ sufficiently small and by sufficiently small we mean smaller that a constant depending on all other parameters (e.g. n, m, α, κ and ϵ) but not on A nor on μ. Typically, if $C = O(1)$ then $C \leq \delta^{-\epsilon}$.

4.2.1 Proof of Theorem 4.1 admitting Theorem 4.4

We deduce Theorem 4.1 from Theorem 4.4.

Proposition 4.17. *Assume that $0 < m < n$, $0 < \alpha < n$, $\kappa > 0$ and $\epsilon > 0$ are parameters that make Theorem 4.4 true. Let A be a subset of $\mathbb{R}^n$ contained in the unit ball. Let μ be a probability measure on $\mathrm{Gr}(\mathbb{R}^n, m)$. Assume that μ satisfies* (4.3) *and A satisfies* (4.2) *and*

$$\mathcal{N}_\delta(A) \geq \delta^{-\alpha+\frac{\epsilon}{2}}. \tag{4.1$'$}$$

Then

$$\mu\big(\mathcal{E}(A, \frac{\epsilon}{3})\big) \leq \delta^{\frac{\epsilon}{2}}. \tag{4.17}$$

The idea is the following. A first application of Theorem 4.4 gives a subset $A' \subset A$ with $\mu(\mathcal{E}(A', \epsilon)) \leq \delta^\epsilon$. Either A' is large enough in which case we are done or we can cut A' out of A and apply Theorem 4.4 again. This will give us another subset A'. Then we iterate until the union of these A's is large enough.

Proof. Let $N \geq 1$ be positive integer. Suppose we have already constructed $A_1, \ldots, A_N$ such that $A_i^{(\delta)}$ are pairwise disjoint and $\mu(\mathcal{E}(A_i, \epsilon)) \leq \delta^\epsilon$ for every $i = 1, \ldots, N$. Either we have

$$\mathcal{N}_\delta\big(A \setminus \bigcup_{i=1}^N A_i^{(2\delta)}\big) \leq \delta^{-\alpha+\epsilon}, \tag{4.18}$$

in which case we stop, or the set $A \setminus \bigcup_{i=1}^N A_i^{(2\delta)}$ satisfies both (4.1) and (4.2). In the latter case Theorem 4.4 gives us $A_{N+1} \subset A \setminus \bigcup_{i=1}^N A_i^{(2\delta)}$ with $\mu(\mathcal{E}(A_{N+1}, \epsilon)) \leq \delta^\epsilon$. By construction, $A_{N+1}^{(\delta)}$ is disjoint with any of $A_i^{(\delta)}$, $i = 1, \ldots, N$.

When this procedure ends write $A_0 = \bigcup_{i=1}^N A_i$. Then (4.18) and (4.1$'$) implies $\mathcal{N}_\delta(A \setminus A_0^{(2\delta)}) \leq \delta^{\frac{\epsilon}{2}} \mathcal{N}_\delta(A)$. Moreover, by the disjointness of $A_1^{(\delta)}, \ldots, A_N^{(\delta)}$, we have

$$\mathcal{N}_\delta(A_0) = \sum_{i=1}^N \mathcal{N}_\delta(A_i).$$

Set $a_i = \frac{\mathcal{N}_\delta(A_i)}{\mathcal{N}_\delta(A_0)}$. We claim that

$$\mathcal{E}(A, \frac{\epsilon}{3}) \subset \bigcup_I \bigcap_{i \in I} \mathcal{E}(A_i, \epsilon),$$

where the index set I runs over subsets of $\{1,\dots,N\}$ with $\sum_{i\in I} a_i \geq \delta^{\frac{\epsilon}{2}}$. The desired upper bound (4.17) then follows immediately from Lemma 4.16.

We now proceed to show the claim. Let $V \in \mathcal{E}(A, \frac{\epsilon}{3})$. By definition, there exists $A' \subset A$ with $\mathcal{N}_\delta(A') \geq \delta^{\frac{\epsilon}{3}}\mathcal{N}_\delta(A)$ and $\mathcal{N}_\delta(\pi_V(A')) \leq \delta^{-\frac{m}{n}\alpha-\frac{\epsilon}{3}}$. Consider the index set I defined as

$$I = \{i \in \{1,\dots,N\} \mid \mathcal{N}_\delta(A'^{(2\delta)} \cap A_i) \geq \delta^\epsilon \mathcal{N}_\delta(A_i)\}.$$

We have, by (2.1) and (2.5),

$$\begin{aligned}\mathcal{N}_\delta(A') - \mathcal{N}_\delta(A \setminus A_0^{(2\delta)}) &\leq \sum_{i=1}^{n} \mathcal{N}_\delta(A' \cap A_i^{(2\delta)}) \\ &\ll \sum_{i\in I} \mathcal{N}_\delta(A_i) + \sum_{i\notin I} \mathcal{N}_\delta(A'^{(2\delta)} \cap A_i) \\ &\ll \sum_{i\in I} a_i \mathcal{N}_\delta(A) + \delta^\epsilon \mathcal{N}_\delta(A)\end{aligned}$$

Hence $\sum_{i\in I} a_i \geq \delta^{\frac{\epsilon}{2}}$. On the other hand, for all $i \in I$, since

$$\mathcal{N}_\delta(\pi_V(A'^{(2\delta)} \cap A_i)) \leq \mathcal{N}_\delta(\pi_V(A')^{(2\delta)}) \ll \mathcal{N}_\delta(\pi_V(A')),$$

we have

$$\mathcal{N}_\delta(\pi_V(A'^{(2\delta)} \cap A_i)) \leq \delta^{-\frac{m}{n}\alpha-\epsilon}.$$

Hence $V \in \mathcal{E}(A_i, \epsilon)$ for all $i \in I$. This finishes the proof of the claim. □

4.2.2 Action of linear transformations

Clearly, all the assumptions and the conclusion of Theorem 4.4 are invariant under the action of the orthogonal group $\mathrm{O}(n)$. The next proposition states that the action of a $\delta^{-\epsilon}$-bi-Lipschitz linear transformation only affects them by a factor of $\delta^{O(\epsilon)}$. Here, while $f \in \mathrm{GL}(\mathbb{R}^n)$ acts on $\mathbb{R}^n$ in the usual way, it acts on the Grassmannian by multiplication by $f^\perp := (f^{-1})^*$ or equivalently, $f^\perp V = (fV^\perp)^\perp$ for all $V \in \mathrm{Gr}(\mathbb{R}^n, m)$.

Lemma 4.18. *Let $0 < m < n$ be dimensions. Let $\epsilon > 0$. Let $f \in \mathrm{GL}(\mathbb{R}^n)$ with $\|f\| + \|f^{-1}\| \leq \delta^{-\epsilon}$. Let A be a bounded subset of $\mathbb{R}^n$ and μ a probability measure on $\mathrm{Gr}(\mathbb{R}^n, m)$.*

(i) For each of the assumptions (4.1)–(4.3) of Theorem 4.4, if it holds for A and μ with the parameters α, κ and ϵ then it also holds for the image set fA and the image measure $f_*^\perp\mu$ with the parameters α, κ and $O(\epsilon)$ in the place of ϵ.

(ii) For all $V \in \mathrm{Gr}(\mathbb{R}^n, m)$, $\mathcal{N}_\delta(\pi_{f^\perp V}(fA)) \leq \delta^{-O(\epsilon)}\mathcal{N}_\delta(\pi_V(A))$.

(iii) We have $\mu(\mathcal{E}(A,\epsilon)) \leq (f_*^\perp\mu)\big(\mathcal{E}(fA, O(\epsilon))\big)$. In particular, if the conclusion of Theorem 4.4 holds for fA and $f_*^\perp\mu$ with some $\epsilon' > 0$ in the place of ϵ then it holds for A and μ with $\epsilon = \frac{\epsilon'}{O(1)}$.

(iv) For all $V \in \mathrm{Gr}(\mathbb{R}^n, m)$ and all $x \in f^\perp V$,

$$\mathcal{N}_\delta\big(fA \cap \pi_{f^\perp V}^{-1}(x^{(\delta)})\big) \leq \delta^{-O(\epsilon)} \max_{y\in V} \mathcal{N}_\delta\big(A \cap \pi_V^{-1}(y^{(\delta)})\big).$$

Proof. The statement about the assumptions (4.1) and (4.2) follows immediately from the inequality (2.4). As for the assumption (4.3), it suffices to prove that for all $W \in \mathrm{Gr}(\mathbb{R}^n, n-m)$ and all $\rho \geq \delta$,

$$f_*^{\perp}\mu\big(\mathcal{V}_{\measuredangle}(f^{\perp}W, \rho)\big) \leq \mu\big(\mathcal{V}_{\measuredangle}(W, \delta^{-O(\epsilon)}\rho)\big). \tag{4.19}$$

From the Cartan decomposition of f, we see easily that $\forall r = 1, \dots, n$, $\|\bigwedge^r f^{\perp}\| + \|(\bigwedge^r f^{\perp})^{-1}\| \leq \delta^{-O(\epsilon)}$. For $V \in \mathrm{Gr}(\mathbb{R}^n, m)$, let $\mathbf{v}$ be the wedge product of an orthonormal basis of V and $\mathbf{w}$ that of W. We have

$$\begin{aligned}
\mathrm{d}_{\measuredangle}(f^{\perp}V, f^{\perp}W) &= \frac{\|(\bigwedge^n f^{\perp})(\mathbf{v} \wedge \mathbf{w})\|}{\|(\bigwedge^m f^{\perp})\mathbf{v}\| \, \|(\bigwedge^{n-m} f^{\perp})\mathbf{w}\|} \\
&\geq \frac{\|(\bigwedge^n f^{\perp})^{-1}\|^{-1}\|\mathbf{v} \wedge \mathbf{w}\|}{\|\bigwedge^m f^{\perp}\| \, \|\bigwedge^{n-m} f^{\perp}\|} \\
&\geq \delta^{O(\epsilon)} \, \mathrm{d}_{\measuredangle}(V, W).
\end{aligned}$$

Hence $f^{\perp}V \in \mathcal{V}_{\measuredangle}(f^{\perp}W, \rho)$ implies $V \in \mathcal{V}_{\measuredangle}(W, \delta^{-O(\epsilon)}\rho)$, which establishes (4.19).

For the second statement, observe that there is a finite set $\tilde{A}$ of cardinality $|\tilde{A}| = \mathcal{N}_{\delta}(\pi_V(A))$ such that

$$A \subset \tilde{A} + V^{\perp} + \mathbf{B}(0, \delta).$$

Applying f and then $\pi_{f^{\perp}V}$ on both sides, we obtain

$$\pi_{f^{\perp}V}(fA) \subset \pi_{f^{\perp}V}(f\tilde{A}) + \mathbf{B}(0, \delta^{1-\epsilon}).$$

This proves that $\mathcal{N}_{\delta^{1-\epsilon}}(\pi_{f^{\perp}V}(fA)) \leq \mathcal{N}_{\delta}(\pi_V(A))$. We conclude by the scale change estimate (2.3).

For the next statement, it suffices to prove that $f^{\perp}V \in \mathcal{E}(fA, O(\epsilon))$ whenever $V \in \mathcal{E}(A, \epsilon)$. Indeed, let $V \in \mathcal{E}(A, \epsilon)$. Then there exists $A' \subset A$ such that $\mathcal{N}_{\delta}(A') \geq \delta^{\epsilon}\mathcal{N}_{\delta}(A)$ and $\mathcal{N}_{\delta}(\pi_V(A')) < \delta^{-\frac{m}{n}\alpha-\epsilon}$. On the one hand, by (2.4), we have $\mathcal{N}_{\delta}(fA') \geq \delta^{O(\epsilon)}\mathcal{N}_{\delta}(fA)$. On the other hand, from (ii) it follows that $\mathcal{N}_{\delta}(\pi_{f^{\perp}V}(fA')) \leq \delta^{-\frac{m}{n}\alpha-O(\epsilon)}$. Hence $f^{\perp}V \in \mathcal{E}(fA, O(\epsilon))$.

For the last statement, it suffices to prove for any $x \in f^{\perp}V$, there exists $y \in V$ such that

$$fA \cap \pi_{f^{\perp}V}^{-1}(x^{(\delta)}) \subset f\big(A \cap \pi_V^{-1}(y^{(\delta^{1-\epsilon})})\big). \tag{4.20}$$

Indeed, if $a \in A$ satisfies $\pi_{f^{\perp}V}(f(a)) \in x^{(\delta)}$, then

$$f(a) \in x + fV^{\perp} + \mathbf{B}(0, \delta).$$

Applying f^{-1} and then π_V on both sides, we obtain

$$\pi_V(a) \in \pi_V(f^{-1}(x)) + \mathbf{B}(0, \delta^{1-\epsilon}).$$

This proves (4.20) with $y = \pi_V(f^{-1}(x))$. □

4.2.3 Non-concentration property for projections

Let A a subset of $\mathbb{R}^n$ as in Theorem 4.4. We want to understand whether a projection of A still satisfies some similar regularity property as A does. More precisely we want a large subset A' of A such that

$$\forall \rho \geq \delta,\ \forall x \in V, \quad \mathcal{N}_\delta(\pi_V(A') \cap x^{(\rho)}) \leq \rho^{\kappa_1} \delta^{-\frac{m}{n}\alpha - \epsilon'},$$

for some κ_1 proportional to κ and ϵ' some multiple of ϵ.

In the special case where m divides n, we have the following result. We will only need this non-concentration result in this special case, although a more general result might be true.

Lemma 4.19. *Let $n = qm$ with $q \geq 2$. For any parameters $0 < \alpha < n$, $\kappa > 0$ and $\epsilon > 0$, the following is true for $\delta > 0$ sufficiently small. If A is a subset of $\mathbb{R}^n$ contained in the unit ball and μ is a probability measure on $\mathrm{Gr}(\mathbb{R}^n, m)$ satisfying the assumptions (4.1)–(4.3) for the parameters α, κ and ϵ, then*

$$\mu(\mathcal{E}(A) \setminus \mathcal{E}_{\mathrm{reg}}(A)) \leq \delta^{2\epsilon},$$

where $\mathcal{E}_{\mathrm{reg}}(A)$ denotes the set of all $V \in \mathcal{E}(A)$ such that $\exists A' \subset A$ with $\mathcal{N}_\delta(A') \geq \delta^{-\alpha+3\epsilon}$ and $\mathcal{N}_\delta(\pi_V(A')) \leq \delta^{-\frac{m}{n}\alpha-\epsilon}$ and

$$\forall \rho \geq \delta,\ \forall x \in V, \quad \mathcal{N}_\delta(\pi_V(A') \cap x^{(\rho)}) \leq \rho^{\frac{\kappa}{2q^2}} \delta^{-\frac{m}{n}\alpha - 11\epsilon}. \tag{4.21}$$

The idea of the proof is the following. When $V \in \mathcal{E}(A)$, there is a large subset A' with small projection to V. We then remove small fibers of the projection $\pi_V : A' \to V$ to get A''. Any large subset in of $\pi_V(A'')$ will have large preimage by π_V. Thus if $V \notin \mathcal{E}_{\mathrm{reg}}(A)$ then there will be a cylinder with axis $V^\perp$ and radius ρ in which A is very dense. If there are a lot of such V we can then intersect these cylinders to get a ball of radius $\rho^{\frac{1}{q}}$ which will contradict the assumption (4.2).

Proof. For conciseness, write $\kappa_1 = \frac{\kappa}{2q^2}$. We claim that if $V \in \mathcal{E}(A) \setminus \mathcal{E}_{\mathrm{reg}}(A)$ then there exists $x \in V$ and $\rho \geq \delta$ such that

$$\mathcal{N}_\delta\big(A \cap \pi_V^{-1}(x^{(\rho)})\big) \geq \rho^{\kappa_1} \delta^{-\alpha - 6\epsilon}. \tag{4.22}$$

Indeed, let $V \in \mathcal{E}(A) \setminus \mathcal{E}_{\mathrm{reg}}(A)$, then there exists $A' \subset A$ with $\mathcal{N}_\delta(A') \geq \delta^{-\alpha+2\epsilon}$ and $\mathcal{N}_\delta(\pi_V(A')) \leq \delta^{-\frac{\alpha}{q}-\epsilon}$. Now we want to remove small fibers of the map π_V restricted to A'. Consider the set

$$B = \Big\{ y \in V \mid \mathcal{N}_\delta\big(A' \cap \pi_V^{-1}(y^{(\delta)})\big) \geq \delta^{-\frac{q-1}{q}\alpha + 4\epsilon} \Big\}$$

and $A'' = A' \cap \pi_V^{-1}(B^{(\delta)})$. We have, for all $y \in V$,

$$\mathcal{N}_\delta\big((A' \setminus A'') \cap \pi_V^{-1}(y^{(\delta)})\big) \leq \delta^{-\frac{q-1}{q}\alpha + 4\epsilon}.$$

for otherwise y would belong to B and the intersection $(A' \setminus A'') \cap \pi_V^{-1}(y^{(\delta)})$ would be empty. Consequently,

$$\mathcal{N}_\delta(A' \setminus A'') \leq \mathcal{N}_\delta(\pi_V(A')) \max_{y \in V} \mathcal{N}_\delta\big((A' \setminus A'') \cap \pi_V^{-1}(y^{(\delta)})\big) \leq \delta^{-\alpha+3\epsilon}.$$

It follows that $\mathcal{N}_\delta(A'') \geq \delta^{-\alpha+3\epsilon}$. Since $V \notin \mathcal{E}_{\mathrm{reg}}(A)$, the non-concentration property (4.21) fails for $\pi_V(A'')$. Hence there exists $x \in V$ and $\rho \geq \delta$ such that

$$\mathcal{N}_\delta\big(\pi_V(A'') \cap x^{(\rho)}\big) \geq \rho^{\kappa_1}\delta^{-\frac{\alpha}{q}-11\epsilon}. \tag{4.23}$$

Let $\tilde{B}$ be a maximal 6δ-separated subset of $\pi_V(A'')\cap x^{(\rho)}$. From (2.1) and (4.23), we have $|\tilde{B}| \gg \rho^{\kappa_1}\delta^{-\frac{\alpha}{q}-11\epsilon}$. Moreover for all $y \in \tilde{B}$, by the definition of A'', $y \in B^{(\delta)}$, hence $\mathcal{N}_\delta\big(A' \cap \pi_V^{-1}(y^{(2\delta)})\big) \geq \delta^{-\frac{q-1}{q}\alpha+4\epsilon}$. Since $\tilde{B}$ is 6δ-separated, all these balls $y^{(2\delta)}$ with center $y \in \tilde{B}$ are 2δ-away from each other. Consequently,

$$\mathcal{N}_\delta\big(A' \cap \pi_V^{-1}(x^{(\rho+2\delta)})\big) \geq \sum_{y\in\tilde{B}} \mathcal{N}_\delta\big(A' \cap \pi_V^{-1}(y^{(2\delta)})\big) \geq \rho^{\kappa_1}\delta^{-\alpha-6\epsilon}.$$

This finishes the proof of the claim.

To obtain a contradiction, suppose that $\mu(\mathcal{E}(A) \setminus \mathcal{E}_{\mathrm{reg}}(A)) \geq \delta^{2\epsilon}$. Note that the radius ρ in the claim depends on V. However, from (4.2) and (4.22) we know that it ranges from δ to $\delta^{5\epsilon}$. For the argument below, we want (4.22) to hold for a lot of $V \in \mathcal{E}(A) \setminus \mathcal{E}_{\mathrm{reg}}(A)$ with a same $\rho \geq \delta$. Indeed, by a simple pigeonhole argument[2], we can find a subset $\mathcal{D} \subset \mathcal{E}(A) \setminus \mathcal{E}_{\mathrm{reg}}(A)$ and a radius $\rho \geq \delta$ such that $\mu(\mathcal{D}) \geq \delta^{3\epsilon}$ and for all $V \in \mathcal{D}$, there exists $x \in V$ such that

$$\mathcal{N}_\delta(A \cap \pi_V^{-1}(x^{(\rho)})) \geq \rho^{2\kappa_1}\delta^{-5\epsilon}\mathcal{N}_\delta(A)$$

and hence, by Lemma 2.1,

$$\lambda(A^{(\delta)} \cap \pi_V^{-1}(x^{(\rho)})) \geq \rho^{2\kappa_1}\delta^{-4\epsilon}\lambda(A^{(\delta)}).$$

Let $V_1, \ldots, V_q$ be random elements of $\mathrm{Gr}(\mathbb{R}^n, m)$ independently distributed according to μ. On the one hand, from Lemma 4.15 applied to the restriction of μ to $\mathcal{D}$, it follows that with probability at least $\frac{1}{2}(\rho^{2\kappa_1}\delta^{-4\epsilon})^q\mu(\mathcal{D})^q \geq \frac{1}{2}\rho^{2q\kappa_1}\delta^{-q\epsilon}$, there exists $x_1 \in V_1, \ldots, x_q \in V_q$ such that

$$\lambda\big(A^{(\delta)} \cap \pi_{V_1}^{-1}(x_1^{(\rho)}) \cap \cdots \cap \pi_{V_q}^{-1}(x_q^{(\rho)})\big) \geq \frac{1}{2}\rho^{2q\kappa_1}\delta^{-4q\epsilon}\lambda(A^{(\delta)}) \tag{4.24}$$

On the other hand, from (4.9) and (4.3), it follows that with probability at least $1-(q-1)\delta^{-\epsilon}\rho^{\frac{\kappa}{q}}$, we have

$$\mathrm{d}_{\measuredangle}(V_1, \ldots, V_q) \geq \rho^{\frac{q-1}{q}}. \tag{4.25}$$

Now with our choice of κ_1, we have $1-(q-1)\delta^{-\epsilon}\rho^{\frac{\kappa}{q}} + \frac{1}{2}\rho^{2q\kappa_1}\delta^{-q\epsilon} > 1$. This means that for some $(V_1, \ldots, V_q)$, both (4.24) and (4.25) hold. By Lemma 4.10, there exists $x \in \mathbb{R}^n$ such that

$$\pi_{V_1}^{-1}(x_1^{(\rho)}) \cap \cdots \cap \pi_{V_q}^{-1}(x_q^{(\rho)}) \subset x^{(\rho')}$$

with $\rho' = q\rho\,\mathrm{d}_{\measuredangle}(V_1, \ldots, V_q)^{-1} \leq q\rho^{\frac{1}{q}}$. Then the non-concentration property (4.2) of A implies that

$$\lambda\big(A^{(\delta)} \cap x^{(\rho')}\big) \ll \delta^{-\epsilon}\rho^{\frac{\kappa}{q}}\lambda(A^{(\delta)}).$$

Combining this with (4.24) yields

$$\rho^{2q\kappa_1}\delta^{-4q\epsilon} \ll \delta^{-\epsilon}\rho^{\frac{\kappa}{q}},$$

which is impossible with our choice of κ_1. □

[2] Arrange different ρ into intervals of the form $[\delta^{2^{-k}}, \delta^{2^{-k-1}}]$, where $0 \leq k \ll -\log(\epsilon)$.

4.2.4 Non-concentration property for slices

We will also consider slices of A, i.e. intersection of A with a δ-neighborhood of a affine subspace. When $n = qm$, we have similar non-concentration results for $(n-m)$-dimensional slices of A.

Lemma 4.20. *Let $n = qm$ with $q \geq 2$ a positive integer. Let $0 < \alpha < n$, $\kappa > 0$ and $\epsilon > 0$ be parameters. If the statement in Theorem 4.4 fails for the set A, then there is a $(n-m)$-dimensional affine subspace $y + W$ and a subset $B \subset A^{(\delta)} \cap (y + W)$ such that*

$$\mathcal{N}_\delta(B) \geq \delta^{-\beta+O(\epsilon)} \text{ and}$$

$$\forall \rho \geq \delta,\ \forall x \in W \quad \mathcal{N}_\delta(B \cap x^{(\rho)}) \leq \rho^{\frac{\kappa}{2q^2}} \delta^{-\beta-O(\epsilon)}, \tag{4.26}$$

where $\beta = \frac{q-1}{q}\alpha$.

Here is an outline of the proof. The negation of Theorem 4.4 to A will imply that a large subset of A is a large subset of a Cartesian product (of q factors). Then, because of Lemma 4.19, the first factor can be chosen to have the non-concentration property. This in turn will imply the non-concentration property of the Cartesian product of the $q-1$ first factors. Now the negation of the projection theorem will give a large slice parallel to the Cartesian product of the $q-1$ first factors. The slice is nearly as big as the Cartesian product. From this we conclude that it has also the non-concentration property.

Proof. Suppose the statement in Theorem 4.4 fails for the set $A \subset \mathbb{R}^n$. This means that for any subset $A' \subset A$, $\mu(\mathcal{E}(A')) > \delta^\epsilon$. In particular, $\mathcal{E}_{\text{reg}}(A)$ is non-empty by Lemma 4.19. Let $V_1 \in \mathcal{E}_{\text{reg}}(A)$. There exists $A_1 \subset A$ with $\mathcal{N}_\delta(A_1) \geq \delta^{-\alpha+3\epsilon}$ and

$$\forall \rho \geq \delta,\ \forall x \in V_1, \quad \mathcal{N}_\delta(\pi_{V_1}(A_1) \cap x^{(\rho)}) \leq \rho^{\frac{\kappa}{2q^2}} \delta^{-\frac{1}{q}\alpha-11\epsilon}. \tag{4.27}$$

Let $\epsilon_1 = \frac{3\epsilon}{\kappa}$. We construct by a simple induction a sequence of subspaces $V_2, \dots, V_q$ and a nested sequence of subsets $A_1 \supset \cdots \supset A_q$ satisfying for any $j = 2, \dots, q$,

$$\begin{aligned} \mathrm{d}_\measuredangle(V_j, V_1 + \cdots + V_{j-1}) &\geq \delta^{\epsilon_1}, \\ \mathcal{N}_\delta(A_j) &\geq \delta^\epsilon \mathcal{N}_\delta(A_{j-1}), \end{aligned} \tag{4.28}$$

$$\mathcal{N}_\delta(\pi_{V_j}(A_j)) \leq \delta^{-\frac{1}{q}\alpha-\epsilon}. \tag{4.29}$$

This is possible since at each step, we have by (4.3),

$$\mu\big(\mathcal{E}(A_{j-1}) \setminus \mathcal{V}_\measuredangle(V_1 + \cdots + V_{j-1}, \delta^{\epsilon_1})\big) \geq \delta^\epsilon - \delta^{2\epsilon} > 0.$$

Since $\mathcal{N}_\delta(A_q) \leq \mathcal{N}_\delta(\pi_{V_q}(A_q)) \max_{y \in V_q} \mathcal{N}_\delta(A_q \cap \pi_{V_q}^{-1}(y^{(\delta)}))$, there exists $y_\star \in V_q$ such that

$$\mathcal{N}_\delta(A_q \cap \pi_{V_q}^{-1}(y_\star^{(\delta)})) \geq \delta^{-\frac{q-1}{q}\alpha+O(\epsilon)}. \tag{4.30}$$

After a translation, we can suppose $y_\star = 0$. We write $V = V_1 + \cdots + V_{q-1}$ and $W = V_q^\perp$ and set $B_0 = A_q \cap W^{(\delta)}$ and $B = \pi_W(B_0)$. We have $\mathcal{N}_\delta(B) \geq \delta^{-\beta+O(\epsilon)}$ by (4.30) and the fact that $B_0 \subset B^{(\delta)}$.

It remains to show the non-concentration property (4.26) for B. Let $\rho \geq \delta$ and $x \in W$. From (4.28), $\mathrm{d}_{\measuredangle}(V, W^{\perp}) = \mathrm{d}_{\measuredangle}(V, V_q) \geq \delta^{O(\epsilon)}$. Hence, by (4.15) in Lemma 4.13,

$$\mathcal{N}_\delta(B \cap x^{(\rho)}) \leq \delta^{-O(\epsilon)} \mathcal{N}_\delta(\pi_V(B) \cap x_0^{(\rho)})$$

where $x_0 = \pi_V(x)$. Moreover $B \subset A_q^{(\delta)}$, hence

$$\mathcal{N}_\delta(B \cap x^{(\rho)}) \leq \delta^{-O(\epsilon)} \mathcal{N}_\delta(\pi_V(A_q) \cap x_0^{(2\rho)}). \tag{4.31}$$

Then Lemma 4.11 applied to the set $\pi_V(A_q) \cap x_0^{(2\rho)}$ in $V = \bigoplus_{j=1}^q V_j$ together with (4.28) yield

$$\mathcal{N}_\delta(\pi_V(A_q) \cap x_0^{(2\rho)}) \leq \delta^{-O(\epsilon)} \mathcal{N}_\delta(\pi_{V_1}(A_q) \cap x_1^{(2\rho)}) \prod_{j=2}^{q-1} \mathcal{N}_\delta(\pi_{V_j}(A_q))$$

where $x_1 = \pi_{V_1}(x_0)$. The required non-concentration property (4.26) then follows from (4.27) and (4.29). □

4.2.5 Without the non-concentration property

As illustrated by the example in the introduction, the non-concentration condition (4.2) on A is crucial to have a gain $\epsilon > 0$ in the conclusion. Without this condition, we still expect $\mathcal{N}_\delta(\pi_V(A))$ to be close to $\mathcal{N}_\delta(A)^{\frac{m}{n}}$ for generic $V \in \mathrm{Gr}(\mathbb{R}^n, m)$. This is the subject of the next proposition.

Proposition 4.21. *Given $0 < m \leq n$, $0 < \alpha < n$ and $\kappa > 0$, there exists $C < +\infty$ such that for all $0 < \epsilon < \frac{1}{C}$, the following is true for all $\delta > 0$ sufficiently small. Let $A \subset \mathbb{R}^n$ be a subset contained in the unit ball and μ a probability measure on $\mathrm{Gr}(\mathbb{R}^n, m)$. Assume that*

$$\mathcal{N}_\delta(A) \geq \delta^{-\alpha - C\epsilon}. \tag{4.32}$$

Further assume the non-concentration property (4.3) for μ if $m < n$. Then $\mu(\mathcal{E}(A)) \leq \delta^\epsilon$.

When m divides n, this follows almost immediately from Lemma 4.11. Then the task is to reduce to this special case. Since it shares the same set of ideas as the proof of Theorem 4.4, the proof below will only be outlined and more details can be found in the next section.

Proof. For $0 < m \leq n$, denote by $\mathcal{P}(n, m)$ the statement we want to show. Note that for all $n \geq 1$, $\mathcal{P}(n, n)$ is trivially true. We will proceed by an induction similar to that in the proof of Theorem 4.4. It suffices to show the following two types of inductive steps. Let $0 < m \leq n$ and $q, r > 0$ be integers.

(i) If $mq \leq n$, then $\mathcal{P}(n, qm)$ implies $\mathcal{P}(n, m)$.

(ii) If $n = q(n - m) + r$ with $0 < r \leq n - m$, then $\mathcal{P}(n, r)$ and $\mathcal{P}(n - r, m)$ imply $\mathcal{P}(n, m)$.

Using the same argument in Proposition 4.17, we see that in order to show $\mathcal{P}(n,m)$, it suffices to show $\mu(\mathcal{E}(A')) \leq \delta^{\epsilon}$ for some subset $A' \subset A$. In other words, if the conclusion of $\mathcal{P}(n,m)$ fails for the set A then for any subset $A' \subset A$, $\mu(\mathcal{E}(A')) \geq \delta^{\epsilon}$.

Proof of (i). Let $V_1, \dots, V_q$ be random elements of $\mathrm{Gr}(\mathbb{R}^n, m)$ independently distributed according to μ. Write $V = V_1 + \cdots + V_q$. When $qm < n$, we have by Lemma 4.25,

$$\mathbb{P}\big[\dim(V) = qm\big] \geq 1 - (q-1)\delta^{\kappa-\epsilon}.$$

Moreover the distribution of V conditional to the event $\dim(V) = qm$ has the corresponding non-concentration property. By $\mathcal{P}(n, qm)$, we know that for any $C' > 0$, if the constant C in (4.32) is large enough (depending on C') then the probability that there exists $A' \subset A$ satisfying

$$\mathcal{N}_\delta(A') \geq \delta^{C'\epsilon}\mathcal{N}_\delta(A) \text{ and } \mathcal{N}_\delta(\pi_V(A')) \leq \delta^{-\frac{qm}{n}\alpha - C'\epsilon}$$

is at most $\delta^{C'\epsilon} + (q-1)\delta^{\kappa-\epsilon}$.

Suppose that $\mathcal{P}(n,m)$ fails for A. Then by a simple induction we show that with probability at least $\delta^{O(\epsilon)}$, we have

$$\mathrm{d}_{\measuredangle}(V_1, \dots, V_q) \geq \delta^{O(\epsilon)}$$

and there exists $A_q \subset A$ such that $\mathcal{N}_\delta(A_q) \geq \delta^{O(\epsilon)}\mathcal{N}_\delta(A)$ and

$$\forall j = 1, \dots, q, \quad \mathcal{N}_\delta(\pi_{V_j}(A_q)) \leq \delta^{-\frac{m}{n}\alpha - \epsilon}$$

and hence, by (4.12) applied to $\pi_V(A)$,

$$\mathcal{N}_\delta(\pi_V(A)) \leq \delta^{-\frac{qm}{n}\alpha - O(\epsilon)}.$$

We obtain a contradiction if C' is large compared to any of the implied constants in the previous Landau notations.

Proof of (ii), Case 1. Assume firstly that A contains large slice of dimension $n-r$. More precisely, assume that there exists $W \in \mathrm{Gr}(\mathbb{R}^n, n-r)$ and $x \in \mathbb{R}^n$ such that

$$\mathcal{N}_\delta\big(A \cap (x + W^{(\delta)})\big) \geq \delta^{-\frac{n-r}{n}\alpha - C'\epsilon}$$

where C' is the constant given by $\mathcal{P}(n-r, m)$ applied to $0 < m \leq n-r$, $\frac{n-r}{n}\alpha$ and κ. Without loss of generality, we can assume that $x = 0$ and that $B = \pi_W(A \cap W^{(\delta)})$ is contained in A. Lemma 4.23 tells us that we can apply $\mathcal{P}(n-r, m)$ to $B \subset W$ with the image measure of μ by π_W. Then we can conclude using Lemma 4.24.

Proof of (ii), Case 2. Otherwise A does not contain any large slice of dimension $n-r$:

$$\forall x \in \mathbb{R}^n,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, n-r), \quad \mathcal{N}_\delta\big(A \cap (x + W^{(\delta)})\big) \leq \delta^{-\frac{n-r}{n}\alpha - O(\epsilon)}. \tag{4.33}$$

Let $V_1, \dots, V_q$ be random elements of $\mathrm{Gr}(\mathbb{R}^n, m)$ independently distributed according to μ. Write $V = V_1 \cap \cdots \cap V_q$. By (4.7) and Lemma 4.25,

$$\mathbb{P}\big[\dim(V) = r\big] \geq 1 - (q-1)\delta^{\kappa-\epsilon}$$

and that the distribution of V conditional to the event $\dim(V) = r$ has a non-concentration property. By $\mathcal{P}(n,r)$, we know that for any $C' > 0$, if the constant

C in (4.32) is large enough (depending on C') then the probability that there exists $A' \subset A$ satisfying

$$\mathcal{N}_\delta(A') \geq \delta^{C'\epsilon}\mathcal{N}_\delta(A) \text{ and } \mathcal{N}_\delta(\pi_V(A')) \leq \delta^{-\frac{r}{n}\alpha - C'\epsilon}$$

is at most $\delta^{C'\epsilon} + (q-1)\delta^{\kappa-\epsilon}$.

Suppose that $\mathcal{P}(n,m)$ fails for A. Again by an induction we show that with probability at least $\delta^{O(\epsilon)}$, we have

$$\mathrm{d}_\angle(V_1^\perp, \ldots, V_q^\perp) \geq \delta^{O(\epsilon)}$$

and there exists $A_q \subset A$ such that $\mathcal{N}_\delta(A_q) \geq \delta^{O(\epsilon)}\mathcal{N}_\delta(A)$ and

$$\forall j = 1, \ldots, q, \quad \mathcal{N}_\delta(\pi_{V_j}(A_q)) \leq \delta^{-\frac{m}{n}\alpha - \epsilon}$$

Together with (4.33), this implies by Proposition 4.26 that there exists $A' \subset A_q$ such that

$$\mathcal{N}_\delta(A') \geq \delta^{O(\epsilon)}\mathcal{N}_\delta(A) \text{ and } \mathcal{N}_\delta(\pi_V(A')) \leq \delta^{-\frac{r}{n}\alpha - O(\epsilon)}.$$

Again we obtain a contradiction if C' is large compared to any of the implied constants in the previous Landau notations. □

4.3 Proof of the main result

In this section, we prove Theorem 4.4 and thus Theorem 4.1. This is done by proving first the base case where $n = 2m$ (Propsoition 4.5) and then the induction steps (Propositions 4.6-4.8). Note that by Proposition 4.17, for a given pair (n,m), if Theorem 4.4 is true for these dimensions then so is Theorem 4.1. Therefore, when we use Theorem 4.4 as induction hypothesis, the conclusion is $\mu(\mathcal{E}(A)) \geq \delta^\epsilon$ while when we prove by contradiction by saying that A is a counterexample for Theorem 4.4, we are assuming $\mu(\mathcal{E}(A')) > \delta^\epsilon$ for all subsets A' of A.

Like in the previous section, all implied constants in Landau and Vinogradov notations in this section may depend on n and κ. Again every statement in this section is true only for $\delta > 0$ smaller than a constant depending on n, m, α, κ and ϵ.

4.3.1 Half dimensional projections

For the special case $n = 2m$, we follow mainly the proof in [7] (which deals with the case $m = 1$) and use a technique in the proof of Proposition 2 in Bourgain-Glibichuk [12].

Proof of Proposition 4.5. Let $A \subset \mathbb{R}^n$ be a counterexample for Proposition 4.5. In particular,

$$\forall A' \subset A, \quad \mu(\mathcal{E}(A')) \geq \delta^\epsilon. \tag{4.34}$$

We will get a contradiction when ϵ is small enough. By Lemma 4.19, there is a subspace V_1 and a subset $A_1 \subset A$ with the following properties:

$$\mathcal{N}_\delta(A_1) \geq \delta^{-\alpha+3\epsilon};$$

$$\mathcal{N}_\delta(\pi_{V_1}(A_1)) \le \delta^{-\alpha/2-\epsilon};$$

$$\forall \rho \ge \delta,\ \forall x \in V_1, \quad \mathcal{N}_\delta(\pi_{V_1}(A_1) \cap x^{(\rho)}) \le \rho^{\frac{\kappa}{8}} \delta^{-\frac{\alpha}{2}-O(\epsilon)}. \tag{4.35}$$

Let $\epsilon_1 = \frac{3\epsilon}{\kappa}$. Then $\mu\big(\mathcal{E}(A_1) \setminus \mathcal{V}_\measuredangle(V_1, \delta^{\epsilon_1})\big) \ge \delta^\epsilon - \delta^{2\epsilon} > 0$ by (4.34) and the non-concentration property (4.3) of μ. Let $V_2 \in \mathcal{E}(A_1) \setminus \mathcal{V}_\measuredangle(V_1, \delta^{\epsilon_1})$ with A_2 such that $\mathcal{N}_\delta(A_2) \ge \delta^{-\alpha+4\epsilon}$ and

$$\mathcal{N}_\delta(\pi_{V_i}(A_2)) \le \delta^{-\frac{\alpha}{2}-\epsilon}, \qquad i = 1, 2. \tag{4.36}$$

Consider $f \in \mathrm{GL}(\mathbb{R}^n)$ which fixes $V_1^\perp$ and sends isometrically $V_2^\perp$ to V_1. Since $\mathrm{d}_\measuredangle(V_1, V_2) \ge \delta^{-O(\epsilon)}$, f is $\delta^{-O(\epsilon)}$-bi-Lipschitz. By definition, $f^\perp V_1 = V_1$ and $f^\perp V_2 = V_1^\perp$. On account of Lemma 4.18, we can suppose without loss of generality that $V_2 = V_1^\perp$.

Put $X = \pi_{V_1}(A_2)$ and $Y = \pi_{V_2}(A_2)$. We have, $\mathcal{N}_\delta(A_2) \ll \mathcal{N}_\delta(X)\mathcal{N}_\delta(Y)$ and this together with the inequalities (4.36) implies

$$\mathcal{N}_\delta(X), \mathcal{N}_\delta(Y) \ge \delta^{-\frac{\alpha}{2}+O(\epsilon)}.$$

Write $\mathcal{D} = \mathcal{E}(A_2) \setminus (\mathcal{V}_\measuredangle(V_1, \delta^{\epsilon_1}) \cup \mathcal{V}_\measuredangle(V_2, \delta^{\epsilon_1}))$. We have, by (4.34) and (4.3), $\mu(\mathcal{D}) \ge \delta^\epsilon - 2\delta^{2\epsilon} \ge \delta^{2\epsilon}$. Let $V \in \mathcal{D}$. By (4.6) and (4.7), we have

$$|\det(\pi_{V|V_1})| = \mathrm{d}_\measuredangle(V_1, V^\perp) = \mathrm{d}_\measuredangle(V_2, V) \ge \delta^{O(\epsilon)}.$$

The same is true for $\pi_{V|V_2}$. Then it follows easily from the Cartan decomposition that

$$\|\pi_{V|V_1}^{-1}\| \le \delta^{-O(\epsilon)} \text{ and } \|\pi_{V|V_2}^{-1}\| \le \delta^{-O(\epsilon)}. \tag{4.37}$$

Since $V \in \mathcal{E}(A_2)$, there is a subset $A_V \subset A_2$ such that $\mathcal{N}_\delta(A_V) \ge \delta^{-\alpha+O(\epsilon)}$ and $\mathcal{N}_\delta(\pi_V(A_V)) \le \delta^{-\frac{\alpha}{2}-\epsilon}$. It follows from (2.7) that

$$\omega_\delta(\pi_V, X+Y) \ge \omega_\delta(\pi_V, A_V) \ge \delta^{-\frac{3\alpha}{2}+O(\epsilon)}.$$

By (4.37), the map $\mathbb{R}^n = V_1 \oplus V_2 \to V \times V$ defined by $v_1 + v_2 \mapsto (\pi_V(v_1), \pi_V(v_2))$ is $\delta^{-O(\epsilon)}$-bi-Lipschitz. Hence, by (2.8), we can bound from below the additive energy between $\pi_V X$ and $\pi_V Y$,

$$\omega_\delta(+, \pi_V X \times \pi_V Y) \ge \delta^{-\frac{3\alpha}{2}+O(\epsilon)} \ge \delta^{O(\epsilon)} \mathcal{N}_\delta(\pi_V X)^{\frac{3}{2}} \mathcal{N}_\delta(\pi_V Y)^{\frac{3}{2}}.$$

That is why we can apply the Balog-Szemerédi-Gowers theorem (Theorem 2.7) to get subsets $X_V \subset X$ and $Y_V \subset Y$ such that

$$\mathcal{N}_\delta(X_V), \mathcal{N}_\delta(Y_V) \ge \delta^{-\frac{\alpha}{2}+O(\epsilon)} \tag{4.38}$$

and

$$\mathcal{N}_\delta(\pi_V X_V + \pi_V Y_V) \le \delta^{-\frac{\alpha}{2}-O(\epsilon)}. \tag{4.39}$$

Applying $\pi_{V|V_1}^{-1}$ to the set in the last inequality and using (2.4), we obtain

$$\mathcal{N}_\delta(X_V + \varphi_V Y_V) \le \delta^{-\frac{\alpha}{2}-O(\epsilon)}, \tag{4.40}$$

where $\varphi_V : V_2 \to V_1$ is $\varphi_V = \pi_{V|V_1}^{-1} \circ \pi_{V|V_2}$. Note that from (4.37), φ_V is $\delta^{-O(\epsilon)}$-bi-Lipschitz.

Let us apply Lemma 4.15 to the collection of subsets $X_V^{(\delta)} \times Y_V^{(\delta)} \subset X^{(\delta)} \times Y^{(\delta)}$ with the restriction of μ to $\mathcal{D}$. We obtain $V_\star \in \mathcal{D}$, $X_\star := X_{V_\star}$ and $Y_\star := Y_{V_\star}$ such that

$$\lambda(X_\star^{(\delta)} \cap X_V^{(\delta)})\lambda(Y_\star^{(\delta)} \cap Y_V^{(\delta)}) \geq \delta^{n-\alpha+O(\epsilon)}$$

whenever $V \in \mathcal{D}'$, where $\mathcal{D}'$ is a subset of $\mathcal{D}$ with

$$\mu(\mathcal{D}') \geq \delta^{O(\epsilon)}\mu(\mathcal{D}) \geq \delta^{O(\epsilon)}. \tag{4.41}$$

By Ruzsa's triangular inequality (Lemma 2.3), (4.40) implies, for all $V \in \mathcal{D}'$

$$\mathcal{N}_\delta(X_V - X_\star^{(\delta)} \cap X_V^{(\delta)}) \ll \mathcal{N}_\delta(X_V - X_V) \leq \delta^{-\frac{\alpha}{2}-O(\epsilon)}.$$

For the same reason $\mathcal{N}_\delta(X_\star - X_\star^{(\delta)} \cap X_V^{(\delta)}) \leq \delta^{-\frac{\alpha}{2}-O(\epsilon)}$. Then by Ruzsa's triangular inequality again, we have

$$\mathcal{N}_\delta(X_\star - X_V) \leq \delta^{-\frac{\alpha}{2}-O(\epsilon)}. \tag{4.42}$$

Similarly, $\mathcal{N}_\delta(Y_\star - Y_V) \leq \delta^{-\frac{\alpha}{2}-O(\epsilon)}$, which implies with (2.4),

$$\mathcal{N}_\delta(\varphi_V Y_\star - \varphi_V Y_V) \leq \delta^{-\frac{\alpha}{2}-O(\epsilon)}. \tag{4.43}$$

Moreover, (4.40) with (2.4) gives

$$\mathcal{N}_\delta(\varphi_V \varphi_\star^{-1} X_\star + \varphi_V Y_\star) \leq \delta^{-\frac{\alpha}{2}-O(\epsilon)}, \tag{4.44}$$

where $\varphi_\star := \varphi_{V_\star}$.

Now successive use of Ruzsa's triangular inequality (recalling (4.42), (4.40), (4.43) and (4.44)) yields that for all $V \in \mathcal{D}'$,

$$\mathcal{N}_\delta(X_\star - \varphi_V \varphi_\star^{-1} X_\star) \leq \delta^{-\frac{\alpha}{2}-O(\epsilon)}. \tag{4.45}$$

Moreover, by the Plünnecke-Ruzsa inequality (Lemma 2.4),

$$\mathcal{N}_\delta(X_\star + X_\star) \leq \delta^{-\frac{\alpha}{2}-O(\epsilon)}. \tag{4.46}$$

Consider the set of endomorphisms $\mathcal{A} = \{-\varphi_V \varphi_\star^{-1} \in \mathrm{End}(V_1) \mid V \in \mathcal{D}'\}$. We claim that the assumptions of Theorem 3.2 are satisfied for $\mathcal{A}$ and $X_\star$ with ϵ replaced by $O(\epsilon)$ and κ replaced by $\frac{\kappa}{8}$. Therefore, when ϵ is small enough, (4.45) and (4.46) contradict Theorem 3.2.

Our claim about the assumptions (i), (iv) and (vi) are clear from what precedes. The assumption (v) follows from (4.35) and (4.38) because for any $\rho \geq \delta$,

$$\mathcal{N}_\delta(X_\star) \leq \mathcal{N}_\rho(X_\star) \max_{x \in V_1} \mathcal{N}_\delta(X_\star \cap x^{(\rho)}).$$

Finally, to prove (ii) and (iii) we use Lemma 4.22 below. For any $\rho \geq \delta$ and any $f \in \mathrm{End}(V_1)$, Lemma 4.22 gives the existence of a subspace $W' \in \mathrm{Gr}(\mathbb{R}^n, m)$ such that $-\varphi_V \varphi_*^{-1} \in \mathbf{B}(f, \rho)$ implies $V \in \mathcal{V}_\measuredangle(W', \delta^{-O(\epsilon)}\rho)$. Hence by (4.3),

$$\mu\big(\{V \in \mathcal{D}' \mid -\varphi_V \varphi_\star^{-1} \in \mathbf{B}(f, \rho)\}\big) \leq \delta^{-O(\epsilon)}\rho^\kappa.$$

Observe that

$$\mu(\mathcal{D}') \le \mathcal{N}_\rho(\mathcal{A}) \max_{f\in \mathrm{End}(V_1)} \mu\big(\{V \in \mathcal{D}' \mid -\varphi_V\varphi_\star^{-1} \in \mathbf{B}(f,\rho)\}\big).$$

Together with (4.41), this gives the assumption (ii), namely,

$$\mathcal{N}_\rho(\mathcal{A}) \ge \delta^{O(\epsilon)}\rho^{-\kappa}.$$

Moreover, for any nonzero proper linear subspace $W \in V_1$, take $w \in W$ some vector with $\|w\| = 1$ and consider

$$\rho_0 = \sup_{V\in\mathcal{D}'} d(-\varphi_V\varphi_\star^{-1}(w), W).$$

By Lemma 4.22 and (4.37), we have $\mathcal{D}' \subset \mathcal{V}_\measuredangle(W', \delta^{-O(\epsilon)}\rho_0)$ for some $W' \in \mathrm{Gr}(\mathbb{R}^n, m)$. In view of (4.41) and (4.3), we have $\delta^{O(\epsilon)} \le \delta^{-O(\epsilon)}\rho_0^\kappa$. Hence $\rho_0 \ge \delta^{O(\epsilon)}$, which establishes (iii). □

Lemma 4.22. *We use the notations in the proof above. For any nonzero vector $v_2 \in V_2$ and any proper linear subspace $W \subset V_1$, there is $W' \in \mathrm{Gr}(\mathbb{R}^n, m)$ such that for all $V \in \mathrm{Gr}(\mathbb{R}^n, m)$,*

$$\mathrm{d}_\measuredangle(V, W') \le \|v_2\|^{-1} d(\varphi_V(v_2), W). \tag{4.47}$$

Proof. Without loss of generality, we can assume that $\dim(W) = m - 1$. For any $V \in \mathrm{Gr}(\mathbb{R}^n, m)$, any $v_2 \in V_2$ and any $w \in W$, by (4.5), we have

$$\mathrm{d}_\measuredangle(V^\perp, \mathbb{R}(v_2 - w)) = \frac{\|\pi_V(v_2 - w)\|}{\|v_2 - w\|}.$$

Note that $\|v_2 - w\| \ge \|v_2\|$ since $v_2 \perp w$ and $\|\pi_V(v_2 - w)\| \le \|\varphi_V(v_2) - w\|$ since $\pi_V(\varphi_V(v_2) - w) = \pi_V(v_2 - w)$. Hence

$$\mathrm{d}_\measuredangle(V^\perp, \mathbb{R}(v_2 - w)) \le \frac{\|\varphi_V(v_2) - w\|}{\|v_2\|}.$$

As w can be any vector in W, we obtain

$$\mathrm{d}_\measuredangle(V^\perp, \mathbb{R}v_2 + W) \le \|v_2\|^{-1} d(\varphi_V(v_2), W).$$

We conclude by setting $W' = (\mathbb{R}v_2 + W)^\perp \in \mathrm{Gr}(\mathbb{R}^n, m)$ and using (4.7). □

4.3.2 Projection of a slice

If the set A contains a relatively large slice of dimension $0 < n' < n$ (a subset $B = A^{(\delta)} \cap (y + W)$ with $\dim(W) = n'$ and $\mathcal{N}_\delta(B) \asymp \delta^{-\frac{n'}{n}\alpha}$) and if it has a correct non-concentration property then we can apply the induction hypothesis to $B - y$ inside W. Instead of projecting to V distributed according to μ, we project to $V' = \pi_W(V)$. The first lemma below shows that V' is not concentrated and the next one shows the relationship between V' being in $\mathcal{E}(B) \cap \mathrm{Gr}(W, m)$ and V being in $\mathcal{E}(B)$. Using this idea we prove Proposition 4.6.

Lemma 4.23. *Let $0 < m < n' < n$ be integers and $\kappa, \epsilon > 0$ be parameters. Let $W \in \mathrm{Gr}(\mathbb{R}^n, n')$ and V be a random element of $\mathrm{Gr}(\mathbb{R}^n, m)$ having the following non-concentration property,*

$$\forall \rho \geq \delta,\ \forall U \in \mathrm{Gr}(\mathbb{R}^n, n-m), \quad \mathbb{P}\big[\mathrm{d}_{\measuredangle}(V, U) \leq \rho\big] \leq \delta^{-\epsilon}\rho^{\kappa}. \tag{4.48}$$

Set $V' = \pi_W(V)$. Then with probability at least $1 - \delta^{\kappa-\epsilon}$, $\dim(V') = m$. Conditional to this event the distribution of V' is a probability measure ν on $\mathrm{Gr}(W, m)$. It satisfies

$$\forall \rho \geq \delta,\ \forall U \in \mathrm{Gr}(W, n'-m), \quad \nu(\mathcal{V}_{\measuredangle}(U, \rho)) \leq \delta^{-2\epsilon}\rho^{\kappa}.$$

Proof. We know that $\dim(V') = m$ if and only if $\mathrm{d}_{\measuredangle}(V, W^{\perp}) > 0$. The first part follows immediately from (4.48). Let us show the non-concentration property for ν. Let U be a $(n'-m)$-dimensional subspace of W. If $\mathrm{d}_{\measuredangle}(V', U) \leq \rho$ then $\mathrm{d}_{\measuredangle}(V, U + W^{\perp}) \leq \rho$ by (4.16). Hence

$$\mathbb{P}\big[\mathrm{d}_{\measuredangle}(V', U) \leq \rho\big] \leq \mathbb{P}\big[\mathrm{d}_{\measuredangle}(V, U + W^{\perp}) \leq \rho\big] \leq \delta^{-\epsilon}\rho^{\kappa}$$

and hence $\nu(\mathcal{V}_{\measuredangle}(U, \rho)) \leq \frac{\delta^{-\epsilon}\rho^{\kappa}}{1-\delta^{\kappa-\epsilon}} \leq \delta^{-2\epsilon}\rho^{\kappa}$. □

Lemma 4.24. *Let $0 < m \leq n' < n$ be integers. Let $0 < \alpha < n$ and $\epsilon > 0$ be parameters. Let $B \subset W$ be a bounded subset in a n'-dimensional linear subspace $W \subset \mathbb{R}^n$. Then*

$$\pi_W\big(\mathcal{E}(B, \epsilon) \setminus \mathcal{V}_{\measuredangle}(W^{\perp}, \delta^{\epsilon})\big) \subset \mathcal{E}(B, O(\epsilon)) \cap \mathrm{Gr}(W, m).$$

Proof. Let $V \in \mathcal{E}(B, \epsilon) \setminus \mathcal{V}_{\measuredangle}(W^{\perp}, \delta^{\epsilon})$. Then there exists $B' \subset B$ such that

$$\mathcal{N}_{\delta}(B') \geq \delta^{\epsilon}\mathcal{N}_{\delta}(B) \text{ and } \mathcal{N}_{\delta}(\pi_V(B')) \leq \delta^{-\frac{m}{n}\alpha-\epsilon}.$$

Denote by V' the projection $\pi_W(V)$. It follows from Lemma 4.13 that

$$\mathcal{N}_{\delta}(\pi_{V'}(B')) \leq \mathrm{d}_{\measuredangle}(V, W^{\perp})^{-O(1)}\mathcal{N}_{\delta}(\pi_V(B')) \leq \delta^{-\frac{m}{n}\alpha-O(\epsilon)}.$$

That is why $V' \in \mathcal{E}(B, O(\epsilon)) \cap \mathrm{Gr}(W, m)$. □

Proof of Proposition 4.6. Let $n = qm$ and suppose that Theorem 4.4 holds for $n' = (q-1)m$ and m. Let A and μ be as in Theorem 4.4 but for which the conclusion fails. By Lemma 4.20, there is an n'-dimensional affine subspace $y + W$ and a subset $B \subset A^{(\delta)} \cap (y + W)$ such that

$$\mathcal{N}_{\delta}(B) \geq \delta^{-\beta+O(\epsilon)} \quad \text{and}$$

$$\forall \rho \geq \delta,\ \forall x \in W, \quad \mathcal{N}_{\delta}(B \cap x^{(\rho)}) \leq \rho^{\frac{\kappa}{2q^2}}\delta^{-\beta-O(\epsilon)}$$

where $\beta = \frac{q-1}{q}\alpha$. Without loss of generality, we can assume $y = 0$ and $B \subset A$.

Let V be a random element of $\mathrm{Gr}(\mathbb{R}^n, m)$ distributed according to μ. Define ν be as in Lemma 4.23. By the lemma, we can apply Theorem 4.4 combined with Proposition 4.17 to $B \subset W$ with the probability measure ν on $\mathrm{Gr}(W, m)$. We obtain a constant $\epsilon' > 0$ depending only on n', β and κ such that when $\epsilon \leq \epsilon'$,

$$\nu\big(\mathcal{E}(B, \epsilon') \cap \mathrm{Gr}(W, m)\big) \leq \delta^{\epsilon'}.$$

Set $\epsilon_1 = \frac{3\epsilon}{\kappa}$. By Lemma 4.24, we have

$$\mu\big(\mathcal{E}(B,\epsilon_1) \setminus \mathcal{V}_\measuredangle(W^\perp, \delta^{\epsilon_1})\big) \leq \nu\big(\mathcal{E}(B, O(\epsilon)) \cap \mathrm{Gr}(W,m)\big).$$

When $\epsilon \leq \frac{\epsilon'}{O(1)}$, the last two inequalities together with (4.3) yield

$$\mu(\mathcal{E}(B,\epsilon)) \leq \mu\big(\mathcal{E}(B,\epsilon_1) \setminus \mathcal{V}_\measuredangle(W^\perp, \delta^{\epsilon_1})\big) + \mu(\mathcal{V}_\measuredangle(W^\perp, \delta^{\epsilon_1})) \leq \delta^{\epsilon'} + \delta^{2\epsilon} \leq \delta^{\epsilon},$$

which finishes the proof of Proposition 4.6. □

4.3.3 Projection to a sum of subspaces

In the situation where $m < \frac{n}{2}$, we consider the sum $V = V_1 + \cdots + V_q$ where q is a positive integer such that $qm < n$ and $V_1, \ldots, V_q$ are m-dimensional subspaces. Using the inequality (4.12), the size of the projection to V can be bounded in terms of the sizes of the projections to each V_j. In the next lemma, we prove that if V_i are independently randomly distributed according to a measure with an appropriate non-concentration property then the distribution of their sum V has a non-concentration property as well. This allows us to apply Theorem 4.4 with the dimensions n and $m' = qm$. This idea leads to the proof of Proposition 4.7.

Lemma 4.25. *Let n, m, q, r be positive integers such that $qm + r = n$. Let $0 < \epsilon < \frac{1}{2}\kappa$ be parameters. Let $V_1, \ldots V_q$ be independent random elements of $\mathrm{Gr}(\mathbb{R}^n, m)$ satisfying $\forall j = 1, \ldots, q$,*

$$\forall \rho \geq \delta,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, n-m) \quad \mathbb{P}\big[\mathrm{d}_\measuredangle(V_j, W) \leq \rho\big] \leq \delta^{-\epsilon}\rho^\kappa.$$

Then with probability at least $1 - (q-1)\delta^{\kappa-\epsilon}$, we have

$$\dim(V_1 + \cdots + V_q) = qm. \tag{4.49}$$

Then the probability measure μ' on $\mathrm{Gr}(\mathbb{R}^n, qm)$ defined as the distribution of $V_1 + \cdots + V_q$ conditional to the event (4.49) satisfies the non-concentration property

$$\forall \rho \geq \delta,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, r), \quad \mu'(\mathcal{V}_\measuredangle(W, \rho)) \leq \delta^{-O(\epsilon)}\rho^{\frac{\kappa}{q}}.$$

Proof. Let $V_1, \ldots, V_q$ be as in the statement. By their independence, for every $j = 2, \ldots, q$,

$$\mathbb{P}\big[\mathrm{d}_\measuredangle(V_j, V_1 + \cdots + V_{j-1}) \leq \delta\big] \leq \delta^{\kappa-\epsilon}.$$

Hence, on account of (4.9), with probability at least $1 - (q-1)\delta^{\kappa-\epsilon}$, we have

$$\mathrm{d}_\measuredangle(V_1, \ldots, V_q) \geq \delta^{(q-1)} > 0$$

and hence $V_1 + \cdots + V_q$ is a direct sum.

Let $\rho \geq \delta$, $W \in \mathrm{Gr}(\mathbb{R}^n, r)$. By (4.10), we know that if

$$\mathrm{d}_\measuredangle(V_1 + \cdots + V_q, W) \leq \rho$$

then for some $j = 1, \ldots, q$,

$$\mathrm{d}_\measuredangle(V_j, V_1 + \cdots + V_{j-1} + W) \leq \rho^{\frac{1}{q}},$$

which happens with probability at most $\delta^{-\epsilon}\rho^{\frac{\kappa}{q}}$. Therefore,

$$\mathbb{P}\big[\mathrm{d}_{\measuredangle}(V_1+\cdots+V_q, W) \le \rho\big] \le q\delta^{-\epsilon}\rho^{\frac{\kappa}{q}}.$$

Hence

$$\mu'(\mathcal{V}_{\measuredangle}(W,\rho)) \le \frac{q\delta^{-\epsilon}\rho^{\frac{\kappa}{q}}}{1-(q-1)\delta^{\kappa-\epsilon}} \le \delta^{-O(\epsilon)}\rho^{\frac{\kappa}{q}}. \qquad \square$$

Proof of Proposition 4.7. Let n, m, q, r be positive integers such that $qm+r = n$. Suppose Theorem 4.4 is true for the dimensions n and $m' = qm$ but it fails for the dimensions n and m with parameters $0 < \alpha < n$, $\kappa > 0$ and $\epsilon > 0$. Let A and μ be a counterexample, i.e. A and μ satisfy (4.1)–(4.3) but $\mu(\mathcal{E}(A')) > \delta^{\epsilon}$ for all subsets $A' \subset A$. We will get a contradiction when ϵ is smaller than a constant depending only on n, α and κ.

Let $V_1, \dots V_q$ be random elements of $\mathrm{Gr}(\mathbb{R}^n, m)$ independently distributed according to μ. Let μ' be the probability measure on $\mathrm{Gr}(\mathbb{R}^n, qm)$ defined as in Lemma 4.25. Thanks to Lemma 4.25, we can apply Theorem 4.4 combined with Proposition 4.17 with dimensions n and $m' = qm$ to the set A and the measure μ'. It gives $\epsilon' = \epsilon'(n, \alpha, \kappa) > 0$ such that if $\epsilon \le \epsilon'$ then the probability that there exists $A' \subset A$ satisfying $\mathcal{N}_\delta(A') \ge \delta^{\epsilon'}\mathcal{N}_\delta(A)$ and

$$\mathcal{N}_\delta(\pi_{V_1+\cdots+V_q}(A')) \le \delta^{-\frac{qm}{n}\alpha-\epsilon'}$$

is at most $\delta^{\epsilon'} + (q-1)\delta^{\kappa-\epsilon}$.

The rest of the proof consist of proving a lower bound for the same probability. First, $V_1 \in \mathcal{E}(A)$ with probability at least δ^{ϵ}. When this happens, there is $A_1 \subset A$ with $\mathcal{N}_\delta(A_1) \ge \delta^{\epsilon}\mathcal{N}_\delta(A)$ and $\mathcal{N}_\delta(\pi_{V_1}(A_1)) \le \delta^{-\frac{m}{n}\alpha-\epsilon}$. Write $\epsilon_1 = \frac{3\epsilon}{\kappa}$. Then conditional to any choice of V_1, we have $V_2 \in \mathcal{E}(A_1) \setminus \mathcal{V}_{\measuredangle}(V_1, \delta^{\epsilon_1})$ with probability at least $\delta^{2\epsilon}$. When this happens, there is $A_2 \subset A_1$ with $\mathcal{N}_\delta(A_2) \ge \delta^{\epsilon}\mathcal{N}_\delta(A_1)$ and $\mathcal{N}_\delta(\pi_{V_2}(A_2)) \le \delta^{-\frac{m}{n}\alpha-\epsilon}$. Then conditional to any choice of V_1 and V_2, the probability that $V_3 \in \mathcal{E}(A_2) \setminus \mathcal{V}_{\measuredangle}(V_1+V_2, \delta^{\epsilon_1})$ is at least $\delta^{2\epsilon}$. Then we continue this construction until we get A_q.

To summarize, we have with probability at least $\delta^{(2q-1)\epsilon}$,

$$\mathrm{d}_{\measuredangle}(V_1, \dots, V_q) \ge \delta^{O(\epsilon)}$$

and there exists a subset $A_q \subset A$ satisfying $\mathcal{N}_\delta(A_q) \ge \delta^{q\epsilon}\mathcal{N}_\delta(A)$ and for every $j = 1, \dots, q$,

$$\mathcal{N}_\delta(\pi_{V_j}(A_q)) \le \delta^{-\frac{m}{n}\alpha-\epsilon}$$

and hence, by Lemma 4.11,

$$\mathcal{N}_\delta(\pi_{V_1+\cdots+V_q}(A_q)) \le \delta^{-\frac{qm}{n}\alpha-O(\epsilon)}.$$

This leads to a contradiction when $\epsilon \le \frac{\epsilon'}{O(1)}$. $\square$

4.3.4 Projection to intersection of subspaces I: a discrete model

When the projections of a set A to subspaces $V_1, \dots, V_q$ are all small, we would like to say that its projection to the intersection $V = V_1 \cap \cdots \cap V_q$ is small as well. This is not true. A typical example is $A = (\mathbb{R}e_1 \oplus \mathbb{R}e_2) \cup \mathbb{R}e_3$ where

(e_1, e_2, e_3) is the standard basis in $\mathbb{R}^3$. While its projections to $\mathbb{R}e_1 \oplus \mathbb{R}e_3$ and to $\mathbb{R}e_2 \oplus \mathbb{R}e_3$ are both small (have dimension 1 in a 2-dimensional space), its projection to $\mathbb{R}e_3$ is full dimensional. In this example, A contains a large slice orthogonal to V. This happens to be the major obstruction.

Proposition 4.26. *Let n, m, q, r be positive integers such that $n = q(n-m)+r$. For any $0 < \alpha < n$ and $\epsilon > 0$, the following is true for sufficiently small $\delta > 0$. Let $A \subset \mathbb{R}^n$ and $V_1, \dots, V_q \in \mathrm{Gr}(\mathbb{R}^n, m)$. Write $V = V_1 \cap \cdots \cap V_q$. Assume that*

(i) $\mathrm{d}_{\measuredangle}(V_1^{\perp}, \dots, V_q^{\perp}) \geq \delta^{\epsilon}$;

(ii) $\delta^{-\alpha+\epsilon} \leq \mathcal{N}_\delta(A) \leq \delta^{-\alpha-\epsilon}$;

(iii) For every $j = 1, \dots, q$, $\mathcal{N}_\delta(\pi_{V_j}(A)) \leq \delta^{-\frac{m}{n}\alpha-\epsilon}$;

(iv) For all $y \in V$, $\mathcal{N}_\delta\big(A \cap \pi_V^{-1}(y^{(\delta)})\big) \leq \delta^{-\frac{n-r}{n}\alpha-\epsilon}$.

Then there exists $A' \subset A$ such that $\mathcal{N}_\delta(A') \geq \delta^{O(\epsilon)}\mathcal{N}_\delta(A)$ and

$$\mathcal{N}_\delta(\pi_V(A')) \leq \delta^{-\frac{r}{n}\alpha-O(\epsilon)}.$$

This proposition is deduced from the following discrete analogue. Let n, m, q, r be as in Proposition 4.26. For $I \subset \{1, \dots, n\}$, we write $\varpi_I \colon \mathbb{Z}^n \to \mathbb{Z}^I$ to denote the discrete projection $(z_i)_{i\in\{1,\dots,n\}} \mapsto (z_i)_{i\in I}$. Consider $I_0 = \{n-r+1, \dots, n\}$ and for $j = 1, \dots, q$

$$I_j = \{1, \dots, n\} \setminus \{(j-1)(n-m)+1, \dots, j(n-m)\}.$$

Proposition 4.27. *We use the notations above. For any parameter $K \geq 1$ and any finite subset $Z \subset \mathbb{Z}^n$. One of the following statements is true.*

(i) There exists $j \in \{1, \dots, q\}$ such that $|\varpi_{I_j}(Z)| \geq K|Z|^{\frac{m}{n}}$.

(ii) There exists $y \in \mathbb{Z}^{I_0}$ such that $|Z \cap \varpi_{I_0}^{-1}(y)| \geq K|Z|^{\frac{n-r}{n}}$.

(iii) There exists $Z' \subset Z$ such that $|Z'| \geq \frac{1}{2K^{q+1}}|Z|$ and $|\varpi_{I_0}(Z')| \leq 2K^q|Z|^{\frac{r}{n}}$.

One of the ingredients is a discrete isoperimetric inequality due to Bollobás-Thomason [5] known as the uniform cover theorem. Let $\mathcal{P}(\{1, \dots, n\})$ denote the set of subsets of $\{1, \dots, n\}$. Recall that a multiset of subsets of $\{1, \dots, n\}$ is a collection of elements of $\mathcal{P}(\{1, \dots, n\})$ which can have repeats. Giving such a multiset is equivalent to giving a map from $\mathcal{P}(\{1, \dots, n\})$ to $\mathbb{N}$. Following Bollobás-Thomason, we say a multiset $\mathcal{C}$ is k-uniform cover of $\{1, \dots, n\}$ if each element $i \in \{1, \dots, n\}$ belongs to exactly k members of $\mathcal{C}$. For exemple, with I_j defined above, $(I_1 \setminus I_0, \dots, I_q \setminus I_0)$ is a $(q-1)$-uniform cover of $\{1, \dots, n\} \setminus I_0$.

Theorem 4.28 (Uniform Cover theorem, Bollobás-Thomason [5])**.** *Let Z be a finite subset of $\mathbb{Z}^n$. Let $\mathcal{C}$ be an k-uniform cover of $\{1, \dots, n\}$. Then we have*

$$|Z|^k \leq \prod_{I\in\mathcal{C}} |\varpi_I(Z)|.$$

For example, if we consider projections onto all canonical m-dimensional subspaces. There is always one which has at least the expected size: there exists $I \subset \{1, \dots, n\}$ such that $|I| = m$ and $|\varpi_I(Z)| \geq |Z|^{m/n}$.

Lemma 4.29. *Let $I_0 \subset \{1,\dots,n\}$. Let Z be a finite subset of $\mathbb{Z}^n$ and $\mathcal{C}$ a k-uniform cover of $\{1,\dots,n\}\setminus I_0$ with q elements. Then*

$$|Z|^{2q-k} \leq \omega(\varpi_{I_0}, Z)^{q-k} \prod_{I\in\mathcal{C}} |\varpi_{I_0\cup I}(Z)|. \tag{4.50}$$

This lemma is a refinement of the uniform cover theorem. Indeed, for $I_0 = \varnothing$, we have $\omega(\varpi_{I_0}, Z) = |Z|^2$ and we recover the uniform cover theorem from (4.50).

Proof. For all $I \in \mathcal{C}$, we have

$$|\varpi_{I_0\cup I}(Z)| = \sum_{y\in\varpi_{I_0}(Z)} |\varpi_I(Z\cap \varpi_{I_0}^{-1}(y))|.$$

Hence, by Hölder's inequality,

$$\sum_{y\in\varpi_{I_0}(Z)} \prod_{I\in\mathcal{C}} |\varpi_I(Z\cap\varpi_{I_0}^{-1}(y))|^{\frac{1}{q}} \leq \prod_{I\in\mathcal{C}} \big(\sum_y |\varpi_I(Z\cap\varpi_{I_0}^{-1}(y))|\big)^{\frac{1}{q}} = \prod_{I\in\mathcal{C}} |\varpi_{I_0\cup I}(Z)|^{\frac{1}{q}}.$$

For each $y \in \varpi_{I_0}(Z)$, we apply the uniform cover theorem (Theorem 4.28) to the set $Z\cap\varpi_{I_0}^{-1}(y)$ seen as a finite subset of $\mathbb{Z}^{\{1,\dots,n\}\setminus I_0}$,

$$|Z\cap\varpi_{I_0}^{-1}(y)|^{\frac{k}{q}} \leq \prod_{I\in\mathcal{C}} |\varpi_I(Z\cap\varpi_{I_0}^{-1}(y))|^{\frac{1}{q}}.$$

From the two inequalities above, we get

$$\|\varpi_{I_0*}\mathbf{1}_Z\|_{\frac{k}{q}}^k \leq \prod_{I\in\mathcal{C}} |\varpi_{I_0\cup I}(Z)|.$$

Finally, Hölder's inequality implies

$$|Z| = \|\varpi_{I_0*}\mathbf{1}_Z\|_1 \leq \|\varpi_{I_0*}\mathbf{1}_Z\|_{\frac{k}{q}}^{\frac{k}{2q-k}} \|\varpi_{I_0*}\mathbf{1}_Z\|_2^{\frac{2q-2k}{2q-k}}.$$

We finish the proof by putting the last two inequalities together and recalling that $\omega(\varpi_{I_0}, Z) = \|\varpi_{I_0*}\mathbf{1}_Z\|_2^2$. □

Proof of Proposition 4.27. We use the notations introduced before Proposition 4.27. By (4.50), we have

$$|Z|^{q+1} \leq \omega(\varpi_{I_0}, Z) \prod_{j=1}^{q} |\varpi_{I_j}(Z)|.$$

If the first statement does not hold. Then we have

$$\omega(\varpi_{I_0}, Z) \geq \frac{1}{K^q} |Z|^{1+\frac{n-r}{n}}.$$

If the second statement fails as well, we can apply Lemma 2.5 with $M = K|Z|^{\frac{n-r}{n}}$ and $K' = K^{q+1}$. The third statement follows immediately. □

Proof of Proposition 4.26. Let $(e_1,\dots,e_n)$ denote the standard basis of $\mathbb{R}^n$. First consider the special case where $V_j^\perp$ is exactly $\mathrm{Span}(e_{(j-1)(n-m)+1},\dots,e_{j(n-m)})$

for each $j = 1, \dots, q$. Then we conclude easily form Proposition 4.27 by setting $K = \delta^{-2\epsilon}$ and

$$Z = \{x \in \mathbb{Z}^n \mid A \cap \delta \cdot (x + [0,1]^n) \neq \varnothing\}.$$

For the general case we consider a map $f \in \mathrm{GL}(\mathbb{R}^n)$ which sends isometrically V to $\mathrm{Span}(e_{n-r+1}, \dots, e_n)$ and $V_j^\perp$ to $\mathrm{Span}(e_{(j-1)(n-m)+1}, \dots, e_{j(n-m)})$ for each $j = 1, \dots, q$. It is easy to see that $\|f^{-1}\| \leq n$ and

$$|\det(f^{-1})| = \mathrm{d}_\measuredangle(V_1^\perp, \dots, V_q^\perp, V) = \mathrm{d}_\measuredangle(V_1^\perp, \dots, V_q^\perp).$$

Therefore f is $\delta^{-O(\epsilon)}$-bi-Lipschitz.

The conclusion for A follows from the special case applied to fA. Indeed, by the inequality (2.4) and Lemma 4.18, the hypotheses are satisfied for fA and $f^\perp V_1, \dots, f^\perp V_q$ with ϵ replaced by $O(\epsilon)$. Moreover, the conclusion for fA and $f^\perp V = f^\perp V_1 \cap \dots \cap f^\perp V_q$ implies that for A and V, again by (2.4) and Lemma 4.18. □

4.3.5 Projection to intersection of subspaces II: concluding proof

Once we have Proposition 4.26, to prove Proposition 4.8, we can use Proposition 4.21 and ideas in Subsection 4.3.2 to rule out the case where A has a very large slice and then apply the arguments in Subsection 4.3.3 to the dual.

Proof of Proposition 4.8. Let n, m, q, r be as in Proposition 4.8. Assume that Theorem 4.4 is true for the dimensions n and $m' = r$ and assume that A and μ are counterexample to Theorem 4.4 for the dimensions n and m with parameters $0 < \alpha < n$, $\kappa > 0$ and $\epsilon > 0$. We begin by making two remarks. Firstly, we can assume that

$$\mathcal{N}_\delta(A) \leq \delta^{-\alpha - O(\epsilon)}, \tag{4.51}$$

for otherwise, we could conclude directly by using Proposition 4.21.

Secondly, we can also assume that A does not contain very large slice of codimension r. More precisely, we can assume that

$$\forall W \in \mathrm{Gr}(\mathbb{R}^n, n-r),\ \forall x \in \mathbb{R}^n, \quad \mathcal{N}_\delta\big(A \cap (x + W^{(\delta)})\big) \leq \delta^{-\frac{n-r}{n}\alpha - O(\epsilon)}. \tag{4.52}$$

Indeed, if (4.52) fails, then put $B = \pi_W\big(A \cap (x + W^{(\delta)})\big)$ and we can apply Proposition 4.21 to $B \subset W$ to obtain that $\mathcal{E}(B) \cap \mathrm{Gr}(W, m)$ does not support any measure with the corresponding non-concentration property in $\mathrm{Gr}(W, m)$. We can conclude as in Subsection 4.3.2 by using Lemma 4.23 and Lemma 4.24.

From now on assume (4.51) and (4.52). Let $V_1, \dots, V_q$ be random elements of $\mathrm{Gr}(\mathbb{R}^n, m)$ independently distributed according to μ. On account of (4.7), the non-concentration property (4.3) implies similar property for the distribution of $V_1^\perp$, namely,

$$\forall \rho \geq \delta,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, m), \quad \mathbb{P}\big[\mathrm{d}_\measuredangle(V_1^\perp, W) \leq \rho\big] \leq \delta^{-\epsilon}\rho^\kappa.$$

From Lemma 4.25 applied to $V_1^\perp, \dots, V_q^\perp$, we know that with probability at least $1 - (q-1)\delta^{\kappa - \epsilon}$, the intersection $V = V_1 \cap \dots \cap V_q$ has dimension r.

Let μ' be the distribution of V conditional to this event. Then μ' has the non-concentration property

$$\forall \rho \geq \delta,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, n-r), \quad \mathbb{P}\big[\mathrm{d}_{\measuredangle}(V,W) \leq \rho\big] \leq \delta^{-O(\epsilon)} \rho^{\frac{\kappa}{q}}.$$

That is why we can apply Theorem 4.4 combined with Proposition 4.17 to the set A and the measure μ' with n and $m' = r$. We obtain $\epsilon' = \epsilon'(n, \alpha, \kappa) > 0$ such that if $\epsilon \leq \epsilon'$ then the probability that there exists $A' \subset A$ satisfying

$$\mathcal{N}_\delta(A') \geq \delta^{\epsilon'} \mathcal{N}_\delta(A) \text{ and } \mathcal{N}_\delta(\pi_V(A')) \leq \delta^{-\frac{r}{n}\alpha - \epsilon'}$$

is at most $\delta^{\epsilon'} + (q-1)\delta^{\kappa - \epsilon}$.

As the conclusion of Theorem 4.4 fails for A, we have $\mu(\mathcal{E}(A')) \geq \delta^\epsilon$ for all subsets $A' \subset A$. Using a similar construction as in the proof of Proposition 4.7, we prove that with probability at least $\delta^{O(\epsilon)}$, we have

$$\mathrm{d}_{\measuredangle}(V_1^\perp, \dots, V_q^\perp) \geq \delta^{O(\epsilon)}$$

and there exists $A_q \subset A$ satisfying $\mathcal{N}_\delta(A_q) \geq \delta^{O(\epsilon)} \mathcal{N}_\delta(A)$ and for all $j = 1, \dots, q$,

$$\pi_{V_j}(A_q) \leq \delta^{-\frac{m}{n}\alpha - \epsilon}.$$

Therefore, all the hypotheses of Proposition 4.26 are satisfied for the set A_q with $O(\epsilon)$ in the place of ϵ. In particular, the assumption (ii) is guaranteed by (4.1) and (4.51) and the assumption (iv) is guaranteed by (4.52). Hence there exists a subset $A' \subset A_q$ such that

$$\mathcal{N}_\delta(A') \geq \delta^{O(\epsilon)} \mathcal{N}_\delta(A) \text{ and } \mathcal{N}_\delta(\pi_V(A')) \leq \delta^{-\frac{r}{n}\alpha - O(\epsilon)}.$$

This leads to a contradiction when $\epsilon \leq \frac{\epsilon'}{O(1)}$. □

4.4 Projection of fractal sets

In this section we derive Corollary 4.2 and Corollary 4.3 from Theorem 4.1.

The ideas of the proof of Corollary 4.2 are contained in that of [7, Theorem 4]. We include the proof here for the sake of completeness.

Proof of Corollary 4.2. Let $0 < m < n$, $0 < \alpha < n$, $\kappa > 0$ be parameters. Let $\epsilon > 0$ be one fourth of the constant given by Theorem 4.1 applied to these parameters. Let A and μ be a counterexample for the corollary with these parameters. Without loss of generality we can assume $A \subset \mathbf{B}(0,1)$.

After normalizing μ we can suppose that it is a probability measure such that

$$\forall \rho > 0,\ \forall W \in \mathrm{Gr}(\mathbb{R}^n, n-m), \quad \mu(\mathcal{V}_{\measuredangle}(W, \rho)) \ll_\mu \rho^\kappa.$$

Thus, the non-concentration condition (4.3) of Theorem 4.1 is satisfied for sufficiently small δ.

By Frostman's lemma (recalled in Theorem 1.1), there is a nonzero Radon measure ν compactly supported on A such that

$$\forall \rho > 0,\ \forall x \in \mathbb{R}^n, \quad \nu(\mathbf{B}(x,\rho)) \leq \rho^{\alpha - \epsilon}. \tag{4.53}$$

For any $V \in \mathrm{Supp}(\mu)$ we have $\dim_{\mathrm{H}}(\pi_V(A)) \leq \eta$ where $\eta = \frac{m}{n}\alpha + 2\epsilon$. By the definition of Hausdorff dimension, for any $k_0 \geq 1$, there is a cover

$$\pi_V(A) \subset \bigcup_{k \geq k_0} B_{V,k}$$

of $\pi_V(A)$ such that each $B_{V,k}$ is a union of at most $2^{k\eta}$ balls of radius 2^{-k} in V.

Set $A_{V,k} = \pi_V^{-1}(B_{V,k})$ for $V \in \mathrm{Supp}(\mu)$ and $k \geq k_0$. Since the sets $A_{V,k}$, $k \geq k_0$, cover A, we have

$$\sum_{k \geq k_0} \nu(A_{V,k}) \gg_\nu 1.$$

Integrating with respect to $\mathrm{d}\mu(V)$ and using Fubini's theorem, we obtain

$$\sum_{k \geq k_0} \int \nu(A_{V,k}) \, \mathrm{d}\mu(V) \gg_\nu 1.$$

This in turn implies that there exists $k \geq k_0$ such that $\mu(\mathcal{E}) \gg_\nu k^{-2}$ where

$$\mathcal{E} = \{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \nu(A_{V,k}) \gg_\nu k^{-2}\}.$$

Now fix this k and set $\delta = 2^{-k}$ so that $\mathcal{N}_\delta(\pi_V(A_{V,k})) \leq \delta^{-\eta}$. Note that as we can choose k_0 arbitrarily large, we can make δ arbitrarily small.

Here we cannot apply Theorem 4.1 directly to the set A because it might not be regular enough. The idea is to partition A into regular parts. Let $\mathcal{Q}$ denotes the set of dyadic cubes in $\mathbb{R}^n$ of side length δ:

$$\mathcal{Q} = \left\{x + [0, \delta[^n \;\middle|\; x \in \delta \cdot \mathbb{Z}^n\right\}.$$

Put $L = \left\lceil \frac{n}{\epsilon} \right\rceil + 1$. For $l = 0, \ldots, L$, let A_l be the union of all cubes $Q \in \mathcal{Q}$ such that

$$\delta^{(l+1)\epsilon} \nu(A) < \nu(Q) \leq \delta^{l\epsilon} \nu(A).$$

It is easy to see that A_l are disjoint and $\sum_{l=0}^{L} \nu(A_l) \geq (1 - \delta^\epsilon)\nu(A)$. Moreover for any $l = 0, \ldots, L$ and any $A' \subset A_l$ which is also a union of cubes in $\mathcal{Q}$, we have

$$\delta^{(l+1)\epsilon} \mathcal{N}_\delta(A') \nu(A) \ll_n \nu(A') \ll_n \delta^{l\epsilon} \mathcal{N}_\delta(A') \nu(A)$$

Hence, if $\nu(A_l) > 0$, then for such A',

$$\delta^\epsilon \frac{\nu(A')}{\nu(A_l)} \ll_n \frac{\mathcal{N}_\delta(A')}{\mathcal{N}_\delta(A_l)} \ll_n \delta^{-\epsilon} \frac{\nu(A')}{\nu(A_l)} \tag{4.54}$$

Consider $\mathcal{L} = \{0 \leq l \leq L \mid \nu(A_l) \geq \delta^\epsilon\}$, the set of levels with sufficient mass. For any $l \in \mathcal{L}$, by (4.53),

$$\mathcal{N}_\delta(A_l) \gg \delta^{-\alpha+\epsilon} \nu(A_l) \geq \delta^{-\alpha+2\epsilon}$$

and from (4.54) and (4.53), for any $\rho \geq \delta$ and any $x \in \mathbb{R}^n$,

$$\frac{\mathcal{N}_\delta(A_l \cap \mathbf{B}(x, \rho))}{\mathcal{N}_\delta(A_l)} \ll_n \delta^{-\epsilon} \frac{\nu(\mathbf{B}(x, \rho + n\delta))}{\nu(A_l)} \ll_n \delta^{-3\epsilon} \rho^\alpha.$$

In other words, the assumptions of Theorem 4.1 are satisfied for A_l.

Now for $l \in \mathcal{L}$ and $V \in \mathcal{E}$, let $A_{V,k,l}$ be the union of $Q \in \mathcal{Q}$ such that $Q \subset A_l$ and $Q \cap A_{V,k} \neq \varnothing$. From the definition of $\mathcal{L}$ and $\mathcal{E}$, we know that for any $V \in \mathcal{E}$

$$\sum_{l \in \mathcal{L}} \nu(A_{V,k,l}) \gg_\nu k^{-2} - (L+1)\delta^\epsilon \gg_\nu k^{-2}.$$

Hence there exists $l \in \mathcal{L}$ such that $\frac{\nu(A_{V,k,l})}{\nu(A_l)} \gg_\nu k^{-2}$. Therefore by setting

$$\mathcal{E}_l = \Big\{ V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \frac{\nu(A_{V,k,l})}{\nu(A_l)} \gg_\nu k^{-2} \Big\},$$

we have $\mathcal{E} = \cup_{l \in \mathcal{L}} \mathcal{E}_l$. Hence there exists $l \in \mathcal{L}$ such that $\mu(\mathcal{E}_l) \gg_{\nu,L} k^{-2}$.

This contradicts Theorem 4.1 applied to the set A_l and the measure μ. Indeed, $\mathcal{E}_l \subset \mathcal{E}(A_l)$. Because if $V \in \mathcal{E}_l$, then by (4.54), $\mathcal{N}_\delta(A_{V,k,l}) \geq \delta^{2\epsilon} \mathcal{N}_\delta(A_l)$ and moreover

$$\mathcal{N}_\delta(\pi_V(A_{V,k,l})) \ll \mathcal{N}_\delta(\pi_V(A_{V,k})) \leq \delta^{-\eta}. \qquad \square$$

Now we deduce Corollary 4.3 from Corollary 4.2.

Proof of Corollary 4.3. We can choose a metric on the Grassmannian $\mathrm{Gr}(\mathbb{R}^n, m)$ which is invariant under the action of the group $\mathrm{O}(n)$. Observe that the exceptional set of directions

$$\{V \in \mathrm{Gr}(\mathbb{R}^n, m) \mid \dim_{\mathrm{H}}(\pi_V(A)) \leq \frac{m}{n}\alpha + \epsilon\}$$

is measurable for the Borel σ-algebra on $\mathrm{Gr}(\mathbb{R}^n, m)$. Suppose that the Hausdorff dimension of the exceptional set is larger than $m(n-m)-1+\kappa$ for some κ. Then by Frostman's Lemma there exists a nonzero Radon measure μ supported on this exceptional set such that for all $\rho > 0$ and all $V \in \mathrm{Gr}(\mathbb{R}^n, m)$, $\mu(\mathbf{B}(V, \rho)) \leq \rho^{m(n-m)-1+\kappa}$. We are going to prove that μ satisfies the non-concentration property forbidden by Corollary 4.2.

We fix $W \in \mathrm{Gr}(\mathbb{R}^n, m)$ and apply the Łojasiewicz inequality (recalled at Theorem 3.8) to the analytic function $\mathrm{d}_\measuredangle(\cdot, W)^2 \colon \mathrm{Gr}(\mathbb{R}^n, m) \to \mathbb{R}$. We conclude that there is a constant $C > 0$ such that for any $0 < \rho \leq 1$, $\mathcal{V}_\measuredangle(W, \rho)$ is contained in the ρ'-neighborhood of the Schubert cell $\mathcal{V}_\measuredangle(W, 0)$ with $\rho' = (C\rho)^{\frac{1}{C}}$. By $\mathrm{O}(n)$-invariance, the constant C is in fact uniform for all $W \in \mathrm{Gr}(\mathbb{R}^n, n-m)$. Since the Schubert cell $\mathcal{V}_\measuredangle(W, 0)$ is a smooth submanifold, we have $\mathcal{N}_{\rho'}(\mathcal{V}_\measuredangle(W, 0)^{(\rho')}) \asymp_n \rho'^{-m(n-m)+1}$. Here again, the estimate is uniform in W thanks to the $\mathrm{O}(n)$-invariance. Therefore,

$$\mu(\mathcal{V}_\measuredangle(W, \rho)) \leq \mathcal{N}_{\rho'}(\mathcal{V}_\measuredangle(W, 0)^{(\rho')}) \sup_{V \in \mathrm{Gr}(\mathbb{R}^n, m)} \mu(\mathbf{B}(V, \rho')) \ll_n \rho^{\frac{\kappa}{C}}.$$

This contradicts Corollary 4.2 if ϵ is sufficiently small and finishes the proof of Corollary 4.3. $\square$

Chapter 5

Product estimates in perfect Lie groups

This last chapter is a joint work with Nicolas de Saxcé. The aim is to generalize Saxcé's discretized product theorem (Theorem 1.7) to perfect Lie groups.

In this chapter, by default, Lie groups and Lie algebras are real. Recall that a Lie group G is said to be perfect if its Lie algebra $\mathfrak{g}$ satisfies the condition $[\mathfrak{g},\mathfrak{g}] = \mathfrak{g}$. The radical R of G is the unique maximal normal closed connected solvable subgroup in G. The Lie algebra of R is the radical $\mathfrak{r}$ of $\mathfrak{g}$. The quotient group G/R is semisimple. If G is simply connected, then G/R is simply connected, hence a direct product of simply connected simple Lie groups. Factors appearing in this direct product are called simple factors of the group G.

Theorem 5.1 (Product theorem for perfect Lie groups)**.** *Let G be a simply-connected perfect Lie group. There is a neighborhood U of the identity in G such that the following holds. Given $\sigma < \dim(G)$ and $\kappa > 0$, there exists $\epsilon > 0$ such that for all $\delta > 0$ sufficiently small, if a subset $A \subset U$ satisfies*

(i) $\mathcal{N}_\delta(A) \leq \delta^{-\sigma-\epsilon}$*;*

(ii) for any simple factor S of G, denote by $\pi_S \colon G \to S$ the canonical projection,

$$\forall \rho \geq \delta, \quad \mathcal{N}_\rho(\pi_S(A)) \geq \delta^\epsilon \rho^{-\kappa};$$

(iii) for any proper closed connected subgroup $H < G$, there exists $a \in A$ with $d(a,H) \geq \delta^\epsilon$;

then $\mathcal{N}_\delta(A^3) \geq \delta^{-\epsilon}\mathcal{N}_\delta(A)$.

For $G = \mathrm{SU}(d)$, $d \geq 2$, this is due to Bourgain-Gamburd [9, 11]. For simple groups, it is due to Saxcé [24].

Definition. A subset A satisfying condition (iii) will be said to be δ^ϵ*-away from closed connected subgroups.*

As an intermediate result, we have a sum-product type estimate for representations of Lie groups which is interesting on its own right. Throughout this chapter, all linear representations are finite-dimensional and over the field of

real numbers. They all come with a norm. Constants in our results depend not only on the algebraic structure but also on the choice of the norm. However, since all norms on a finite-dimensional linear space are all equivalent, if we have the result for one norm, we have it for any norm at a loss of a constant factor. In particular, we can always assume that the norm is Euclidean.

Let V be a representation of G. Let A a subset of G and X a bounded subset of V. For $s \geq 1$, recall that $\langle A, X \rangle_s$ stands for the set of elements in V that can be obtained as sums, differences and products of at most s elements of A and X.

Definition. For $\rho > 0$, we say X is *ρ-away from submodules* if for any proper submodule $W < V$, there is $x \in X$ such that $d(x, W) \geq \rho$.

Theorem 5.2 (Sum-product estimates in representations)**.** *Let G be a Lie group and V finite-dimensional linear representation of G. There exists a neighborhood U of the identity in G depending only on V such that the following holds. Given $\epsilon_0, \kappa > 0$, there exists $s \geq 1$ and $\epsilon > 0$ such that for all $\delta > 0$ sufficiently small, if $A \subset U$ and $X \subset \mathbf{B}_V(0,1)$ satisfy the following :*

(i) There is a Jordan-Hölder sequence $0 = V_0 < \ldots < V_l = V$ such that for every $i = 1, \ldots, l$

$$\forall \rho \geq \delta, \quad \mathcal{N}_\rho(\pi_{V_i/V_{i-1}}(A)) \geq \delta^\epsilon \rho^{-\kappa}.$$

where $\pi_{V_i/V_{i-1}} \colon G \to \mathrm{GL}(V_i/V_{i-1})$ denotes the representation of G on V_i/V_{i-1}.

(ii) A is δ^ϵ-away from closed connected subgroups.

(iii) X is δ^ϵ-away from submodules.

Then,

$$\mathbf{B}_V(0, \delta^{\epsilon_0}) \subset \langle A, X \rangle_s + \mathbf{B}_V(0, \delta).$$

In particular, considering $\mathbb{R}^*$ acting on $\mathbb{R}$, we can recover Bourgain's discretized sum-product theorem (Theorem 1.4). Similarly, we can recover discretized sum-product estimates for $\mathbb{C}$ or $\mathbb{H}$.

5.0.1 Outline of the proof

The proof of Theorem 5.1 goes as follows. In view of Lemma 2.13, we can investigate the growth of $(A \cup \{1\} \cup A^{-1})^s$, $s \geq 1$. The strategy is to prove that after a bounded number of steps (depending on ϵ_0), it becomes δ-dense in a ball of radius δ^{ϵ_0}, for any prescribed $\epsilon_0 > 0$.

Theorem 5.3. *Let G be a simply connected perfect Lie group. There exists a neighborhood U of the identity in G such that the following holds. Given $\kappa > 0$ and $\epsilon_0 > 0$, there exists $\epsilon > 0$ and $s \geq 1$ such that for all $\delta > 0$ sufficiently small, if A is a subset of U such that*

(i) for any simple factor S_i of G denote by $\pi_{S_i} \colon G \to S_i$ the projection,

$$\forall \rho \geq \delta, \quad \mathcal{N}_\rho(\pi_{S_i}(A)) \geq \delta^\epsilon \rho^{-\kappa};$$

(ii) A is δ^ϵ-away from closed connected subgroups;

then

$$\mathbf{B}_G(1, \delta^{\epsilon_0}) \subset (A \cup \{1\} \cup A^{-1})^s \, \mathbf{B}_G(1, \delta).$$

This statement immediately implies Theorem 5.1, just like Theorem 3.4 implies Theorem 3.1 in the sum-product setting.

The first step of the proof of Theorem 5.3 is to reduce to the case where the radical of G is abelian. In this reduction, we exploit the elementary algebraic fact that if R is a nilpotent group and R_l the last nontrivial term in its lower central series and if $H < G$ is a subgroup whose image under the projection R/R_l is full then $H = G$. For the special case we use the Theorem 5.2 to show growth of $\log A \subset \mathfrak{g}$ under addition and action of A through the adjoint representation. To get back inside G, we use Campbell-Hausdorff formula to show that we can approximate sums in $\mathfrak{g}$ by products in G.

The proof of Theorem 5.2 is by induction on the length of the Jordan-Hölder sequence. The base case where V is an irreducible representation is a variant of Theorem 3.2. The induction step exploits the fact that X is δ^ϵ-away from submodules hence generates V as G-module. We use systematically the Łojasiewicz inequality (recalled at Theorem 3.8) to extract quantitative estimates from algebraic facts.

5.1 Sum-product estimates in representations

In this section, we introduce a class of representations called $\mathcal{P}(G)$, show some basic properties of these representations and then prove Theorem 5.2.

5.1.1 Representations without trivial simple quotients

Definition. Let G be a connected Lie group. A representation V is said to be of class $\mathcal{P}(G)$ if there exists a sequence $0 = V_0 < V_1 < \ldots < V_l = V$ of subrepresentations of V such that, for each $i = 1, \ldots, l$, the quotient representation V_i/V_{i-1} is nontrivial and irreducible.

Equivalently, V is in $\mathcal{P}(G)$ if the trivial representation does not appear as a simple quotient in a Jordan-Hölder decomposition of V. This property, of course, does not depend on the choice of the Jordan-Hölder decomposition. We now list some elementary properties of representations in $\mathcal{P}(G)$.

Proposition 5.4. *Let V be a representation of a connected Lie group G.*

(i) (Subrepresentations and quotients representations) If W a subrepresentation of V, then V belongs to $\mathcal{P}(G)$ if and only if both W and V/W belong to $\mathcal{P}(G)$.

(ii) If H is a closed subgroup of G and $V \in \mathcal{P}(H)$ as a representation of H, then $V \in \mathcal{P}(G)$.

(iii) Let H be a normal subgroup of G. If the representation $G \to \mathrm{GL}(V)$ factors through G/H, then $V \in \mathcal{P}(G/H)$ as a representation of G/H if and only if $V \in \mathcal{P}(G)$ as a representation of G.

Proof. Indeed, (i) follows from the fact that the set of simple quotients of the Jordan-Hölder decomposition of V is the union of those of W and V/W. For (ii), note that a Jordan-Hölder sequence of G-submodules in V can be refined to a Jordan-Hölder sequence of H-submodules, and that if there is a trivial quotient in the first sequence there must be also one in the refined sequence. Finally, (iii) is clear, since Jordan-Hölder decompositions of V into G-modules coincide with Jordan-Hölder decompositions into G/H-modules. □

Remark. The class $\mathcal{P}(G)$ is the smallest class of finite-dimensional representations of G containing all non-trivial irreducible representation of G is in $\mathcal{P}(G)$ and such that if W and V' are in $\mathcal{P}(G)$, and $0 \to W \to V \to V' \to 0$ is a short exact sequence of G-modules, then V is in $\mathcal{P}(G)$.

Example.
- If a representation V contains the trivial representation, then it is not in $\mathcal{P}(G)$. Similarly, if V admits the trivial representation as a quotient, then it is not in $\mathcal{P}(G)$.
- The representation of $G = \mathbb{R}_+^*$ on $\mathbb{R}^n$ given by $g \cdot v = gv$ (scalar multiplication) is in $\mathcal{P}(G)$.
- The adjoint representation of a semisimple Lie group G is in $\mathcal{P}(G)$.
- Condition (i) of Theorem 5.2 implies $V \in \mathcal{P}(G)$.

Whenever V is a normed vector space, we endow the space of linear endomorphisms $\mathrm{End}(V)$ with the associated operator norm. Moreover, the norm on V induces a norm on each submodule W. There is also a natural norm on each quotient $V' = V/W$, given by the formula

$$\forall v \in V,\ \|\pi(v)\| = d(v, W),$$

where $\pi\colon V \to V'$ is the canonical projection. Throughout this chapter, quotients will be equipped with this norm.

For example, with this convention, it is easy to see that if $\pi\colon V \to V'$ is a quotient map then $\pi(X)$ is ρ-away from submodules in V' whenever X is ρ-away from submodules in V.

The proof of Theorem 5.2 goes by induction on the length of the Jordan-Hölder decomposition of V. We shall prove the base case, where V is a nontrivial irreducible representation, in the next subsection. The induction step will then be carried out in subsection 5.1.3.

5.1.2 Irreducible representations

In the case V is an irreducible representation of G, Theorem 5.2 is a variant of Theorem 3.2. We state it and prove it here for completeness.

Theorem 5.5 (Base case: irreducible representations)**.** *Let G be a Lie group and $\pi\colon G \to \mathrm{GL}(V)$ an irreducible representation. There is a neighborhood U of the identity in G such that the following holds. Given $\epsilon_0, \kappa > 0$, there exist $s \geq 1$ and $\epsilon > 0$ such that for all $\delta > 0$ sufficiently small, if $A \subset U$ and $X \subset \mathbf{B}_V(0,1)$ satisfy*

(i) For all $\rho \geq \delta$, $\mathcal{N}_\rho(\pi(A)) \geq \delta^\epsilon \rho^{-\kappa}$;

(ii) A is δ^ϵ-away from closed connected subgroups;

(iii) There exists $v \in X$ such that $\|v\| \geq \delta^\epsilon$;

then

$$\mathbf{B}_V(0, \delta^{\epsilon_0}) \subset \langle A, X \rangle_s + \mathbf{B}_V(0, \delta).$$

Observe that condition (i) implies that the representation V is non-trivial.

Recall that we say a subset $A \subset \mathrm{End}(V)$ *acts ρ-irreducibly on V* for some $\rho > 0$ if for any proper linear subspace $W \subset V$, there is $a \in A$, $w \in \mathbf{B}_W(0,1)$ such that $d(aw, W) \geq \rho$.

Proposition 5.6. *Let V be a finite-dimensional normed vector space. Given $\epsilon_0, \kappa > 0$, there exist $s \geq 1$ and $\epsilon > 0$ such that for all $\delta > 0$ sufficiently small, if $A \subset \mathbf{B}_{\mathrm{End}(V)}(0, \delta^{-\epsilon})$ and $v \in V$ satisfy*

(i) for any $\rho \geq \delta$, $\mathcal{N}_\rho(A) \geq \delta^\epsilon \rho^{-\kappa}$;

(ii) A acts δ^ϵ-irreducibly on V;

(iii) $\delta^\epsilon \leq \|v\| \leq \delta^{-\epsilon}$;

then

$$\mathbf{B}_V(0, \delta^{\epsilon_0}) \subset \langle A \rangle_s \cdot v + \mathbf{B}_V(0, \delta)$$

In the following proof, we use the notations in Chapter 3 : if K is a division algebra over $\mathbb{R}$ then $\mathrm{End}(K^n)$ denotes the space of real endomorphisms of K^n and $\mathcal{M}_n(K)$ denotes the space of $n \times n$ matrices with coefficients in K. Furthermore, we identify $\mathcal{M}_n(K)$ with the subspace of $\mathrm{End}(K^n)$ consisting of K-linear maps.

Proof. Given $\epsilon_1 > 0$, it follows from Proposition 3.25 that there exists $c > 0$ such that, provided $\epsilon > 0$ is small enough, there exists a $\delta^{-O(\epsilon)}$-bi-Lipschitz linear bijection $f \colon V \to K^n$, where K is $\mathbb{R}$, $\mathbb{C}$ or the quaternions $\mathbb{H}$, n is $\frac{\dim V}{\dim K}$ and K^n is endowed with its usual L^2 norm, and a scale δ_1 with $\delta \leq \delta_1 \leq \delta^c$ such that

$$fAf^{-1} \subset \mathcal{M}_n(K) + \mathbf{B}_{\mathrm{End}(K^n)}(0, \delta_1)$$

and for every proper subalgebra $F < \mathcal{M}_n(K)$,

$$\exists a \in A : \ d(faf^{-1}, F) \geq \delta_1^{\epsilon_1}.$$

Choosing ϵ_1 small enough in terms of ϵ_0 and κ, we may then apply Theorem 3.4 to conclude that, provided $\epsilon > 0$ is sufficiently small, then for some integer s,

$$\mathbf{B}_{\mathcal{M}_n(K)}(0, \delta_1^{\epsilon_0}) \subset f\langle A \rangle_s f^{-1} + \mathbf{B}_{\mathrm{End}(K^n)}(0, \delta_1).$$

Therefore, without loss of generality, we may assume that $V = K^n$ and

$$\mathbf{B}_{\mathcal{M}_n(K)}(0, \delta_1^{\epsilon_0}) \subset A + \mathbf{B}_{\mathrm{End}(V)}(0, \delta_1). \tag{5.1}$$

We can further assume that $\|v\| = 1$. Then

$$\mathbf{B}_V(0, \delta_1^{\epsilon_0}) \subset A \cdot v + \mathbf{B}_V(0, \delta_1).$$

In other words, the conclusion of the proposition holds at scale δ_1. It remains to bring the scale back to δ. To do this, we note that from (5.1), we have in particular

$$\delta_1^{\frac{1}{2}} \operatorname{id} \in A + \mathbf{B}_{\mathrm{End}(V)}(0, \delta_1).$$

Hence, starting from (5.1), we may multiply both sides by $\delta_1^{\frac{1}{2}} \operatorname{id}$ to obtain

$$\mathbf{B}_V(0, \delta_1) \subset \mathbf{B}_V(0, \delta_1^{\epsilon_0 + \frac{1}{2}}) \subset \langle A \rangle_2 \cdot v + \mathbf{B}_V(0, 2\delta_1^{\frac{3}{2}}),$$

and iterating this procedure, we get, for all $k \geq 2$, for some integers s_k depending on k,

$$\mathbf{B}_V(0, s_k \delta_1^{\frac{k}{2}}) \subset \langle A \rangle_s \cdot v + \mathbf{B}_V(0, s_{k+1} \delta_1^{\frac{k+1}{2}}).$$

for some integers $k \geq 0$ and $s \geq 1$. Choosing $k > \frac{2}{c}$ and combining all these inclusions, we find, for $s = s_1 + \cdots + s_k$,

$$\mathbf{B}_V(0, \delta_1^{\epsilon_0}) \subset \langle A \rangle_s \cdot v + \mathbf{B}_V(0, \delta),$$

which proves the proposition. □

The above proposition readily implies Theorem 5.5.

Proof of Theorem 5.5. It suffices to apply Proposition 5.6 to the set $\pi(A) \subset \mathrm{End}(V)$. By assumption on A, conditions (i) and (iii) of the proposition are satisfied for the set $\pi(A)$. That condition (ii) is also satisfied is a consequence of Lemma 5.7 below. □

Lemma 5.7. *Let $0 < \rho < \frac{1}{2}$ be a parameter. Let $\pi \colon G \to \mathrm{GL}(V)$ be a nontrivial irreducible representation. There is a neighborhood U of 1 in G such that if $A \subset U$ is ρ-away from closed connected subgroups then $\pi(A)$ acts $\rho^{O_\pi(1)}$-irreducibly on V.*

The proof of this lemma is an application of Łojasiewicz's inequality, but first, it is convenient to reduce to the case where A is finite. For that, we use another lemma. Let $\pi \colon G \to \mathrm{GL}(V)$ be a representation and $0 < \rho < \frac{1}{2}$ a parameter. We say that a subset $A \subset G$ is *ρ-away from proper stabilizers* if for any linear subspace W of V which is not a G-submodule there exists an element a in A whose distance to the stabilizer of W, i.e.,

$$\mathrm{Stab}_G(W) = \{g \in G \mid g \cdot W \subset W\}$$

is at least ρ.

Lemma 5.8. *Let $0 < \rho < \frac{1}{2}$ be a parameter. Let $\pi \colon G \to \mathrm{GL}(V)$ be a representation. There is a neighborhood U of 1 in G such that if $A \subset U$ is ρ-away from closed connected subgroups then A contains a subset of cardinality at most $\dim G$ which is $\rho^{O_\pi(1)}$-away from proper stabilizers.*

Here by $O_\pi(1)$ we mean a constant depending on the algebraic structure of G and π and the distance on G.

Remark. Note that in this lemma, the neighborhood U depends on the representation π, and not only on G. This is readily seen by considering $G = \mathbb{R}$, $V = \mathbb{C} \simeq \mathbb{R}^2$, and $\pi(x)v = e^{inx}v$ where n is some integer.

Proof. The representation π differentiates to a representation of the Lie algebra $\mathfrak{g}$ of G, which we denote by $T_1\pi\colon \mathfrak{g} \to \mathrm{End}(V)$. Then the stabilizer of W in $\mathfrak{g}$,

$$\mathrm{Stab}_{\mathfrak{g}}(W) = \{x \in \mathfrak{g} \mid T_1\pi(x)W \subset W\}$$

is the Lie algebra of $\mathrm{Stab}_G(W)$. In particular, its image under the exponential map exp is contained in $\mathrm{Stab}_G(W)^\circ$, the identity component of $\mathrm{Stab}_G(W)$.

We may assume that exp induces a diffeomorphism from U to its image, and denote the inverse map by log. We say that $\log(A)$ is *ρ-away from proper stabilizers* in $\mathfrak{g}$ if for any linear subspace $W < V$ which is not a G-submodule, there exists $a \in A$ such that $d(\log(a), \mathrm{Stab}_{\mathfrak{g}}(W)) \geq \rho$. The argument in [24, Lemma 2.5] shows that if $\log(A)$ is ρ-away from proper stabilizers then $\log(A)$ contains a subset of cardinality at most $\dim \mathfrak{g}$ which is $\rho^{O_{\dim(\mathfrak{g})}(1)}$-away from stabilizers.

We claim that there is a neighborhood U of 1 in G such that if $A \subset U$ is ρ-away from identity components of proper stabilizers then $\log(A)$ is $\frac{\rho}{C}$-away from proper stabilizers in $\mathfrak{g}$ and conversely if $\log(A)$ is ρ-away from proper stabilizers then A is $\frac{\rho}{C}$-away from proper stabilizers. This finishes the proof of the lemma.

Let us prove this claim. Indeed, from the identity $\pi(\exp(x)) = e^{T_1\pi(x)}$, we can express $T_1\pi(x)$ as an absolutely convergent series

$$T_1\pi(x) = -\sum_{n\geq 1} \frac{1}{n}\big(\mathrm{id}_V - \pi(\exp(x))\big)^n$$

whenever $\|\pi(\exp(x)) - \mathrm{id}_V\| < 1$. It follows that there is $r > 0$ depending only on π such that

$$\mathrm{Stab}_G(W) \cap \mathbf{B}_G(1, r) \subset \exp(\mathrm{Stab}_{\mathfrak{g}}(W)).$$

Let $U = \mathbf{B}_G(1, \frac{r}{2})$. Then for any $g \in U$ and any proper linear subspace W,

$$\frac{1}{C} d(g, \mathrm{Stab}_G(W)^\circ) \leq d(\log(g), \mathrm{Stab}_{\mathfrak{g}}(W)) \leq C d(g, \mathrm{Stab}_G(W))$$

where $C > 0$ is some constant depending only on the the representation. This finishes the proof of the claim. □

Proof of Lemma 5.7. Let U be the neighborhood given by Lemma 5.8. On account of this lemma we can assume that A is finite of cardinality $n \leq \dim G$ and ρ-away from proper stabilizers. Let $0 < k < \dim(V)$. Consider the analytic map $f\colon G^n \times \mathrm{Gr}(V, k) \to \mathbb{R}$ defined by

$$f(g_1, \ldots, g_n; W) = \sum_{i=1}^{n} \int_{\mathbf{B}_W(0,1)} d(g_i \cdot w, W)^2 \, \mathrm{d}w.$$

The zero set of f is exactly

$$Z = \{(\mathbf{g}, W) \in G^n \times \mathrm{Gr}(V, k) \mid \forall i, g_i \in \mathrm{Stab}_G(W)\}.$$

By Łojasiewicz inequality (Theorem 3.8) applied to $\bar{U}^n \times \mathrm{Gr}(V, k)$, there is a constant $C > 0$ depending only on π such that for any $(\mathbf{g}, W) \in U^n \times \mathrm{Gr}(V, k)$,

$$f(\mathbf{g}, W) \geq \frac{1}{C} d\big((\mathbf{g}, W), Z\big)^C.$$

Assuming that $\pi(A)$ does not act $\frac{1}{nC}\rho^C$-irreducibly on V, we can find $W \in \mathrm{Gr}(V,k)$ such that for all $a \in A$ and all $w \in \mathbf{B}_W(0,1)$, $\pi(a)w \in W + \mathbf{B}_V(0, \frac{1}{C}\rho^C)$. Then it follows from the inequality above that there exists $W' \in \mathrm{Gr}(V,k)$ such that for all $a \in A$, $d(a, \mathrm{Stab}_G(W')) \leq \rho$, so that A is not ρ-away from proper stabilizers. □

5.1.3 Induction step

The core of the induction step in the proof of Theorem 5.2 is the following lemma.

Lemma 5.9. *Let G be a Lie group acting on a finite-dimensional normed vector space V. There exists a neighborhood U of the identity in G and a constant $C \geq 1$ such that for any parameters δ, ρ_1, ρ_2 with $0 < \delta \leq \rho_1 \leq \rho_2^4 \leq 1$, the following holds when ρ_1 is sufficiently small.*

Let $\pi : V \to V'$ be a quotient morphism of G-submodules with kernel $V_1 < V$. Let $A \subset U$ and $X \subset \mathbf{B}_V(0,1)$ and assume that

(i) $\langle A, X\rangle_3 \cap V_1^{(\delta)} \subset \mathbf{B}_V(0, \rho_1)$,

(ii) $\mathbf{B}_{V'}(0,\rho_2) \subset \pi(X)$,

(iii) *A is $\rho_1^{\frac{1}{C}}$-away from closed connected subgroups.*

Then there exists a submodule $W < V$ supplementary to V_1 such that

$$\text{the restriction } \pi_{|W} : W \to V' \text{ is } 3\rho_2^{-2}\text{-bi-Lipschitz}; \tag{5.2}$$

$$\mathbf{B}_W(0, \frac{1}{2}\rho_2) \subset X^{(\rho_1^{\frac{1}{C}})} \text{ and } X \cap V_1^{(\rho_2)} \subset W^{(\rho_1^{\frac{1}{C}})}. \tag{5.3}$$

Proof. On account of Lemma 5.8, we may assume that A is finite of cardinal $n \leq \dim(G)$ and is $\rho_1^{\frac{1}{C}}$-away from proper stabilizers. Shrinking again the neighborhood U if necessary, we can further assume that the action on V of any element in A is 2-bi-Lipschitz.

Let $\sigma : \mathbf{B}_{V'}(0,\rho_2) \to X$ be a section of the projection π, i.e. for any $y \in \mathbf{B}_{V'}(0,\rho_2)$,

$$\pi \circ \sigma(y) = y.$$

Let $x \in X \cap V_1^{(\rho_2)}$, $y, z \in \mathbf{B}_{V'}(0,\rho_2)$ and $a \in A$. The assumption (i) yields

$$\begin{aligned} \|x - \sigma(\pi(x))\| &\leq \rho_1; \\ \|\sigma(y)\| &\leq \rho_1 \quad \text{if } y \in \mathbf{B}_{V'}(0,\delta); \\ \|\sigma(y) + \sigma(z) - \sigma(y+z)\| &\leq \rho_1 \quad \text{if } y + z \in \mathbf{B}_{V'}(0,\rho_2); \\ \|a \cdot \sigma(y) - \sigma(a \cdot y)\| &\leq \rho_1 \quad \text{if } a \cdot y \in \mathbf{B}_{V'}(0,\rho_2). \end{aligned}$$

because we have respectively $x - \sigma(\pi(x)) \in (X - X) \cap V_1$, $\sigma(y) \in X \cap V_1^{(\delta)}$, $\sigma(y) + \sigma(z) - \sigma(y+z) \in 3X \cap V_1$ and $a \cdot \sigma(y) - \sigma(a \cdot y) \in (A \cdot X - X) \cap V_1$.

It follows from Lemma 5.10 that there exists a linear section $\varphi : V' \to V$ of π (i.e. $\pi \circ \varphi = \mathrm{Id}_{V'}$) such that for all $y \in \mathbf{B}_{V'}(0,\rho_2)$,

$$\|\varphi(y) - \sigma(y)\| \ll (-\log\delta)\rho_1.$$

The properties of σ imply the following properties for φ. For all $y \in V'$ and all $a \in A$,

$$\|y\| \leq \|\varphi(y)\| \leq 2\rho_2^{-1}\|y\|;$$
$$\|a \cdot \varphi(y) - \varphi(a \cdot y)\| \leq \rho_1^{\frac{1}{2}} \quad \text{if } y,\ a \cdot y \in \mathbf{B}_{V'}(0, \rho_2).$$

Let W_0 be the image subspace of φ. From the above, it follows that

$$\text{the restriction } \pi_{|W_0} \colon W_0 \to V' \text{ is } 2\rho_2^{-2}\text{-bi-Lipschitz;} \tag{5.4}$$

$$X \cap V_1^{(\rho_2)} \subset W_0 + \mathbf{B}_V(0, \rho_1^{\frac{1}{2}}); \tag{5.5}$$

$$\mathbf{B}_{W_0}(0, \rho_2) \subset X + \mathbf{B}_V(0, \rho_1^{\frac{1}{2}}); \tag{5.6}$$

$$\forall a \in A,\ \forall w \in \mathbf{B}_{W_0}(0,1), \quad d(a \cdot w, W_0)^2 \leq 4\rho_1\rho_2^{-2} \leq 4\rho_1^{\frac{1}{2}}. \tag{5.7}$$

Let $a_1, \ldots, a_n$ be the elements of A and write $\mathbf{a} = (a_1, \ldots, a_n)$. Consider the analytic function defined on $G^{\times n} \times \mathrm{Gr}(V, \dim(V'))$ defined by

$$f(g_1, \ldots, g_n; W) = \sum_{i=1}^{n} \int_{\mathbf{B}_W(0,1)} d(g_i \cdot w, W)^2 \, \mathrm{d}w.$$

By (5.7), we have $f(\mathbf{a}, W_0) \leq \rho_1^{\frac{1}{3}}$. By Łojasiewicz's inequality (Theorem 3.8) applied to the compact set $\bar{U}^{\times d} \times \mathrm{Gr}(V, \dim(V'))$, there exists a constant C depending only on the representation V such that for all $\mathbf{g} = (g_1, \ldots, g_n) \in U^{\times n}$ and $W \in \mathrm{Gr}(V, \dim(V'))$,

$$f(\mathbf{g}, W) \geq \frac{1}{C} d((\mathbf{g}, W), Z)^C,$$

where Z is the zero set of f. Therefore there exists $\mathbf{b} = (b_1, \ldots, b_n)$ and $W \in \mathrm{Gr}(V, \dim(V'))$ such that $f(\mathbf{b}, W) = 0$ and $d((\mathbf{a}, W_0), (\mathbf{b}, W)) \leq \rho_1^{\frac{1}{C}}$. The equality $f(\mathbf{b}, W) = 0$ exactly means that each b_i belongs to the stabilizer $\mathrm{Stab}_G(W)$, and hence

$$A \subset \mathrm{Stab}_G(W)^{(\rho_1^{\frac{1}{C}})}$$

But A is $\rho_1^{\frac{1}{C}}$-away from proper stabilizers, hence W must be a G-submodule. Finally, conclusions (5.2) and (5.3) follow from (5.4), (5.5), (5.6) and the fact that W is $\rho_1^{\frac{1}{C}}$-close to W_0. □

In the above proof, we made use of the following elementary lemma, a discretized version of the fact that any continuous additive map between two vector spaces is automatically linear.

Lemma 5.10 (Almost additive maps)**.** *Let $0 < \delta \leq \rho_1 \leq \rho_2 \leq 1$ be parameters. Let V' and V be finite dimensional normed vector spaces. If $\sigma \colon \mathbf{B}_{V'}(0, \rho_2) \to V$ satisfies*

(i) $\sigma(\mathbf{B}_{V'}(0,\delta)) \subset \mathbf{B}_V(0,\rho_1)$ *and*

(ii) for all $x, y \in \mathbf{B}_{V'}(0,\rho_2)$*, if* $x+y \in \mathbf{B}_{V'}(0,\rho_2)$ *then*

$$\sigma(x)+\sigma(y)-\sigma(x+y) \in \mathbf{B}_V(0,\rho_1).$$

Then there is a linear map $\varphi\colon V' \to V$ *such that for all* $x \in \mathbf{B}_{V'}(0,\rho_2)$*,*

$$\|\sigma(x)-\varphi(x)\| \ll_{V'} (-\log\delta+1)\rho_1.$$

Proof. Rescaling by a factor ρ_2^{-1} in V', we can suppose without loss of generality that $\rho_2 = 1$.

We first consider the special case where $V' = \mathbb{R}$. In this case define $\varphi\colon \mathbb{R} \to V$ to be the unique linear map such that $\varphi(1) = \sigma(1)$. From the assumption (ii), it follows that

$$\forall x \in [0,\frac{1}{2}], \quad \|2\sigma(x)-\sigma(2x)\| \le \rho_1.$$

Using this and a simple induction, we prove that

$$\forall n \in \mathbb{N}, \quad \|\sigma(2^{-n})-\varphi(2^{-n})\| \le \rho_1. \tag{5.8}$$

Let N be the integer such that $2^{-N} \le \delta < 2^{-N+1}$. It follows from (5.8) and the assumption (i) that

$$\|\varphi(2^{-N})\| \le 2\rho_1 \tag{5.9}$$

For any $x \in [0,1]$, let $(x_1,\dots,x_N) \in \{0,1\}^N$ be the N first digits in its binary expansion, i.e. for some $r \in [0,\delta]$, $x = \sum_{n=1}^N x_n 2^{-n} + r$. Then by the assumption (ii), (5.8) and (5.9),

$$\begin{aligned}\|\sigma(x)-\varphi(x)\| &\le \sum_{n=1}^N x_n\|\sigma(2^{-n})-\varphi(2^{-n})\| + \|\sigma(r)\| + 2^N r\|\varphi(2^{-N})\| + N\rho_1\\ &\le (2N+5)\rho_1.\end{aligned}$$

Consequently,

$$\begin{aligned}\|\sigma(-x)-\varphi(-x)\| &\le \|\varphi(x)-\sigma(x)\| + \|\sigma(-x)+\sigma(x)-\sigma(0)\| + \|\sigma(0)\|\\ &\le (2N+7)\rho_1.\end{aligned}$$

This finishes the proof for the special case. For general normed vector space V', pick a basis $(u_1,\dots,u_d)$ consisting of vectors of unit length then apply the special case to each partial function $t \mapsto \sigma(tu_i)$, $i = 1,\dots,d$. Then by (ii), we have the desired inequality for any vector in $\mathbf{B}_{V'}(0,1) \cap ([-1,1]u_1 + \cdots + [-1,1]u_d)$. This domain contains a ball $\mathbf{B}_{V'}(0,\frac{1}{k})$ where $k \in \mathbb{N}$ depends only on V' and the choice of the basis. We conclude by using k times the almost additivity (ii). □

We can now prove Theorem 5.2.

Proof of Theorem 5.2. The proof goes by induction on the length l of the Jordan-Hölder decomposition of V. The base case $l = 1$, where V is a nontrivial irreducible representation, corresponds to Theorem 5.5, and is proved above.

Assume that the result holds for all representations of length less than $l \geq 2$, let $V \in \mathcal{P}(G)$ be a representation of length l, and suppose $A \subset G$ and $X \subset V$ satisfy conditions (i)-(iii) of the theorem, for some small $\epsilon > 0$, to be specified later. Let $0 = V_0 < \ldots < V_l = V$ be the Jordan-Hölder sequence given by assumption (i). Write $V' = V/V_1$ and denote by $\pi \colon V \to V'$ the projection. Then the module V' has length $l-1$ and conditions (i) and (iii) are satisfied for A acting on $\pi(X) \subset V'$.

Note that it suffices to show that for some constant $C > 0$, $\langle A, X \rangle_s$ is $C\delta$-dense in $\mathbf{B}_V(0, \delta^{\epsilon_0})$ because we can repeat the argument at the smaller scale $\frac{\delta}{C}$. Consequently, we can replace X by $\langle A, X \rangle_s^{(\delta)}$ at any moment in the proof as long as s is a constant which depends only on the representation V and the parameters ϵ_0 and κ.

In what follows, we will use two methods to prove a subset is away from submodules. First, if X is ρ-away from submodules then its projections to any quotient module is also ρ-away from submodules. Secondly, if $X \subset \mathbf{B}_V(0,1)$ satisfies $\mathcal{N}_\delta(X) \geq \delta^{-\dim(V)+\epsilon}$ then X is $\delta^{O(\epsilon)}$-away from submodules.

First step: We first prove that there exists $\epsilon_1 > 0$ and $s_1 \geq 1$ depending on V, ϵ_0 and κ such that

$$\mathbf{B}_{V'}(0, \delta^{\epsilon_0}) \subset \pi(\langle A, X \rangle_{s_1} \cap \mathbf{B}_V(0, \delta^{\epsilon_1})) + \mathbf{B}_{V'}(0, \delta).$$

Let $\epsilon_1 > 0$ be a small parameter, whose precise value will be specified at the end of this step. By applying the induction hypothesis to V', whose length is at most $l-1$ and replacing X by $\langle A, X \rangle_s$, we can suppose that $\mathbf{B}_{V'}(0, \delta^{\epsilon_1}) \subset \pi(X)^{(\delta)}$. Cover X with $\delta^{-O(\epsilon_1)}$ balls of radius δ^{ϵ_1}, pick a ball B such that $\mathcal{N}_\delta(\pi(B \cap X))$ is maximal and translate it back to the origin to get

$$\mathcal{N}_\delta(\pi(X')) \geq \delta^{-\dim(V')+O(\epsilon_1)},$$

with $X' = (X - X) \cap \mathbf{B}_V(0, \delta^{\epsilon_1})$. This lower size bound ensures that $\pi(X')$ is $\delta^{O(\epsilon_1)}$-away from proper linear subspaces in V'. The induction hypothesis, applied to the subset $\pi(X') \subset V'$, with acting set A, yields the desired inclusion provided that ϵ_1 is small enough.

Second step: Assuming $X^{(\delta)} \cap V_1$ contains a large vector.

Let $s_2, \epsilon_2 > 0$ be the quantities given by Theorem 5.5 applied to the representation V_1, with constants κ, ϵ_1, and assume that there exists $v \in X^{(\delta)} \cap V_1$ with $\|v\| \geq \delta^{\epsilon_2}$. Then, using the base case for the action of G on the irreducible module V_1, we find that

$$\mathbf{B}_{V_1}(0, \delta^{\epsilon_1}) \subset \langle A, X^{(\delta)} \rangle_{s_2} + \mathbf{B}_V(0, \delta). \tag{5.10}$$

Recalling the inclusion obtained in the first step and setting $s = s_1 + O_{s_1}(s_2)$, we find

$$\mathbf{B}_V(0, \delta^{\epsilon_0}) \subset \langle A, X \rangle_s + \mathbf{B}_V(0, O_{s_1, s_2}(\delta)).$$

This finishes the proof of the theorem in this case.

Third step: Finally, we prove that there exists $s_3 \geq 1$ depending on V, ϵ_0 and κ such that $\langle A, X \rangle_{s_3}^{(\delta)} \cap V_1$ contains a vector of length at least δ^{ϵ_2}, which allows to conclude by using the second step.

Let $0 < \epsilon_3 < \frac{1}{4}\epsilon_2$ be a parameter whose value will be chosen later according to ϵ_2. Let $0 < \epsilon_4 < \epsilon_3$ be a parameter whose value will be chosen later according

to ϵ_3. Using the induction hypothesis for the representation V' with ϵ_4 and κ, we obtain

$$\mathbf{B}_{V'}(0,\delta^{\epsilon_4}) \subset \pi(\langle A, X\rangle_s^{(\delta)}),$$

where $s \geq 1$ is a constant depending only on V', ϵ_4 and κ. Replacing $\langle A, X\rangle_s^{(\delta)}$ by X, we may assume without loss of generality that

$$\mathbf{B}_{V'}(0,\delta^{\epsilon_4}) \subset \pi(X).$$

Lemma 5.9 applied with $\rho_1 = \delta^{\epsilon_2}$ and $\rho_2 = \delta^{\epsilon_4}$ gives a submodule $W < V$ such that $\pi_{|W}$ is $3\delta^{-\epsilon_4}$-bi-Lipschitz and

$$\mathbf{B}_W(0,\delta_1^{O(\frac{\epsilon_4}{\epsilon_2})}) \subset X^{(\delta_1)} \tag{5.11}$$

where $\delta_1 = \delta^{\frac{\epsilon_2}{C}}$. We want to apply base case Theorem 5.5 to the nontrivial irreducible representation V/W at scale δ_1 with ϵ_3 and κ. Observe that $\pi_{|W}$ being $3\delta^{-\epsilon_4}$-bi-Lipschitz implies that $\pi_{V/W|V_1} : V_1 \to V/W$ is $4\delta^{-\epsilon_4}$-bi-Lipschitz. Hence

$$\forall \rho \geq \delta, \quad \mathcal{N}_\rho(\pi_{V/W}(A)) \geq \delta^{O(\epsilon_4)}\mathcal{N}_\rho(\pi_{V_1}(A)).$$

Therefore when ϵ_4 and ϵ are small enough (according to V_1, ϵ_3 and κ),

$$\mathbf{B}_{V/W}(0,\delta_1^{\epsilon_3}) \subset \pi_{V/W}(\langle A, X\rangle_s) + \mathbf{B}_{V/W}(0,\delta_1),$$

where $s \geq 1$ is a constant depending only on V_1, ϵ_3 and κ. Together with inclusion (5.11), this implies that

$$\mathcal{N}_{\delta_1}(\langle A, X\rangle_{s+1}) \geq \delta_1^{-\dim V + O(\frac{\epsilon_3}{\epsilon_2})}.$$

Cutting $\langle A, X\rangle_{s+1}$ into cylinders of axis V_1 and diameter δ^{ϵ_3} and picking the part with largest size, we see that

$$\mathcal{N}_{\delta_1}(\langle A, X\rangle_{2s+2} \cap V_1^{(\delta^{\epsilon_3})}) \geq \delta_1^{-\dim V + O(\frac{\epsilon_3}{\epsilon_2})},$$

which ensures that $\langle A, X\rangle_{2s+2} \cap V_1^{(\delta^{\epsilon_3})}$ is $\delta_1^{O(\frac{\epsilon_3}{\epsilon_2})} = \delta^{O(\epsilon_3)}$-away from submodules.

Applying the induction hypothesis to A acting on $\pi(X) \subset V'$, we know that

$$\mathbf{B}_{V'}(0,\delta^{\epsilon_3}) \subset \pi(\langle A, X\rangle_{s'}^{(\delta)}),$$

where $s' \geq 1$ is a constant depending only on V', ϵ_3 and κ. Let $s'' = \max\{s', 2s+2\}$ and set $X' = \langle A, X\rangle_{s''} \cap V_1^{(\delta^{\epsilon_3})}$ so that X' is $\delta^{O(\epsilon_3)}$-away from submodules and

$$\mathbf{B}_{V'}(0,\delta^{\epsilon_3}) \subset \pi(X'^{(\delta)}).$$

At this stage apply Lemma 5.9 to the set $X'^{(\delta)}$ with $\rho_1 = \delta^{\epsilon_2}$ and $\rho_2 = \delta^{\epsilon_3}$. If ϵ_3 is chosen sufficiently small compared to ϵ_2, conclusion (5.3) fails. Therefore there must be $v \in \langle A, X'^{(\delta)}\rangle_3 \cap V_1^{(\delta)}$ with $\|v\| \geq \delta^{\epsilon_2}$.

This proves that $\langle A, X^{(\delta)}\rangle_{3s''} \cap V_1^{(\delta)}$ contains a vector of length at least δ^{ϵ_2}. If we did this argument at scale $\frac{1}{C}\delta$ with C a large constant depending on s'', we would get the claim of the third step. This finishes the proof of the theorem. $\square$

5.2 Product theorem for perfect Lie groups

In this section we prove Theorem 5.3.

Here, we adhere to the convention that the metric d on G is left-invariant and on any quotient G/N, the distance is defined by $x, y \in G$,

$$d(\bar{x}, \bar{y}) = \inf_{n,n' \in N} d(xn, yn') = d(y, xN) = d(x^{-1}y, N).$$

In particular the metric on G/N is left invariant and moreover if N' is another normal closed subgroup of G such that $N' < N$ then the canonical projection $G/N' \to G/N$ is 1-Lipschitz.

Unless the contrary is stated, throughout this section, G denotes a simply connected perfect Lie group and $\mathfrak{g}$ its Lie algebra. Moreover R denotes the radical of G and $\mathfrak{r}$ the radical of $\mathfrak{g}$.

5.2.1 Perfect Lie algebras and Lie groups

We will list some elementary and standard facts about perfect Lie groups and Lie algebras.

First, recall that $\mathfrak{r}$ is nilpotent since $\mathfrak{g}$ is perfect, see e.g. [2, Lemma 2.4]. Let $\mathfrak{s}$ be a Levi factor of $\mathfrak{g}$, i.e. a Lie subalgebra $\mathfrak{s}$ such that $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{r}$, see e.g. [54, Corollary 1, p. 49]. Note that the group theoretic Levi decomposition holds for G, i.e. there is a closed connected subgroup S of Lie algebra $\mathfrak{s}$ such that $G = S \ltimes R$, because since G/R is simply connected, a section of $\mathfrak{g} \to \mathfrak{g}/\mathfrak{r}$ integrates to a section of $G \to G/R$ (cf. [54, Theorem 1, p. 152]).

Lemma 5.11. *The image of a proper ideal of $\mathfrak{g}$ under the map $\mathfrak{g} \to \mathfrak{g}/\mathfrak{r}$ is a proper ideal. In particular, the image of a maximal proper ideal is a maximal proper ideal.*

Recall also that ideals in a semisimple Lie algebra are sums of its simple factors. Hence a maximal ideal is a sum of all simple factors except one.

Proof. Suppose the contrary, then $\mathfrak{n} + \mathfrak{r} = \mathfrak{g}$ for some proper ideal $\mathfrak{n} < \mathfrak{g}$. Denote by $\mathrm{D}^i\,\mathfrak{r}$, $i \geq 0$ the derived series of $\mathfrak{r}$, i.e. $\mathrm{D}^0\,\mathfrak{r} = \mathfrak{r}$ and $\mathrm{D}^{i+1}\,\mathfrak{r} = [\mathrm{D}^i\,\mathfrak{r}; \mathrm{D}^i\,\mathfrak{r}]$, $\forall i \geq 0$. We show by induction that $\forall i \geq 0$,

$$\mathfrak{g} = \mathfrak{n} + \mathrm{D}^i\,\mathfrak{r}, \tag{5.12}$$

which is impossible since $\mathfrak{r}$ is solvable and $\mathfrak{n}$ is proper. Indeed, (5.12) is true for $i = 0$. Suppose that it is true for some $i \geq 0$ then from $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$ follows that

$$\mathfrak{g} = [\mathfrak{n}, \mathfrak{n}] + [\mathfrak{n}, \mathrm{D}^i\,\mathfrak{r}] + [\mathrm{D}^i\,\mathfrak{r}, \mathrm{D}^i\,\mathfrak{r}] \subset \mathfrak{n} + \mathrm{D}^{i+1}\,\mathfrak{r},$$

since $\mathfrak{n}$ is an ideal. This is exactly (5.12) for $i + 1$ and finishes the proof of the lemma. □

Lemma 5.12 (Perfect abelian extension of a semisimple group)**.** *If $\mathfrak{r}$ is abelian, then the adjoint representation of G is of class $\mathcal{P}$.*

Remark. If G is not perfect, then $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is non-zero, and G acts trivially on $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$, so that the adjoint representation does not belong to $\mathcal{P}(G)$.

Proof of Lemma 5.12. We have an exact sequence of G-modules

$$0 \to \mathfrak{r} \to \mathfrak{g} \to \mathfrak{g}/\mathfrak{r} \to 0,$$

and by Proposition 5.4(i), all we need to check is that both $\mathfrak{r}$ and $\mathfrak{g}/\mathfrak{r}$ belong to $\mathcal{P}(G)$.

On the one hand, the representation $G \to \mathrm{GL}(\mathfrak{g}/\mathfrak{r})$ factors through G/R and the adjoint representation of the semisimple group G/R belongs to $\mathcal{P}(G/R)$. By Proposition 5.4(iii) , $\mathfrak{g}/\mathfrak{r}$ is of class $\mathcal{P}$ as a representation of G.

On the other hand, $\mathfrak{r}$ is totally reducible under the action of the semisimple group S, and moreover,

$$\mathfrak{r} = [\mathfrak{s}, \mathfrak{r}],$$

because $\mathfrak{g}$ is perfect and $\mathfrak{r}$ abelian. This implies that $\mathfrak{r}$ is a representation of class $\mathcal{P}$ for S, and therefore for G by Proposition 5.4(ii). □

Remark. It is not true in general that the adjoint representation of a perfect connected Lie group is of class $\mathcal{P}$; indeed, there exist perfect Lie algebras with nontrivial centers. For instance, let $G = \mathrm{SL}(2,\mathbb{R}) \ltimes \mathcal{F}_{2,2}$, where $\mathcal{F}_{2,2}$ is the free 2-nilpotent Lie algebra over 2 generators x, y, and the action of $\mathrm{SL}(2,\mathbb{R})$ is by linear substitution. The adjoint representation of G on its Lie algebra $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{R}) \ltimes \mathcal{F}_{2,2}$ is not of class $\mathcal{P}$, because G acts trivially on the center of $\mathfrak{g}$, generated by $[\mathsf{x}, \mathsf{y}]$.

5.2.2 Abelian extensions of semisimple groups

Here, we prove Theorem 5.3 in the case where the radical $\mathfrak{r}$ is abelian. In some sense, this is the most important case of the result, as we will see in 5.2.3 that the general case follows from this one.

To prove Theorem 5.3 in this case, the idea is to apply Theorem 5.2 to the adjoint representation of G on its Lie algebra, and then to use the Campbell-Hausdorff formula. Before that, we note that condition (i) in Theorem 5.3 automatically implies non-concentration for the image of A under any non-trivial group homomorphism.

Lemma 5.13. *Let H be another Lie group. Let $\varphi\colon G \to H$ be a nontrivial homomorphism. There exists a neighborhood U of the identity in G such that the following holds. Let $\epsilon > 0$ and $\kappa > 0$ be parameters and let $A \subset U$ be a subset satisfying condition (i) of Theorem 5.3. Then*

$$\forall \rho \geq \delta, \quad \mathcal{N}_\rho(\varphi(A)) \gg_\varphi \delta^\epsilon \rho^{-\kappa}.$$

Proof. The isomorphism $G/\ker\varphi \to \varphi(G)$ is bi-Lipschitz when restricted to compact neighborhoods. Hence without loss of generality, we can suppose that $H = G/\ker\varphi$. Since $\ker\varphi$ is closed, there exists a neighborhood U of the identity in G, such that $\forall x, y$, $d(x^{-1}y, \ker\varphi) = d(x^{-1}y, (\ker\varphi)^\circ)$. This allows us to further assume that $\ker\varphi$ is connected.

Let $\mathfrak{n}$ be a maximal proper ideal of $\mathfrak{g}$ containing the Lie algebra of $\ker\varphi$. By [54, Theorem 4, p. 154], $\mathfrak{n}$ is the Lie algebra of a normal closed connected subgroup $N < G$. This subgroup N contains $\ker\varphi$. By Lemma 5.11, $\mathfrak{n}$ is exactly the kernel of the projection of $\mathfrak{g}$ to one of the simple factors of $\mathfrak{g}/\mathfrak{r}$. It follows that G/N is one of the simple factors of G. We deduce the desired estimate

from condition (i) of Theorem 5.3 by using the fact that $G/\ker\varphi \to G/N$ is 1-Lipschitz. □

Proof of Theorem 5.3, case where $\mathfrak{r}$ is abelian. We start by the case where the radical $\mathfrak{r}$ of $\mathfrak{g}$ is abelian. In this proof, implied constants in Landau and Vinogradov notations depend G and the parameter κ.

By Lemma 5.12, the adjoint representation of G on $\mathfrak{g}$ is of class $\mathcal{P}$. We would like to apply Theorem 5.2 to A and $X = \log(AA^{-1} \cap \mathbf{B}_G(1, \delta^\epsilon))$. The hypotheses of Theorem 5.2 are all met with ϵ replaced by $O(\epsilon)$: assumption (i) is guaranteed by Lemma 5.13; A being a δ^ϵ-away from subgroups is exactly assumption (ii) of Theorem 5.3.

It remains to check that X is $\delta^{O(\epsilon)}$-away from any proper submodule W in $\mathfrak{g}$. We can assume W maximal. Hence it is a maximal proper ideal of $\mathfrak{g}$, which by Lemma 5.11 is the Lie algebra of $\ker\pi_{S_i}$ where S_i is a simple factor of G. In particular, there are only finitely many such W. Shrinking the neighborhood U if necessary, it suffices to check that $AA^{-1} \cap \mathbf{B}_G(1, \delta^\epsilon)$ is $\delta^{O(\epsilon)}$-away from $\ker\pi_{S_i}$. By assumption (i), for any $\rho \geq \delta$,

$$\begin{aligned}\mathcal{N}_\rho\big(\pi_{S_i}(AA^{-1} \cap \mathbf{B}_G(1, \delta^\epsilon))\big) &\geq \max_g \mathcal{N}_\rho\big(\pi_{S_i}(A \cap \mathbf{B}_G(g, \delta^\epsilon))\big)\\ &\geq \delta^{O(\epsilon)}\mathcal{N}_\rho(\pi_{S_i}(A))\\ &\geq \delta^{O(\epsilon)}\rho^{-\kappa}.\end{aligned}$$

The last quantity is larger than 1 if we choose $\rho = \delta^{C\epsilon}$ with a large $C = O(1)$. This shows that $AA^{-1} \cap \mathbf{B}_G(1, \delta^\epsilon)$ is $\delta^{O(\epsilon)}$-away from $\ker\pi_{S_i}$.

That is why we can use Theorem 5.2 to get an integer $s \geq 1$ such that

$$\mathbf{B}_\mathfrak{g}(0, \delta^{\epsilon_0}) \subset \langle A, X\rangle_s + \mathbf{B}_\mathfrak{g}(0, \delta) \tag{5.13}$$

when ϵ is small enough.

The idea is now to apply the Campbell-Hausdorff formula at an order l such that the error term is of size at most δ. We identify an element of the free group F_s generated by s elements and the word map $G^{\times s} \to G$ it induces. If x, y are elements in $\mathfrak{g}$, we want to approximate e^{x+y} by a word in e^x, e^y. For example, with a remainder term of order 2, $e^{x+y} = e^x e^y e^{O(\|x\|^2+\|y\|^2)}$. In order to get a remainder term of order 3, it is esasier to approximate $e^{2(x+y)}$, and then, we get $e^{2(x+y)} = (e^x)^2(e^y)^2(e^y)^2e^x(e^y)^{-2}(e^x)^{-1}e^{O(\|x\|^3+\|y\|^3)}$. We shall use the following lemma, which generalizes these elementary computations, and follows from the Campbell-Hausdorff formula.

Lemma 5.14. *Let* $\exp\colon \mathfrak{g} \to G$ *denote the exponential map of a Lie group. We fix a norm on* $\mathfrak{g}$ *and a distance on* G*. For all integers* $s \geq 1$ *and* $l \geq 1$*, there exists an integer* $C \geq 1$*, a word map* $w \in F_s$ *and a neighborhood* U *of* 0 *in* $\mathfrak{g}$ *such that for all* $x_1, \dots, x_s \in U$*,*

$$d\big(\exp(Cx_1 + \cdots + Cx_s), w(\exp x_1, \dots, \exp x_s)\big) \ll (\|x_1\| + \cdots + \|x_s\|)^l.$$

This lemma is a metric analogue of [1, Lemma 3.5] and the proofs are similar.

Proof. Consider $\mathfrak{g}$-valued functions f defined on a neighborhood of 0 in $\mathfrak{g}^{\times s}$ that can be written as a sum of a convergent series

$$f(x_1, \dots, x_s) = \sum_{k=1}^{+\infty} f_k(x_1, \dots, x_s)$$

where for each k, $f_k(x_1,\dots,x_s)$ is a $\mathbb{Q}$-linear combination of repeated brackets $[x_{i_1},\dots,x_{i_k}] = [x_{i_1},[x_{i_2},\dots,[x_{i_{l-1}},x_{i_k}]]]$ of length k. The series converges on $\mathbf{B}_{\mathfrak{g}}(0,r)^{\times s}$ for some $r>0$ in the sense that the numerical series obtained by replacing each repeated bracket of length k by r^k and each coefficient by its absolute value is convergent. Identifying two such functions if they agree on a neighborhood of 0, we get a linear space $\mathcal{G}_s$ over $\mathbb{Q}$. Equipped with its obvious Lie bracket, $\mathcal{G}_s$ is a graded Lie algebra over $\mathbb{Q}$. For $l\geq 1$, we write $O(d^\circ \geq l)$ to denote an unspecified element in $\mathcal{G}_s$ of valuation at least l.

By the Baker-Campbell-Hausdorff formula (cf. [26]), the map defined by $(x,y)\mapsto x*y=\log(\exp(x)\exp(y))$ belongs to $\mathcal{G}_2$ and moreover,

$$x*y = x+y+\frac{1}{2}[x,y]+O(d^\circ\geq 3). \tag{5.14}$$

From that we deduce, by an induction on s, that

$$x_1*\cdots*x_s = x_1+\cdots+x_s+O(d^\circ\geq 2). \tag{5.15}$$

We denote by $[x,y]_*$ the group commutator $x*y*(-x)*(-y)$ and by $[x_1,\dots,x_s]_*$ the repeated group commutator $[x_1,[x_2\dots,[x_{s-1},x_s]_*]_*]_*$. We have by (5.14),

$$[x,y]_* = [x,y]+O(d^\circ\geq 3)$$

and again by an induction on s,

$$[x_1,\dots,x_s]_* = [x_1,\dots,x_s]+O(d^\circ\geq s+1). \tag{5.16}$$

Now we prove by an induction l, that there exists an integer C_l and a word $w_l\in F_s$ such that

$$x_1+\cdots+x_s = w_l^*\big(\frac{x_1}{C_l},\dots,\frac{x_s}{C_l}\big)+O(d^\circ\geq l), \tag{5.17}$$

where w_l^* is the word map induced by w_l which is well-defined on a neighborhood of 0 in $\mathfrak{g}^{\times s}$. For $l=2$, this is given by (5.15). Suppose it is true for l and we will prove it for $l+1$. Let f be the sums of terms of degree l in the $R_{\geq l}$ on the right-hand side of (5.17). Since f has rational coefficients, there is an integer $C\geq 1$ such that we can write

$$f(x_1,\dots,x_s)=\sum_{i=1}^{N} m_i\big(\frac{x_1}{C},\dots,\frac{x_s}{C}\big)$$

where each m_i is a repeated bracket of length l. Hence by (5.16) and (5.15), there is $w'\in F_s$ a product of repeated commutators such that

$$f(x_1,\dots,x_s) = w'^*\big(\frac{x_1}{C},\dots,\frac{x_s}{C}\big)+O(d^\circ\geq l+1).$$

Hence

$$\begin{aligned} x_1+\cdots+x_s &= w_l^*\big(\frac{x_1}{C_l},\dots,\frac{x_s}{C_l}\big)+w'^*\big(\frac{x_1}{C},\dots,\frac{x_s}{C}\big)+O(d^\circ\geq l+1)\\ &= w_l^*\big(\frac{x_1}{C_l},\dots,\frac{x_s}{C_l}\big)*w'^*\big(\frac{x_1}{C},\dots,\frac{x_s}{C}\big)+O(d^\circ\geq l+1). \end{aligned}$$

In the last step we used the fact that $w'^*\big(\frac{x_1}{C},\dots,\frac{x_s}{C}\big)$ has valuation at least l. This finishes the proof of the induction step and concludes the proof of the lemma. □

We choose $l > \frac{1}{\epsilon}$ and apply the lemma to x_i of the form $x_i = \mathrm{Ad}(a_i)y_i$ with $a_i \in A^s$ and $y_i \in X$. By definition $X \subset \mathbf{B}_{\mathfrak{g}}(0, \delta^{\epsilon})$, so the error term is indeed of size $O_s(\delta^{l\epsilon}) = O(\delta)$, and therefore,

$$\begin{aligned}\exp[C\,\mathrm{Ad}(a_1)y_1 + \cdots + C\,\mathrm{Ad}(a_s)y_s] &\in w(a_1 e^{y_1} a_1^{-1}, \ldots, a_s e^{y_s} a_s^{-1})\,\mathbf{B}_G(1, O(\delta)) \\ &\in (A \cup \{1\} \cup A^{-1})^{s'}\,\mathbf{B}_G(1, O(\delta)),\end{aligned}$$

for some $s' = O_{s,l}(1)$. Recalling (5.13), we obtain

$$\begin{aligned}\mathbf{B}_G(1, \delta^{\epsilon_0}) &\subset \exp[C \cdot \mathbf{B}_{\mathfrak{g}}(0, \delta^{\epsilon_0})] \\ &\subset \exp[C \cdot \langle A, X\rangle_s + \mathbf{B}_{\mathfrak{g}}(0, C\delta))] \\ &\subset A^{s'}\,\mathbf{B}_G(1, O(\delta)).\end{aligned}$$

This finishes the proof of the theorem in the case $\mathfrak{r}$ is abelian. □

5.2.3 Proof of the product theorem, general case

We now have to explain how to deal with the case where the radical $\mathfrak{r}$ is nilpotent but not necessarily abelian. This will follow from the previous case, together with a quantitative version of the following fact: If R is a nilpotent Lie group, a subset $A \subset R$ generates the group R if and only if $A \mod [R, R]$ generates $R/[R, R]$.

For A and B subsets of G, we shall write $[A, B]$ to denote the set of all commutators $[a, b]$, $a \in A$, $b \in B$. This notation is in conflict with the group theoretic commutator which is the subgroup generated by all commutators. Despite this inconvenience, it will be clear from the context what $[A, B]$ means.

The precise lemma that we shall use is as follows.

Lemma 5.15. *Let R_i, $i \geq 1$ denote the lower central series of R then for $i \geq 1$ there is $k \geq 1$ such that for all $\rho > 0$ small enough,*

$$\mathbf{B}_{R_{i+1}}(1, \rho^2) \subset [\mathbf{B}_R(1, \rho), \mathbf{B}_{R_i}(1, \rho)]^k.$$

Proof. Denote by $\mathfrak{r}_i$, $i \geq 1$ the lower central series of the lie algebra $\mathfrak{r}$. Let $(z_1, \ldots, z_m)$ be a basis of $\mathfrak{r}_{i+1}$ consisting of commutators $z_j = [x_j, y_j]$ with $x_j \in \mathfrak{r}$ and $y_j \in \mathfrak{r}_i$, $\forall j$. Consider the map $f_j \colon \mathbb{R} \to R_{i+1}$ defined as

$$f_j(t) = \begin{cases} [\exp(\sqrt{t}x_j), \exp(\sqrt{t}y_j)] & \text{if } t \geq 0 \\ [\exp(\sqrt{-t}y_j), \exp(\sqrt{-t}x_j)] & \text{if } t < 0 \end{cases}$$

and further define $f \colon \mathbb{R}^m \to R_{i+1}$ by $f(t_1, \ldots, t_m) = f_1(t_1) \cdots f_m(t_m)$. Thus f is of class C^1 and its differential at 0 is

$$T_0 f(h_1, \ldots, h_m) = h_1 z_1 + \cdots + h_m z_m.$$

Hence f is a C^1-diffeomorphism on a neighborhood of 0. We conclude by using the fact that the exponential maps is a also bi-Lipschitz on a neighborhood of 0. □

We are now ready to finish the proof of Theorem 5.3.

Proof of Theorem 5.3, general case. Here again implied constants in Landau and Vinogradov notations depend on G and κ.

Let R_i, $i \geq 1$ denote the lower central series of the group R, i.e. $R_1 = R$ and $R_{i+1} = [R, R_i]$. Recall that since G is simply connected, these normal subgroups R_i , $i \geq 1$ are all closed and connected. The Lie algebra of R_i is exactly the i-th term in the lower central series of $\mathfrak{r}$, see e.g. [47, Theorem 5.7, p. 55]. The quotients G/R_i are simply connected. We proceed by induction on the nilpotency class l of R. We already see that Theorem 5.3 holds if $l \leq 1$. Now suppose that R has nilpotency class equal to l and that Theorem 5.3 has been proved if the nilpotency class is strictly less that l.

We first remark that the assumptions of Theorem 5.3 are preserved when projecting to a quotient. The nilpotency class of the radical of G/R_l is $l-1$. So by the induction hypothesis, we have when ϵ is small enough compared to ϵ_1,

$$\mathbf{B}_G(1, \delta^{\epsilon_1}) \subset (A \cup \{1\} \cup A^{-1})^s \, \mathbf{B}_G(1, \delta) R_l$$

where $\epsilon_1 > 0$ is a constant which we will choose according to ϵ_0 and $s \geq 1$ is an integer depending on κ and ϵ_1. Without loss of generality, we can replace $(A \cup \{1\} \cup A^{-1})^s \, \mathbf{B}_G(1, \delta)$ by A. In particular,

$$\mathbf{B}_R(1, \delta^{\epsilon_1}) \subset (R \cap A) R_l \quad \text{and} \quad \mathbf{B}_{R_{l-1}}(1, \delta^{\epsilon_1}) \subset (R_{l-1} \cap A) R_l.$$

By Lemma 5.15, we have

$$\mathbf{B}_{R_l}(1, \delta^{2\epsilon_1}) \subset [\mathbf{B}_R(1, \delta^{\epsilon_1}), \mathbf{B}_{R_{l-1}}(1, \delta^{\epsilon_1})]^{O(1)}.$$

From these inclusions and the fact that R_l is in the center of R, it follows that

$$\mathbf{B}_{R_l}(1, \delta^{2\epsilon_1}) \subset A^{O(1)} \, \mathbf{B}_G(1, O(\delta)). \tag{5.18}$$

At this stage replace $A^{O(1)} \, \mathbf{B}_G(1, O(\delta))$ by A. The fact that $\mathbf{B}_G(1, \delta^{\epsilon_1}) \subset A R_l$ and $\mathbf{B}_{R_l}(1, \delta^{2\epsilon_1}) \subset A$ does not prove what we want yet but gives the lower bound

$$\mathcal{N}_\delta(A^2) \gg_G \mathcal{N}_\delta(\pi_{G/R_l}(A)) \mathcal{N}_\delta(R_l \cap A) \gg_G \delta^{-\dim(G) + O(\epsilon_1)},$$

where $\pi_{G/R_l} \colon G \to G/R_l$ denote the canonical projection.

Covering A^2 by balls of radius $\frac{1}{2}\delta^{3\epsilon_1}$, we obtain

$$\mathcal{N}_\delta(A^2 A^{-2} \cap \mathbf{B}_G(1, \delta^{3\epsilon_1})) \gg_G \delta^{-\dim(G) + O(\epsilon_1)}.$$

Write $A' = A^2 A^{-2} \cap \mathbf{B}_G(1, \delta^{3\epsilon_1})$. Then A' satisfies the assumptions of Theorem 5.3 with $\kappa = 1$ and $\epsilon = O(\epsilon_1)$. Hence if ϵ_1 is small enough compared to ϵ_0, then by the induction hypothesis again,

$$\mathbf{B}_G(1, \delta^{\epsilon_0}) \subset A'^s \, \mathbf{B}_G(1, \delta) R_l$$

for some s depending on ϵ_0. Since any element in R_l involving in this inclusion is within distance $\delta^{2\epsilon_1}$ from the identity, we can conclude using (5.18) that

$$\mathbf{B}_G(1, \delta^{\epsilon_0}) \subset A'^{O(1)} \, \mathbf{B}_G(1, \delta) A.$$

This finishes the proof of Theorem 5.3. □

Bibliography

[1] M. Aka, E. Breuillard, L. Rosenzweig, and N. de Saxcé. Diophantine properties of nilpotent Lie groups. *Compos. Math.*, 151(6):1157–1188, 2015.

[2] Y. Benoist and N. de Saxcé. Convolution in perfect Lie groups. *Math. Proc. Cambridge Philos. Soc.*, 161(1):31–45, 2016.

[3] Y. Benoist and N. de Saxcé. A spectral gap theorem in simple Lie groups. *Invent. Math.*, 205(2):337–361, 2016.

[4] Y. Benoist and J.-F. Quint. Mesures stationnaires et fermés invariants des espaces homogènes. *Ann. of Math. (2)*, 174(2):1111–1162, 2011.

[5] B. Bollobás and A. Thomason. Projections of bodies and hereditary properties of hypergraphs. *Bull. London Math. Soc.*, 27(5):417–424, 1995.

[6] J. Bourgain. On the Erdős-Volkmann and Katz-Tao ring conjectures. *Geom. Funct. Anal.*, 13(2):334–365, 2003.

[7] J. Bourgain. The discretized sum-product and projection theorems. *J. Anal. Math.*, 112:193–236, 2010.

[8] J. Bourgain, A. Furman, E. Lindenstrauss, and S. Mozes. Stationary measures and equidistribution for orbits of nonabelian semigroups on the torus. *J. Amer. Math. Soc.*, 24(1):231–280, 2011.

[9] J. Bourgain and A. Gamburd. On the spectral gap for finitely-generated subgroups of SU(2). *Invent. Math.*, 171(1):83–121, 2008.

[10] J. Bourgain and A. Gamburd. Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$. *Ann. of Math. (2)*, 167(2):625–642, 2008.

[11] J. Bourgain and A. Gamburd. A spectral gap theorem in $\mathrm{SU}(d)$. *J. Eur. Math. Soc. (JEMS)*, 14(5):1455–1511, 2012.

[12] J. Bourgain and A. Glibichuk. Exponential sum estimates over a subgroup in an arbitrary finite field. *J. Anal. Math.*, 115:51–70, 2011.

[13] J. Bourgain, N. Katz, and T. Tao. A sum-product estimate in finite fields, and applications. *Geom. Funct. Anal.*, 14(1):27–57, 2004.

[14] J. Bourgain and P. P. Varjú. Expansion in $SL_d(\mathbf{Z}/q\mathbf{Z})$, q arbitrary. *Invent. Math.*, 188(1):151–173, 2012.

[15] J. Bourgain and A. Yehudayoff. Expansion in $\mathrm{SL}_2(\mathbb{R})$ and monotone expanders. *Geom. Funct. Anal.*, 23(1):1–41, 2013.

[16] R. Boutonnet, A. Ioana, and A. S. Golsefidy. Local spectral gap in simple lie groups and applications. *Invent. math.*, 2016.

[17] E. Breuillard. Lectures on approximate groups. IHP, Paris, February-March 2011, available at `http://www.math.u-psud.fr/~breuilla/ClermontLectures.pdf`.

[18] E. Breuillard, B. Green, and T. Tao. Approximate subgroups of linear groups. *Geom. Funct. Anal.*, 21(4):774–819, 2011.

[19] E. Breuillard, B. Green, and T. Tao. The structure of approximate groups. *Publ. Math. Inst. Hautes Études Sci.*, 116:115–221, 2012.

[20] M.-C. Chang. A polynomial bound in Freiman's theorem. *Duke Math. J.*, 113(3):399–419, 2002.

[21] M.-C. Chang. Additive and multiplicative structure in matrix spaces. *Combin. Probab. Comput.*, 16(2):219–238, 2007.

[22] N. de Saxcé. Subgroups of fractional dimension in nilpotent or solvable Lie groups. *Mathematika*, 59(2):497–511, 2013.

[23] N. de Saxcé. Trou dimensionnel dans les groupes de Lie compacts semisimples via les séries de Fourier. *J. Anal. Math.*, 120:311–331, 2013.

[24] N. de Saxcé. A product theorem in simple Lie groups. *Geom. Funct. Anal.*, 25(3):915–941, 2015.

[25] N. de Saxcé. Borelian subgroups of simple Lie groups. *Duke Math. J.*, 166(3):573–604, 2017.

[26] E. B. Dynkin. Calculation of the coefficients in the Campbell-Hausdorff formula. *Doklady Akad. Nauk SSSR (N.S.)*, 57:323–326, 1947.

[27] G. A. Edgar and C. Miller. Borel subrings of the reals. *Proc. Amer. Math. Soc.*, 131(4):1121–1129, 2003.

[28] P. Erdős and E. Szemerédi. On sums and products of integers. In *Studies in pure mathematics*, pages 213–218. Birkhäuser, Basel, 1983.

[29] P. Erdős and B. Volkmann. Additive Gruppen mit vorgegebener Hausdorffscher Dimension. *J. Reine Angew. Math.*, 221:203–208, 1966.

[30] A. Eskin, S. Mozes, and H. Oh. On uniform exponential growth for linear groups. *Invent. Math.*, 160(1):1–30, 2005.

[31] K. Falconer, J. Fraser, and X. Jin. Sixty years of fractal projections. In *Fractal geometry and stochastics V. Selected papers of the 5th conference, Tabarz, Germany, March 24–29, 2014*, pages 3–25. Cham: Springer, 2015.

[32] K. J. Falconer. Hausdorff dimension and the exceptional set of projections. *Mathematika*, 29(1):109–115, 1982.

[33] H. Furstenberg. Stiffness of group actions. In *Lie groups and ergodic theory (Mumbai, 1996)*, volume 14 of *Tata Inst. Fund. Res. Stud. Math.*, pages 105–117. Tata Inst. Fund. Res., Bombay, 1998.

[34] W. T. Gowers. Quasirandom groups. *Combin. Probab. Comput.*, 17(3):363–387, 2008.

[35] P. Griffiths and J. Harris. *Principles of algebraic geometry*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1994. Reprint of the 1978 original.

[36] H. A. Helfgott. Growth and generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$. *Ann. of Math. (2)*, 167(2):601–623, 2008.

[37] H. A. Helfgott. Growth in $\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})$. *J. Eur. Math. Soc. (JEMS)*, 13(3):761–851, 2011.

[38] N. H. Katz and T. Tao. Some connections between Falconer's distance set conjecture and sets of Furstenburg type. *New York J. Math.*, 7:149–187, 2001.

[39] Y. Katznelson. *An introduction to harmonic analysis*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2004.

[40] S. Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002.

[41] E. Lindenstrauss and N. de Saxcé. Hausdorff dimension and subgroups of $SU(2)$. *Israel J. Math.*, 209(1):335–354, 2015.

[42] S. Łojasiewicz. Ensembles semi-analytiques, 2006. Notes from a course given in Orsay, available at `https://perso.univ-rennes1.fr/michel.coste/Lojasiewicz.pdf`.

[43] J. M. Marstrand. Some fundamental geometrical properties of plane sets of fractional dimensions. *Proc. London Math. Soc. (3)*, 4:257–302, 1954.

[44] P. Mattila. Hausdorff dimension, orthogonal projections and intersections with planes. *Ann. Acad. Sci. Fenn. Ser. A I Math.*, 1(2):227–244, 1975.

[45] P. Mattila. *Geometry of sets and measures in Euclidean spaces*, volume 44 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. Fractals and rectifiability.

[46] P. Mattila. *Fourier analysis and Hausdorff dimension*. Cambridge: Cambridge University Press, 2015.

[47] A. Onishchik and E. Vinberg. Foundations of Lie theory. In *Lie groups and Lie algebras I. Foundations of Lie theory. Lie transformation groups. Transl. from the Russian by A. Kozlowski*, page 1. Berlin: Springer-Verlag, 1988.

[48] T. Orponen. A discretised projection theorem in the plane. *ArXiv e-prints 1407.6543*, July 2014.

[49] G. Petridis. New proofs of Plünnecke-type estimates for product sets in groups. *Combinatorica*, 32(6):721–733, 2012.

[50] R. S. Pierce. *Associative algebras*, volume 88 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982. Studies in the History of Modern Science, 9.

[51] L. Pyber and E. Szabó. Growth in finite simple groups of Lie type. *J. Amer. Math. Soc.*, 29(1):95–146, 2016.

[52] P. Sarnak and X. X. Xue. Bounds for multiplicities of automorphic representations. *Duke Math. J.*, 64(1):207–227, 1991.

[53] J.-P. Serre. Applications algébriques de la cohomologie des groupes. ii : théorie des algèbres simples. In *Séminaire Henri Cartan, 1950/51, Exp. 6*, pages 1–9. Secrétariat mathématique, Paris, 1950/1951. Available at `http://www.numdam.org/item?id=SHC_1950-1951__3__A6_0`.

[54] J.-P. Serre. *Lie algebras and Lie groups. 1964 lectures, given at Harvard University. 2nd ed.* Berlin etc.: Springer-Verlag, 2nd ed. edition, 1992.

[55] P. Shmerkin. Projections of self-similar and related fractals: a survey of recent developments. In *Fractal geometry and stochastics V. Selected papers of the 5th conference, Tabarz, Germany, March 24–29, 2014*, pages 53–74. Cham: Springer, 2015.

[56] P. Shmerkin. On Furstenberg's intersection conjecture, self-similar measures, and the L^q norms of convolutions. *ArXiv e-prints 1609.07802*, Sept. 2016.

[57] T. Tao. Product set estimates for non-commutative groups. *Combinatorica*, 28(5):547–594, 2008.

[58] T. Tao. The sum-product phenomenon in arbitrary rings. *Contrib. Discrete Math.*, 4(2):59–82, 2009.

[59] T. Tao and V. H. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[60] P. P. Varjú. Expansion in $SL_d(\mathcal{O}_K/I)$, I square-free. *J. Eur. Math. Soc. (JEMS)*, 14(1):273–305, 2012.

[61] M. Wu. A proof of Furstenberg's conjecture on the intersections of $\times p$ and $\times q$-invariant sets. *ArXiv e-prints 1609.08053*, Sept. 2016.

université PARIS-SACLAY | École doctorale de mathématiques Hadamard (EDMH)

Titre : Sommes, produits et projections des ensembles discrétisés

Mots Clefs : Phénomène somme-produit, entropie métrique, théorème de projection, groupes de Lie, théorème produit

Résumé : Dans le cadre discrétisé, la taille d'un ensemble à l'échelle δ est évaluée par son nombre de recouvrement par δ-boules (également connu sous le nom de l'entropie métrique). Dans cette thèse, nous étudions les propriétés combinatoires des ensembles discrétisés sous l'addition, la multiplication et les projections orthogonales. Il y a trois parties principales. Premièrement, nous démontrons un théorème somme-produit dans les algèbres de matrices, qui généralise un théorème somme-produit de Bourgain concernant l'anneau des réels. On améliore aussi des estimées somme-produit en dimension supérieure obtenues précédemment par Bougain et Gamburd. Deuxièmement, on étudie les projections orthogonales des sous-ensembles de l'espace euclidien et étend ainsi le théorème de projection discrétisé de Bourgain aux projections de rang supérieur. Enfin, dans un travail en commun avec Nicolas de Saxcé, nous démontrons un théorème produit dans les groupes de Lie parfaits. Ce dernier résultat généralise les travaux antérieurs de Bourgain-Gamburd et de Saxcé.

Title : Sums, products and projections of discretized sets

Keys words : Sum-product phenomenon, metric entropy, projection theorem, Lie groups, product theorem

Abstract : In the discretized setting, the size of a set is measured by its covering number by δ-balls (a.k.a. metric entropy), where δ is the scale. In this document, we investigate combinatorial properties of discretized sets under addition, multiplication and orthogonal projection. There are three parts. First, we prove sum-product estimates in matrix algebras, generalizing Bourgain's sum-product theorem in the ring of real numbers and improving higher dimensional sum-product estimates previously obtained by Bourgain-Gamburd. Then, we study orthogonal projections of subsets in the Euclidean space, generalizing Bourgain's discretized projection theorem to higher rank situations. Finally, in a joint work with Nicolas de Saxcé, we prove a product theorem for perfect Lie groups, generalizing previous results of Bourgain-Gamburd and Saxcé.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France