



Weikun He · Nicolas de Saxcé

## Sum-product for real Lie groups

Received June 8, 2018

**Abstract.** We prove a discretized sum-product theorem for representations of Lie groups whose Jordan–Hölder decomposition does not contain the trivial representation. This expansion result is used to derive a product theorem in perfect Lie groups.

**Keywords.** Additive combinatorics, covering numbers, expansion

### 1. Introduction

Throughout this paper,  $G$  will denote a connected real Lie group, endowed with a left-invariant Riemannian metric. For  $x \in G$  and  $\rho > 0$ , we denote by  $B_G(x, \rho)$  the ball of center  $x$  and radius  $\rho$  in  $G$ . For  $A \subset G$  and  $\rho > 0$ ,  $A^{(\rho)}$  stands for the  $\rho$ -neighborhood of  $A$  and  $N(A, \rho)$  stands for the covering number of  $A$  by  $\rho$ -balls, i.e.

$$N(A, \rho) = \min \left\{ N \in \mathbb{N} \mid \exists x_1, \dots, x_N \in G, A \subset \bigcup_{i=1}^N B_G(x_i, \rho) \right\}.$$

The same notation is used for other metric spaces.

#### 1.1. Sum-product theorem in representations of Lie groups

In the first part of this paper, we study the sum-product phenomenon in representations of Lie groups. We shall work with some linear representation of  $G$  over some finite-dimensional real vector space  $V$ , endowed with some norm. We shall also refer to representations of  $G$  as  $G$ -modules. For  $A \subset G$ ,  $X \subset V$  and  $s \geq 1$ , we denote by  $\langle A, X \rangle_s$  the set of elements in  $V$  that can be obtained as combinations of sums, differences and products of at most  $s$  elements from  $A$  and  $X$ .

Note that the distance on  $G$  induces a natural distance on each of its quotients. Let  $N \triangleleft G$  be a closed normal subgroup. We denote by  $\pi_{G/N}: G \rightarrow G/N$  the canonical

W. He: Korea Institute for Advanced Study,  
Seoul 02455, Republic of Korea; e-mail: heweikun@kias.re.kr

N. de Saxcé: LAGA, UMR 7539, CNRS, Université Paris Nord,  
93430 Villetaneuse, France; e-mail: desaxce@math.univ-paris13.fr

*Mathematics Subject Classification (2020):* 22E30, 22E45, 05B99

projection. Then there is a unique distance on  $G/N$  satisfying  $d(\pi_{G/N}(x), \pi_{G/N}(y)) = d(x^{-1}y, N)$  for all  $x, y \in G$ . Throughout this paper, all quotients  $G/N$  will be endowed with this distance.

Following [23], we say that a subset  $A \subset G$  is  $\rho$ -away from closed connected subgroups for some parameter  $\rho > 0$  if for any proper closed connected subgroup  $H < G$ , there exists  $a \in A$  with  $d(a, H) > \rho$ . Similarly, we say that a subset  $X \subset V$  is  $\rho$ -away from submodules if for any proper  $G$ -submodule  $W < V$ , there exists  $x \in X$  with  $d(x, W) > \rho$ .

We say that a  $G$ -module  $V$  is in  $\mathcal{P}(G)$  if the trivial representation does not appear as a simple quotient in the Jordan–Hölder decomposition of  $G$ —see Definition 2.1.

**Theorem 1.1** (Sum-product theorem in representations of class  $\mathcal{P}$ ). *Let  $G$  be a connected real Lie group and  $V \in \mathcal{P}(G)$ . There exists a neighborhood  $U$  of the identity in  $G$  such that, for every  $\varepsilon_0, \kappa > 0$ , there exist  $s \geq 1$  and  $\varepsilon > 0$  such that the following holds for any  $\delta > 0$  sufficiently small. Assume  $A \subset U$  and  $X \subset B_V(0, 1)$  satisfy*

(i) *for any proper closed connected normal subgroup  $N \triangleleft G$ ,*

$$\forall \rho \geq \delta, \quad N(\pi_{G/N}(A), \rho) \geq \delta^\varepsilon \rho^{-\kappa};$$

- (ii)  *$A$  is  $\delta^\varepsilon$ -away from closed connected subgroups;*
- (iii)  *$X$  is  $\delta^\varepsilon$ -away from submodules.*

Then

$$B_V(0, \delta^{\varepsilon_0}) \subset \langle A, X \rangle_s^{(\delta)}.$$

This is a bounded generation statement which can be viewed as a generalization of previous sum-product results in rings. For example, applying to  $G = \mathbb{R}^*$  acting on  $V = \mathbb{R}$ , one recovers a version Bourgain’s discretized sum-product theorem [3, 4]. Note that this is not a new proof of Bourgain’s theorem as the latter is used as an ingredient in the proof of Theorem 1.1.

In Section 2, Theorem 1.1 will be proved in a more precise form where the conditions (i) and (ii) can be slightly relaxed; see Theorem 2.3.

### 1.2. Product theorem in perfect Lie groups

In the second part of this paper, we use Theorem 1.1 to derive a product theorem in perfect Lie groups. For subsets  $A, B \subset G$  of a Lie group  $G$ , we denote by  $AB$  their product set, i.e.

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

For  $k \geq 2$ , we denote by  $A^k$  the  $k$ -fold product set of  $A$  with itself,  $A \cdots A$ . To avoid confusion with Cartesian products between sets, we write  $A^{\times k}$  for the Cartesian power  $A \times \cdots \times A$ .

Recall that a Lie group is perfect if its Lie algebra  $\mathfrak{g}$  is perfect, i.e.  $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$ .

**Theorem 1.2** (Product theorem in perfect Lie groups). *Let  $G$  be a connected perfect Lie group. There exists a neighborhood  $U$  of the identity in  $G$  such that given  $\kappa > 0$ , there exists  $\varepsilon > 0$  such that the following holds for  $\delta > 0$  sufficiently small. Let  $A$  be a subset of  $U$  such that*

- (i)  $N(A, \delta) \leq \delta^{-\dim G + \kappa}$ ;
- (ii) for any proper closed connected normal subgroup  $N \triangleleft G$ ,

$$\forall \rho \geq \delta, \quad N(\pi_{G/N}(A), \rho) \geq \delta^\varepsilon \rho^{-\kappa};$$

- (iii)  $A$  is  $\delta^\varepsilon$ -away from closed connected subgroups in  $G$ .

Then

$$N(AAA, \delta) \geq \delta^{-\varepsilon} N(A, \delta).$$

For  $G = \text{SU}(2)$ , the above theorem was proved by Bourgain–Gamburd [5], and for a general simple Lie group by the second author [23], borrowing many ideas from the work of Bourgain and Gamburd [6] on the spectral gap property in  $\text{SU}(d)$ .

It is not difficult to see that the assumption of perfectness is optimal for such a product theorem to hold, in the following sense.

**Proposition 1.3.** *Let  $G$  be a simply connected Lie group which is not perfect with Lie algebra  $\mathfrak{g}$ . Write  $d = \dim \mathfrak{g} - \dim [\mathfrak{g}, \mathfrak{g}]$ . For any neighborhood  $U$  of the identity in  $G$ , for any  $\kappa \in (0, 1)$  and for any  $\delta > 0$  small enough, there exists  $A \subset U$  such that*

- (i)  $N(A, \delta) \approx_U \delta^{-\dim G + d(1-\kappa)}$ ;
- (ii) for any proper closed connected normal subgroup  $N \triangleleft G$ ,

$$\forall \rho \geq \delta, \quad N(\pi_{G/N}(A), \rho) \gg_U \rho^{-\kappa};$$

- (iii)  $A$  is  $\frac{1}{O_U(1)}$ -away from closed connected subgroups in  $G$ ,

but

$$N(AAA, \delta) \ll_U N(A, \delta).$$

Note that in a closely related setting, Salehi Golsefidy and Varjú [20] have already observed that perfectness is a sufficient and necessary condition for an expansion result to hold. In fact, at different places, our arguments share some conceptual similarities with the recent work of Salehi Golsefidy [18, 19] on super-approximation. Also, these examples of approximate subgroups can be seen as discretized analogues of measurable subgroups of intermediate dimension whose existence is known in abelian Lie groups [10] and solvable Lie groups [21].

We shall prove Theorem 1.2 and Proposition 1.3 in Section 3.

### 1.3. Applications

We conclude this introduction by mentioning several applications to Theorems 1.1 and 1.2 above. The first is that the spectral gap property discovered by Bourgain and Gamburd [5, 6] in the setting of  $\text{SU}(d)$ ,  $d \geq 2$ , generalizes to all compact semisimple Lie

groups. The proof follows the strategy developed by Bourgain and Gamburd, as presented for instance in [2], where the theorem is proved for compact simple Lie groups, but we now use Theorem 1.2 above instead of the second author's Product Theorem in simple Lie groups [23].

**Theorem 1.4.** *Let  $G$  be a compact connected semisimple Lie group, with Lie algebra  $\mathfrak{g}$ , and write  $L_0^2(G)$  for the space of zero-mean square-integrable functions on  $G$ . Let  $\mu$  be a probability measure on  $G$  whose support generates a dense subgroup in  $G$ . Assume moreover that in some basis for  $\mathfrak{g}$ , for every  $g \in \text{Supp } \mu$ , the matrix of  $\text{Ad } g$  has algebraic entries. Then the convolution operator*

$$T_\mu : L_0^2(G) \rightarrow L_0^2(G), \quad f \mapsto f * \mu,$$

*satisfies  $\|T_\mu\| < 1$ .*

The local spectral gap property introduced by Boutonnet, Ioana and Salehi Golsefidy [8] for non-compact Lie groups can also be generalized to a general connected perfect Lie group, but in order to keep statements as elementary as possible, we do not make this precise here.

Originally, discretized expansion statements were introduced by Katz and Tao [12] and used by Bourgain [3] to study Hausdorff dimensions of sum-sets and product-sets in  $\mathbb{R}$  and give a quantitative solution to the Erdős–Volkman conjecture: If  $A$  is any Borel measurable subset of  $\mathbb{R}$  with  $\dim_H A \in (0, 1)$ , then  $\dim_H(A + AA) > \dim_H A$ . The theorems proven here have similar consequences on the Hausdorff dimension of sum and product sets in semisimple algebras or perfect Lie groups. In particular, it should be possible to generalize the results of the second author presented in [22] to the setting of perfect Lie groups; we hope to address these matters in another paper.

Another nice application of Theorem 1.1 is the very recent work of Li [14, 15] on the regularity of the Furstenberg measure associated to a random walk on a semisimple Lie group.

Finally, our results could be used to construct new family of expanders, in the spirit of the works of Bourgain–Yehudayoff [7] or Vigolo [26].

## 2. Sum-product theorem in representations of $G$

The goal of this section is to prove Theorem 1.1 from the introduction. In fact, our proof will yield a slightly more precise version—see Theorem 2.3.

### 2.1. Representations without trivial simple quotients

We now define the class of representations to which our theorem will apply, and gather some elementary properties. Then, we state the refined version of Theorem 1.1 which will be proved here, Theorem 2.3.

**Definition 2.1.** Let  $G$  be a connected Lie group. We let  $\mathcal{P}(G)$  denote the class of finite-dimensional linear representations  $V$  of  $G$  such that there exists a sequence  $\{0\} = V_0 < V_1 < \dots < V_\ell = V$  of subrepresentations of  $V$  such that, for each  $i = 0, \dots, \ell - 1$ , the quotient representation  $V_{i+1}/V_i$  is *non-trivial* 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 2.2.** *Let  $V$  be a representation of a connected Lie group  $G$ .*

- (i) *If  $W$  is 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  $G \rightarrow \text{GL}(V)$  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 those of  $V/W$ . 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 1.** The class  $\mathcal{P}(G)$  is the smallest class of finite-dimensional  $G$ -modules that contains all non-trivial irreducible representations of  $G$  and is closed under extension (i.e., if  $W$  and  $V'$  are in  $\mathcal{P}(G)$ , and  $0 \rightarrow W \rightarrow V \rightarrow V' \rightarrow 0$  is a short exact sequence of  $G$ -modules, then  $V$  is in  $\mathcal{P}(G)$ ).

**Example 1.** • 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  $\mathcal{P}(G)$ .

- Let  $n$  be a positive integer. 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  $\mathcal{P}(G)$ .

Throughout this article, we shall consider representations of  $G$  as normed vector spaces: By *normed  $G$ -module*, we mean a  $G$ -module endowed with a norm which makes the underlying linear space a normed vector space.

Whenever  $V$  is a normed vector space, and  $W \leq V$  is a linear subspace, we shall always consider on  $W$  the norm induced by the norm on  $V$ , and on the quotient space  $V' = V/W$  the norm given by the formula

$$\forall v \in V, \quad \|\pi(v)\| = d(v, W),$$

where  $\pi : V \rightarrow V'$  is the canonical projection. Finally, we endow the space  $\text{End}(V)$  of linear endomorphisms with the associated operator norm.

Let  $\rho \in (0, 1/2)$  be a parameter and  $V$  a normed  $G$ -module. We say that a subset  $X \subset V$  is  $\rho$ -away from submodules if for every proper submodule  $W < V$ , there exists  $x \in X$  such that  $d(x, W) \geq \rho$ . Similarly, a subset  $A \subset G$  is said to be  $\rho$ -away from closed connected subgroups if for every proper closed connected subgroup  $H$ , there exists  $a \in A$  such that  $d(a, H) \geq \rho$ . Finally, a subset  $A \subset G$  is said to be  $\rho$ -away from identity components of proper stabilizers if for any subspace  $W \subset V$  which is not a  $G$ -submodule, there exists  $a \in A$  such that  $d(a, (\text{Stab}_G W)^\circ) \geq \rho$ , where  $(\text{Stab}_G W)^\circ$  denotes the identity component of the stabilizer  $\text{Stab}_G W$  of  $W$  in  $G$ .

We shall prove the following.

**Theorem 2.3** (Sum-product in representations of class  $\mathcal{P}$ ). *Let  $G$  be a connected real Lie group and  $V$  a normed  $G$ -module. There exists a neighborhood  $U$  of the identity in  $G$  such that, for every  $\varepsilon_0, \kappa > 0$ , there exist  $s \geq 1$  and  $\varepsilon > 0$  such that the following holds for any  $\delta > 0$  sufficiently small. Assume  $A \subset U$  and  $X \subset B_V(0, 1)$  satisfy*

- (i) *there is a Jordan–Hölder sequence  $0 = V_0 < \dots < V_\ell = V$  such that for every  $i = 0, \dots, \ell - 1$ ,*

$$\forall \rho \geq \delta, \quad N(p_{V_{i+1}/V_i}(A), \rho) \geq \delta^\varepsilon \rho^{-\kappa},$$

*where  $p_{V_{i+1}/V_i} : G \rightarrow \text{GL}(V_{i+1}/V_i)$  denotes the representation of  $G$  on  $V_{i+1}/V_i$ ;*

- (ii)  *$A$  is  $\delta^\varepsilon$ -away from identity components of proper stabilizers;*  
 (iii)  *$X$  is  $\delta^\varepsilon$ -away from submodules.*

Then

$$B_V(0, \delta^{\varepsilon_0}) \subset \langle A, X \rangle_s + B_V(0, \delta).$$

Note that assumption (i) implies that  $V$  is of class  $\mathcal{P}(G)$ . The proof goes by induction on the length of  $V$  (i.e. the length of any Jordan–Hölder decomposition of  $V$ ). We shall prove the base case, where  $V$  is a non-trivial irreducible representation, in the next subsection. The induction step will then be carried out in Subsection 2.3.

## 2.2. Irreducible representations

In the case  $V$  is an irreducible representation of  $G$ , the above theorem is a variant of [11, Theorem 3]. For clarity, we restate our theorem in this particular case. Then, we shall explain how to derive it from the first author's sum-product theorem in simple algebras [11, Theorem 2].

**Theorem 2.4** (Base case: irreducible representations). *Let  $G$  be a connected real Lie group and  $\pi_V : G \rightarrow \text{GL}(V)$  a non-trivial irreducible representation. There exists a neighborhood  $U$  of the identity in  $G$  such that, for every  $\varepsilon_0, \kappa > 0$ , there exist  $s \geq 1$  and  $\varepsilon > 0$  such that the following holds for any  $\delta > 0$  sufficiently small. Assume  $A \subset U$  and  $X \subset B_V(0, 1)$  satisfy*

- (i) for every  $\rho \geq \delta$ ,  $N(\pi_V(A), \rho) \geq \delta^\varepsilon \rho^{-\kappa}$ ;
- (ii)  $A$  is  $\delta^\varepsilon$ -away from identity components of proper stabilizers;
- (iii) there exists  $v \in X$  such that  $\|v\| \geq \delta^\varepsilon$ .

Then

$$B_V(0, \delta^{\varepsilon_0}) \subset \langle A, X \rangle_s + B_V(0, \delta).$$

The proof of this theorem is based on Proposition 2.5 below, a sum-product statement for matrix representations, which is essentially contained in [11]. Below,  $A$  denotes a subset of  $\text{End}(V)$  for some real vector space  $V$ , and  $\langle A \rangle_s$  denotes the set of elements in  $\text{End}(V)$  that can be obtained as combinations of sums and products of at most  $s$  elements in  $A$ . If  $V$  is a real vector space and  $A$  is a subset of  $\text{End } V$ , and if  $\rho \in (0, 1/2)$  is a parameter, we say that  $A$  acts  $\rho$ -irreducibly on  $V$  if for every non-trivial proper linear subspace  $W < V$ , there exist  $v \in B_W(0, 1)$  and  $a \in A$  such that  $d(a \cdot v, W) \geq \rho$ .

**Proposition 2.5** (Sum-product in irreducible representations). *Let  $V$  be a finite-dimensional normed vector space. Given  $\varepsilon_0, \kappa > 0$ , there exist  $s \geq 1$  and  $\varepsilon > 0$  such that the following holds. Let  $A \subset B_{\text{End}(V)}(0, \delta^{-\varepsilon})$  and  $v \in V$ . Assume that*

- (i) for every  $\rho \geq \delta$ ,  $N(A, \rho) \geq \delta^\varepsilon \rho^{-\kappa}$ ;
- (ii)  $A$  acts  $\delta^\varepsilon$ -irreducibly on  $V$ ;
- (iii)  $\delta^\varepsilon \leq \|v\| \leq \delta^{-\varepsilon}$ .

Then

$$B_V(0, \delta^{\varepsilon_0}) \subset \langle A \rangle_s \cdot v + B_V(0, \delta).$$

*Proof.* Given  $\varepsilon_1 > 0$ , it follows from [11, Proposition 31] that there exists  $c > 0$  such that, provided  $\varepsilon > 0$  is small enough, there exists a  $\delta^{-O(\varepsilon)}$ -bi-Lipschitz linear bijection  $f: V \rightarrow K^n$ , 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  $\delta_1$  with  $\delta \leq \delta_1 \leq \delta^c$  such that

$$f A f^{-1} \subset \text{Mat}_n(K) + B(0, \delta_1)$$

and such that for every proper real subalgebra  $F < \text{End}(K^n)$ ,

$$\exists a \in A : d(f a f^{-1}, F) \geq \delta_1^{\varepsilon_1}.$$

Choosing  $\varepsilon_1$  small enough in terms of  $\varepsilon_0$  and  $\kappa$ , we may then apply [11, Theorem 5] to conclude that, provided  $\varepsilon > 0$  is sufficiently small, for some integer  $s$ ,

$$B_{\text{Mat}_n(K)}(0, \delta_1^{\varepsilon_0}) \subset f \langle A \rangle_s f^{-1} + B_{\text{Mat}_n(K)}(0, \delta_1).$$

Therefore, without loss of generality, we may assume that  $V = K^n$  and

$$B_{\text{Mat}_n(K)}(0, \delta_1^{\varepsilon_0}) \subset A + B_{\text{Mat}_n(K)}(0, \delta_1). \tag{1}$$

We can further assume that  $\|v\| = 1$ . Then

$$B_V(0, \delta_1^{\varepsilon_0}) \subset A \cdot v + B_V(0, \delta_1).$$

In other words, the conclusion of the proposition holds at scale  $\delta_1$ . It remains to bring the scale back to  $\delta$ . To do this, we note that from (1), we have in particular

$$\delta_1^{1/2} \text{id} \in A + B_{\text{Mat}_n(K)}(0, \delta_1).$$

Hence, starting from (1), we may multiply both sides by  $\delta_1^{1/2} \text{id}$  to obtain

$$B_V(0, \delta_1) \subset B_V(0, \delta_1^{\varepsilon_0+1/2}) \subset \langle A \rangle_2 \cdot v + B_V(0, 2\delta_1^{3/2}),$$

and iterating this procedure, we get a sequence of integers  $s_2 = 1, s_3 = 2, \dots$  such that for any  $k \geq 2$ ,

$$B_V(0, s_k \delta_1^{k/2}) \subset \langle A \rangle_{s_{k+1}} \cdot v + B_V(0, s_{k+1} \delta_1^{(k+1)/2}).$$

Choose  $k > 2/c$  so that  $s_k \delta_1^{k/2} \leq \delta$ . Combining all these inclusions, we find, for  $s = s_2 + \dots + s_k$ ,

$$B_V(0, \delta^{\varepsilon_0}) \subset \langle A \rangle_s \cdot v + B_V(0, \delta),$$

which proves the proposition. □

The above proposition readily implies Theorem 2.4.

*Proof of Theorem 2.4.* It suffices to apply Proposition 2.5 to the set  $\pi_V(A) \subset \text{End}(V)$ . By the assumption on  $A$ , conditions (i) and (iii) of the proposition are satisfied for the set  $\pi_V(A)$ . That condition (ii) is also satisfied is a consequence of Lemma 2.6 below. □

**Lemma 2.6.** *Let  $0 < \rho < 1/2$  be a parameter. Let  $\pi : G \rightarrow \text{GL}(V)$  be a non-trivial irreducible representation. There is a neighborhood  $U$  of 1 in  $G$  such that if  $A \subset U$  is  $\rho$ -away from identity components of proper stabilizers then  $\pi(A)$  acts  $\rho^{O_\pi(1)}$ -irreducibly on  $V$ .*

It might be interesting to understand how the implied constant in  $O_\pi(1)$  depends on  $\pi$  and  $G$ , especially if one wants to have explicit constants in Theorems 1.1 and 1.2. Our proof does not provide any insight on this matter since we rely on an application of Łojasiewicz’s inequality and effectivizing Łojasiewicz’s inequality can be difficult and involves mathematics far away from the scope of the present paper (see for example [13] and references therein).

The proof of this lemma is an application of Łojasiewicz’s inequality (which we recall below as Theorem 2.8). First, it is convenient to reduce to the case where  $A$  is finite. This reduction is the subject of the next lemma. Given a representation  $\pi : G \rightarrow \text{GL}(V)$  of  $G$ , a subset  $A \subset G$  and a parameter  $\rho \in (0, 1/2)$ , we say that  $A$  is  $\rho$ -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  $\text{Stab}_G W$  is at least  $\rho$ .

**Lemma 2.7.** *Let  $0 < \rho < 1/2$  be a parameter. Let  $\pi : G \rightarrow \text{GL}(V)$  be a representation. There is a neighborhood  $U$  of 1 in  $G$  such that if  $A \subset U$  is  $\rho$ -away from identity components of proper stabilizers then  $A$  is  $\rho^{O_\pi(1)}$ -away from proper stabilizers. In fact,  $A$  contains a subset of cardinality at most  $\dim G$  which is  $\rho^{O_\pi(1)}$ -away from proper stabilizers.*



*Proof.* The representation  $\pi$  differentiates to a representation of the Lie algebra  $\mathfrak{g}$  of  $G$ , which we denote by  $T_1\pi : \mathfrak{g} \rightarrow \text{End}(V)$ . The stabilizer of  $W$  in  $\mathfrak{g}$ ,

$$\text{Stab}_{\mathfrak{g}} W = \{x \in \mathfrak{g} \mid T_1\pi(x)W \subset W\},$$

is the Lie algebra of  $\text{Stab}_G W$ . In particular, its image under the exponential map is contained in  $(\text{Stab}_G W)^\circ$ , the identity component of  $\text{Stab}_G W$ . We may assume that  $\exp$  induces a diffeomorphism from  $U$  to its image, and denote the inverse map by  $\log$ . Say that  $\log A$  is  $\rho$ -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  $d(\log a, \text{Stab}_{\mathfrak{g}} W) \geq \rho$ .

We claim that there is a neighborhood  $U$  of 1 in  $G$  such that if  $A \subset U$  is  $\rho$ -away from identity components of proper stabilizers then  $\log A$  is  $\rho/C$ -away from proper stabilizers in  $\mathfrak{g}$  and conversely if  $\log A$  is  $\rho$ -away from proper stabilizers then  $A$  is  $\rho/C$ -away from proper stabilizers.

Let us prove this claim. Let  $x \in \mathfrak{g}$ . From the identity  $\pi(e^x) = e^{T_1\pi(x)}$ , we can express  $T_1\pi(x)$  as an absolutely convergent series

$$T_1\pi(x) = -\sum_{n \geq 1} \frac{1}{n} (\text{id}_V - \pi(e^x))^n$$

whenever  $\|\pi(e^x) - \text{id}_V\| < 1$ . Therefore, if  $e^x \in \text{Stab}_G W$  is such that the above series converges, then  $x \in \text{Stab}_{\mathfrak{g}} W$ . It follows that there is  $r > 0$  depending only on  $\pi$  such that

$$B_G(1, r) \cap \text{Stab}_G W \subset \exp(\text{Stab}_{\mathfrak{g}} W).$$

Let  $U = B_G(1, r/2)$ . Then for any  $g \in U$  and any proper linear subspace  $W$ ,

$$\frac{1}{C} d(g, (\text{Stab}_G W)^\circ) \leq d(\log g, \text{Stab}_{\mathfrak{g}} W) \leq C d(g, \text{Stab}_G W)$$

where  $C > 0$  is some constant depending only on the representation. This proves our claim, and the first part of the lemma.

For the second part, one can reproduce the argument in [23, Lemma 2.5] to show that if  $\log A$  is  $\rho$ -away from proper stabilizers then  $\log A$  contains a subset of cardinality at most  $\dim \mathfrak{g}$  which is  $\rho^{O(\dim \mathfrak{g})}$ -away from stabilizers. □

**Remark 2.** Note that in the above lemma, the neighborhood  $U$  depends on the representation  $\pi$ , and not only on  $G$ . This is readily seen by considering  $G = \mathbb{R}$ ,  $V = \mathbb{C} \simeq \mathbb{R}^2$ , and  $\pi(x)v = e^{inx}v$ ,  $n \in \mathbb{N}$ .

Let us recall Łojasiewicz’s inequality [16, Théorème 2, p. 62] before we proceed to prove Lemma 2.6.

**Theorem 2.8** (Łojasiewicz’s inequality). *Let  $M$  be a real-analytic manifold endowed with a Riemannian distance  $d$  and let  $f : M \rightarrow \mathbb{R}$  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  $x \in K$ ,*

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

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

*Proof of Lemma 2.6.* Let  $U$  be the neighborhood given by Lemma 2.7. On account of that lemma we may assume that  $A$  is finite of cardinality  $n \leq \dim G$  and  $\rho$ -away from proper stabilizers.

Let  $0 < k < \dim V$  and consider the real-analytic map  $f : G^{\times n} \times \text{Grass}(k, V) \rightarrow \mathbb{R}$  defined by

$$f(g_1, \dots, g_n, W) = \sum_{i=1}^n \int_{B_W(0,1)} d(g_i \cdot w, W)^2 dw.$$

The integration is with respect to the  $k$ -dimensional Lebesgue measure on each  $W \in \text{Grass}(k, V)$ . To see that this map is real-analytic, observe that we can locally trivialize the tautological bundle over  $\text{Grass}(k, V)$ , i.e. for any  $W_0 \in \text{Grass}(k, V)$  there is a neighborhood  $\mathcal{U}$  of  $W_0$  in  $\text{Grass}(k, V)$  and a map  $\varphi : \mathcal{U} \times \mathbb{R}^k \rightarrow V$  such that for all  $W \in \text{Grass}(k, V)$ , the map  $\varphi(W, \cdot) : \mathbb{R}^k \rightarrow V$  is linear and has  $W$  as image. Moreover, we can choose  $\varphi$  to be real-analytic and such that  $\varphi(W, \cdot) : \mathbb{R}^k \rightarrow W$  is an isometry for the Euclidean structures, for any  $W \in \text{Grass}(k, V)$ . Thus, for all  $W \in \text{Grass}(k, V)$  and all  $i = 1, \dots, n$ ,

$$\int_{B_W(0,1)} d(g_i \cdot w, W)^2 dw = \int_{B_{\mathbb{R}^k}(0,1)} \|(g_i \cdot \varphi(W, v)) \wedge \varphi(W, e_1) \wedge \dots \wedge \varphi(W, e_k)\|^2 dv$$

where  $(e_1, \dots, e_k)$  is the standard basis of  $\mathbb{R}^k$ . The right-hand side is obviously real-analytic.

The zero set of  $f$  is exactly

$$Z = \{(\mathbf{g}, W) \in G^{\times n} \times \text{Grass}(k, V) \mid \forall i, g_i \in \text{Stab}_G W\}.$$

By Łojasiewicz’s inequality (Theorem 2.8) applied on  $\bar{U}^{\times n} \times \text{Grass}(k, V)$ , there is a constant  $C > 0$  such that for any  $(\mathbf{g}, W) \in U^{\times n} \times \text{Grass}(k, V)$ ,

$$f(\mathbf{g}, W) \geq \frac{1}{C} d((\mathbf{g}, W), Z)^C.$$

Assuming that  $A$  does not act  $\frac{1}{C}\rho^C$ -irreducibly on  $V$ , we can find  $W \in \text{Grass}(k, V)$  such that for all  $a \in A$  and all  $w \in B_W(0, 1)$ ,  $\pi(a)w \in W + B_V(0, \frac{1}{C}\rho^C)$ . So  $f(a_1, \dots, a_n, W) \leq \frac{1}{C}\rho^C$ , and by the inequality above there exists  $W' \in \text{Grass}(k, V)$  such that for all  $a \in A$ ,  $d(a, \text{Stab}_G W') \leq \rho$ , so that  $A$  is not  $\rho$ -away from proper stabilizers.  $\square$

### 2.3. Induction step

The core of the induction step in the proof of Theorem 2.3 is the following lemma. It is a quantitative discretized version of the following elementary fact: let  $V$  be a  $G$ -module, and  $V_1, X$  two submodules of  $V$ ; if  $\pi : V \rightarrow V/V_1$  maps  $X$  onto  $V/V_1$  and if  $X \cap V_1 = \{0\}$ , then  $V = X \oplus V_1$ . Once more, the proof relies on Łojasiewicz’s inequality.

**Lemma 2.9.** *Let  $G$  be a connected Lie group and  $V$  a normed  $G$ -module. There exist a neighborhood  $U$  of the identity in  $G$  and a constant  $C \geq 1$  such that for any parameters  $0 < \eta < \tau < 1$ , the following holds when  $\delta$  is sufficiently small. Let  $V_1$  be a proper submodule of  $V$  and  $\pi : V \rightarrow V/V_1$  the canonical projection. Let  $A \subset U$  and  $X \subset B_V(0, 1)$  and assume that*

- (i)  $\langle A, X \rangle_3 \cap V_1^{(\delta)} \subset B_V(0, \delta^{C\tau})$ ;
- (ii)  $\pi(X) = B_{V/V_1}(0, \delta^\eta)$ ,
- (iii)  $A$  is  $\delta^\tau$ -away from identity components of proper stabilizers.

Then there exists a submodule  $W < V$  such that

- (a) the restriction  $\pi|_W : W \rightarrow V/V_1$  is  $3\delta^{-\eta}$ -bi-Lipschitz;
- (b)  $B_W(0, \delta^\eta) \subset X^{(\delta^\tau)}$  and  $X \subset W^{(\delta^\tau)}$ .

*Proof.* For convenience, we write  $V' = V/V_1$ . On account of Lemma 2.7, which gives us the neighborhood  $U$ , we may assume that  $A$  is finite of cardinality  $n \leq \dim G$  and is  $\delta^{C_1\tau}$ -away from proper stabilizers, where  $C_1 \geq 2$  is a constant depending only on  $V$ . Shrinking again the neighborhood  $U$  if necessary, we can ensure that the action on  $V$  of any element in  $A$  is 2-bi-Lipschitz.

Assumption (ii) allows us to pick a section  $\sigma : B_{V'}(0, \delta^\eta) \rightarrow X$  of the projection  $\pi$ , i.e. for any  $y \in B_{V'}(0, \delta^\eta)$ ,

$$\pi \circ \sigma(y) = y.$$

The choice of such a  $\sigma$  is arbitrary. In fact, thanks to assumption (i), different choices only differ by at most  $\delta^{C\tau}$ . Indeed, for any  $x \in X$ , we have  $x - \sigma(\pi(x)) \in (X - X) \cap V_1$  and therefore, by assumption (i),

$$\|x - \sigma(\pi(x))\| \leq \delta^{C\tau}. \tag{2}$$

Again from assumption (i), it follows that  $\sigma$  is almost a morphism of  $G$ -modules, in the sense that for all  $y, z \in B_{V'}(0, \delta^\eta)$  and all  $a \in A$ ,

$$\|\sigma(y)\| \leq \delta^{C\tau} \quad \text{if } y \in B_{V'}(0, \delta); \tag{3}$$

$$\|\sigma(y) + \sigma(z) - \sigma(y + z)\| \leq \delta^{C\tau} \quad \text{if } y + z \in B_{V'}(0, \delta^\eta); \tag{4}$$

$$\|a \cdot \sigma(y) - \sigma(a \cdot y)\| \leq \delta^{C\tau} \quad \text{if } a \cdot y \in B_{V'}(0, \delta^\eta). \tag{5}$$

Indeed, we have, respectively,  $\sigma(y) \in X \cap V_1^{(\delta)}$ ,  $\sigma(y) + \sigma(z) - \sigma(y + z) \in 3X \cap V_1$  and  $a \cdot \sigma(y) - \sigma(a \cdot y) \in (A \cdot X - X) \cap V_1$ .

In particular, (3) and (4) say that  $\sigma$  is almost additive; by Lemma 2.10 below,  $\sigma$  is close to a genuine linear map. More precisely, there exists a linear section  $\varphi : V' \rightarrow V$  of  $\pi$  (i.e.  $\pi \circ \varphi = \text{Id}_{V'}$ ) such that for all  $y \in B_{V'}(0, \delta^\eta)$ ,

$$\|\varphi(y) - \sigma(y)\| \leq \delta^{(C-1)\tau} \tag{6}$$

provided  $\delta$  is small enough. From the linearity of  $\varphi$ , the fact that  $X \subset B_V(0, 1)$ , and (2), (5) and (6), we deduce that for all  $y \in V'$ , all  $a \in A$  and all  $x \in X$ ,

$$\begin{aligned} \|\varphi(y)\| &\leq 2\delta^{-\eta}\|y\|; \\ \|a \cdot \varphi(y) - \varphi(a \cdot y)\| &\leq \delta^{(C-3)\tau}\|y\|; \\ \|x - \varphi(\pi(x))\| &\leq \delta^{(C-2)\tau}. \end{aligned}$$

Let  $W_0$  be the image subspace of  $\varphi$ . From the above, it follows that

the restriction  $\pi|_{W_0} : W_0 \rightarrow V'$  is  $2\delta^{-\eta}$ -bi-Lipschitz; (7)

$$X \subset W_0 + B_V(0, \delta^{(C-2)\tau}); \tag{8}$$

$$B_{W_0}(0, \delta^\eta/2) \subset \varphi(B_{V'}(0, \delta^\eta)) \subset X + B_V(0, \delta^{(C-1)\tau}); \tag{9}$$

$$\forall a \in A, \forall w \in B_{W_0}(0, 1), \quad d(a \cdot w, W_0) \leq \delta^{(C-3)\tau}. \tag{10}$$

The inequality (10) says that  $W_0$  is almost invariant under the action of  $A$ . We now use Łojasiewicz’s inequality to show that it is close to a  $G$ -submodule. Let  $a_1, \dots, a_n$  be the elements of  $A$  and write  $\mathbf{a} = (a_1, \dots, a_n)$ . Consider the real-analytic function on  $G^{\times n} \times \text{Grass}(\dim V', V)$  defined by

$$f(g_1, \dots, g_n, W) = \sum_{i=1}^n \int_{B_W(0,1)} d(g_i \cdot w, W)^2 dw.$$

From (10) it follows that  $f(\mathbf{a}, W_0) \leq \delta^{(2C-7)\tau}$  provided  $\delta$  is small enough. By Łojasiewicz’s inequality (Theorem 2.8) applied to the compact set  $\bar{U}^{\times n} \times \text{Grass}(\dim V', V)$ , there exists a constant  $C_2$  depending only on the representation  $V$  such that for all  $\mathbf{g} = (g_1, \dots, g_n) \in U^{\times n}$  and  $W \in \text{Grass}(\dim V', V)$ ,

$$f(\mathbf{g}, W) \geq \frac{1}{C_2} d((\mathbf{g}, W), Z)^{C_2}$$

where  $Z$  is the zero set of  $f$ . Therefore, there exist  $\mathbf{b} = (b_1, \dots, b_n) \in G^{\times n}$  and  $W \in \text{Grass}(\dim V', V)$  such that  $f(\mathbf{b}, W) = 0$  and

$$d((\mathbf{a}, W_0), (\mathbf{b}, W)) \leq \delta^{C_1\tau}$$

provided  $2C - 7 \geq (C_1 + 1)C_2$ . The equality  $f(\mathbf{b}, W) = 0$  exactly means that each  $b_i$  belongs to the stabilizer  $\text{Stab}_G W$ , and hence

$$A \subset (\text{Stab}_G W)^{(\delta^{C_1\tau})}$$

But  $A$  is  $\delta^{C_1\tau}$ -away from proper stabilizers, hence  $W$  must be a  $G$ -submodule. Finally, conclusions (a) and (b) follow from (7)–(9) and the fact that  $W$  is  $\delta^{C_1\tau}$ -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 2.10** (Almost additive maps). *Let  $0 < \delta < \rho_1 < \rho_2 \leq 1$  be parameters. Let  $V$  and  $V'$  be finite-dimensional normed vector spaces. If  $\sigma : B_{V'}(0, \rho_2) \rightarrow V$  satisfies*

- (i)  $\sigma(B_{V'}(0, \delta)) \subset B_V(0, \rho_1)$ ;
- (ii) *for all  $x, y \in B_{V'}(0, \rho_2)$ , if  $x + y \in B_{V'}(0, \rho_2)$  then*

$$\sigma(x) + \sigma(y) - \sigma(x + y) \in B_V(0, \rho_1),$$

*then there is a linear map  $\varphi : V' \rightarrow V$  such that for all  $x \in B_{V'}(0, \rho_2)$ ,*

$$\|\sigma(x) - \varphi(x)\| \ll_{V'} (-\log \delta + 1)\rho_1.$$

*Moreover, if there are linear maps  $\pi : V \rightarrow V''$  and  $\psi : V' \rightarrow V''$  such that  $\pi \circ \sigma = \psi$  on  $B_{V'}(0, \rho_2)$ , then we may also ensure that  $\pi \circ \varphi = \psi$  on  $V'$ .*

*Proof.* We first consider the special case where  $\rho_2 = 1$  and  $V' = \mathbb{R}$ . In this case define  $\varphi : \mathbb{R} \rightarrow V$  to be the unique linear map such that  $\varphi(1) = \sigma(1)$ . From assumption (ii), it follows that

$$\forall x \in [0, 1/2], \quad \|2\sigma(x) - \sigma(2x)\| \leq \rho_1.$$

Using this and a simple induction, we prove that

$$\forall n \in \mathbb{N}, \quad \|\sigma(2^{-n}) - \varphi(2^{-n})\| \leq \rho_1. \tag{11}$$

Let  $N$  be the integer such that  $2^{-N} \leq \delta < 2^{-N+1}$ . It follows from (11) and assumption (i) that

$$\|\varphi(2^{-N})\| \leq 2\rho_1. \tag{12}$$

For any  $x \in [0, 1]$ , let  $(x_1, \dots, x_N) \in \{0, 1\}^N$  be the first  $N$  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 assumption (ii), (11) and (12),

$$\begin{aligned} \|\sigma(x) - \varphi(x)\| &\leq \sum_{n=1}^N x_n \|\sigma(2^{-n}) - \varphi(2^{-n})\| + \|\sigma(r)\| + 2^N r \|\varphi(2^{-N})\| + N\rho_1 \\ &\leq (2N + 5)\rho_1. \end{aligned}$$

Consequently,

$$\begin{aligned} \|\sigma(-x) - \varphi(-x)\| &\leq \|\varphi(x) - \sigma(x)\| + \|\sigma(-x) + \sigma(x) - \sigma(0)\| + \|\sigma(0)\| \\ &\leq (2N + 7)\rho_1. \end{aligned}$$

This proves the lemma in the case  $V' = \mathbb{R}$  and  $\rho_2 = 1$ . For a general normed vector space  $V'$ , in the case  $\rho_2 = 1$ , pick a basis  $(u_1, \dots, u_d)$  consisting of vectors of unit length and apply the special case to each function  $\sigma_i : t \mapsto \sigma(tu_i)$ ,  $i = 1, \dots, d$ . This yields linear maps  $\varphi_1, \dots, \varphi_d : \mathbb{R} \rightarrow V$ , and we define  $\varphi : V' \rightarrow V$  by  $\varphi(t_1u_1 + \dots + t_du_d) = \varphi_1(t_1) + \dots + \varphi_d(t_d)$ . Then by (ii), we have the desired inequality for any vector in  $B_{V'}(0, 1) \cap ([-1, 1]u_1 + \dots + [-1, 1]u_d)$ . This domain contains a ball  $B_{V'}(0, 1/k)$  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).

The general case  $\rho_2 \leq 1$  follows from the case  $\rho_2 = 1$ , by considering the map  $\sigma' : V' \rightarrow V$  defined by  $\sigma'(x) = \sigma(\rho_2 x)$ .

The “moreover” part is clear from the proof. □

We are now ready to prove Theorem 2.3. The main idea is to use induction on the length of the module. Note that among the assumptions of Theorem 2.3, (ii) is preserved by passing to any submodule or any quotient of  $V$  and (iii) is preserved by passing to any quotient. Finally, (i) passes to the quotient  $V/V_1$  of  $V$  by the first submodule  $V_1$  in the Jordan–Hölder decomposition. Thus by the induction hypothesis, it is easy to produce a large ball in  $V/V_1$ . Then it can be proved (this is done in the third step of the proof below) that we can produce a large vector in  $V_1$  and hence a large ball in  $V_1$  by the base case. Then a technical difficulty arises: a large ball in  $V/V_1$  and a large ball in  $V_1$  do not add up to a ball in  $V$ . To deal with this difficulty we need to produce the large ball in  $V/V_1$  using only vectors of controlled length (this is done in the first step of the proof below). Another technical difficulty is in the third step where we want to produce a vector in  $V_1$  of length  $\geq \delta^{\varepsilon_2}$  for any given  $\varepsilon_2 > 0$ . The idea is that, otherwise we could apply Lemma 2.9 to conclude that  $X$  is trapped in a submodule, which would contradict assumption (iii).

*Proof of Theorem 2.3.* The proof goes by induction on the length  $\ell$  of the module  $V$ . The base case  $\ell = 1$ , where  $V$  is a non-trivial irreducible representation, corresponds to Theorem 2.4, and is proved above. Assume that the result holds for all representations of length less than  $\ell \geq 2$ , let  $V \in \mathcal{P}(G)$  be a representation of length  $\ell$ , and suppose  $A \subset G$  and  $X \subset V$  satisfy conditions (i)–(iii) of the theorem, for some small  $\varepsilon > 0$ . Let  $0 = V_0 < \dots < V_\ell = V$  be the Jordan–Hölder sequence given by assumption (i). Write  $V' = V/V_1$  and denote by  $\pi_{V'}: V \rightarrow V'$  the projection. Then the module  $V'$  has length  $\ell - 1$  and as noted above, the conditions in Theorem 2.3 are satisfied for  $A$  acting on  $\pi_{V'}(X) \subset V'$ .

*First step.* We first prove that there exist  $\varepsilon_1 > 0$  and  $s_1 \geq 1$  depending on  $V$ ,  $\varepsilon_0$  and  $\kappa$  such that

$$B_{V'}(0, \delta^{\varepsilon_0}) \subset \pi_{V'}(\langle A, X \rangle_{s_1} \cap B(0, \delta^{\varepsilon_1})) + B_{V'}(0, \delta).$$

Let  $\varepsilon_1 \in (0, \varepsilon_0)$  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  $\ell - 1$ , and replacing  $X$  by  $\langle A, X \rangle_s$ , we may assume that  $B_{V'}(0, \delta^{\varepsilon_1}) \subset \pi_{V'}(X)^{(\delta)}$ . Cover  $X$  with  $\delta^{-O(\varepsilon_1)}$  balls of radius  $\delta^{2\varepsilon_1}$ , pick a ball  $B$  such that  $N(\pi_{V'}(B \cap X), \delta)$  is maximal, and translate it back to the origin to get

$$N(\pi_{V'}(X'), \delta) \geq \delta^{-\dim V' + O(\varepsilon_1)}$$

with  $X' = (X - X) \cap B_{V'}(0, \delta^{2\varepsilon_1})$ . This lower bound ensures that  $\pi_{V'}(X')$  is  $\delta^{O(\varepsilon_1)}$ -away from proper linear subspaces in  $V'$ . The induction hypothesis, applied to the subset  $\pi_{V'}(X') \subset V'$ , with acting set  $A$ , yields the desired inclusion provided that  $\varepsilon_1$  is small enough.

*Second step.* We assume  $X^{(\delta)} \cap V_1$  contains a large vector.

Let  $s_2, \varepsilon_2 > 0$  be the quantities given by Theorem 2.4 applied to the representation  $V_1$ , with constants  $\kappa, \varepsilon_1$ . We may choose  $s_2$  and  $\varepsilon_2$  uniformly over all choices for  $V_1$ ; indeed, up to a  $(\dim V)$ -bi-Lipschitz isomorphism of  $G$ -modules, there are only finitely many

choices for  $V_1$ . And assume that there exists  $v \in X^{(\delta)} \cap V_1$  with  $\|v\| \geq \delta^{\varepsilon_2}$ . Then, using the base case for the action of  $G$  on the irreducible module  $V_1$ , we find that

$$B_{V_1}(0, \delta^{\varepsilon_1}) \subset \langle A, X^{(\delta)} \rangle_{s_2} + B_V(0, \delta). \tag{13}$$

Now let  $z \in B_V(0, \delta^{\varepsilon_0})$ . By the first step, we may find  $y \in \langle A, X \rangle_{s_1} \cap B_V(0, \delta^{\varepsilon_1})$  and  $t \in V_1$  such that  $z = y + t + O(\delta)$ . Necessarily,  $\|t\| < 2\delta^{\varepsilon_1}$ , so that by (13),  $t \in \langle A, X \rangle_{2s_2} + O(\delta)$ . All in all, setting  $s = s_1 + O_{s_1}(s_2)$ , we find

$$B_V(0, \delta^{\varepsilon_0}) \subset \langle A, X \rangle_s + B_V(0, O_{s_1, s_2}(\delta)).$$

This finishes the proof of the theorem in this case.

*Third step.* Finally, we prove that there exists  $s_3 \geq 1$  depending on  $V$ ,  $\varepsilon_0$  and  $\kappa$  such that  $\langle A, X \rangle_{s_3}^{(\delta)} \cap V_1$  always contains a vector of length at least  $\delta^{\varepsilon_2}$ , which allows us to conclude the proof using the second step.

Let  $C$  be the constant given by Lemma 2.9. Let  $0 < \varepsilon_3 < \varepsilon_2/C$  be a parameter whose value will be chosen later according to  $\varepsilon_2$ . Let  $0 < \varepsilon_4 < \varepsilon_3$  be a parameter whose value will be chosen later according to  $\varepsilon_3$ . Using the induction hypothesis for the representation  $V'$  with  $\varepsilon_4$  and  $\kappa$ , and replacing  $\langle A, X \rangle_s^{(\delta)}$  by  $X$ , we may assume without loss of generality that

$$B_{V'}(0, \delta^{\varepsilon_4}) \subset \pi_{V'}(X). \tag{14}$$

Either  $\langle A, X \rangle_3 \cap V_1^{(\delta)}$  contains a vector of length  $\geq \delta^{\varepsilon_2}$ , in which case we are done, or  $\langle A, X \rangle_3 \cap V_1^{(\delta)} \subset B_V(0, \delta^{\varepsilon_2})$ . In the latter case, Lemma 2.9 applied with  $\tau = \varepsilon_2/C$  and  $\eta = \varepsilon_4$  gives a submodule  $W < V$  such that the restriction of  $\pi_{V'}$  to  $W$  is  $3\delta^{-\varepsilon_4}$ -bi-Lipschitz and

$$B_W(0, \frac{1}{2}\delta^{\varepsilon_4}) \subset X^{(\delta_1)} \tag{15}$$

where  $\delta_1 = \delta^{\varepsilon_2/C}$ . Now we apply the base case, Theorem 2.4, to the non-trivial irreducible representation  $V/W$  with  $\varepsilon_3$  and  $\kappa$ . Observe that  $\pi_{V'/W}$  being  $3\delta^{-\varepsilon_4}$ -bi-Lipschitz implies that  $\pi_{V/W|V_1} : V_1 \rightarrow V/W$  is  $4\delta^{-\varepsilon_4}$ -bi-Lipschitz. Hence, for the projections  $p_{V/W} : G \rightarrow \text{End}(V/W)$  and  $p_{V_1} : G \rightarrow \text{End}(V_1)$ , we have

$$\forall \rho \geq \delta, \quad N(p_{V/W}(A), \rho) \geq \delta^{O(\varepsilon_4)} N(p_{V_1}(A), \rho) \geq \delta^{O(\varepsilon_4)+\varepsilon} \rho^\kappa.$$

Therefore, provided  $\varepsilon_4$  and  $\varepsilon$  are small enough in terms of  $V_1$ ,  $\varepsilon_3$  and  $\kappa$ , Theorem 2.4 yields some constant  $s \geq 1$  depending only on  $V/W$ ,  $\kappa$  and  $\varepsilon_3$  such that

$$B_{V/W}(0, \delta^{\varepsilon_3}) \subset \pi_{V/W}(\langle A, X \rangle_s) + B_{V/W}(0, \delta).$$

Together with inclusion (15), this implies that

$$N(\langle A, X \rangle_s + X, \delta_1) \gg (\delta_1^{-1} \delta^{\varepsilon_3})^{\dim V/W} (\delta_1^{-1} \delta^{\varepsilon_4})^{\dim W} \geq \delta_1^{-\dim V} \delta^{O(\varepsilon_3)}.$$

Cutting  $\langle A, X \rangle_{s+1}$  into cylinders of axis  $V_1$  and diameter  $\delta^{\varepsilon_3}$  and picking the part with largest size, we see that

$$N(\langle A, X \rangle_{2s+2} \cap V_1^{(\delta^{\varepsilon_3})}, \delta_1) \geq \delta_1^{-\dim V} \delta^{O(\varepsilon_3)},$$

which ensures that  $X' := \langle A, X \rangle_{2s+2} \cap V_1^{(\delta^{\varepsilon_3})}$  is  $\delta^{O(\varepsilon_3)}$ -away from proper linear subspaces and a fortiori from submodules. Remembering (14), we know that

$$\pi_{V'}(X') = B_{V'}(0, \delta^{\varepsilon_3}).$$

At this stage apply Lemma 2.9 to the set  $X'$  with  $\tau = \varepsilon_2/C$  and  $\eta = \varepsilon_3$ . If  $\varepsilon_3$  is chosen sufficiently small compared to  $\varepsilon_2$ , conclusion (b) fails while all assumptions except (i) are satisfied. So there must be  $v \in \langle A, X' \rangle_3 \cap V_1^{(\delta)}$  with  $\|v\| > \delta^{\varepsilon_2}$ . This concludes the proof of the theorem. □

### 3. A product theorem for perfect Lie groups

The goal of this section is to use Theorem 2.3 to prove Theorem 1.2. More precisely, we prove the following essentially equivalent version of Theorem 1.2, which is a bounded generation statement.

**Theorem 3.1** (Product theorem in perfect Lie groups). *Let  $G$  be a connected perfect Lie group. There exists a neighborhood  $U$  of the identity in  $G$  such that given  $\kappa > 0$  and  $\varepsilon_0 > 0$ , there exist  $\varepsilon > 0$  and  $s \geq 1$  such that the following holds for  $\delta > 0$  sufficiently small. Let  $A$  be a subset of  $U$  such that*

(i) *for any projection  $\pi_i : G \rightarrow G/H_i$  to a simple factor,*

$$\forall \rho \geq \delta, \quad N(\pi_i(A), \rho) \geq \delta^\varepsilon \rho^{-\kappa};$$

(ii)  *$A$  is  $\delta^\varepsilon$ -away from closed connected subgroups in  $G$ .*

*Then*

$$B_G(1, \delta^{\varepsilon_0}) \subset (A \cup \{1\} \cup A^{-1})^s B_G(1, \delta).$$

Theorem 1.2 follows immediately from Theorem 3.1 in combination with a Ruzsa-type inequality [25, Theorem 6.8].

The proof of Theorem 3.1 goes as follows. We shall first prove the special case where the radical of our perfect Lie group  $G$  is abelian. In this case, the adjoint representation of  $G$  belongs to  $\mathcal{P}(G)$ , as we shall see in Lemma 3.4 below. So Theorem 2.3 applies and shows that we can produce in the Lie algebra  $\mathfrak{g}$  of  $G$  a large ball using addition and the adjoint action of  $G$ :  $B_{\mathfrak{g}}(0, \delta^{\varepsilon_0}) \subset \langle A, \log A \rangle_s^{(\delta)}$ . Then we want to exponentiate this inclusion to the level of the group  $G$ . For that, we use the Campbell–Hausdorff formula, which allows us to approximate sums in  $\mathfrak{g}$  by products in  $G$  with any desired precision; this is the content of Lemma 3.6. Finally, to deduce the general case from the special case, we shall use an induction on the nilpotency class of the radical of  $G$ .

#### 3.1. Perfect Lie algebras and Lie groups

We begin by recording some elementary facts about perfect Lie groups and Lie algebras.

Let  $G$  be a connected Lie group with Lie algebra  $\mathfrak{g}$ . Using Levi’s decomposition theorem [24, Corollary 1, p. 49], we may write  $\mathfrak{g}$  as a semidirect product  $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{r}$  of a



semisimple Lie algebra  $\mathfrak{s}$  and a solvable radical  $\mathfrak{r}$ . Writing  $\mathfrak{s} = \mathfrak{s}_1 \oplus \dots \oplus \mathfrak{s}_k$  as a sum of simple ideals, one sees that for each  $i$  in  $\{1, \dots, k\}$ ,  $\mathfrak{h}_i = (\bigoplus_{j \neq i} \mathfrak{s}_j) \rtimes \mathfrak{r}$  is an ideal in  $\mathfrak{g}$ . The Lie algebra  $\mathfrak{h}_i$  is the Lie algebra of a closed normal subgroup  $H_i \triangleleft G$ . The projection maps  $\pi_i : G \rightarrow G/H_i$  are the *projections of  $G$  to its simple factors*. Note that any left-invariant Riemannian metric  $d$  on  $G$  induces a left-invariant metric on  $G/H_i$ . Indeed, if  $N \triangleleft G$  is any closed normal subgroup, one defines a distance on the quotient  $G/N$  by

$$\forall x, y \in G, \quad d(\bar{x}, \bar{y}) = \inf_{n, n' \in N} d(xn, yn') = d(y, xN) = d(x^{-1}y, N).$$

For later use, we now list three elementary and standard lemmas about perfect Lie algebras.

**Lemma 3.2.** *If  $\mathfrak{g}$  is a perfect Lie algebra, then its solvable radical  $\mathfrak{r}$  is nilpotent. In particular,  $\mathfrak{g}$  can be written as a semidirect product  $\mathfrak{g} = \mathfrak{s} \rtimes \mathfrak{r}$  of a semisimple Lie algebra  $\mathfrak{s}$  with a nilpotent ideal  $\mathfrak{r}$ .*

*Proof.* See for instance [1, Lemma 2.4]. □

**Lemma 3.3.** *Let  $\mathfrak{g}$  be a perfect Lie algebra with Levi decomposition  $\mathfrak{g} = \mathfrak{s} \rtimes \mathfrak{r}$ . The image of a proper ideal of  $\mathfrak{g}$  under the map  $\mathfrak{g} \rightarrow \mathfrak{g}/\mathfrak{r}$  is a proper ideal. In particular, the image of a maximal proper ideal is a maximal proper ideal.*

*Proof.* Let  $\mathfrak{n}$  be an ideal in  $\mathfrak{g}$  such that  $\mathfrak{n} + \mathfrak{r} = \mathfrak{g}$ . We want to show that  $\mathfrak{n} = \mathfrak{g}$ . Denote by  $D^i \mathfrak{r}$ ,  $i \geq 0$ , the derived series of  $\mathfrak{r}$ , i.e.  $D^0 \mathfrak{r} = \mathfrak{r}$  and  $D^{i+1} \mathfrak{r} = [D^i \mathfrak{r}, D^i \mathfrak{r}]$  for  $i \geq 0$ . We show by induction that for all  $i \geq 0$ ,

$$\mathfrak{g} = \mathfrak{n} + D^i \mathfrak{r}. \tag{16}$$

Indeed, (16) is true for  $i = 0$ . Suppose that it is true for some  $i \geq 0$ ; then it follows from  $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$  that

$$\mathfrak{g} = [\mathfrak{n}, \mathfrak{n}] + [\mathfrak{n}, D^i \mathfrak{r}] + [D^i \mathfrak{r}, D^i \mathfrak{r}] \subset \mathfrak{n} + D^{i+1} \mathfrak{r},$$

because  $\mathfrak{n}$  is an ideal in  $\mathfrak{g}$ . Since  $\mathfrak{r}$  is solvable, we may take  $i$  such that  $D^i \mathfrak{r} = 0$  to conclude that  $\mathfrak{n} = \mathfrak{g}$ . □

**Lemma 3.4** (Perfect abelian extension of a semisimple group). *Let  $G$  be a perfect Lie group with Lie algebra  $\mathfrak{g}$ . If the radical  $\mathfrak{r}$  of  $\mathfrak{g}$  is abelian, then the adjoint representation of  $G$  is of class  $\mathcal{P}$ .*

*Proof.* We have an exact sequence of  $G$ -modules

$$0 \rightarrow \mathfrak{r} \rightarrow \mathfrak{g} \rightarrow \mathfrak{g}/\mathfrak{r} \rightarrow 0,$$

and by Proposition 2.2(i), all we need to check is that both  $\mathfrak{r}$  and  $\mathfrak{g}/\mathfrak{r}$  belong to  $\mathcal{P}(G)$ . Let  $R$  be the solvable radical of  $G$ ; it is equal to the closed connected subgroup of  $G$  with Lie algebra  $\mathfrak{r}$ . The Lie group  $G/R$  is semisimple, so its adjoint representation belongs to  $\mathcal{P}(G/R)$ . By Proposition 2.2(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  $S = G/R$ , 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 2.2(iii). □

**Remark 3.** 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  $\mathcal{P}(G)$ .

**Remark 4.** 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 non-trivial centers. For instance, let  $\mathcal{F}_{2,2}$  denote the free 2-nilpotent Lie algebra over two generators  $x, y$ . It is the Lie algebra of the Heisenberg group  $H_3(\mathbb{R})$ . The action of  $SL(2, \mathbb{R})$  on  $\mathcal{F}_{2,2}$  by linear substitution integrates to an action of  $SL(2, \mathbb{R})$  on  $H_3(\mathbb{R})$  by group automorphisms. This allows us to define the Lie group  $G = SL(2, \mathbb{R}) \times H_3(\mathbb{R})$ . Its Lie algebra  $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{R}) \times \mathcal{F}_{2,2}$  is perfect. However, the adjoint representation of  $G$  is not of class  $\mathcal{P}$ , because  $G$  acts trivially on the center of  $\mathfrak{g}$ , generated by  $[x, y]$ .

### 3.2. Abelian extensions of semisimple groups

Here, we prove Theorem 3.1 in the case where the Lie algebra of  $G$  can be written as a semidirect product  $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{r}$  with  $\mathfrak{r}$  abelian. We shall see in 3.3 that the general case follows from this one.

We fix a connected perfect Lie group  $G$  with Lie algebra  $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{r}$ , where  $\mathfrak{s}$  is semisimple and  $\mathfrak{r}$  is an abelian ideal. To prove Theorem 3.1 in this case, the idea is to apply Theorem 2.3 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 3.1 automatically implies non-concentration for the image of  $A$  under any non-trivial group homomorphism.

**Lemma 3.5.** *Let  $G$  be a perfect connected Lie group. Given a non-trivial homomorphism  $\varphi: G \rightarrow H$  to some connected Lie group  $H$ , there exists a neighborhood  $U$  of the identity in  $G$  such that the following holds. Let  $\varepsilon > 0$  and  $\kappa > 0$  be parameters and let  $A \subset U$  be a subset satisfying condition (i) of Theorem 3.1. Then*

$$\forall \rho \geq \delta, \quad N(\varphi(A), \rho) \gg_{\varphi} \delta^{\varepsilon} \rho^{-\kappa}.$$

*Proof.* The isomorphism  $G/\ker \varphi \rightarrow \varphi(G)$  is bi-Lipschitz when restricted to compact neighborhoods. Hence without loss of generality, we may assume that  $H = G/\ker \varphi$ . Since  $\ker \varphi$  is closed, there exists a neighborhood  $U$  of the identity in  $G$  such that for all  $x, y, d(x^{-1}y, \ker \varphi) = d(x^{-1}y, (\ker \varphi)^{\circ})$ . 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 Lemma 3.3,  $\mathfrak{n}$  is exactly the kernel of the projection of  $\mathfrak{g}$  to one of its simple factors. It follows that  $\mathfrak{n}$  is the Lie algebra of a proper closed normal subgroup  $N \triangleleft G$ , with  $G/N$  one of the simple factors of  $G$ . We deduce the desired estimate from condition (i) of Theorem 3.1 by using the fact that  $G/\ker \varphi \rightarrow G/N$  is 1-Lipschitz. □

*Proof of Theorem 3.1 in the case where  $\mathfrak{v}$  is abelian.* In this proof, implied constants in Landau and Vinogradov notations depend on  $G$  and on the parameter  $\kappa$ .

By Lemma 3.4, the adjoint representation of  $G$  on  $\mathfrak{g}$  is of class  $\mathcal{P}$ . Set  $X = \log(A^{-1}A \cap B_G(1, \delta^\varepsilon)) \subset \mathfrak{g}$ . Then the hypotheses of Theorem 2.3 are all met with  $\varepsilon$  replaced by  $O(\varepsilon)$ . Indeed, assumption (i) is guaranteed by Lemma 3.5, and  $A$  being  $\delta^\varepsilon$ -away from subgroups is exactly assumption (ii) of Theorem 3.1. So it remains to check that  $X$  is  $\delta^{O(\varepsilon)}$ -away from any proper submodule  $W$  in  $\mathfrak{g}$ . We may assume that  $W$  is maximal. Then it is a maximal proper ideal of  $\mathfrak{g}$ , which by Lemma 3.3 is equal to the kernel  $H_i$  of some projection  $\pi_i : \mathfrak{g} \rightarrow \mathfrak{g}/\mathfrak{h}_i$  of  $G$  on a simple factor. In particular, there are only finitely many such  $W$ . Shrinking the neighborhood  $U$  if necessary, it suffices to check that  $A^{-1}A \cap B_G(1, \delta^\varepsilon)$  is  $\delta^{O(\varepsilon)}$ -away from  $H_i$ . By assumption (i), for any  $\rho \geq \delta$ ,

$$\begin{aligned} N(\pi_i(A^{-1}A \cap B_G(1, \delta^\varepsilon)), \rho) &\geq \max_{\mathfrak{g}} N(\pi_i(A \cap B_G(g, \delta^\varepsilon)), \rho) \\ &\geq \delta^{O(\varepsilon)} N(\pi_i(A), \rho) \geq \delta^{O(\varepsilon)} \rho^{-\kappa}. \end{aligned}$$

The last quantity is larger than 1 if we choose  $\rho = \delta^{C\varepsilon}$  with a large constant  $C = O(1)$ . This shows that  $A^{-1}A \cap B_G(1, \delta^\varepsilon)$  is  $\delta^{O(\varepsilon)}$ -away from  $\ker \pi_i$ .

Thus, we can apply Theorem 2.3 to get an integer  $s \geq 1$  such that

$$B_{\mathfrak{g}}(0, \delta^{\varepsilon_0}) \subset \langle A, X \rangle_s + B_{\mathfrak{g}}(0, \delta) \tag{17}$$

when  $\varepsilon$  is small enough.

The idea is now to apply the Campbell–Hausdorff formula at an order  $\ell$  such that the error term is of size at most  $\delta$ . We identify an element  $w$  of the free group  $F_s$  generated by  $s$  elements and the word map  $w : G^{\times s} \rightarrow G$  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 easier to approximate  $e^{2(x+y)}$ , and then we get  $e^{2(x+y)} = (e^x)^2 (e^y)^2 (e^y)^2 e^x (e^y)^{-2} (e^x)^{-1} e^{O(\|x\|^3 + \|y\|^3)}$ . We shall use the following lemma, which generalizes these elementary computations, and follows from the Campbell–Hausdorff formula.

**Lemma 3.6.** *Let  $\exp : \mathfrak{g} \rightarrow G$  denote the exponential map of a Lie group. Fix a Euclidean norm on  $\mathfrak{g}$  and endow  $G$  with the associated left-invariant Riemannian metric. For all integers  $s \geq 1$  and  $\ell \geq 1$ , 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  $x_1, \dots, x_s \in U$ ,*

$$d(\exp(Cx_1 + \dots + Cx_s), w(\exp x_1, \dots, \exp x_s)) \ll_{\ell} (\|x_1\| + \dots + \|x_s\|)^{\ell}.$$

*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  $[x_{i_1}, \dots, x_{i_k}] = [x_{i_1}, [x_{i_2}, \dots, [x_{i_{\ell-1}}, x_{i_k}]\dots]]$  of length  $k$ . The series converges on

$B_{\mathfrak{g}}(0, r)^{\times s}$  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  $\ell \geq 1$ , we write  $O(d^\circ \geq \ell)$  to denote an unspecified element in  $\mathcal{G}_s$  of valuation at least  $\ell$ .

By the Baker–Campbell–Hausdorff formula [9], the map defined by  $(x, y) \mapsto x * y = \log(\exp(x) \exp(y))$  belongs to  $\mathcal{G}_2$ , and moreover

$$x * y = x + y + \frac{1}{2}[x, y] + O(d^\circ \geq 3). \tag{18}$$

From that we deduce, by induction on  $s$ , that

$$x_1 * \cdots * x_s = x_1 + \cdots + x_s + O(d^\circ \geq 2). \tag{19}$$

We denote by  $[x, y]_*$  the group commutator  $x * y * (-x) * (-y)$  and by  $[x_1, \dots, x_s]_*$  the repeated group commutator  $[x_1, [x_2, \dots, [x_{s-1}, x_s]_* \dots]_*]_*$ . By (18) we have

$$[x, y]_* = [x, y] + O(d^\circ \geq 3),$$

and again by induction on  $s$ ,

$$[x_1, \dots, x_s]_* = [x_1, \dots, x_s] + O(d^\circ \geq s + 1). \tag{20}$$

Now we prove by induction on  $\ell$  that there exists an integer  $C_\ell$  and a word  $w_\ell \in F_s$  such that

$$x_1 + \cdots + x_s = w_\ell^*(x_1/C_\ell, \dots, x_s/C_\ell) + O(d^\circ \geq \ell), \tag{21}$$

where  $w_\ell^*$  is the word map induced by  $w_\ell$ , which is well defined on a neighborhood of 0 in  $\mathfrak{g}^{\times s}$ . For  $\ell = 2$ , this is given by (19). Suppose the result has been proved for some  $\ell \geq 2$ . Let  $f$  be the sum of terms of degree  $\ell$  in the remainder term  $O(d^\circ \geq \ell)$  on the right-hand side of (21). Since  $f$  has rational coefficients, there is an integer  $C \geq 1$  such that we can write

$$f(x_1, \dots, x_s) = \sum_{i=1}^N m_i(x_1/C, \dots, x_s/C)$$

where each  $m_i$  is a repeated bracket of length  $\ell$ . Therefore, by (20) and (19), there is a product  $w' \in F_s$  of repeated commutators such that

$$f(x_1, \dots, x_s) = w'^*(x_1/C, \dots, x_s/C) + O(d^\circ \geq \ell + 1).$$

Thus,

$$\begin{aligned} x_1 + \cdots + x_s &= w_\ell^*(x_1/C_\ell, \dots, x_s/C_\ell) + w'^*(x_1/C, \dots, x_s/C) + O(d^\circ \geq \ell + 1) \\ &= w_\ell^*(x_1/C_\ell, \dots, x_s/C_\ell) * w'^*(x_1/C, \dots, x_s/C) + O(d^\circ \geq \ell + 1). \end{aligned}$$

In the last step we used the fact that  $w'^*(x_1/C, \dots, x_s/C)$  has valuation at least  $\ell$ . This finishes the proof of the induction step and concludes the proof of the lemma.  $\square$

To conclude the proof of Theorem 3.1 in case  $\mathfrak{r}$  is abelian, we choose  $\ell > 1/\varepsilon$  and apply Lemma 3.6 to elements  $x_i$  of the form  $x_i = \text{Ad}(a_i)y_i$  with  $a_i \in A^s$  and  $y_i \in X$ . By definition  $X \subset B_{\mathfrak{g}}(0, \delta^\varepsilon)$ , so the error term is indeed of size  $O_s(\delta^{\ell\varepsilon}) = O(\delta)$ , and therefore

$$\begin{aligned} \exp[C \text{Ad}(a_1)y_1 + \dots + C \text{Ad}(a_s)y_s] &\in w(a_1 e^{y_1} a_1^{-1}, \dots, a_s e^{y_s} a_s^{-1}) B_G(1, O(\delta)) \\ &\in (A \cup \{1\} \cup A^{-1})^{s'} B_G(1, O(\delta)) \end{aligned}$$

for some  $s' = O_{s,\ell}(1)$ . Recalling (17), we obtain

$$B_G(1, \delta^{\varepsilon_0}) \subset \exp[C \cdot B_{\mathfrak{g}}(0, \delta^{\varepsilon_0})] \subset \exp[C \cdot \langle A, X \rangle_s + B_{\mathfrak{g}}(0, C\delta)] \subset A^{s'} B_G(1, O(\delta)).$$

This finishes the proof of the theorem for  $\mathfrak{r}$  abelian. □

### 3.3. Proof of the product theorem: general case

We now explain how to deal with a perfect Lie group  $G$  with Lie algebra  $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{r}$ , where  $\mathfrak{r}$  is nilpotent by Lemma 3.2 but not abelian. This will follow from the previous case, together with a quantitative version of the following fact: If  $R$  is a nilpotent Lie group, then a subset  $A \subset R$  generates the group  $R$  if and only if  $A \bmod [R, R]$  generates  $R/[R, R]$ .

For  $A$  and  $B$  subsets of a group  $G$ , we shall write  $[A, B]$  for 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 3.7.** *Let  $R$  be a connected nilpotent Lie group with descending central series  $R_i$ ,  $i \geq 1$ , i.e.  $R_1 = R$  and for  $i \geq 1$ ,  $R_{i+1} = [R, R_i]$ . For each  $i \geq 1$  there is  $k \geq 1$  such that for all  $\rho > 0$  small enough,*

$$B_{R_{i+1}}(1, \rho^2) \subset [B_R(1, \rho), B_{R_i}(1, \rho)]^k.$$

*Proof.* Denote by  $\mathfrak{r}_i$ ,  $i \geq 1$ , the descending central series of the Lie algebra  $\mathfrak{r}$ . Let  $(z_1, \dots, z_m)$  be a basis of  $\mathfrak{r}_{i+1}$  consisting of commutators  $z_j = [x_j, y_j]$  with  $x_j \in \mathfrak{r}$  and  $y_j \in \mathfrak{r}_i$ . For each  $j$ , consider the map  $f_j: \mathbb{R} \rightarrow R_{i+1}$  defined by

$$f_j(t) = \begin{cases} [\exp(\sqrt{t} x_j), \exp(\sqrt{t} y_j)] & \text{if } t \geq 0, \\ [\exp(\sqrt{-t} y_j), \exp(\sqrt{-t} x_j)] & \text{if } t < 0, \end{cases}$$

and further define  $f: \mathbb{R}^m \rightarrow R_{i+1}$  by  $f(t_1, \dots, t_m) = f_1(t_1) \cdots f_m(t_m)$ . The function  $f$  is of class  $C^1$  and its differential at 0 is

$$T_0 f(h_1, \dots, h_m) = h_1 z_1 + \dots + h_m z_m,$$

so it is a  $C^1$ -diffeomorphism on a neighborhood of 0. This implies that for some constant  $c > 0$  depending only on  $R$ ,

$$B_{R_{i+1}}(1, c\rho^2) \subset f(B_{\mathbb{R}}(0, \rho)) \subset [B_R(1, \rho), B_{R_i}(1, \rho)]^m.$$

This finishes the proof, because for  $\rho$  small enough,  $B_{R_{i+1}}(1, c\rho^2) \cdot B_{R_{i+1}}(1, c\rho^2) \supset B_{R_{i+1}}(1, 2c\rho^2)$ . □

We are now ready to finish the proof of Theorem 3.1.

*Proof of Theorem 3.1 in the general case.* Here again implied constants in Landau and Vinogradov notations depend on  $G$  and  $\kappa$ . Write the Lie algebra of  $G$  as a semidirect product  $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{t}$ , with  $\mathfrak{s}$  semisimple and  $\mathfrak{t}$  a nilpotent ideal, and let  $R$  be the nilpotent radical of  $G$ , i.e. the closed connected normal subgroup of  $G$  with Lie algebra  $\mathfrak{t}$ . The proof goes by induction on the nilpotency class  $\ell$  of  $R$ .

We have already seen that Theorem 3.1 holds if  $\ell \leq 1$ . Now suppose that  $R$  has nilpotency class  $\ell \geq 2$  and that Theorem 3.1 has been proved if the nilpotency class is strictly less than  $\ell$ .

Let  $R_i, i \geq 1$ , denote the lower central series of the group  $R$ . Each  $R_i, i \geq 1$ , is closed and connected, and the Lie algebra of  $R_i$  is exactly the  $i$ -th term in the lower central series of  $\mathfrak{t}$  (see e.g. [17, Theorem 5.7, p. 55]). We first remark that the assumptions of Theorem 3.1 are preserved when projecting to a quotient. The nilpotency class of the radical of  $G/R_\ell$  is  $\ell - 1$ . Let  $\varepsilon_1 > 0$  be some constant, whose value will be specified later. By the induction hypothesis applied to  $G/R_\ell$ , provided  $\varepsilon$  is small enough compared to  $\varepsilon_1$ , for some integer  $s$  depending on  $\kappa$  and  $\varepsilon_1$ ,

$$B_G(1, \delta^{\varepsilon_1}) \subset (A \cup \{1\} \cup A^{-1})^s B_G(1, \delta) R_\ell$$

Without loss of generality, we may replace  $(A \cup \{1\} \cup A^{-1})^s B_G(1, \delta)$  by  $A$ , and assume that

$$B_R(1, \delta^{\varepsilon_1}) \subset (R \cap A) R_\ell \quad \text{and} \quad B_{R_{\ell-1}}(1, \delta^{\varepsilon_1}) \subset (R_{\ell-1} \cap A) R_\ell.$$

By Lemma 3.7, we also have

$$B_{R_\ell}(1, \delta^{2\varepsilon_1}) \subset [B_R(1, \delta^{\varepsilon_1}), B_{R_{\ell-1}}(1, \delta^{\varepsilon_1})]^{O(1)}.$$

From these inclusions and the fact that  $R_\ell$  is in the center of  $R$ , it follows that

$$B_{R_\ell}(1, \delta^{2\varepsilon_1}) \subset A^{O(1)} B_G(1, O(\delta)). \tag{22}$$

At this stage replace  $A^{O(1)} B_G(1, O(\delta))$  by  $A$ . The fact that  $B_R(1, \delta^{\varepsilon_1}) \subset AR_\ell$  and  $B_{R_\ell}(1, \delta^{2\varepsilon_1}) \subset A$  does not prove what we want yet but gives the lower bound

$$N(A^2, \delta) \gg_G \delta^{-\dim G + O(\varepsilon_1)}.$$

Covering  $A^2$  by balls of radius  $\frac{1}{2}\delta^{3\varepsilon_1}$ , we obtain

$$N(A^{-2}A^2 \cap B_G(1, \delta^{3\varepsilon_1}), \delta) \gg_G \delta^{-\dim G + O(\varepsilon_1)}.$$

Write  $A' = A^{-2}A^2 \cap B_G(1, \delta^{3\varepsilon_1})$ . Then  $A'$  satisfies the assumptions of Theorem 3.1 with  $\kappa = 1$  and  $\varepsilon = O(\varepsilon_1)$ . Hence if  $\varepsilon_1$  is small enough compared to  $\varepsilon_0$ , then by the induction hypothesis again,

$$B_G(1, \delta^{\varepsilon_0}) \subset A'^s B_G(1, \delta) R_\ell$$

for some  $s$  depending on  $\varepsilon_0$ . Since any element in  $R_\ell$  involved in this inclusion is within distance  $\delta^{2\varepsilon_1}$  from the identity, we can conclude using (22) that

$$B_G(1, \delta^{\varepsilon_0}) \subset A'^{O(1)} B_G(1, \delta) A.$$

This finishes the proof of Theorem 3.1. □

### 3.4. Approximate subgroups in non-perfect Lie groups

Here we prove Proposition 1.3. First, observe that in a non-trivial abelian Lie group, generalized arithmetic progressions (i.e. sums of arithmetic progressions) are the prototypes of approximate subgroups. Then in a non-perfect Lie group  $G$ , it suffices to lift a generalized arithmetic progression in its abelianization  $G/[G, G]$  to obtain an approximate subgroup with the desired properties.

*Proof of Proposition 1.3.* First consider the abelian case  $G = \mathbb{R}^{\times d}$ , with  $d \geq 1$ . Let  $\kappa \in (0, 1]$ . Given a neighborhood  $U$  of  $0 \in \mathbb{R}^{\times d}$ , let  $r > 0$  be such that  $B_{\mathbb{R}^d}(0, r) \subset U$ . Define

$$P = \{\delta^\kappa x \in \mathbb{R}^{\times d} \mid x \in \mathbb{Z}^{\times d} \cap [-\delta^{-\kappa} r, \delta^{-\kappa} r]^{\times d}\}.$$

It is easy to check that  $P$  has the required properties.

Now let  $G$  be a simply connected non-perfect Lie group. Then  $G/[G, G] \simeq \mathbb{R}^{\times d}$  where  $d = \dim \mathfrak{g} - \dim [\mathfrak{g}, \mathfrak{g}]$ . Let  $\pi : G \rightarrow \mathbb{R}^{\times d}$  be the projection. Given a neighborhood  $U$  of  $1_G \in G$ , let  $r > 0$  be such that  $B_G(1_G, 2r) \subset U$  and  $B_{\mathbb{R}^d}(0, r) \subset \pi(U)$ . Let  $P$  be defined as above and put  $A = B_G(1_G, 2r) \cap \pi^{-1}(P)$ .

On the one hand,

$$N(A, \delta) \approx_{G,r} \delta^{-\dim[\mathfrak{g}, \mathfrak{g}]} N(\pi(A), \delta) \approx_{G,r} \delta^{-\dim[\mathfrak{g}, \mathfrak{g}]} N(P, \delta) \approx_{G,r} \delta^{-\dim[\mathfrak{g}, \mathfrak{g}] - d\kappa},$$

and for similar reasons,

$$N(AAA, \delta) \ll_{G,r} \delta^{-\dim[\mathfrak{g}, \mathfrak{g}]} N(P + P + P, \delta) \ll_{G,r} N(A, \delta).$$

On the other hand, when  $\delta$  is so small that  $\delta^\kappa < r$ , then  $A$  is  $\delta^\kappa$ -dense in  $B_G(1_G, r)$ , that is,

$$B_G(1_G, r) \subset A^{(\delta^\kappa)}.$$

It follows immediately that for any connected normal subgroup  $N \triangleleft G$ ,  $\pi_{G/N}(A)$  is  $\delta^\kappa$ -dense in  $B_{G/N}(1_{G/N}, r)$  and hence

$$\forall \rho \geq \delta, \quad N(\pi_{G/N}(A), \rho) \gg_{G,r} \rho^{-\kappa}.$$

Moreover, it is not difficult to see that given a simply connected Lie group  $G$  and  $r > 0$ , there is  $c = c(G, r) > 0$  such that no proper closed connected subgroup is  $c$ -dense in

$B_G(1_G, r)$ . From this we deduce that  $A$  is  $(c - \delta^k)$ -away from proper closed connected subgroups.  $\square$

*Acknowledgments.* This project started during the preparation of W.H.'s doctoral thesis at Université Paris 11, in Orsay, under the supervision of Emmanuel Breuillard and Péter Varjú. We thank them for their support and advice. We also thank the referee for some useful comments on an earlier version of the paper.

W.H. was partially supported by ERC grant ErgComNum 682150.

## References

- [1] Benoist, Y., de Saxcé, N.: Convolution in perfect Lie groups. *Math. Proc. Cambridge Philos. Soc.* **161**, 31–45 (2016) [Zbl 1371.22015](#) [MR 3505667](#)
- [2] Benoist, Y., de Saxcé, N.: A spectral gap theorem in simple Lie groups. *Invent. Math.* **205**, 337–361 (2016) [Zbl 1357.22003](#) [MR 3529116](#)
- [3] Bourgain, J.: On the Erdős–Volkmann and Katz–Tao ring conjectures. *Geom. Funct. Anal.* **13**, 334–365 (2003) [Zbl 1115.11049](#) [MR 1982147](#)
- [4] Bourgain, J.: The discretized sum-product and projection theorems. *J. Anal. Math.* **112**, 193–236 (2010) [Zbl 1234.11012](#) [MR 2763000](#)
- [5] Bourgain, J., Gamburd, A.: On the spectral gap for finitely-generated subgroups of  $SU(2)$ . *Invent. Math.* **171**, 83–121 (2008) [Zbl 1135.22010](#) [MR 2358056](#)
- [6] Bourgain, J., Gamburd, A.: A spectral gap theorem in  $SU(d)$ . *J. Eur. Math. Soc.* **14**, 1455–1511 (2012) [Zbl 1254.43010](#) [MR 2966656](#)
- [7] Bourgain, J., Yehudayoff, A.: Expansion in  $SL_2(\mathbb{R})$  and monotone expanders. *Geom. Funct. Anal.* **23**, 1–41 (2013) [Zbl 1268.05103](#) [MR 3037896](#)
- [8] Boutonnet, R., Ioana, A., Salehi Golsefidy, A. S.: Local spectral gap in simple Lie groups and applications. *Invent. Math.* **208**, 715–802 (2017) [Zbl 1366.22004](#) [MR 3648974](#)
- [9] Dynkin, E. B.: Calculation of the coefficients in the Campbell–Hausdorff formula. *Dokl. Akad. Nauk SSSR (N.S.)* **57**, 323–326 (1947) (in Russian) [Zbl 0029.24507](#) [MR 0021940](#)
- [10] Erdős, P., Volkmann, B.: Additive Gruppen mit vorgegebener Hausdorffscher Dimension. *J. Reine Angew. Math.* **221**, 203–208 (1966) [Zbl 0135.10202](#) [MR 0186782](#)
- [11] He, W.: Discretized sum-product estimates in matrix algebras. *J. Anal. Math.* **139**, 637–676 (2019) [Zbl 07172359](#) [MR 4041116](#)
- [12] Katz, N. H., Tao, T.: Some connections between Falconer’s distance set conjecture and sets of Furstenberg type. *New York J. Math.* **7**, 149–187 (2001) [Zbl 0991.28006](#) [MR 1856956](#)
- [13] Kollár, J.: An effective Łojasiewicz inequality for real polynomials. *Period. Math. Hungar.* **38**, 213–221 (1999) [Zbl 0973.26012](#) [MR 1756239](#)
- [14] Li, J.: Discretized sum-product and Fourier decay in  $\mathbb{R}^n$ . [arXiv:1811.06852v1](#) (2018)
- [15] Li, J.: Fourier decay, renewal theorem and spectral gaps for random walks on split semisimple Lie groups. [arXiv:1811.06484v1](#) (2018)
- [16] Łojasiewicz, S.: Ensembles semi-analytiques. Preprint IHES (1965); a version revised by M. Coste (2006) available at <https://perso.univ-rennes1.fr/michel.coste>
- [17] Onishchik, A. L., Vinberg, E. B.: Foundations of Lie theory. In: *Lie Groups and Lie Algebras I. Foundations of Lie Theory. Lie Transformation Groups*. Springer, Berlin, 1–94 (1993) [Zbl 0781.22003](#) [MR 0950862](#)
- [18] Salehi Golsefidy, A.: Super-approximation, I:  $p$ -adic semisimple case. *Int. Math. Res. Notices* **2017**, 7190–7263 [Zbl 1405.22025](#) [MR 3802123](#)



- 
- [19] Salehi Golsefidy, A.: Super-approximation, II: the  $p$ -adic and bounded powers of square-free integers cases. *J. Eur. Math. Soc.* **21**, 2163–2232 (2019) [Zbl 1422.22013](#) [MR 3959861](#)
  - [20] Salehi Golsefidy, A., Varjú, P. P.: Expansion in perfect groups. *Geom. Funct. Anal.* **22**, 1832–1891 (2012) [Zbl 1284.20044](#) [MR 3000503](#)
  - [21] de Saxcé, N.: Subgroups of fractional dimension in nilpotent or solvable Lie groups. *Mathematika* **59**, 497–511 (2013) [Zbl 1272.22006](#) [MR 3081784](#)
  - [22] de Saxcé, N.: Borelian subgroups of simple Lie groups. *Duke Math. J.* **166**, 573–604 (2017) [Zbl 1362.22010](#) [MR 3606726](#)
  - [23] de Saxcé, N.: A product theorem in simple Lie groups. *Geom. Funct. Anal.* **25**, 915–941 (2015) [Zbl 1321.22013](#) [MR 3361775](#)
  - [24] Serre, J.-P.: *Lie Algebras and Lie Groups*. 2nd ed., Springer, Berlin (1992) [Zbl 0132.27803](#) [MR 1176100](#)
  - [25] Tao, T.: Product set estimates for non-commutative groups. *Combinatorica* **28**, 547–594 (2008) [Zbl 1254.11017](#) [MR 2501249](#)
  - [26] Vigolo, F.: Measure expanding actions, expanders and warped cones. *Trans. Amer. Math. Soc.* **371**, 1951–1979 (2019) [Zbl 1402.05205](#) [MR 3894040](#)