

RESEARCH ARTICLE OPEN ACCESS

# Random Lie bracket on $\mathfrak{sl}_2(\mathbf{F}_p)$

Urban Jezernik | Matevž Miščič 

Faculty of Mathematics and Physics, University of Ljubljana and Institute of Mathematics, Physics, and Mechanics, Ljubljana, Slovenia

**Correspondence:** Matevž Miščič ([matevz.miscic@imfm.si](mailto:matevz.miscic@imfm.si))**Received:** 1 April 2025 | **Revised:** 24 July 2025 | **Accepted:** 10 December 2025

## ABSTRACT

We study a random walk on the Lie algebra  $\mathfrak{sl}_2(\mathbf{F}_p)$  where new elements are produced by randomly applying adjoint operators of two generators. Focusing on the generic case where the generators are selected at random, we analyze the limiting distribution of the random walk and the speed at which it converges to this distribution. These questions reduce to the study of a random walk on a cyclic group. We show that, with high probability, the walk exhibits a pre-cutoff phenomenon after roughly  $p$  steps. Notably, the limiting distribution need not be uniform, and it depends on the prime divisors of  $p - 1$ . Furthermore, we prove that by incorporating a simple random twist into the walk, we can embed a well-known affine random walk on  $\mathbf{F}_p$  into the modified random Lie bracket, allowing us to show that the entire Lie algebra is covered in roughly  $\log p$  steps in the generic case.

## 1 | Introduction

Random walks on finite groups are a powerful tool for analyzing their structure and properties (see Diaconis (1988); Hildebrand (2005) and references therein). Of particular significance is their connection to growth and expansion in Cayley graphs of finite simple groups (see, for example, Bourgain and Gamburd (2008)). However, extending these techniques, and ultimately the results, to the closely related setting of finite simple Lie algebras is not straightforward. Unlike groups, Lie algebras have two operations: addition and the Lie bracket. The fact that the Lie bracket is non-invertible introduces additional challenges.

In this paper, we introduce and study a random walk on the Lie algebra  $\mathfrak{sl}_2(\mathbf{F}_p)$  that focuses solely on the Lie bracket structure. This walk is defined as follows. Let  $A$  and  $B$  be a generating pair of the Lie algebra  $\mathfrak{sl}_2(\mathbf{F}_p)$  with  $p \geq 3$ . Let  $Z_k$ , for  $k \in \mathbf{N}$ , be independent random variables uniformly distributed in the set  $\{A, B\}$ . Starting with  $X_0 = [A, B]$ , the random walk evolves according to

$$X_k = [Z_k, X_{k-1}] \quad (k \in \mathbf{N}).$$

This process generates a sequence  $(X_k)_{k \in \mathbf{N}_0}$  that forms a Markov chain on the set  $\mathfrak{sl}_2(\mathbf{F}_p)$ . New elements are produced by iteratively applying the adjoint operators associated with the generators, a process we refer to as the *random Lie bracket*.

This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Random Structures & Algorithms* published by Wiley Periodicals LLC.

**Example.** Let us look at a simulation of the random Lie bracket. Suppose

$$p = 11, \quad A = \begin{pmatrix} 7 & 5 \\ 2 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 8 & 8 \\ 10 & 3 \end{pmatrix}, \quad X_0 = [A, B] = \begin{pmatrix} 1 & 10 \\ 2 & 10 \end{pmatrix}.$$

Randomly sample the sequence  $(B, A, B, B, B)$  from  $\{A, B\}$  to obtain the values

$$X_1 = \begin{pmatrix} 4 & 1 \\ 10 & 7 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 4 & 7 \\ 8 & 7 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 5 & 4 \\ 7 & 6 \end{pmatrix}, \quad X_4 = \begin{pmatrix} 5 & 6 \\ 10 & 6 \end{pmatrix}, \quad X_5 = \begin{pmatrix} 9 & 5 \\ 6 & 2 \end{pmatrix}.$$

Observe that  $X_2 = 4X_0$ ,  $X_4 = 5X_0$ .

Our goal is to analyze the limiting distribution of this random walk and to study the rate at which it converges to that distribution, particularly when the elements  $A$  and  $B$  are chosen uniformly at random. We prove the following somewhat surprising results.

The random walk  $(X_k)_{k \in \mathbb{N}_0}$  exhibits different behavior depending on whether the number of steps is even or odd (as hinted at in the example above). In fact, as we will explain, it suffices to focus on what happens after an even number of steps. We first show that the distribution of  $X_{2k}$  is supported on the line  $\text{Lin}\{[A, B]\}$  and converges (in total variation distance)<sup>1</sup> to a distribution  $\nu_{2k}$  on this line in a *pre-cutoff* manner at  $k \approx p$ .

**Theorem 1.1.** For every  $\delta, \epsilon > 0$ , there are constants  $C, c > 0$  such that the following holds. Let  $p > C$ , let  $A, B \in \mathfrak{sl}_2(\mathbb{F}_p)$  be uniformly random, and let  $(X_k)_{k \in \mathbb{N}_0}$  be the associated random Lie bracket. Let  $\mu_{2k}$  be the distribution of  $X_{2k}$ . There is a distribution  $\nu_{2k}$  that is uniform on a coset of a subgroup of  $\mathbb{F}_p^* = \text{Lin}\{[A, B]\} \setminus \{0\}$  such that with probability at least  $1 - \delta$ :

1. For  $k < cp$ , we have  $d(\mu_{2k}, \nu_{2k}) > 1 - \epsilon$ .
2. For  $k > Cp$ , we have  $d(\mu_{2k}, \nu_{2k}) < \epsilon$ .

After  $2k + 1$  of steps, the distribution  $\nu_{2k}$  equally splits into two distributions, one supported on the line  $\text{Lin}\{[A, [A, B]]\}$  and the other on the line  $\text{Lin}\{[B, [A, B]]\}$ .

**Example.** Consider the example above. Take  $k = 50$  and generate 1000 independent samples of the random Lie bracket  $X_{100}$ . The following table shows the sample distribution:

$$\begin{pmatrix} [A, B] & 5[A, B] & 4[A, B] & 3[A, B] & 9[A, B] \\ 189 & 185 & 193 & 209 & 224 \end{pmatrix}$$

In this case, the distribution  $\nu_{100}$  (indeed, any  $\nu_{2k}$  with  $k > 0$ ) is uniform over the five listed multiples of  $[A, B]$ . The total variation distance between the sample distribution and  $\nu_{100}$  is 0.033.

For randomly chosen  $A, B$ , the distribution  $\nu_{2k}$  is uniform on the whole of  $\text{Lin}\{[A, B]\} \setminus \{0\}$  with a certain probability that depends on the prime divisors of  $p - 1$ .

**Theorem 1.2.** Let  $(s, t)$  be a subinterval of

$$\mathcal{I} = \left[ \frac{1}{\zeta(2)}, \frac{2}{3} \right] \cup \left[ \frac{9}{8\zeta(2)}, \frac{3}{4} \right] \approx [0.6079, 0.6667] \cup [0.6839, 0.7500].$$

Then there is a set of primes  $p$  of positive lower density with the following property. Let  $A, B$  be uniformly random in  $\mathfrak{sl}_2(\mathbb{F}_p)$ , and let  $\nu$  be the uniform distribution on  $\text{Lin}\{[A, B]\} \setminus \{0\}$ . Then  $\mathbf{P}_{A,B}(\nu_{2k} = \nu) \in (s, t)$  for all  $k \geq 0$ .<sup>2</sup>

Based on our inspection of the random Lie bracket on  $\mathfrak{sl}_2(\mathbb{F}_p)$  driven by random generators  $A$  and  $B$ , we can prove an additional result in this setting. Although the random walk requires roughly  $p$  steps to approach its limiting distribution, the two elements  $A$  and  $B$  generate the entire Lie algebra  $\mathfrak{sl}_2(\mathbb{F}_p)$  much more quickly. In fact, we will show that, with high probability, every element of  $\mathfrak{sl}_2(\mathbb{F}_p)$  can be reached after only  $O(\log p)$  steps.

We measure steps in a weighted way using the notion of the diameter of a Lie algebra as defined in Dona (2023). Let  $S$  be a subset of a Lie algebra  $\mathfrak{g}$ . The set  $S$  is called *symmetric* if  $0 \in S$  and  $S = -S$ . For a symmetric set  $S$ , we inductively define the set of weighted  $k$ -balls as

$$S^1 = S, \quad S^k = \bigcup_{0 < j < k} ((S^j + S^{k-j}) \cup [S^j, S^{k-j}]) \quad (k \geq 2),$$

where we have denoted  $X + Y = \{x + y \mid x \in X, y \in Y\}$  and  $[X, Y] = \{[x, y] \mid x \in X, y \in Y\}$  for subsets  $X, Y \subseteq \mathfrak{g}$ . If  $S$  is a symmetric generating set of a Lie algebra  $\mathfrak{g}$  over  $\mathbf{F}_p$ , then the *diameter*  $\text{diam}(\mathfrak{g}, S)$  of  $\mathfrak{g}$  with respect to  $S$  is the smallest  $k$  such that  $S^k = \mathfrak{g}$ .

For any classical Lie algebra  $\mathfrak{g}$  over  $\mathbf{F}_p$ , Dona proves in Dona (2023) that there is a constant  $C = O(\dim(\mathfrak{g})^2 \log \dim(\mathfrak{g}))$  such that

$$\text{diam}(\mathfrak{g}, S) \leq (\log |\mathfrak{g}|)^C$$

for any symmetric generating set  $S$  of  $\mathfrak{g}$ . For the case of  $\mathfrak{sl}_2(\mathbf{F}_p)$ , this gives a bound of the form  $(\log p)^{O(1)}$ . We will now improve this bound to linear in  $\log p$  for the case when the generating set is chosen uniformly at random.

**Theorem 1.3.** *There is a (possibly empty) small set<sup>3</sup> of primes  $T$  with the following property. For any  $\epsilon > 0$  there is a  $C$  such that for all primes  $p > C$  not in  $T$ , we have*

$$\mathbf{P}_{A,B}(\text{diam}(\mathfrak{sl}_2(\mathbf{F}_p), S) \leq 120 \log p + 8) \geq 1 - \epsilon,$$

where  $S = \{0, \pm A, \pm B\}$  for uniformly random  $A, B \in \mathfrak{sl}_2(\mathbf{F}_p)$ .

The situation is very similar to what happens in groups: the diameter of  $\text{SL}_2(\mathbf{F}_p)$  is  $(\log p)^{O(1)}$  for any generating set, and it is  $O(\log p)$  for a random generating set (Helfgott 2008, Corollary 6.5).

## 1.1 | Reader's Guide

The ideas behind these results can be outlined as follows. First, we analyze the initial steps of the random walk and observe that odd and even steps exhibit distinct behaviors. This observation allows us to reduce the problem to a random walk on the cyclic group  $\mathbf{F}_p^*$  of order  $n = p - 1$  (see Section 2). The key parameters of this random walk are the coefficients of the Gram matrix associated with the Hilbert–Schmidt inner product on  $\mathfrak{sl}_2(\mathbf{F}_p)$ . We then study the distribution of these parameters and demonstrate that they are nearly uniform (see Section 3). As a result, the problem reduces to analyzing a random walk on the cyclic group  $\mathbf{Z}_n$  with almost uniformly random generators, which can be approached using Fourier analysis on cyclic groups (see Section 4). We identify the characters that contribute to the limiting distribution  $\nu_{2k}$  and relate the probability of converging to the uniform distribution to a  $\zeta(2)$ –type condition concerning the prime divisors of  $n = p - 1$  (see Section 5). After that, we establish the pre-cutoff<sup>4</sup> phenomenon for the random Lie bracket (see Section 6). Upper bounds on the total variation distance in the generic case were first obtained in Hildebrand (1994), and lower bounds were established in Hildebrand (2005) (see Theorem 2 and the exercises following it). Here, we provide simpler proofs tailored to our specific requirements. For the upper bound, we adapt recent elementary methods based on Fourier analysis from Abrams et al. (2022). These techniques alone yield only a  $1/2 - \epsilon$  lower bound, as we discuss. We therefore streamline the probabilistic argument from Hildebrand (2005) in our context to obtain the stronger  $1 - \epsilon$  bound. Lastly, we discuss the linear bound of  $O(\log p)$  on the diameter with respect to generic generators (see Section 7). Our result here relies on embedding a variant of the well-known Chung–Diaconis–Graham affine random walk on  $\mathbf{F}_p$  Chung et al. (1987) into the random Lie bracket, which allows us to leverage the results of Breuillard and Varjú (2022) on covering times of such walks.

## 2 | Translating the Walk to a Cyclic Group

### 2.1 | Inner Product and Adjoint Operators

The vector space  $\mathfrak{sl}_2(\mathbf{F}_p)$  comes equipped with the inner product  $\langle A, B \rangle = \text{tr}(AB)$ . Note that  $\langle A, A \rangle = \text{tr}(A^2)$ , and since  $A^2 + \det(A) = 0$  for a matrix in  $\mathfrak{sl}_2(\mathbf{F}_p)$ , we have  $\langle A, A \rangle = -2 \det(A)$ .<sup>5</sup>

**Lemma 2.1** (compare with Proposition 6.2 in Cantor et al. (2025)). *For any two elements  $A, B \in \mathfrak{sl}_2(\mathbb{F}_p)$ , we have*

$$[A, B, B] = 2\langle B, B \rangle A - 2\langle A, B \rangle B.$$

*Proof.* The characteristic polynomial of  $A$  is  $A^2 + \det A = 0$ , hence  $A^3 - \frac{1}{2}\langle A, A \rangle A = 0$ . Multilinearize this identity by considering it in the Lie algebra  $\mathfrak{sl}_2(\mathbb{F}_p) \otimes_{\mathbb{F}_p} \mathbb{F}_p[T]/(T^2)$  and using it with the element  $TA + B$ . Inspecting the  $T$ -term, we obtain

$$AB^2 + BAB + B^2A - \frac{1}{2}(2\langle A, B \rangle B + \langle B, B \rangle A) = 0.$$

Taking into account that  $B^2 = \frac{1}{2}\langle B, B \rangle$ , we get

$$\langle B, B \rangle A - 2\langle A, B \rangle B + 2BAB = 0.$$

The last term of the sum is  $2BAB = 2B[A, B] + 2B^2A = 2B[A, B] + \langle B, B \rangle A$ , hence

$$2\langle B, B \rangle A - 2\langle A, B \rangle B = -2B[A, B]. \quad (1)$$

Finally, we have  $B[A, B] = BAB - B^2A = BAB - AB^2 = [B, A]B$ , therefore

$$[A, B, B] = [A, B]B - B[A, B] = -2B[A, B],$$

which is the right-hand side of (1).  $\square$

We clearly have  $\text{ad}_A A = 0$ ,  $\text{ad}_A B = [A, B]$ , and it follows from Lemma 2.1 that  $\text{ad}_A[A, B] = -2\langle A, B \rangle A + 2\langle A, A \rangle B$ . Symmetrically, we have  $\text{ad}_B A = -[A, B]$ ,  $\text{ad}_B B = 0$ , and  $\text{ad}_B[A, B] = -2\langle B, B \rangle A + 2\langle A, B \rangle B$ . Thus, if  $A, B$  is a generating pair of the Lie algebra  $\mathfrak{sl}_2(\mathbb{F}_p)$ , the matrices  $A, B, [A, B]$  generate it as a vector space. In this basis, the adjoint operators  $\text{ad}_A$  and  $\text{ad}_B$  can be represented by the matrices

$$\text{ad}_A = \begin{pmatrix} 0 & 0 & -2\langle A, B \rangle \\ 0 & 0 & 2\langle A, A \rangle \\ 0 & 1 & 0 \end{pmatrix}, \quad \text{ad}_B = \begin{pmatrix} 0 & 0 & -2\langle B, B \rangle \\ 0 & 0 & 2\langle A, B \rangle \\ -1 & 0 & 0 \end{pmatrix}.$$

To simplify notation we shall write  $\alpha = 2\langle A, B \rangle$ ,  $\beta = 2\langle A, A \rangle$  and  $\gamma = 2\langle B, B \rangle$ .

## 2.2 | Odd/Even Number of Steps

The random walk  $(X_k)_{k \in \mathbb{N}_0}$  begins at  $X_0 = [A, B]$ . After the first step,  $X_1$  is uniformly distributed among the elements  $\text{ad}_A X_0 = -\alpha A + \beta B$  and  $\text{ad}_B X_0 = -\gamma A + \alpha B$ . At the second step,  $X_2$  is uniformly distributed among  $\beta[A, B]$ ,  $\alpha[A, B]$ ,  $\alpha[A, B]$ , and  $\gamma[A, B]$ .

By induction, we observe that for all even indices,  $X_{2k} \in \text{Lin}\{[A, B]\}$ , where  $\text{Lin}\{[A, B]\}$  denotes the line spanned by  $[A, B]$ . Similarly, for all odd indices,  $X_{2k+1} \in \text{Lin}\{-\alpha A + \beta B\} \cup \text{Lin}\{-\gamma A + \alpha B\}$ . Given the distribution of  $X_{2k}$  on the line  $\text{Lin}\{[A, B]\}$ , it follows that  $X_{2k+1}$  is distributed proportionally (scaled by a factor of 1/2) across the lines  $\text{Lin}\{-\alpha A + \beta B\}$  and  $\text{Lin}\{-\gamma A + \alpha B\}$ .

Thus, to understand the behavior of our random walk, it suffices to focus on the distribution after an even number of steps. This allows us to reduce the problem to analyzing how the distribution evolves along the line  $\text{Lin}\{[A, B]\}$ .

The sequence  $(X_{2k})_{k \in \mathbb{N}}$  is equivalent to a random walk on  $\mathbb{F}_p$ , where we start at the element  $1 \in \mathbb{F}_p$  and, at each step, multiply by  $\alpha$  with probability 1/2, or by  $\beta$  or  $\gamma$  each with probability 1/4. Provided that  $\alpha, \beta$ , and  $\gamma$  are nonzero, this random walk occurs on the group  $\mathbb{F}_p^*$ . By fixing an isomorphism  $\log : \mathbb{F}_p^* \rightarrow \mathbb{Z}_n$  where  $n = p - 1$ , we can translate the problem to a random walk on the cyclic group  $\mathbb{Z}_n$ .

### 3 | Distribution of the Gram Matrix

In order to analyze our random walk, we need to determine how the parameters  $\alpha = 2\langle A, B \rangle$ ,  $\beta = 2\langle A, A \rangle$ , and  $\gamma = 2\langle B, B \rangle$  are distributed in  $\mathbf{F}_p$  as  $A, B$  vary across  $\mathfrak{sl}_2(\mathbf{F}_p)$ . These parameters are essentially coefficients of the Gram matrix of the inner product on  $\mathfrak{sl}_2(\mathbf{F}_p)$ :

$$G_{A,B} = \begin{pmatrix} \langle A, A \rangle & \langle A, B \rangle \\ \langle A, B \rangle & \langle B, B \rangle \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \beta & \alpha \\ \alpha & \gamma \end{pmatrix}.$$

We thus consider the Gram map to the set of all symmetric matrices,

$$G : \mathfrak{sl}_2(\mathbf{F}_p)^2 \rightarrow \text{Sym}_2(\mathbf{F}_p), \quad (A, B) \mapsto G_{A,B}.$$

The orthogonal group  $O(3)$  over  $\mathbf{F}_p$  acts naturally on the space  $\mathfrak{sl}_2(\mathbf{F}_p)$  (see Clark (2013)), with Witt decompositon  $\langle e, f \rangle \oplus \langle h \rangle$ , where  $e = E_{12}$ ,  $f = E_{21}$ , and  $h = E_{11} - E_{22}$ . The discriminant of the form is  $-2$ . This action induces an action on pairs  $(A, B) \in \mathfrak{sl}_2(\mathbf{F}_p)^2$ . The Gram map  $G$  is invariant under this action.

#### 3.1 | Fibers of the Gram Map

**Lemma 3.1.** *The number of elements  $A \in \mathfrak{sl}_2(\mathbf{F}_p)$  with  $\langle A, A \rangle = a$  is equal to  $p(p + \eta(a/2))$ , where  $\eta$  is the quadratic character on  $\mathbf{F}_p$ .*

*Proof.* Write  $A = xe + yf + zh$  with  $x, y, z \in \mathbf{F}_p$ . The condition  $\langle A, A \rangle = a$  is equivalent to  $2xy + 2z^2 = a$ . This is an affine quadric. If  $y \neq 0$ , we can uniquely solve for  $x$ , giving  $(p-1)p$  solutions. If  $y = 0$ , we have  $z^2 = a/2$ , which has  $1 + \eta(a/2)$  solutions for  $z$ , and  $x$  is arbitrary, giving  $p(1 + \eta(a/2))$  solutions.  $\square$

**Proposition 3.2.** *Let  $X \in \text{Sym}_2(\mathbf{F}_p)$ . Then*

$$|G^{-1}(X)| = \begin{cases} p^3 - p & \text{rank } X = 2 \\ 2p^3 + p^2 - p & \text{rank } X = 1, \eta(X_{11}/2), \eta(X_{22}/2) \geq 0 \\ p^2 - p & \text{rank } X = 1, \eta(X_{11}/2) = -1 \text{ or } \eta(X_{22}/2) = -1 \\ p^3 + p^2 - p & \text{rank } X = 0. \end{cases}$$

*Proof.* Each fiber of  $G$  is a disjoint union of orbits under the action of  $O(3)$ . The size of the orthogonal group is  $|O(3)| = 2p(p^2 - 1)$ .

Suppose first that  $\text{rank } X = 2$ . Each pair  $(A, B) \in \mathfrak{sl}_2(\mathbf{F}_p)^2$  with  $G_{A,B} = X$  determines a nondegenerate quadratic subspace  $\langle A, B \rangle$  of  $\mathfrak{sl}_2(\mathbf{F}_p)$ . For any other pair  $(A', B')$  with  $G_{A',B'} = X$ , we have an isometry mapping  $A \mapsto A'$ ,  $B \mapsto B'$ . By Witt's extension lemma, this isometry can be extended to an action of  $O(3)$  on  $\mathfrak{sl}_2(\mathbf{F}_p)$  that maps  $(A, B)$  to  $(A', B')$ . Thus, the fiber of  $G$  over  $X$  is a single orbit of  $O(3)$ . The stabilizer  $\text{Stab}_{O(3)}(A, B)$  fixes  $\langle A, B \rangle$  pointwise, hence it also preserves the line  $\langle A, B \rangle^\perp$ , on which it can only act as  $\pm \text{id}$ . Hence  $|\text{Stab}_{O(3)}(A, B)| = 2$  and so the orbit of  $O(3)$  on  $(A, B)$  is of size  $p^3 - p$ .

Assume now that  $\text{rank } X = 0$ . Suppose  $(A, B)$  is a pair with  $G_{A,B} = X$ . Since the Witt index of the quadratic space  $\mathfrak{sl}_2(\mathbf{F}_p)$  is 1, we must have  $\dim \langle A, B \rangle < 2$ . Assume  $(A, B) \neq (0, 0)$ , which forms a single orbit of size 1. Thus  $(A, B) = (\lambda Z, \mu Z)$  for some  $(\lambda, \mu) \in \mathbf{F}_p^2 \setminus \{0\}$  and  $0 \neq Z \in \mathfrak{sl}_2(\mathbf{F}_p)$  with  $\langle Z, Z \rangle = 0$ . By Witt extension lemma, this pair maps to a pair with  $Z = e$  under  $O(3)$ . Two such pairs can further be mapped to each other by an  $O(3)$  map if and only if their  $(\lambda, \mu)$  vectors are dependent. Thus, a set of orbit representatives is parameterized by vectors in  $\mathbf{P}^1$ , so there are  $p+1$  orbits in total. Each such orbit has size  $|O(3).e|$ , which is the number of nonzero isotropic vectors in  $\mathfrak{sl}_2(\mathbf{F}_p)$ . There are a total of  $p^2 - 1$  of these by the previous lemma. In total, we obtain  $1 + (p+1)(p^2 - 1) = p^3 + p^2 - p$  pairs in the fiber over 0.

Consider finally the option when  $\text{rank } X = 1$ . Suppose  $(A, B)$  is a pair with  $G_{A,B} = X$ .

$\dim \langle A, B \rangle = 1$ : This means  $(A, B) = (xZ, yZ)$  for some  $Z \in \mathfrak{sl}_2(\mathbf{F}_p)$  and  $x, y$  not both 0. The Gram matrix of this pair is

$$X = \langle Z, Z \rangle \begin{pmatrix} x^2 & xy \\ xy & y^2 \end{pmatrix},$$

so  $\langle Z, Z \rangle \neq 0$ . Let  $(C, D)$  be another pair with  $G_{C,D} = X$ . Thus  $(C, D) = (\alpha W, \beta W)$  for some  $\alpha, \beta$  not both zero. Comparing the Gram matrices, it follows that  $\langle W, W \rangle / \langle Z, Z \rangle$  is a nonzero square  $u^2 \in \mathbf{F}_p$ . Hence  $W$  can be mapped by an element of  $O(3)$  to  $uZ$ , and the pair  $(C, D)$  gets mapped to  $(\alpha u Z, \beta u Z)$ . Again, comparing the Gram matrices gives  $\alpha u = \pm x$  and  $\beta u = \pm y$ , where nonequal signs can only occur if  $\alpha\beta = 0$ . Hence  $(C, D) = (xZ, yZ) = (A, B)$  or  $(C, D) = (-A, -B)$  when  $\alpha\beta \neq 0$ , and  $(C, D) = (0, \pm yZ) = (0, \pm B)$  or  $(\pm xZ, 0) = (\pm A, 0)$  when  $\alpha\beta = 0$ . In all cases, an extra application of an element of  $O(3)$  identifies elements with distinct signs. Therefore  $(C, D)$  is in the orbit of  $(A, B)$ , and hence  $(A, B)$  has a single orbit under  $O(3)$ . The stabilizer of any such point is the transformations in  $O(3)$  that fix  $Z$ . Since  $Z$  is anisotropic, we have two options for the complement: either it is split or non-split, depending on whether the discriminant  $\Delta(\langle Z \rangle^\perp) = -2/\langle Z, Z \rangle$  is  $-1$  modulo squares. In the split case, the stabilizer has size  $2(p-1)$ , and in the non-split case, it has size  $2(p+1)$ . Thus, the stabilizer is of size  $2(p-\eta(\langle Z, Z \rangle/2))$ . The number of pairs in the fiber over  $X$  is thus  $p(p+\eta(\langle Z, Z \rangle/2))$ . These orbits appear whenever  $X$  is of rank 1.

$\dim \langle A, B \rangle = 2$ : In this case, we can find a nonzero vector  $C \in \langle A, B \rangle$  that is orthogonal to  $A$  and  $B$ . Hence  $C$  is isotropic, and we can map it to  $e$  under  $O(3)$ . Then  $A, B \in \langle e \rangle^\perp = \langle e, h \rangle$ . Thus  $A = x_0 e + xh$  and  $B = y_0 e + yh$  with not both  $x, y$  zero. Assuming  $y \neq 0$ , an orthogonal transformation maps this pair into  $(e+xh, yh)$ . On the other hand, if  $y = 0$ , we can map it to  $(xh, e)$ . In both cases, the values of  $x, y$  are uniquely determined by  $X$  (the option  $-x, -y$  is equivalent to  $(A, B)$  by the flip along  $\langle e, f \rangle$ ). Hence, we again have a unique orbit over any  $X$ . Since the stabilizer of  $(e, f)$  is trivial, the orbit is of size  $2p(p^2-1)$ . These orbits only appear when  $\eta(X_{11}/2), \eta(X_{22}/2) \geq 0$  (not both zero).

Summing up the contributions from the two cases completes the proof.  $\square$

**Corollary 3.3.** *Let  $A, B$  be uniformly random in  $\mathfrak{sl}_2(\mathbf{F}_p)$ . Then  $G_{A,B}$  has all entries nonzero with probability at least  $1 - 9/p$ .*

*Proof.* Each fiber of the Gram map is of size at most  $3p^3$ . The probability that  $G_{A,B}$  has a zero entry can thus be upper bounded by the union bound as

$$\frac{1}{p^6} \sum_{X \in \text{Sym}_2(\mathbf{F}_p) \setminus \text{Sym}_2(\mathbf{F}_p^*)} |G^{-1}(X)| \leq \frac{3p^3 \cdot 3p^2}{p^6} = \frac{9}{p}.$$

$\square$

### 3.2 | Distribution of Parameters Governing the Random Walk

Let  $\mathcal{U} = G^{-1}(\text{Sym}_2(\mathbf{F}_p^*))$  be the set of pairs  $(A, B) \in \mathfrak{sl}_2(\mathbf{F}_p)^2$  whose Gram matrix  $G_{A,B}$  has all entries nonzero. For any  $(A, B) \in \mathcal{U}$ , the random walk on  $\mathfrak{sl}_2(\mathbf{F}_p)$  after an even number of steps reduces to a random walk on the cyclic group  $\mathbf{Z}_{p-1}$  as in Section 2. In this case, the behavior of the walk is determined by two parameters  $a, b$  extracted from  $G_{A,B}$  via the map

$$P : \text{Sym}_2(\mathbf{F}_p^*) \rightarrow \mathbf{Z}_{p-1}^2, \quad X = \frac{1}{2} \begin{pmatrix} \beta & \alpha \\ \alpha & \gamma \end{pmatrix} \mapsto (\log(\alpha/\beta), \log(\alpha/\gamma)).$$

Let  $F = P \circ G : \mathcal{U} \rightarrow \mathbf{Z}_{p-1}^2$  be the composite map, so that  $(a, b) = F(A, B)$ . Let  $\pi$  be the pushforward measure of the uniform distribution on  $\mathfrak{sl}_2(\mathbf{F}_p)^2$  under  $F$ . In other words,  $\pi(S) = |F^{-1}(S)|/p^6 = \mathbf{P}_{A,B}(P(G_{A,B}) \in S)$  for any subset  $S \subseteq \mathbf{Z}_{p-1}^2$ . Denote the uniform distribution on  $\mathbf{Z}_{p-1}^2$  by  $\nu_{\mathbf{Z}_{p-1}^2}$ . As  $p$  grows to infinity, the distributions  $\pi$  and  $\nu_{\mathbf{Z}_{p-1}^2}$  become close to each other.

**Proposition 3.4.** *For any subset  $S \subseteq \mathbf{Z}_{p-1}^2$ , we have*

$$|\pi(S) - \nu_{\mathbf{Z}_{p-1}^2}(S)| \leq \frac{6}{p}.$$

To prove this proposition, we first examine the sizes of fibers of the map  $F$ .

**Lemma 3.5.** *For any  $x = (a, b) \in \mathbf{Z}_{p-1}^2$  we have*

$$\left| |F^{-1}(x)| - \frac{p^6}{(p-1)^2} \right| \leq 6p^3.$$

*Proof.* We have  $|P^{-1}(x)| = p - 1$ . First, suppose that  $a + b = 0$ . In this case we have  $\alpha^2 = \beta\gamma$ , so all matrices  $X \in P^{-1}(x)$  have rank 1. Half of these satisfy  $\eta(X_{11}/2) = \eta(X_{22}/2) = 1$ , and the other half satisfy  $\eta(X_{11}/2) = \eta(X_{22}/2) = -1$ . It follows from Proposition 3.2 that

$$|F^{-1}(x)| = \frac{p-1}{2}(2p^3 + p^2 - p) + \frac{p-1}{2}(p^2 - p) = p^4 - 2p^3 + p.$$

We compute  $0 \leq p^6 - (p^4 - 2p^3 + p)(p-1)^2 \leq 6p^3(p-1)^2$ . Dividing by  $(p-1)^2$  completes the proof for this case.  $\square$

Now, suppose that  $a + b \neq 0$ . In this scenario, all matrices  $X \in P^{-1}(x)$  have rank 2. Again, by Proposition 3.2, the size of the fiber  $F^{-1}(x)$  is  $(p-1)(p^3 - p) = p^4 - p^3 - p^2 + p$ . As in the previous case, this quantity is within  $6p^3$  of  $p^6/(p-1)^2$ .  $\square$

*Proof of Proposition 3.4.* We have

$$|\pi(S) - \nu_{\mathbf{Z}_{p-1}^2}(S)| = \left| \frac{1}{p^6} \sum_{x \in S} |F^{-1}(x)| - \frac{|S|}{(p-1)^2} \right| \leq \frac{1}{p^6} \sum_{x \in S} \left| |F^{-1}(x)| - \frac{p^6}{(p-1)^2} \right|.$$

The last sum can be upper bounded by the last lemma as  $6p^3|S|/p^6 \leq 6/p$ .  $\square$

We will need the following corollary multiple times in the rest of the paper.

**Corollary 3.6.** Let  $S \subseteq \mathfrak{sl}_2(\mathbf{F}_p)^2$  and  $T \subseteq \mathbf{Z}_{p-1}^2$  with  $S \cap \mathcal{U} = F^{-1}(T)$ . Then

$$\left| \frac{|S|}{p^6} - \frac{|T|}{(p-1)^2} \right| < 15/p.$$

*Proof.* Since  $F^{-1}(T) \subseteq S \subseteq F^{-1}(T) \cup (\mathfrak{sl}_2(\mathbf{F}_p)^2 \setminus \mathcal{U})$ , we have

$$\left| \frac{|S|}{p^6} - \frac{|F^{-1}(T)|}{p^6} \right| \leq \frac{|\mathfrak{sl}_2(\mathbf{F}_p)^2 \setminus \mathcal{U}|}{p^6} \leq \frac{9}{p}$$

by Corollary 3.3. On the other hand, we have

$$\left| \frac{|F^{-1}(T)|}{p^6} - \frac{|T|}{(p-1)^2} \right| = |\pi(T) - \nu_{\mathbf{Z}_{p-1}^2}(T)| \leq \frac{6}{p}$$

by Proposition 3.4. The result now follows by the triangle inequality.  $\square$

We shall apply the above corollary as follows. Suppose that for every  $(A, B) \in \mathcal{U}$ , the random walk on  $\mathfrak{sl}_2(\mathbf{F}_p)$  exhibits a property  $\mathcal{P}$  if and only if the corresponding parameters  $(a, b) = F(A, B) \in \mathbf{Z}_{p-1}^2$  satisfy a condition  $\mathcal{C}$ . Then, for uniformly random  $A, B \in \mathfrak{sl}_2(\mathbf{F}_p)$ , the probability that the walk exhibits property  $\mathcal{P}$  differs by at most  $15/p$  from the probability that uniformly random  $(a, b) \in \mathbf{Z}_{p-1}^2$  satisfy  $\mathcal{C}$ .

## 4 | A Brief Overview of Fourier Analysis on Cyclic Groups

We give a brief overview of Fourier analysis on cyclic groups, following the exposition in Diaconis (1988).

### 4.1 | Setup

Let  $S$  be a subset of a cyclic group  $\mathbf{Z}_n$  equipped with a probability measure  $p : S \rightarrow \mathbf{C}$ . Let  $(Y_k)_{k \in \mathbf{N}_0}$  be the random walk on  $\mathbf{Z}_n$  starting at 0 with transition probabilities given by  $p(x, y) = p(y - x)$ . Let  $\sigma_k$  be the probability distribution of the walk after  $k$  steps, that is,  $\sigma_k(x) = \mathbf{P}(Y_k = x)$  for  $x \in \mathbf{Z}_n$ . This distribution is a function in the vector space  $L^2(\mathbf{Z}_n)$  of all complex-valued functions on  $\mathbf{Z}_n$ .

## 4.2 | Markov Operator

Let  $M : L^2(\mathbf{Z}_n) \rightarrow L^2(\mathbf{Z}_n)$  be the operator defined by

$$(Mf)(x) = \sum_{s \in S} p(s)f(x - s).$$

The distribution  $\sigma_k$  can be written as

$$\sigma_k(x) = \sum_{s \in S} p(s)\sigma_{k-1}(x - s) = (M\sigma_{k-1})(x),$$

and so  $\sigma_k = M^k\sigma_0 = M^k1_0$ . In order to understand high powers of the Markov operator, we need to analyze its eigenvalues and eigenfunctions. For any character  $\chi$  of the group  $\mathbf{Z}_n$ , we have

$$(M\chi)(x) = \left( \sum_{s \in S} p(s)\chi(-s) \right) \chi(x),$$

so  $\chi$  is an eigenfunction of  $M$  with eigenvalue  $\lambda = \sum_{s \in S} p(s)\chi(-s)$ . Characters of  $\mathbf{Z}_n$  are of the form  $\chi_j(x) = \omega^{jx}$  for  $j \in \mathbf{Z}_n$ , where  $\omega = e^{2\pi i/n}$ , and they form a basis of  $L^2(\mathbf{Z}_n)$ .

## 4.3 | Limiting Distribution

The initial distribution  $1_0$  can be expressed in terms of the basis of characters as  $1_0 = \frac{1}{n} \sum_{j=0}^{n-1} \chi_j$ . It follows that the distribution after  $k$  steps is given by

$$\sigma_k = \frac{1}{n} \sum_{j=0}^{n-1} \lambda_j^k \chi_j,$$

where  $\lambda_j = \sum_{s \in S} p(s)\chi_j(-s)$  is the eigenvalue corresponding to  $\chi_j$ . We can decompose  $\sigma_k$  into two components  $\sigma_k = \phi_k + \omega_k$ , where

$$\phi_k = \frac{1}{n} \sum_{|\lambda_j|=1} \lambda_j^k \chi_j, \quad \omega_k = \frac{1}{n} \sum_{|\lambda_j|<1} \lambda_j^k \chi_j.$$

As  $k$  grows to infinity, the contribution of  $\omega_k$  diminishes to zero, making  $\sigma_k$  increasingly close to  $\phi_k$ .<sup>6</sup> This proximity can be quantified in terms of the spectral radius  $\rho = \max\{|\lambda_j| \mid |\lambda_j| < 1\}$  of the Markov operator.

**Lemma 4.1** (Lemma 3 in Abrams et al. (2022)). *The distance to the limiting distribution satisfies*

$$\rho^{2k} \leq \|\sigma_k - \phi_k\|_1^2 \leq \sum_{|\lambda_j|<1} |\lambda_j|^{2k}.$$

*Proof.* The characters  $\chi_j$  form an orthogonal basis of  $L^2(\mathbf{Z}_n)$  and have norm  $\sqrt{n}$ . Thus

$$\|\omega_k\|_1^2 \leq n \|\omega_k\|_2^2 = \frac{1}{n} \sum_{|\lambda_j|<1} \|\lambda_j^k \chi_j\|_2^2 = \sum_{|\lambda_j|<1} |\lambda_j|^{2k},$$

proving the second inequality. For the first inequality, let  $\chi$  be a character of  $\mathbf{Z}_n$  corresponding to an eigenvalue with absolute value  $\rho$ . Since  $\|\chi\|_\infty = 1$ , we have

$$\|\omega_k\|_1 \geq |\langle \omega_k, \chi \rangle| = \frac{1}{n} \rho^k |\langle \chi, \chi \rangle| = \rho^k.$$

□

## 5 | The Limiting Distribution

We are now ready to analyze our random Lie bracket in more detail. We have already seen that the distribution of the random walk after an even number of steps is equivalent to a random walk on the cyclic group  $\mathbf{F}_p^*$  starting at 1 and evolving by multiplication with  $\alpha, \beta, \gamma$  (all nonzero) with probabilities  $1/2, 1/4, 1/4$ .

## 5.1 | Log

Fix an isomorphism  $\log : \mathbf{F}_p^* \rightarrow \mathbf{Z}_n$  and let  $\alpha' = \log \alpha$ ,  $\beta' = \log \beta$ , and  $\gamma' = \log \gamma$ . Under this mapping, the random walk on  $\mathbf{F}_p^*$  translates to a random walk on  $\mathbf{Z}_n$ , starting at  $0 = \log 1 \in \mathbf{Z}_n$ . At each step, the walk adds  $\alpha'$  with probability  $1/2$ , or  $\beta'$  or  $\gamma'$  each with probability  $1/4$ . Let  $S' = \{\alpha', \beta', \gamma'\}$  be the set of steps of the walk with probability measure  $p(\alpha') = 1/2$ ,  $p(\beta') = p(\gamma') = 1/4$ . The limiting distribution on  $\mathbf{Z}_n$  is given by

$$\phi_k = \frac{1}{n} \sum_{|\lambda_j|=1} \lambda_j^k \chi_j.$$

## 5.2 | Contributing Characters

Let us examine which characters  $\chi_j$  appear in the sum  $\phi_k$ . Certainly, the trivial character  $\chi_0$  will always contribute to the limiting distribution. The character  $\chi_j$  will contribute if and only if

$$|\lambda_j| = \left| \sum_{s \in S'} p(s) \chi_j(-s) \right| = 1.$$

As  $|\chi_j(-s)| = 1$  for each  $s \in S'$ , this holds if and only if  $\chi_j(-\alpha') = \chi_j(-\beta') = \chi_j(-\gamma') = 1$ , which is equivalent to  $\chi_j(\alpha' - \beta') = \chi_j(\alpha' - \gamma') = 1$ . Introduce parameters  $a = \alpha' - \beta'$  and  $b = \alpha' - \gamma'$ . Hence a character  $\chi_j$  appears in the limiting distribution if and only if  $a, b \in \ker \chi_j$ , and the corresponding eigenvalue is  $\lambda_j = \overline{\chi_j(\alpha')}$ . The random walk is thus essentially determined by the parameters  $a, b$ .

**Lemma 5.1.**  $\ker \chi_j = (n/\gcd(n, j))\mathbf{Z}_n$ .

*Proof.* An element  $x \in \mathbf{Z}_n$  satisfies  $x \in \ker \chi_j$  if and only if  $n$  divides  $jx$ . We can write  $j = dj'$  and  $n = dn'$ , where  $d = \gcd(j, n)$ . By canceling  $d$ , we see that this is equivalent to  $n'$  dividing  $j'x$  and hence to  $n'$  dividing  $x$ . Therefore  $\ker \chi_j = \frac{n}{d}\mathbf{Z}_n$  and  $|\ker \chi_j| = d$ .  $\square$

**Lemma 5.2.** Let  $a = \alpha' - \beta'$  and  $b = \alpha' - \gamma'$ . Then the characters appearing in  $\phi_k$  are precisely those  $\chi_j$  for which  $j$  is divisible by  $n/\gcd(a, b, n)$ . The number of contributing characters is thus  $\gcd(a, b, n)$ .

*Proof.* The character  $\chi_j$  contributes to the limiting distribution if and only if  $a, b \in \ker \chi_j = (n/\gcd(n, j))\mathbf{Z}_n$ . This is equivalent to  $n/\gcd(n, j)|a, b$ , which is the same as  $n/\gcd(a, b, n)|\gcd(n, j)$ , and this is further equivalent to  $j$  being divisible by  $n/\gcd(a, b, n)$ .  $\square$

**Corollary 5.3.** The limiting distribution  $\phi_k$  is uniform if and only if  $\gcd(a, b, n) = 1$ .

In fact, we always get a uniform distribution on  $\gcd(a, b, n)\mathbf{Z}_n$  provided the number of steps  $k$  is divisible by  $\gcd(a, b, n)$ , and more generally a uniform distribution on a coset of  $\gcd(a, b, n)\mathbf{Z}_n$ .

**Lemma 5.4.** The distribution  $\phi_k$  is uniform on the coset  $k\alpha' + \gcd(a, b, n)\mathbf{Z}_n$ .

*Proof.* Let  $d = \gcd(a, b, n)$ . We have

$$\phi_k = \frac{1}{n} \sum_{\frac{n}{d}|j} \overline{\chi_j(\alpha')}^k \chi_j.$$

Any  $x \in k\alpha' + d\mathbf{Z}_n$  can be written as  $x = k\alpha' + md$  for some  $m \in \mathbf{Z}$ . Thus

$$\overline{\chi_j(\alpha')}^k \chi_j(x) = \chi_j(-k\alpha') \chi_j(k\alpha' + md) = \chi_j(md) = 1$$

for any  $j$  appearing in the sum. Therefore  $\phi_k(x) = d/n$ , and so  $\phi_k$  is uniform on  $k\alpha' + d\mathbf{Z}_n$ .  $\square$

### 5.3 | Randomly Chosen Parameters

Suppose that the parameters  $\alpha', \beta', \gamma'$  are chosen uniformly at random from  $\mathbf{Z}_n$ . In this case, the parameters  $a = \alpha' - \beta'$  and  $b = \alpha' - \gamma'$  are also uniformly distributed in  $\mathbf{Z}_n$ . Thus the probability that a character  $\chi_j$  contributes to the limiting distribution is equal to  $(|\ker \chi_j|/n)^2$ . Taking  $j = 2$ , we see that the probability that  $\chi_{n/2}$  contributes is  $1/4$ . It follows that the limiting distribution  $\phi_k$  is uniform with probability at most  $3/4$ . We will show that for any given  $\epsilon > 0$ , there is a positive proportion of primes for which the probability that any other character contributes to the limiting distribution is less than  $\epsilon$ . In fact, we can compute the exact probability that the limiting distribution is uniform in terms of the prime divisors of  $n$ .

**Lemma 5.5.** *Let  $a, b \in \mathbf{Z}_n$  be uniformly random. Then*

$$\mathbf{P}_{a,b}(\gcd(a, b, n) = 1) = \prod_{q|n} \left(1 - \frac{1}{q^2}\right),$$

where the product is over all primes  $q$  dividing  $n$ .

*Proof.* For each prime  $q|n$ , the probability that  $q$  divides both  $a$  and  $b$  is  $1/q^2$ . Since the events for distinct primes are independent, the probability that no prime divisor of  $n$  divides both  $a$  and  $b$  (which is exactly the probability that  $\gcd(a, b, n) = 1$ ) is

$$\prod_{q|n} \left(1 - \frac{1}{q^2}\right).$$

□

When the only small prime dividing  $n$  is 2, the probability that the limiting distribution is uniform is close to  $3/4$ . Conversely, when  $n$  is divisible by all the primes up to some large number (“divisible by all primes”), the probability that the limiting distribution is uniform is close to

$$\prod_{p \in \mathbf{P}} \left(1 - \frac{1}{p^2}\right) = \frac{1}{\zeta(2)} \approx 0.6079. \quad (2)$$

We will now give more precise estimates for the probability that the limiting distribution is uniform and show it can come arbitrarily close to almost all numbers between  $1/\zeta(2)$  and  $3/4$  apart from some subinterval.

### 5.4 | Probability of Convergence to the Uniform Distribution

**Lemma 5.6.** *Let  $X$  and  $Y$  be two finite disjoint sets of odd primes. There exists a positive lower density<sup>7</sup> subset of primes  $p \in \mathbf{P}$  that satisfy*

1. All primes from  $X$  divide  $p - 1$ .
2. No prime from  $Y$  divides  $p - 1$ .

*Proof.* Let  $x = \prod_{q \in X} q$  and  $y = \prod_{q \in Y} q$ . Since  $x$  and  $y$  are coprime, the conditions  $p \equiv 1 \pmod{x}$  and  $p \equiv 2 \pmod{y}$  are equivalent to  $p \equiv c \pmod{xy}$  for some  $c \in \mathbf{N}$  by the Chinese remainder theorem. As  $c$  is coprime to  $xy$ , the proportion of primes congruent to  $c$  modulo  $xy$  is  $1/\phi(xy) > 0$  (see (Serre 2012, Section VI.4)). For any prime  $p$  with this property,  $p - 1$  is divisible by  $x$  and is coprime to  $y$ . □

**Lemma 5.7.** *Let  $X = \sum_{j \geq 1} x_j$  be the sum of a convergent series of positive numbers, where  $x_k < \sum_{j \geq k+1} x_j$  for each  $k \in \mathbf{N}$ . Then for any open interval  $(s, t)$  contained in  $(0, X)$ , there is a finite subset  $S \subseteq \mathbf{N}$  such that  $\sum_{j \in S} x_j \in (s, t)$ .*

*Proof.* Inductively define subsets  $S_n \subseteq \mathbf{N}$  by

$$S_1 = \begin{cases} \{x_1\} & x_1 < t \\ \emptyset & \text{otherwise,} \end{cases} \quad S_{n+1} = \begin{cases} S_n \cup \{n+1\} & x_{n+1} + \sum_{j \in S_n} x_j < t \\ S_n & \text{otherwise.} \end{cases}$$

Take  $S = \bigcup_{n \geq 1} S_n$ . We may assume that  $t < X$  after possibly shortening the interval. Thus we can find some  $k \in \mathbf{N}$  with  $k \notin S$ . Then

$$\sum_{j \geq k+1} x_j + \sum_{j \in S_{k-1}} x_j > x_k + \sum_{j \in S_{k-1}} x_j \geq t,$$

so we can find some  $k' > k$  with  $k' \notin S$ . Therefore, there are arbitrarily large numbers that are not in  $S$ . Since  $\lim_{j \rightarrow \infty} x_j = 0$ , there is an  $m \in \mathbf{N} \setminus S$  with  $x_m < t - s$ . Since  $m \notin S$ , we have  $\sum_{j \in S_{m-1}} x_j + x_m \geq t$  hence  $\sum_{j \in S_{m-1}} x_j \geq t - x_m > s$ . Hence  $S_{m-1}$  is the desired finite subset.  $\square$

We shall apply the previous lemma with  $x_j = -\log(1 - 1/p_j^2)$ . The terms  $x_j$  are positive and  $\sum_{j \geq 1} x_j = \log \zeta(2)$  by (2). However, the condition that  $x_k < \sum_{j \geq k+1} x_j$  fails to hold for all  $k \in \mathbf{N}$ . Here is how the numbers look like for the first few primes:

$$\begin{array}{llll} x_1 \approx 0.2877 & x_2 \approx 0.1178 & x_3 \approx 0.0408 & x_4 \approx 0.0206 \\ \sum_{j \geq 2} x_j \approx 0.2100 & \sum_{j \geq 3} x_j \approx 0.0922 & \sum_{j \geq 4} x_j \approx 0.0514 & \sum_{j \geq 5} x_j \approx 0.0308 \end{array}$$

The inequality fails with  $k = 1, 2$ , and works with  $k = 3, 4$ . Let us verify that the condition does hold from that point on, so we can use the previous lemma with the sequence  $(x_k)_{k \geq 3}$ .

**Lemma 5.8.** *Let  $p_j$  be the  $j$ -th prime and let  $x_j = -\log(1 - 1/p_j^2)$ . Then  $x_k < \sum_{j \geq k+1} x_j$  for all  $k \geq 5$ .*

*Proof.* We first claim that  $x_{j+1} > x_j/2$  for any  $j \geq 5$ . This is equivalent to

$$\left(1 - \frac{1}{p_{j+1}^2}\right)^2 < 1 - \frac{1}{p_j^2}. \quad (3)$$

It follows from Axler (2019) that

$$j(\log j + \log \log j - 3/2) < p_j < j(\log j + \log \log j - 1/2)$$

for all  $j \geq 20$ . Basic analysis then gives  $p_{j+1} < 14p_j/10$  for any  $j \geq 28$ , and hence

$$\left(1 - \frac{1}{p_{j+1}^2}\right)^2 = 1 - \frac{2}{p_{j+1}^2} + \frac{1}{p_{j+1}^4} < 1 - \frac{200}{144} \frac{1}{p_j^2} + \frac{1}{p_j^4} < 1 - \frac{1}{p_j^2}.$$

It can be checked with a computer that (3) also holds for  $j \in \{5, 6, \dots, 27\}$ . We obtain, for any  $k \geq 5$ ,

$$\sum_{j \geq k+1} x_j > x_k \sum_{j \geq 1} \frac{1}{2^j} = x_k,$$

as required.  $\square$

**Proposition 5.9.** *Let  $(s, t)$  be a subinterval of*

$$\mathcal{I} = \left[ \frac{1}{\zeta(2)}, \frac{2}{3} \right] \cup \left[ \frac{9}{8\zeta(2)}, \frac{3}{4} \right].$$

*Then there is a set of primes  $p$  of positive lower density for which*

$$\prod_{q|p-1} \left(1 - \frac{1}{q^2}\right) \in (s, t).$$

*Proof.* Let  $x_j = -\log(1 - 1/p_j^2)$ . It follows by combining the previous two lemmas that the sums  $\sum_{j \in S} x_j$  with  $S$  a finite subset of  $\mathbf{N} \setminus \{1, 2\}$  form a dense subset of the interval  $(0, \log \zeta(2) + \log(3/4) + \log(8/9)) = (0, \log(2\zeta(2)/3))$ . Equivalently, the products  $\prod_{j \in S} e^{-x_j} = \prod_{j \in S} (1 - 1/p_j^2)$  form a dense subset of the interval  $(3/(2\zeta(2)), 1)$ .

Suppose first that  $(s, t) \subseteq (1/\zeta(2), 2/3)$ . By the argument above, we can find a set  $S$  with  $\prod_{j \in S} (1 - 1/p_j^2) \in (3s/2, 3t/2)$ . Let  $S' = S \cup \{1, 2\}$ , so that  $\prod_{j \in S'} (1 - 1/p_j^2) \in (s, t)$ . We can thus find a  $k$  larger than all elements in  $S'$  such that

$$\prod_{j \in S'} \left(1 - \frac{1}{p_j^2}\right) > s \cdot \exp\left(\sum_{j \geq k} x_j\right) = s \cdot \prod_{j \geq k} \left(1 - \frac{1}{p_j^2}\right)^{-1}.$$

Take  $X = \{p_j \mid j \in S'\}$  and  $Y = \{p_j \mid j < k\} \setminus X$ . Then any prime  $p$  with the property that all primes from  $X$  divide  $p - 1$  and no prime from  $Y$  divides  $p - 1$  satisfies

$$t > \prod_{q \mid p-1} \left(1 - \frac{1}{q^2}\right) > \prod_{j \in S' \setminus Y, j \geq k} \left(1 - \frac{1}{p_j^2}\right) > s.$$

The set of primes  $p$  with this property has positive lower density by Lemma 5.6.

The case when  $(s, t) \subseteq (9/(8\zeta(2)), 3/4)$  can be handled in a similar way by taking  $S' = S \cup \{1\}$ . The interval  $(3/(2\zeta(2)), 1)$  transforms under multiplication by  $3/4$  to  $(9/(8\zeta(2)), 3/4)$ .  $\square$

**Corollary 5.10.** *Let  $(s, t)$  be a subinterval of  $\mathcal{I}$ . Then there is a set of primes  $p$  of positive lower density for which the following holds. Let  $\alpha', \beta', \gamma'$  be uniformly random in  $\mathbf{Z}_n$ , and let  $\nu$  be the uniform distribution on  $\mathbf{Z}_n$ , where  $n = p - 1$ . Then  $\mathbf{P}_{\alpha', \beta', \gamma'}(\phi_k = \nu) \in (s, t)$  for all  $k \geq 0$ .*

We now transport this result to the original setting of the random Lie bracket.

**Corollary 5.11.** *Let  $(s, t)$  be a subinterval of  $\mathcal{I}$ . Then there is a set of primes  $p$  of positive lower density with the following property. Let  $A, B$  be uniformly random in  $\mathfrak{sl}_2(\mathbf{F}_p)$ , let  $\phi_k$  be the limiting distribution of the corresponding random walk on  $\mathbf{Z}_n$  after  $k$  steps,<sup>8</sup> let  $\nu_{2k}$  be the corresponding distribution on  $\text{Lin}\{[A, B]\} \setminus \{0\}$ , and let  $\nu$  be the uniform distribution on  $\text{Lin}\{[A, B]\} \setminus \{0\}$ . Then  $\mathbf{P}_{A, B}(\nu_{2k} = \nu) \in (s, t)$  for all  $k \geq 0$ .*

*Proof.* Take any  $A, B \in \mathfrak{sl}_2(\mathbf{F}_p)$  such that  $G_{A, B}$  has nonzero entries. Then the random Lie bracket satisfies  $\nu_{2k} = \nu$  if and only if the parameters  $(a, b) = F(A, B)$  satisfy  $\gcd(n, a, b) = 1$ . By Corollary 3.6 we have  $|\mathbf{P}_{A, B}(\nu_{2k} = \nu) - \mathbf{P}_{a, b}(\gcd(n, a, b) = 1)| < 15/p$  and by the previous corollary, we have

$$\mathbf{P}_{a, b}(\gcd(n, a, b) = 1) = \nu_{\mathbf{Z}_n^3}(\{(\alpha', \beta', \gamma') \in \mathbf{Z}_n^3 \mid \phi_k = \nu_{\mathbf{Z}_n}\}) \in (s + 15/p, t - 15/p)9$$

for a set of primes  $p$  of positive lower density. Combining these gives us that  $\mathbf{P}_{A, B}(\nu_{2k} = \nu) \in (s, t)$  for a set of primes  $p$  of positive lower density.  $\square$

In particular, for any  $x \in \mathcal{I}$  and any  $\epsilon > 0$ , there is a set of primes of positive lower density such that the random Lie bracket converges to the uniform distribution on a line after an even number of steps with probability in  $(x - \epsilon, x + \epsilon)$ .

## 6 | Pre-Cutoff

Let  $A, B \in \mathfrak{sl}_2(\mathbf{F}_p)$  be uniformly random. Let  $\mu_{2k}$  be the distribution of the random Lie bracket after  $2k$  steps, and let  $\nu_{2k}$  be the associated limiting distribution, supported on  $\text{Lin}\{[A, B]\}$ . We shall now prove that with high probability, the distance  $d(\mu_{2k}, \nu_{2k})$  quickly transitions from 1 to 0 at around  $k \approx p$ . This is the phenomenon of pre-cutoff, and we establish it by providing sharp upper and lower bounds on  $d(\mu_{2k}, \nu_{2k})$  in terms of  $k$  and the parameters  $a, b$ . We further inspect what happens for generic  $a, b$ .

The distribution of the random Lie bracket after  $2k$  steps is the same as the distribution of the random walk on  $\mathbf{F}_p^*$  with parameters  $\alpha, \beta, \gamma$  after  $k$  steps. Supposing these parameters are all nonzero, the random walk is equivalent to a random walk on  $\mathbf{Z}_n$  with parameters  $\alpha', \beta', \gamma'$ . Let  $\sigma_k$  be the distribution of this random walk after  $k$  steps. The Markov operator has eigenvalues

$$\lambda_j = \frac{1}{2} \chi_j(-\alpha') + \frac{1}{4} \chi_j(-\beta') + \frac{1}{4} \chi_j(-\gamma')$$

of absolute value

$$|\lambda_j| = \left| \frac{1}{2} + \frac{1}{4} \chi_j(a) + \frac{1}{4} \chi_j(b) \right|,$$

where  $a = \alpha' - \beta'$  and  $b = \alpha' - \gamma'$ .

## 6.1 | Upper Bound

For  $x \in \mathbf{R}$ , let  $\langle x \rangle$  be the unique value in the interval  $(-1/2, 1/2]$  such that  $x - \langle x \rangle$  is an integer. Let us associate to each eigenvalue  $\lambda_j$  the vector

$$u_j = \left( \left\langle \frac{ja}{n} \right\rangle, \left\langle \frac{jb}{n} \right\rangle \right) \in \mathbf{R}^2.$$

The norm  $\|u_j\|$  controls  $|\lambda_j|$  in the following way.

**Lemma 6.1.**  $|\lambda_j| \leq \exp\left(-\frac{1}{2}\|u_j\|^2\right)$ .

*Proof.* The triangle inequality gives

$$|\lambda_j| \leq \frac{1}{4}(|1 + \chi_j(a)| + |1 + \chi_j(b)|) = \frac{1}{2}|\cos(\pi ja/n)| + \frac{1}{2}|\cos(\pi jb/n)|.$$

Note that  $|\cos(\pi x)| = |\cos(\pi \langle x \rangle)|$ . Using the bound  $\cos(\pi x) \leq \exp(-2x^2)$ , which holds for all  $x \in [-3/2, 3/2]$ , we obtain

$$|\lambda_j| \leq \frac{1}{2}(\exp(-2\langle ja/n \rangle^2) + \exp(-2\langle jb/n \rangle^2)).$$

By Jensen's inequality for the function  $x \mapsto \exp(-2x^2)$ , which is concave on the interval  $[-1/2, 1/2]$ , the latter is at most

$$\exp\left(-\frac{1}{2}(|\langle ja/n \rangle| + |\langle jb/n \rangle|)^2\right) \leq \exp\left(-\frac{1}{2}\|u_j\|^2\right)$$

and the proof is complete.  $\square$

The characters contributing to the limiting distribution  $\phi_k$  are those  $\chi_j$  for which  $|\lambda_j| = 1$ . Note that this happens if and only if  $a, b \in \ker \chi_j$ , which is the same as  $u_j = 0$ . In order to control the residual distribution  $\sigma_k - \phi_k = \omega_k$ , let  $\Lambda$  be the plane lattice

$$\Lambda = \bigcup_{j \in \mathbf{Z}_n} (u_j + \mathbf{Z}^2) = \left(\frac{a}{n}, \frac{b}{n}\right) \mathbf{Z} + \mathbf{Z}^2 \subseteq \mathbf{R}^2$$

and let  $\Delta$  be the minimal distance between two distinct points in  $\Lambda$ . When  $a, b \neq 0$ , the lattice  $\Lambda$  is nontrivial, and we have

$$\Delta = \min \{ \|u_j\| \mid j \in \mathbf{Z}_n, u_j \neq 0 \}.$$

**Proposition 6.2.** For all  $n$  and all  $a, b \in \mathbf{Z}_n$ , we have

$$d(\sigma_k, \phi_k)^2 \leq 11 \gcd(a, b, n) e^{-k\Delta^2}.$$

*Proof.* Use Lemma 4.1 with the last lemma to bound<sup>10</sup>

$$\|\sigma_k - \phi_k\|_1^2 \leq \sum_{|\lambda_j| < 1} |\lambda_j|^{2k} \leq \sum_{u_j \neq 0} \exp(-k\|u_j\|^2). \quad (4)$$

Let us upper bound this sum in terms of the lattice  $\Lambda$ . Note that the union in the definition of  $\Lambda$  might not be disjoint, since we might have  $u_i = u_j$  for some  $i, j$ . This happens precisely when  $n$  divides  $a(j - i)$  and  $b(j - i)$ , which is equivalent

to  $n$  dividing  $\gcd(a, b)(j - i)$ , and this is the same as saying that  $n/\gcd(a, b, n)$  divides  $j - i$ . Hence, every value of  $u_j$  occurs precisely  $\gcd(a, b, n)$  times. We can thus bound the exponential sum in (4) by

$$\frac{\gcd(a, b, n)}{\sum_{u \in \Lambda \setminus \{0\}} \exp(-k\|u\|_2^2)}.$$

The sum of norms over the whole lattice  $\Lambda$  can be upper bounded as in (Abrams et al. 2022, Lemma 6) by

$$9 \sum_{i \geq 1} (i+1)^2 e^{-k\Delta^2 i^2} \leq 18 \sum_{i \geq 1} (i+1) e^{-k\Delta^2 i} = 18 e^{-k\Delta^2} \frac{2 - e^{-k\Delta^2}}{(1 - e^{-k\Delta^2})^2}.$$

For  $e^{-k\Delta^2} \leq 1/10$ , the value of the fraction is less than  $12/5$ , so we get the overall bound

$$\|\sigma_k - \phi_k\|_1^2 \leq 44 \gcd(a, b, n) e^{-k\Delta^2}.$$

For  $e^{-k\Delta^2} > 1/10$ , the same bound holds since we always have  $\|\sigma_k - \phi_k\|_1^2 \leq 4$ .  $\square$

The minimal norm  $\Delta$  can be estimated as follows.

**Lemma 6.3.** *We have*

$$\frac{\gcd(a, b, n)}{n} \leq \Delta \leq \frac{2}{\sqrt{\pi}} \sqrt{\frac{\gcd(a, b, n)}{n}}.$$

*Proof.* The lattice  $\Lambda$  is contained in  $(\gcd(a, b, n)/n)\mathbf{Z}^2$ , hence  $\Delta \geq \gcd(a, b, n)/n$ . For the upper bound, project  $\mathbf{R}^2$  to the torus  $\mathbf{R}^2/\mathbf{Z}^2$ . Observe that the image of the lattice  $\Lambda$  contains exactly  $n/\gcd(a, b, n)$  points. Since the minimal distance between two distinct points of  $\Lambda$  is  $\Delta$ , the open discs of radius  $\Delta/2$  in the torus  $\mathbf{R}^2/\mathbf{Z}^2$  around these points are disjoint. Comparing areas, we thus obtain

$$\frac{n}{\gcd(a, b, n)} \pi \left( \frac{\Delta}{2} \right)^2 \leq 1.$$

Rearranging terms gives the claimed upper bound.  $\square$

The bounds in the lemma are sharp with respect to  $n$ . If  $a = 1$  and  $b = 0$ , we have  $\gcd(a, b, n) = 1$  and  $\Delta = 1/n$ , matching the lower bound. For the upper bound, consider  $n = m^2$ ,  $a = m$ , and  $b = 1$  for some  $m \in \mathbf{N}$ . If  $\|u_j\| < 1/m$  for some  $j \in \mathbf{Z}$ , then, since the second coordinate of  $u_j$  is less than  $1/m$  in absolute value, we must have  $|j| < m$ . However, this implies  $j = 0$ , as the same condition holds for the first coordinate. Thus  $u_j = 0$ . This demonstrates that  $\Delta \geq 1/m = 1/\sqrt{n}$ , and the upper bound is also sharp up to a constant factor.

## 6.2 | Randomly Chosen Parameters

Suppose the parameters  $a, b \in \mathbf{Z}_n$  of the random walk are chosen uniformly at random. We shall now show that in this case,  $\gcd(a, b, n)$  is not large and  $\Delta$  is of order  $1/\sqrt{n}$  with high probability.

**Lemma 6.4.** *Let  $a, b$  be uniformly random in  $\mathbf{Z}_n$ . Then for any  $M > 1$ , we have*

$$\mathbf{P}_{a,b}(\gcd(a, b, n) \geq M) < 1/(M-1).$$

*Proof.* If  $\gcd(a, b, n) \geq M$ , then there is a  $d \geq M$  that divides  $a, b, n$ . By the union bound, we thus get

$$\mathbf{P}_{a,b}(\gcd(a, b, n) \geq M) \leq \sum_{\substack{d|n \\ d \geq M}} \mathbf{P}_{a,b}(a, b \in d\mathbf{Z}_n) < \sum_{d \geq M} \frac{1}{d^2} < \int_{M-1}^{\infty} \frac{1}{x^2} dx = \frac{1}{M-1}.$$

**Lemma 6.5.** Let  $a, b \in \mathbf{Z}_n$  be uniformly random. For any  $0 \neq j \in \mathbf{Z}_n$  and  $r > 0$ , we have

$$\mathbf{P}_{a,b}(0 < \|u_j\| < r) < 8r^2.$$

*Proof.* Let  $d = \gcd(n, j)$  and  $m = n/d$ . Let  $L$  be the additive subgroup of  $\mathbf{R}^2$  generated by  $(d/n, 0)$  and  $(0, d/n)$ , and consider  $\langle L \rangle = L \cap (-1/2, 1/2]^2$ . For uniformly random  $a, b$ , the point  $u_j$  is uniformly distributed on  $\langle L \rangle$ . Note that  $\langle L \rangle$  consists of  $m^2$  points and at most  $(2\lfloor rm \rfloor + 1)^2$  of them are of norm at most  $r$ . The probability that  $0 < \|u_j\| < r$  is thus bounded by  $((2\lfloor rm \rfloor + 1)^2 - 1)/m^2$ . This is 0 when  $m < 1/r$ , and for  $m \geq 1/r$  it is at most  $4r^2 + 4r/m \leq 8r^2$ .  $\square$

**Lemma 6.6.** Let  $a, b \in \mathbf{Z}_n$  be uniformly random. For any  $\epsilon > 0$ , we have

$$\mathbf{P}_{a,b}(\Delta < \epsilon/\sqrt{n}) < 8\epsilon^2.$$

*Proof.* The event  $\Delta < \epsilon/\sqrt{n}$  implies that we must have  $0 < \|u_j\| < \epsilon/\sqrt{n}$  for some  $0 \neq j \in \mathbf{Z}_n$ . The claim now follows by the union bound and the previous lemma.  $\square$

**Theorem 6.7.** For every  $\epsilon, \delta > 0$  there is a constant  $C$  such that the following holds. Let  $a, b \in \mathbf{Z}_n$  be uniformly random. Then for all  $k > Cn$ , we have

$$\mathbf{P}_{a,b}(d(\sigma_k, \phi_k) < \epsilon) > 1 - \delta.$$

*Proof.* With probability larger than  $1 - \delta$  we have, by Lemma 6.4 and Lemma 6.6, that both  $\gcd(a, b, n) < 1 + 2/\delta < 3/\delta$  and  $\Delta^2 \geq \delta/16n$  hold. Using these with Proposition 6.2 gives

$$d(\sigma_k, \phi_k)^2 \leq \frac{33}{\delta} \exp(-\delta k/16n).$$

For  $k > Cn$ , this is less than  $\epsilon^2$ , completing the proof.  $\square$

**Corollary 6.8.** For every  $\epsilon, \delta > 0$  there is a constant  $C$  such that the following holds. Let  $A, B \in \mathfrak{sl}_2(\mathbf{F}_p)$  be uniformly random. Then for all  $k > Cp$ , we have

$$\mathbf{P}_{A,B}(d(\mu_{2k}, \nu_{2k}) < \epsilon) > 1 - \delta.$$

*Proof.* Take  $A, B \in \mathfrak{sl}_2(\mathbf{F}_p)$  such that  $G_{A,B}$  has nonzero entries. Then the random Lie bracket satisfies  $d(\mu_{2k}, \nu_{2k}) < \epsilon$  if and only if the parameters  $(a, b) = F(A, B)$  satisfy  $d(\sigma_k, \phi_k) < \epsilon$ . By the previous theorem, we have  $\mathbf{P}_{a,b}(d(\sigma_k, \phi_k) < \epsilon) > 1 - \delta/2$  for  $k > Cp$ . It now follows from Corollary 3.6 that  $\mathbf{P}_{A,B}(d(\mu_{2k}, \nu_{2k}) < \epsilon) > 1 - \delta$ .  $\square$

### 6.3 | Lower Bound in $L^1$ -Distance

We now prove the corresponding lower bound on the distance between  $\sigma_k$  and  $\nu_k$ . Using Fourier analysis, we can easily obtain a lower bound  $1 - \epsilon$  on the  $L^1$ -distance (but not on the total variation distance, this will be done in the following section) in terms of the vectors  $u_j$  and minimal norm  $\Delta$  as in the previous section. We tightly follow the argument in Abrams et al. (2022).

**Lemma 6.9** (Lemma 4 in Abrams et al. (2022)). If  $\|u_j\| \leq 1/2\pi$ , then

$$|\lambda_j| \geq \exp(-\pi^2 \|u_j\|^2).$$

*Proof.* We have

$$|\lambda_j| \geq \operatorname{Re}\left(\frac{1}{2} + \frac{1}{4}\chi_j(a) + \frac{1}{4}\chi_j(b)\right) = \frac{1}{2} + \frac{1}{4}\cos(2\pi\langle ja/n \rangle) + \frac{1}{4}\cos(2\pi\langle jb/n \rangle).$$

This can be lower bounded using  $\cos x \geq \exp(-x^2)$ , which holds for all  $|x| \leq 1$ , by

$$\frac{1}{2} + \frac{1}{4} \exp(-4\pi^2 \langle ja/n \rangle^2) + \frac{1}{4} \exp(-4\pi^2 \langle jb/n \rangle^2),$$

and applying Jensen's inequality for the convex function  $x \mapsto \exp(-x)$  gives

$$|\lambda_j| \geq \exp(-\pi^2 \langle ja/n \rangle^2 - \pi^2 \langle jb/n \rangle^2) = \exp(-\pi^2 \|u_j\|^2).$$

□

**Theorem 6.10.** *For every  $\epsilon, \delta > 0$  there are constants  $C, c > 0$  such that the following holds. Let  $n > C$  and let  $a, b \in \mathbf{Z}_n$  be uniformly random. For all  $k < cn$ , we have*

$$\mathbf{P}_{a,b}(\mathbf{d}(\sigma_k, \phi_k) > 1/2 - \epsilon) > 1 - \delta.$$

*Proof.* With probability at least  $1 - \delta$ , we have  $\gcd(a, b, n) \leq 1 + 1/\delta$  by Lemma 6.4. Let  $C$  be such that

$$2\sqrt{\frac{1 + 1/\delta}{\pi C}} \leq \frac{1}{2\pi}.$$

As long as  $n \geq C$ , we thus have  $\Delta \leq 1/2\pi$  by Lemma 6.3. Let  $u_j$  be the vector with  $\Delta = \|u_j\|$ . We have  $|\lambda_j| \neq 1$  since  $\|u_j\| \neq 0$ , and so Lemma 6.9 gives  $\rho \geq |\lambda_j| \geq \exp(-\pi^2 \Delta^2)$ . Hence

$$\mathbf{d}(\sigma_k, \phi_k) \geq \rho^k/2 \geq \exp(-4\pi(1 + 1/\delta)k/n)/2$$

by Lemma 4.1. For  $k < cn$ , this is more than  $1/2 - \epsilon$ , completing the proof. □

## 6.4 | Lower Bound in Total Variation Distance

Let us now prove the corresponding stronger lower bound of  $1 - \epsilon$  in *total variation distance*. We do this by refining the argument in (Hildebrand 2005, Theorem 2) (itself a modification of (Greenhalgh 1989, Theorem 6.1.1)), which proves a bound of  $1/2 - \epsilon$  in total variation distance (so as strong as the one in the previous section).

**Theorem 6.11.** *For every  $\epsilon, \delta > 0$  there are constants  $C, c > 0$  such that the following holds. Let  $n > C$  and let  $a, b \in \mathbf{Z}_n$  be uniformly random. For all  $k < cn$ , we have*

$$\mathbf{P}_{a,b}(\mathbf{d}(\sigma_k, \phi_k) > 1 - \epsilon) > 1 - \delta.$$

*Proof.* Let  $\lambda > 0$  be a parameter to be determined later. Let  $C, c > 0$  be such that

$$(1 + 1/\delta)(2\lambda\sqrt{c} + 1/\sqrt{C})^2 < \epsilon.$$

Assume  $n > C$  and  $k < cn$ . We will exhibit a set  $S \subseteq \mathbf{Z}_n$  that satisfies  $\phi_k(S) < \epsilon$  and  $\sigma_k(S) > 1 - 2\epsilon$ , which in turn gives  $\mathbf{d}(\sigma_k, \phi_k) > 1 - 3\epsilon$ . Let  $\phi : \mathbf{Z}^3 \rightarrow \mathbf{Z}_n$  be the homomorphism defined by  $\phi(1, 0, 0) = \alpha', \phi(0, 1, 0) = \beta', \phi(0, 0, 1) = \gamma'$ , and let

$$T = \left\{ (k - x - y, x, y) \in \mathbf{Z}^3 \mid |x - k/4|, |y - k/4| \leq \lambda\sqrt{k} \right\} \quad \text{and} \quad S = \phi(T).$$

The distribution  $\phi_k$  is uniform on a subset of size  $n/\gcd(a, b, n)$  by Lemma 5.4. With probability at least  $1 - \delta$ , we have  $\gcd(a, b, n) \leq 1 + 1/\delta$  by Lemma 6.4, and hence

$$\phi_k(S) \leq |S| \gcd(a, b, n)/n \leq (1 + 1/\delta)(2\lambda\sqrt{k} + 1)^2/n < \epsilon$$

by the assumption on  $n, k$ . On the other hand, let  $X, Y$  be the number of times the generators  $\beta', \gamma'$  are chosen in the random walk. Note that  $X, Y$  are distributed as a sum of  $k$  independent Bernoulli random variables with success probability  $1/4$ . Thus

$$\sigma_k(S) \geq \mathbf{P}(|X - k/4| \leq \lambda\sqrt{k}, |Y - k/4| \leq \lambda\sqrt{k}) \geq 1 - 2\mathbf{P}(|X - k/4| > \lambda\sqrt{k}).$$

By Hoeffding's inequality, the last probability is at most  $2 \exp(-2\lambda^2)$ . Let  $\lambda$  be such that this is less than  $\epsilon$ . Then  $\sigma_k(S) > 1 - 2\epsilon$ , as required.  $\square$

**Corollary 6.12.** *For every  $\epsilon, \delta > 0$  there are constants  $C, c > 0$  such that the following holds. Let  $p > C$  and let  $A, B \in \mathfrak{sl}_2(\mathbf{F}_p)$  be uniformly random. For all  $k < cp$ , we have*

$$\mathbf{P}_{A,B}(\mathbf{d}(\mu_{2k}, \nu_{2k}) > 1 - \epsilon) > 1 - \delta.$$

*Proof.* For matrices  $A, B \in \mathfrak{sl}_2(\mathbf{F}_p)$  such that  $G_{A,B}$  has nonzero entries, the random Lie bracket satisfies  $\mathbf{d}(\mu_{2k}, \nu_{2k}) > 1 - \epsilon$  if and only if the parameters  $(a, b) = F(A, B)$  satisfy  $\mathbf{d}(\sigma_k, \phi_k) > 1 - \epsilon$ . By the previous theorem, we have  $\mathbf{P}_{a,b}(\mathbf{d}(\sigma_k, \phi_k) > 1 - \epsilon) > 1 - \delta/2$  for  $k < cp$ . We can assume  $C$  is large enough so that  $15/C < \delta/2$ , so Corollary 3.6 implies that  $\mathbf{P}_{A,B}(\mathbf{d}(\mu_{2k}, \nu_{2k}) < 1 - \epsilon) > 1 - \delta$ .  $\square$

Corollaries 6.8 and 6.12 together prove Theorem 1.1.

## 7 | Diameter

Here, we prove the diameter bound of  $O(\log p)$  for the random Lie bracket in  $\mathfrak{sl}_2(\mathbf{F}_p)$ , as stated in Theorem 1.3. The key point behind the argument is based on the analysis of the affine random walk on  $\mathbf{F}_p$  evolving according to  $x_{n+1} = ax_n + b_n$  for fixed  $a \in \mathbf{F}_p^*$  and i.i.d. random  $b \in \mathbf{F}_p$ . This walk has been studied many times in the literature, the starting point being Chung et al. (1987). The most relevant for us are the strong results of Breuillard and Varjú (2022) which exhibit rapid mixing of the walk for almost all values of  $a$ . In order to explain this more precisely, let, for any  $a \in \mathbf{F}_p$  and  $n \in \mathbf{N}$ ,

$$P_n(a) = \left\{ \sum_{i=0}^n b_i a^i \mid b_i \in \{-1, 0, 1\} \right\} \subseteq \mathbf{F}_p$$

be the set of all polynomials in  $a$  of degree at most  $n$  with coefficients in  $\{-1, 0, 1\}$ .

**Theorem 7.1** (Proposition 13 together with Theorem 2 in Breuillard and Varjú (2022)). *There is a (possibly empty) small set of primes  $T$  with the following property. For any  $\epsilon > 0$  there is a  $C$  such that for all primes  $p > C$  not in  $T$  and at least  $(1 - \epsilon)p$  values of  $a \in \mathbf{F}_p$ , we have  $P_n(a) = \mathbf{F}_p$  for some  $n \leq 10 \log p$ .*

**Remark 7.2.** Alternatively, we can use Theorem 1 from Breuillard and Varjú (2022) instead of Theorem 2. This yields a weaker bound of  $n \leq C_\epsilon \log p \log \log p$ , but it holds for all primes  $p$ .

We now show how to do deduce the stated diameter bound from this. First of all, using basic properties of the random Lie bracket, we show that obtaining  $P_n(a)$ -multiples of  $[A, B] \in \mathfrak{sl}_2(\mathbf{F}_p)$  grows linearly with  $n$ .

**Lemma 7.3.** *Let  $A, B \in \mathfrak{sl}_2(\mathbf{F}_p)$  and  $S = \{0, \pm A, \pm B\}$ . Let  $\alpha = 2\langle A, B \rangle$ . Then for any  $n \in \mathbf{N}$ , we have*

$$P_n(\alpha) \cdot [A, B] \subseteq S^{4n+2}.$$

*Proof.* We prove this by induction on  $n$ . The statement is trivial for  $n = 0$ . Suppose it holds for  $n$  and take any  $x \in P_{n+1}(\alpha)$ . Then  $x = \alpha y + z$  for some  $y \in P_n(\alpha)$  and  $z \in \{-1, 0, 1\}$ . By the induction hypothesis, we have  $y[A, B] \in S^{4n+2}$ , and so  $\alpha y[A, B] = [B, [A, y[A, B]]] \in S^{4n+4}$  by Lemma 2.1. Then

$$x[A, B] = \alpha y[A, B] + z[A, B] \in S^{4n+6} = S^{4(n+1)+2}.$$

Secondly, we show that, with high probability, the value  $\alpha = 2\langle A, B \rangle$  belongs to any large enough set.

**Lemma 7.4.** *Let  $\mathcal{A} \subseteq \mathbf{F}_p$  be a subset of  $\mathbf{F}_p$  with  $|\mathcal{A}| \geq (1 - \epsilon)p$  and let  $A, B \in \mathbf{F}_p$  be uniformly random. Then*

$$\mathbf{P}_{A,B}(\alpha \notin \mathcal{A}) \leq 3\epsilon.$$

*Proof.* Let  $\mathcal{X} \subseteq \text{Sym}_2(\mathbf{F}_p)$  be the set of symmetric matrices  $X$  with  $2X_{12} \notin \mathcal{A}$ . By assumption, we have  $|\mathcal{X}| \leq \epsilon p^3$ . By Proposition 3.2, we then get the bound  $|G^{-1}(\mathcal{X})| \leq 3\epsilon p^6$ . The result follows.  $\square$

We are now ready to deduce Theorem 1.3.

*Proof of Theorem 1.3.* Let  $T$  be the exceptional set of primes from Theorem 7.1. Take any  $\epsilon > 0$ . Then there is a  $C$  such that for all primes  $p > C$  not in  $T$  and at least  $(1 - \epsilon)p$  values of  $a \in \mathbf{F}_p$ , we have  $P_n(a) = \mathbf{F}_p$  for  $n = \lfloor 10 \log p \rfloor$ . Hence, by Lemma 7.4, for  $A, B \in \mathfrak{sl}_2(\mathbf{F}_p)$  uniformly random we have  $P_n(a) = \mathbf{F}_p$  with probability at least  $1 - 3\epsilon$ . This implies, by Lemma 7.3, that  $\text{Lin}\{[A, B]\} \subseteq S^{4n+2}$ .

We can assume that  $C$  (and therefore  $p$ ) is large enough that with probability at least  $1 - \epsilon$ , the matrices  $A, B$  and  $[A, B]$  span  $\mathfrak{sl}_2(\mathbf{F}_p)$  as a vector space.<sup>12</sup> In that case, we can write any element of  $\mathfrak{sl}_2(\mathbf{F}_p)$  as  $xA + yB + z[A, B]$  for some  $x, y, z \in \mathbf{F}_p$ . The Gram matrix  $G_{A,B}$  is invertible with probability at least  $1 - 4/p > 1 - \epsilon$ . Thus there are  $r, s \in \mathbf{F}_p$  such that  $(r, s) \cdot G_{A,B} = (y/4, -x/4)$ , meaning that  $-2\alpha r - 2\gamma s = x$  and  $2\beta r + 2\alpha s = y$ . Then, by Lemma 2.1,

$$xA + yB + z[A, B] = [A, r[A, B]] + [B, s[A, B]] + z[A, B] \in S^{12n+8}.$$

Hence  $S^{12n+8} = \mathfrak{sl}_2(\mathbf{F}_p)$ , and so with probability at least  $1 - 5\epsilon$ , we have the desired bound on the diameter.  $\square$

## Funding

Both authors are supported by the Slovenian Research and Innovation Agency research program P1-0222. UJ is also supported by the Slovenian Research and Innovation Agency grants J1-50001, J1-4351, J1-3004, and N1-0217.

## Acknowledgments

We thank Daniele Dona for suggesting the use of the results in Breuillard and Varjú (2022) to establish diameter bounds. We also thank Martin Hildebrand for drawing our attention to the total variation distance bounds for random walks on cycles in Hildebrand (2005). Finally, we thank the anonymous referee for their many helpful comments and suggestions.

## Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

## Endnotes

<sup>1</sup> The total variation distance between probability measures  $\mu$  and  $\nu$  is  $d(\mu, \nu) = \frac{1}{2} \|\mu - \nu\|_1$ .

<sup>2</sup> The probability  $\mathbf{P}_{A,B}(v_{2k} = v)$  is independent of  $k$ , as  $v_{2k}$  is supported on cosets of the same subgroup of  $\text{Lin}\{[A, B]\} \setminus \{0\}$ . Let  $v(p)$  denote this probability. The theorem and its proof imply that the set of limit points of the sequence  $v(p)$  as  $p$  varies is exactly  $\mathcal{I}$ .

<sup>3</sup> A small set is a set of positive integers whose sum of reciprocals converges.

<sup>4</sup> The term *cutoff* describes the stronger property that  $t_{\text{mix}}(\epsilon)/t_{\text{mix}}(1 - \epsilon) \rightarrow 1$  for any  $0 < \epsilon < 1/2$ , where  $t_{\text{mix}}(x) = \min\{k \in \mathbf{N} \mid d(\mu_{2k}, \nu_{2k}) < x\}$ . This phenomenon cannot occur on a cyclic group with a bounded number of generators by Abrams et al. (2022).

<sup>5</sup> Throughout the paper we assume  $p > 2$ .

<sup>6</sup> Note that  $\phi_k$  may not converge as  $k \rightarrow \infty$ , nor is it guaranteed to be uniform.

<sup>7</sup> The lower density of a subset  $Z$  of  $\mathbf{P}$  is  $\liminf_{n \rightarrow \infty} |Z \cap \mathbf{P}_{\leq n}|/|\mathbf{P}_{\leq n}|$ .

<sup>8</sup> If  $\alpha\beta\gamma = 0$ , there is no corresponding walk on  $\mathbf{Z}_n$ . In that case, just take  $\phi_k = 0$ .

<sup>10</sup> When  $a, b = 0$ , both sums are empty, and we have  $d(\sigma_k, \phi_k) = 0$ .

<sup>11</sup> We can assume  $p$  is large enough so that  $15/p < \delta/2$ .

<sup>12</sup> Write  $A = (a_{ij})$ ,  $B = (b_{ij})$ . Then  $A, B, [A, B]$  span  $\mathfrak{sl}_2(\mathbf{F}_p)$  if and only if the  $3 \times 4$  matrix of vectorized  $A, B, [A, B]$  is of full rank, and so some minor is nonzero. By the Schwartz-Zippel lemma, this occurs with probability at least  $1 - 4/p$ .

## References

Abrams, A., E. Babson, H. Landau, Z. Landau, and J. Pommersheim. 2022. “No Cutoff for Circulants: An Elementary Proof.” *Probability and Mathematical Statistics* 42: 97–108.

Axler, C. 2019. “New Estimates for the  $n$ -th Prime Number.” *Journal of Integer Sequences* 22, no. 2: 3.

Bourgain, J., and A. Gamburd. 2008. "Expansion and Random Walks in  $SL_d(\mathbb{Z})$ : I." *Journal of the European Mathematical Society* 10, no. 4: 987–1011.

Breuillard, E., and P. P. Varjú. 2022. "Cut-Off Phenomenon for the  $AX+B$  Markov Chain Over a Finite Field." *Probability Theory and Related Fields* 184, no. 1: 85–113.

Cantor, O., U. Jezernik, and A. Zozaya. 2025. "Two-Generation of Traceless Matrices Over Finite Fields." *Linear Algebra and Its Applications* 713: 1–17.

Chung, F. R. K., P. Diaconis, and R. L. Graham. 1987. "Random Walks Arising in Random Number Generation." *Annals of Probability* 15: 1148–1165.

Clark, P. L. 2013. *Quadratic Forms Chapter i: Witt's Theory*. University of Georgia.

Diaconis, P. 1988. "Group Representations in Probability and Statistics." *Lecture Notes-Monograph Series* 11: i-192.

Dona, D. 2023. "A Sum-Bracket Theorem for Simple Lie Algebras." *Journal of Algebra* 631: 658–694.

Greenhalgh, A. S. 1989. *Random Walks on Groups With Subgroup Invariance Properties*. Stanford University.

Helfgott, H. A. 2008. "Growth and Generation in  $SL_2(\mathbb{Z}/p\mathbb{Z})$ ." *Annals of Mathematics* 167: 601–623.

Hildebrand, M. 1994. "Random Walks Supported on Random Points of  $\mathbb{Z}/n\mathbb{Z}$ ." *Probability Theory and Related Fields* 100: 191–203.

Hildebrand, M. 2005. "A Survey of Results on Random Random Walks on Finite Groups." *Probability Surveys* 2: 33–63.

Serre, J.-P. 2012. *A Course in Arithmetic, Volume 7*. Springer Science & Business Media.