Université Grenoble Alpes

ANNALES DE L'INSTITUT FOURIER

Zhen-Qing Chen, Takashi Kumagai, Laurent Saloff-Coste, Jian Wang & Tianyi Zheng

Long range random walks and associated geometries on groups of polynomial growth

Tome 72, n° 3 (2022), p. 1249-1304.

https://doi.org/10.5802/aif.3515

Les *Annales de l'Institut Fourier* sont membres du Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org e-ISSN : 1777-5310

Ann. Inst. Fourier, Grenoble
72, 3 (2022) 1249-1304

LONG RANGE RANDOM WALKS AND ASSOCIATED GEOMETRIES ON GROUPS OF POLYNOMIAL GROWTH

by Zhen-Qing CHEN, Takashi KUMAGAI,
Laurent SALOFF-COSTE, Jian WANG & Tianyi ZHENG (*)

Abstract. — In the context of countable groups of polynomial volume growth, we consider a large class of random walks that are allowed to take long jumps along multiple subgroups according to power law distributions. For such a random walk, we study the large time behavior of its probability of return at time n in terms of the key parameters describing the driving measure and the structure of the underlying group. We obtain assorted estimates including near-diagonal two-sided estimates and the Hölder continuity of the solutions of the associated discrete parabolic difference equation. In each case, these estimates involve the construction of a geometry adapted to the walk.

Résumé. — Dans le contexte des groupes finiment engendrés à croissance polynomiale du volume, nous considérons une large classe de marches aléatoires à sauts de longue portée distribués suivant des lois puissances dans la direction de plusieurs sous-groupes. Pour de telles marches, nous déterminons la probabilité de retour au temps n en fonction de la distribution des sauts et de la structure algébrique du groupe. Nous obtenons des estimations autour de la diagonale ainsi que la continuité Hölderienne des solutions de l'équation de la chaleur discrète associée. Dans chaque cas, ces estimations utilisent la géométrie associée à la marche.

1. Introduction

1.1. Random walks and word-length

Given a probability measure μ on a discrete group G with identity element e, a random walk driven by μ with initial measure ν_0 is a G-valued

Keywords: long range random walk, group, return probability, Pseudo-Poincaré inequality, Hölder continuity.
2020 *Mathematics Subject Classification:* 60G50, 20F65, 60B15.
(*) The research of ZC is supported by Simons Foundation Grant 520542, a Victor Klee Faculty Fellowship at UW, and NNSFC grant 11731009, TK by the Grant-in-Aid for Scientific Research (A) 17H01093, Japan, LSC by NSF grants DMS-1404435 and DMS-1707589, and JW by NNSFC grant 11831014 and 12071076, and the Education and Research Support Program for Fujian Provincial Agencies.

stochastic process $(X_n)_0^\infty$ such that X_0 has law ν_0 and $X_{n+1} = X_n\xi_{n+1}$, where $(\xi_i)_1^\infty$ is a G-valued i.i.d sequence with ξ_i distributed according to μ. This discrete Markov process has transition kernel

$$p(x,y) = \mathbf{P}(X_{n+1} = y \mid X_n = x) = \mu(x^{-1}y)$$

and satisfies

$$\mathbf{P}(X_n = x) = \nu_0 * \mu^{(n)}(x),$$

where $u * v(x) = \sum u(y)v(y^{-1}x)$ and $\mu^{(n)}$ stands for the n-fold convolution of μ with itself. Understanding the behavior of the function of the discrete time parameter n,

$$n \longmapsto \mu^{(n)}(e),$$

which represents the return probability to the starting point after n steps, is one of the key questions in the study of random walks. When μ is symmetric (i.e., $\mu(g^{-1}) = \mu(g)$ for all $g \in G$), it is an easy exercise to check that

$$n \longmapsto \mu^{(2n)}(e) = \|\mu^{(2n)}\|_\infty = \max_{x\in G} \mu^{(2n)}(x)$$

is a non-increasing function of n. The aim of this article is to study, in the context of finitely generated groups of polynomial volume growth, a natural class of random walks that allow for long range jumps. General random walks on countable groups were first considered in Harry Kesten's 1958 Ph.D. dissertation published as [17]. For further background information, see [16, 21, 27].

The most natural and best studied random walks on a finitely generated group G are driven by finitely supported symmetric measures, and it is then natural to assume that the support of the measure generates the group G (otherwise, we can restrict attention to the subgroup generated by the support). In the study of these random walks, the word-length distance and associated geometry are very useful. Given a finite symmetric generating set S, the associated word-length of an element g in G, $|g| = |g|_{G,S}$, is the least number of generators needed to express g as a product over S in G (by convention, $|e|_S = 0$). The associated (left-invariant) distance between two elements $x, y \in G$ is

$$d(x,y) = d_{G,S}(x,y) = |x^{-1}y|.$$

The volume growth function of the pair (G,S) is the counting function

$$V(r) = \#\{g : |g| \leqslant r\}.$$

We will use the notation $f_1 \asymp f_2$ between two real valued functions defined on an abstract domain D (often omitted) to indicate that there are constants $c_1, c_2 \in (0,\infty)$ such that

$$\forall\, x \in D,\;\; c_1 f_1(x) \leqslant f_2(x) \leqslant c_2 f(x).$$

We will also use the notation $f_1 \simeq f_2$ between two positive real functions defined on an appropriate domain $D \subset \mathbb{R}$ (typically, $D = [1,\infty)$ or $D = (0,1]$ or also $D = \{0,1,2,\dots\}$) to indicate that there are constants c_i, $1 \leqslant i \leqslant 4$, such that

$$\forall\, x \in D,\;\; c_1 f_1(c_2 x) \leqslant f_2(x) \leqslant c_3 f_1(c_4 x)$$

(in each case, c_2x and c_4x should be understood appropriately. Specifically, when $D = [1,\infty)$, $D = (0,1]$ and $D = \{0,1,2,\dots\}$, c_2x and c_4x should be understood as $(c_2x) \vee 1$ and $(c_4x) \vee 1$, as $(c_2x) \wedge 1$ and $(c_4x) \wedge 1$, and as $\lfloor c_2x \rfloor$ and $\lceil c_4x \rceil$, respectively. Here for $a,b \in \mathbb{R}$, $a \vee b := \max\{a,b\}$, $a \wedge b := \min\{a,b\}$, and $\lfloor a \rfloor$ denotes the largest integer not exceeding a).

Typically, we assume that at least one of these functions is monotone (otherwise, this notion is not very practical). Similarly, we define the associated order relations $\preceq$ and $\succeq$ so that $f \preceq g$ means that $f(x) \leqslant c_1 g(c_2 x)$, and so on. For instance, when $|\cdot|_1$ and $|\cdot|_2$ are word-length functions associated to two finite symmetric generating sets S_1, S_2 of the same group G then, for all $x \in G$, $|x|_1 \asymp |x|_2$. If V_1, V_2 are the associated volume growth functions then $V_1 \simeq V_2$. In particular, up to the $\simeq$ equivalence relation, the volume growth function of a finitely generated group G does not depend on the choice of the finite generating set S, see, e.g., [14].

Definition 1.1. — *A finitely generated group has polynomial volume growth of degree d if, $V(r) \asymp r^d$ for $r \in [1,\infty)$.*

By a celebrated theorem of M. Gromov, it suffices that

$$\liminf_{r\to\infty} r^{-A} V(r) < \infty$$

with some constant A for the group G to have polynomial volume growth of degree d for some integer $d = d(G) \in \{0,1,\dots\}$. In this context, the tight relation between volume growth and random walk behavior is illustrated by the following result (See also [15, 26, 27]).

Theorem 1.2 (N. Varopoulos, [25]). — *Let G be a finitely generated group of polynomial volume growth of degree d and let μ be a finitely supported symmetric probability measure on G with generating support. Then, for all $n \in \{1,2,\dots\}$,*

$$\mu^{(2n)}(e) \asymp \frac{1}{V(\sqrt{n})} \asymp n^{-d/2}.$$

In fact, this result can be generalized in two significant directions by allowing μ to have finite second moment and by estimating $\mu^{(2n)}(g)$ for a range of g that depends on n.

Theorem 1.3. — *Let G be a finitely generated group of polynomial volume growth of degree d and let μ be a symmetric probability measure on G with generating support and with finite second moment, that is, $\sum_g |g|^2\mu(g) < \infty$. For simplicity, assume that $\mu(e) > 0$. Then, for any fixed $A > 0$, we have*

$$\forall\, g \in G,\ n \in \{1,2,\dots,\} \text{ with } |g| \leqslant A\sqrt{n},\ \ \mu^{(n)}(g) \asymp \frac{1}{V(\sqrt{n})} \asymp n^{-d/2}.$$

See, e.g., [15, 20] and the references therein. This type of estimate is often called a near diagonal estimate. In the result above, the range of order $\sqrt{n}$ is optimal. To close this short review and emphasize the importance of the word-length geometry in this context, let us mention briefly two more sophisticated results, namely, the parabolic Harnack inequality and Hölder continuity for solutions $(n,x) \mapsto u_n(x)$ of the parabolic difference equation

$$u_{n+1} - u_n = u_n * (\mu - \delta_e) \text{ or, equivalently, } u_{n+1} = u_n * \mu. \tag{1.1}$$

This discrete time evolution equation is parabolic because it resembles the classical heat equation with the operator $f \mapsto f * (\mu - \delta_e)$ playing the role of the Laplace operator (note that $f \mapsto f * (\mu - \delta_e)$ is non-positive definite on $L^2(G)$). The function

$$(n,x) \longmapsto \mu^{(n)}(x)$$

is a global solution of this equation. Note that for equation (1.1) to make sense and hold in a given subset A, it is necessary that u_n be defined, not only in A but over a set containing $A(\mathrm{support}(\mu))^{-1}$. In the next theorem, μ is symmetric and has finite support S. In such cases, whenever we say that u_n is solution of (1.1) in $[0,T] \times A$, we tacitly assume that $(k,x) \mapsto u_k(x)$ is defined for all $(k,x) \in [0,T] \times A(S \cup \{e\})$.

Theorem 1.4 ([10, special case]). — *Assume that G has polynomial volume growth and the measure μ is symmetric, finitely supported with generating support S containing the identity element, e. Then there are constants C and $\alpha > 0$ such that the following two properties hold.*

Parabolic Harnack Inequality: Any positive solution u of the difference equation (1.1) in the discrete time cylinder $Q = [0,N^2] \times \{x \in G : |x| \leqslant N\}$ satisfies

$$u_m(y) \leqslant C u_n(z)$$

for all $m \in [N^2/8, N^2/4]$, $n \in [N^2/2, N^2]$ *and* $y, z \in \{x \in G : |x| \leqslant N/2\}$.

Hölder Estimate: Any bounded solution u *of* (1.1) *in the discrete time cylinder* $Q = [0, N^2] \times \{x \in G : |x| \leqslant N\}$ *satisfies*

$$|u_n(z) - u_m(y)| \leqslant C \left[\left(|m-n|^{1/2} + |y^{-1}z| \right) /N \right]^{\alpha} \sup_{Q}\{|u|\} \tag{1.2}$$

for all $m, n \in [N^2/8, N^2/2]$ *and* $y, z \in \{x \in G : |x| \leqslant N/2\}$.

In these two statements, the constants C and α are independent of N and of the solution u (which can thus be translated both in time and in space if one so desires). In the context of parabolic differential equations, these estimates are the highlight of the celebrated De Giorgi–Nash–Moser theory. Informally, the first property (Parabolic Harnack Inequality) is the strongest as it (relatively easily) implies the second property (Hölder Estimate). The parabolic Harnack inequality also easily implies the near diagonal two-sided estimate of Theorem 1.3.

The goal of this work is to develop results such as Theorems 1.3 and 1.4 for walks on countable groups of polynomial volume growth when the measures driving the walks allow for a wide variety of long range jumps and have infinite second moments. For such random walks, it is known that a statement analogous to the above parabolic Harnack inequality cannot hold true. See, e.g., [1]. (Some integral version of the Harnack inequality, called a weak Harnack inequality, may hold in such cases; see [5].) However, we will be able to prove a version of the near diagonal two-sided estimate of Theorem 1.3 and a Hölder estimate (1.2) for globally bounded solutions of (1.1). In both cases, the word-length geometry must be replaced by a geometry adapted to the long jump probability measure driving the random walk. See Theorems 4.3–5.5 and 6.8.

1.2. Random walks with long range jumps

In a finitely generated group of polynomial volume growth of degree d, all subgroups are finitely generated and have polynomial volume growth of degree at most d. This work focuses on a natural family of symmetric probability measures defined as follows. For a book treatment of the notion of regular variation, see [4].

Definition 1.5. — *Let* G *be a finitely generated group of polynomial volume growth. Say a probability measure* μ *is in* $\mathcal{P}(G, \mathrm{reg})$, *if there is an* *integer $k \geqslant 0$ such that μ can be written in the form*

$$\mu = \sum_{i=0}^{k} p_i \mu_i, \quad \sum_{i=0}^{k} p_i = 1, \quad p_0 \geqslant 0, \quad p_i > 0, \ i = 1, \dots, k,$$

where each μ_i, $0 \leqslant i \leqslant k$, is a symmetric probability measure on G such that:

- *The probability measure μ_0 is finitely supported.*
- *For each $1 \leqslant i \leqslant k$, there exists a subgroup H_i of G, equipped with a word-length $|\cdot|_i$ and of polynomial volume growth of degree d_i, and a function, $\phi_i : [0, \infty) \to (0, \infty)$, positive, increasing and of regular variation of positive index at infinity such that*

$$\mu_i(h) \asymp \begin{cases} \left[\phi_i(1 + |h|_i)(1 + |h|_i)^{d_i}\right]^{-1} & \text{if } h \in H_i, \\ 0 & \text{otherwise.} \end{cases} \tag{1.3}$$

- *There is an $\varepsilon > 0$ such that the finite set $\{g : \mu(g) > \varepsilon\}$ generates G and contains the identity element e.*

Remark 1.6. — When considering a measure μ in $\mathcal{P}(G, \mathrm{reg})$, we will always assume that μ is given in the form $\mu = \sum_{i=0}^{k} p_i \mu_i$ where the measures μ_i, $1 \leqslant i \leqslant k$, are described as in (1.3). Hence, for any such μ, we are given the subgroups H_i and increasing regularly varying functions ϕ_i, $1 \leqslant i \leqslant k$, that are implicit in the fact that μ is in $\mathcal{P}(G, \mathrm{reg})$. By convention, we set $H_0 = G$ so that we have a well defined subgroup H_i for each $i \in \{0, \dots, k\}$.

Remark 1.7. — A measure μ in $\mathcal{P}(G, \mathrm{reg})$ can be finitely supported if $k = 0$ or if $k \geqslant 1$ and each subgroup H_i is a finite subgroup of G (and so $d_i = 0$). When $k \geqslant 1$, the condition that $\mu(e) > 0$ is automatically satisfied.

The set $\mathcal{P}(G, \mathrm{reg})$ includes all (non-degenerated) convex combinations of finitely many probability measures of the power-law type

$$\mu_{H,\alpha}(h) \asymp \begin{cases} (1 + |h|_{H,S_H})^{-(d_H + \alpha_H)} & \text{if } h \in H, \quad \alpha_H > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Here, H is a subgroup of G with intrinsic volume growth of degree d_H. The subgroup H and the positive real α_H can both vary freely and independently. Note that our notion of "power-law type" is defined in reference to an intrinsic word-length $|\cdot|_{H,S_H}$ for the subgroup H (here, S_H is a fixed but arbitrary symmetric finite generating set for H).

More generally, simple examples of increasing functions of regular variation are

$$\phi(t) = (1+t)^{\alpha}[1 + \log(1+t)]^{\beta_1}[1 + \log(1 + \log(1+t))]^{\beta_2},$$

where $\alpha > 0$ is the index of regular variation and $\beta_1, \beta_2 \in \mathbb{R}$. We refer the reader to [4] for a detailed treatment of the notion of regular variation. Some readers may prefer to restrict their attention to the simplest case $\phi(t) = (1+t)^\alpha$ as in the following theorem which illustrates one of the main results of this paper.

Theorem 1.8. — *Let G be a finitely generated group of polynomial volume growth. Let μ be a symmetric probability measure on G which belongs to $\mathcal{P}(G, \mathrm{reg})$ with $\phi_i(t) = t^{\alpha_i}$, $\alpha_i \in (0,2)$, $1 \leqslant i \leqslant k$. Then there exists a real $d = d(G,\mu) \geqslant 0$ such that*

$$\forall\, n \in \{0,1,2,\dots,\}, \quad \mu^{(n)}(e) \asymp \frac{1}{(1+n)^d}.$$

In fact we will prove a stronger version of this theorem which deals with all measures in $\mathcal{P}_{\preceq}(G, \mathrm{reg})$. This is a subset of $\mathcal{P}(G, \mathrm{reg})$ whose definition involves a minor technical additional assumption regarding the functions ϕ_i, $i \in \{1,\dots,k\}$ (see Definition 3.9). In this more general version,

$$\mu^{(n)}(e) \asymp \frac{1}{\mathbf{F}(n)},$$

where $\mathbf{F}$ is a regularly varying function which has positive index when G is infinite. Our results allow for the explicit computation of the index d (more generally, $\mathbf{F}$) in terms of the data describing the measure μ and the structure of the group G. This is done by introducing quasi-norms on G that generalize the word-length (see Definitions 2.1-2.4). Different measures typically call for different quasi-norms and for each measure μ in $\mathcal{P}_{\preceq}(G, \mathrm{reg})$, we construct an adapted quasi-norm $\|\cdot\|$. Using this adapted quasi-norm, we prove a near diagonal two-sided estimate for $\mu^{(n)}$ and show that the bounded solutions of the associated parabolic difference equation are Hölder continuous.

The results proved here extend in significant ways those obtained in [22] by two of the authors. First, [22] only deals with nilpotent groups. It is one of the main goals of this paper to treat the larger and more natural class of group of polynomial volume growth. Second, the measures considered in [22] are convex combination of measures supported on *one parameter* discrete subgroups, i.e., subgroups of the type $\{g = s^m : m \in \mathbb{Z}\}$, $s \in G$. Here, we consider measures supported on general subgroups. Even when G is nilpotent and the subgroups H_i appearing in the definition of μ are one parameter subgroups, the present paper treats cases that where left aside in [22] (e.g., power laws with arbitrary positive exponents). Nevertheless, some of the main technical results of [22] are used here again in a crucial way to pass from nilpotent groups to groups of polynomial volume growth.

1.3. Dirichlet forms and spectral profiles

We will make use of well established techniques based on Dirichlet forms and the notion of spectral profile. Let μ be a symmetric probability measure on a finitely generated group G. We do not necessarily assume that the support of μ generates G. The symmetric probability measure μ determines a Dirichlet form given by

$$\mathcal{E}_{G,\mu}(f,f)=\frac{1}{2}\sum_{x,y\in G}|f(xy)-f(x)|^2\mu(y),\quad f\in L^2(G).$$

Here, $L^2(G)$ is the Hilbert space with norm

$$\|f\|_2=\left(\sum_{x\in G}|f(x)|^2\right)^{1/2}.$$

The spectral profile of the measure μ, $\Lambda_{2,G,\mu}$, is the function defined over $[1,\infty)$ by

$$\Lambda_{2,G,\mu}(v)=\min\left\{\mathcal{E}_{G,\mu}(f,f)/\|f\|_2^2:\ 1\leqslant \#\,\mathrm{support}(f)\leqslant v\right\}.$$

It can also be defined by considering each non-empty finite set $A\subset G$ of volume at most v, minimizing the Raleigh quotient of functions f supported in A to obtain the lowest eigenvalue $\lambda_\mu(A)$ of (minus) the discrete Laplacian, $f\mapsto f*(\delta_e-\mu)$, with Dirichlet boundary condition outside A, and taking the minimum of $\lambda_\mu(A)$ over all such finite sets A. (Note that the discrete Laplacian $f\mapsto f*(\mu-\delta)$ is non-positive definite.)

In the cases of interest here, we expect inverse power-function estimates for $\Lambda_{2,G,\mu}$. The well established relation between the spectral profile of μ and the decay of $\mu^{(2n)}(e)$ indicates that, for any $\gamma>0$,

- $\forall\, v\geqslant 1,\ \ \Lambda_{2,G,\mu}(v)\succeq v^{-1/\gamma}$, is equivalent to $\mu^{(2n)}(e)\preceq n^{-\gamma}$.
- $\forall\, v\geqslant 1,\ \ \Lambda_{2,G,\mu}(v)\preceq v^{-1/\gamma}$, is equivalent to $\mu^{(2n)}(e)\succeq n^{-\gamma}$.

More generally, if F is a positive monotone function of regular variation of index $\gamma>0$ (at infinity) and F^{-1} is its inverse (hence a function of regular variation of index $1/\gamma$), then

- $\Lambda_{2,G,\mu}\succeq 1/F^{-1}$ is equivalent to $\mu^{(2n)}(e)\preceq 1/F(n)$.
- $\Lambda_{2,G,\mu}\preceq 1/F^{-1}$ is equivalent to $\mu^{(2n)}(e)\succeq 1/F(n)$.

For details, see [8] and [23, Section 2.1].

Another key property that we will use without further comment throughout is the fact that for any two symmetric probability measures μ_1,μ_2, the inequality

$$\mathcal{E}_{G,\mu_1}\leqslant A\mathcal{E}_{G,\mu_2}$$

implies that

$$\mu_2^{(2n)}(e) \preceq \mu_1^{(2n)}(e);$$

that is, there exist A_1, A_2 such that

$$\forall\, n = \{1, 2, \dots\},\ \ \mu_2^{(2A_1 n)}(e) \leqslant A_2 \mu_1^{(2n)}(e).$$

In particular, if $\mu_1 \asymp \mu_2$ on G then $\mu_1^{(2n)}(e) \simeq \mu_2^{(2n)}(e)$. Whenever, in addition, $\mu_1(e)\mu_2(e) > 0$, the conclusion easily extends to $\mu_1^{(n)}(e) \simeq \mu_2^{(n)}(e)$. For background information on these notions and techniques, we refer the reader to the books [26, 27] and to [8, 15, 20, 23].

1.4. Guide to the reader

The paper is organized as follows. Subsection 2.1 introduces the quasi-norms and geometries that are key to the study of the walks driven by measures in $\mathcal{P}_{\preceq}(G, \mathrm{reg})$. See Definition 3.9. Each of these geometries is associated with a generating tuples $\Sigma = (s_1, \dots, s_k)$ of elements of the group G and a weight function system $\mathfrak{F} = \{F_s, s \in \Sigma\}$. In the study of random walks, the structure of a given measure μ in $\mathcal{P}_{\preceq}(G, \mathrm{reg})$ will determine in large part how to choose Σ and $\mathfrak{F}$.

Subsection 2.2 describes results from [22] concerning the case of nilpotent groups which play a key role in the rest of the paper. See Theorem 2.14.

Section 3 discusses how geometric results (existence of coordinate-like systems and volume growth) leads to lower bounds on the spectral profile and upper bounds on the probability of return of measures in $\mathcal{P}_{\preceq}(G, \mathrm{reg})$. Subsection 3.2 applies these results to nilpotent groups. Subsection 3.3, one of the most important parts of the paper, explains how to obtain sharp results in the case of groups of polynomial volume growth. Given a group of polynomial growth and a measure $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$, we explain the construction of a well adapted geometry on G based on the (well-known) existence of a nilpotent group N with finite index in G. In fact, we construct geometries on N and on G which are closely related to each other and well adapted to the given measure μ on G. Some explicit examples are given.

Section 4 provides matching upper-bounds on the spectral profiles and the corresponding lower bounds on the probability of return. This is done by providing appropriate test functions which are defined using the quasi-norms of Section 2. See Theorem 4.3.

Section 5 contains one of the main theorems, Theorem 5.5, which gather the main properties of the iterated convolution $\mu^{(n)}$ and the associated random walk when $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$ and G has polynomial volume growth.

Section 6 proves the Hölder estimate for solutions of the corresponding discrete parabolic equation (see, Theorem 6.8). The main results of the paper are in Theorems 4.3, 5.5 and 6.8.

2. Geometries for random walks with long range jumps

As noted in the introduction, the word-length associated to a finite symmetric generating set S is a key element in developing an understanding of the behavior of the random walks driven by symmetric finitely supported measures. The question arises as to what are the natural geometries that might help us understand random walks that allow for long range jumps. This section introduces such geometries.

2.1. Weight systems and quasi-norms

First, let us give a more formal definition of the word-length associated with a finite set of generator. Fix a finite alphabet $\Sigma = \{s_1, \dots, s_k\}$ and adjoin to it the formal inverses (new letters) $\Sigma^{-1} = \{s_1^{-1}, \dots, s_k^{-1}\}$. A finite word w over $\Sigma \cup \Sigma^{-1}$ is a formal product (i.e., a finite sequence) $w = \sigma_1 \dots \sigma_m$ with $\sigma_i \in \Sigma \cup \Sigma^{-1}$, $1 \leqslant i \leqslant m$. Equivalently, we can write $w = \sigma_1^{\varepsilon_1} \dots \sigma_m^{\varepsilon_m}$ with $\sigma_i \in \Sigma$ and $\varepsilon_i \in \{\pm 1\}$, $1 \leqslant i \leqslant m$. If G is a group which contains elements called $s_1, \dots, s_k$, we say that the word $w = \sigma_1 \dots \sigma_m$ over $\Sigma \cup \Sigma^{-1}$ is equal to $g \in G$, if $\sigma_1 \dots \sigma_m = g$ when reading this product in G. Formally, one should denote the letters by $\mathbf{s}_i$, the corresponding group elements by s_i, and introduce the map $\pi : \bigcup_{q=0}^{\infty} (\Sigma \cup \Sigma^{-1})^q \to G$ defined by $\pi(\boldsymbol{\sigma}_1 \dots \boldsymbol{\sigma}_m) = \sigma_1 \dots \sigma_m$. With this notation the word-length $|g|$ of an element $g \in G$ with respect to the k-tuple of generators $(s_1, \dots, s_k)$ and their inverses is

$$|g| = \inf\{m : \exists\, w \in (\Sigma \cup \Sigma^{-1})^m,\ g = w \text{ in } G\}.$$

By convention, $|e| = 0$ (e can be obtained as the empty word). For illustrative purpose, we introduce the following variant

$$\|g\| = \inf\left\{\max_{s\in\Sigma}\{\deg_s(w)\} : w \in \bigcup_0^{\infty} (\Sigma \cup \Sigma^{-1})^m,\ \ g = w \text{ in } G\right\},$$

where, for each $s \in \Sigma$ and $w \in \bigcup_0^{\infty} (\Sigma \cup \Sigma^{-1})^m$, $w = s_{j_1}^{\varepsilon_1} \dots s_{j_m}^{\varepsilon_m}$, we set

$$\deg_s(w) = \#\{\ell \in \{1, \dots, m\} : s_{j_\ell} = s\}.$$

In words, $\deg_s(w)$ is the number of times the letter s is used (in the form s or s^{-1}) in the word w. Obviously,

$$\|g\| \leqslant |g| \leqslant k\|g\|, \quad \text{where} \quad k = \#\Sigma.$$

The reader should note that when defining $\deg_s$, we think of s as a letter in the alphabet Σ (two distinct words consisting of letters in $\Sigma \cup \Sigma^{-1}$ might become equal as an element in G). In addition, $\deg_s$ counts the occurrences of both s and s^{-1}. For instance, consider the word $w = s_1 s_2 s_1^{-1} s_2^{-1} s_3 s_1^{-1}$. The degrees are as follows:

$$\deg_{s_1}(w) = 3, \quad \deg_{s_2}(w) = 2, \quad \deg_{s_3}(w) = 1.$$

This is the case even if it happens, as it may, that $s_3 = s_1^{-1}$ in G.

Definition 2.1. — *We say that a map $N : G \to [0, \infty)$ is a norm, if*

$$N(gh) \leqslant N(g) + N(h).$$

We say that it is a quasi-norm, if there exist a constant A such that

$$N(gh) \leqslant A(N(g) + N(h)).$$

Remark 2.2. — The quasi-norms constructed in this paper have two additional properties. They are symmetric ($N(g) = N(g^{-1})$, $g \in G$) and $N(e) = 0$.

Example 2.3. — The maps $g \mapsto |g|$ and $g \mapsto \|g\|$ associated to a generating tuple $(s_1, \ldots, s_k)$ as above are norms.

Now, we introduce a (potential) quasi-norm $\|\cdot\|_{\mathfrak{F}}$ associated with a family $\mathfrak{F}$ of continuous and strictly increasing functions on $[0, \infty)$. This will be a quasi-norm under some additional technical assumptions on the family $\mathfrak{F}$. The basic data for such a function $\|\cdot\|_{\mathfrak{F}}$ consists of a group G, a tuple $\Sigma = (s_1, \ldots, s_k)$ (abusing notation, we will consider each s_i both as an abstract symbol (letter) and as a group element in G) and a family $\mathfrak{F}$ of continuous and strictly increasing functions

$$F_s : [0, \infty) \to [0, \infty), \quad s \in \Sigma, \quad F_s(t) \asymp t \text{ on } [0, 1],$$

with the property that for $s, s' \in \Sigma$,

$$\text{either } F_s \preceq F_{s'} \text{ or } F_{s'} \preceq F_s \text{ on a neighborhood of infinity.} \tag{2.1}$$

With proper care and technical modifications, condition (2.1) can probably be removed but we will assume it holds throughout this paper. Each of the function F_s is invertible, and we denote by F_s^{-1} its inverse. A good example to keep in mind is the case when, for each $s \in \Sigma$, we are given a positive real $\omega(s)$ and $F_s(t) = t\mathbf{1}_{[0,1]}(t) + t^{\omega(s)}\mathbf{1}_{(1,\infty)}(t)$ (or, more or less equivalently,

$F_s(t) = (1+t)^{\omega(s)} - 1)$. We think of F_s as a weight function assigned to $s \in \Sigma$.

Definition 2.4. — *Given G, Σ and $\mathfrak{F}$ as above, for each element $g \in G$, set*

$$\|g\|_{\mathfrak{F}} = \inf \left\{ \max_{s\in\Sigma} \{F_s^{-1}(\deg_s(w))\} : w \in \bigcup_0^\infty (\Sigma \cup \Sigma^{-1})^m, \quad g = w \text{ in } G \right\}.$$

By convention, $\|e\|_{\mathfrak{F}} = 0$. If g cannot be represented as a finite word over $\Sigma \cup \Sigma^{-1}$, set $\|g\|_{\mathfrak{F}} = \infty$.

In other word, $\|g\|_{\mathfrak{F}}$ is the least R such that there exists a finite word w such that $w = g$ in G and

$$\deg_s(w) \leqslant F_s(R) \text{ for each } s \in \Sigma.$$

This last inequality indicates that each letter s (in the form s or s^{-1}) in Σ is used at most $F_s(R)$ times in the word w.

Remark 2.5. — In the context of nilpotent groups, this definition of $\|\cdot\|_{\mathfrak{F}}$ appears in [22, Definition 2.8].

Remark 2.6. — If each F_s satisfies $F_s^{-1}(t_1+t_2) \leqslant A(F_s^{-1}(t_1)+F_s^{-1}(t_2))$, then $\|\cdot\|_{\mathfrak{F}}$ is a quasi-norm (a norm, if $A=1$). In particular, $\|\cdot\|_{\mathfrak{F}}$ is a norm, if each F_s is convex.

Remark 2.7. — If each F_s is replaced by $\widetilde{F}_s = F_s \circ F^{-1}$ for some continuous and strictly increasing function $F : [0,\infty) \to [0,\infty)$ with $F(t) \asymp t$ on $[0,1]$, then

$$\|g\|_{\widetilde{\mathfrak{F}}} = F(\|g\|_{\mathfrak{F}}) \text{ for each } g \in G.$$

Remark 2.8. — Say that a non-negative function f defined on $(0,\infty)$ is doubling, if there exists a constant $A_f > 0$ such that

$$\forall\, t > 0, \quad f(2t) \leqslant A_f f(t) \text{ and } 2f(t) \leqslant f(A_f t). \tag{2.2}$$

Over the class of doubling functions, the equivalence relations $\simeq$ and $\asymp$ coincide. Suppose that we have two weight functions systems $\mathfrak{F}$ and $\mathfrak{F}'$ define over Σ, and that all functions F_s and F'_s are doubling. Suppose further that for each $s \in \Sigma$, $F_s \simeq F'_s$. Then we can conclude that $\|g\|_{\mathfrak{F}} \asymp \|g\|_{\mathfrak{F}'}$ over G.

Example 2.9. — Let $G = \mathbb{Z}^k$ with the canonical generators $(s_1,\ldots,s_k)$. For $t \geqslant 1$, let $F_{s_i}(t) = t^{\omega_i}$ with $\omega_i > 0$. Then, for $x = (x_1,\ldots,x_k) = \sum x_i s_i$,

$$\|(x_1,\ldots,x_k)\|_{\mathfrak{F}} = \max_i \{|x_i|^{1/\omega_i}\}.$$

Example 2.10 (Heisenberg group). — Let $G = (\mathbb{Z}^3, \bullet)$ with

$$g \bullet g' = (x_1 + x_1', x_2 + x_2', x_3 + x_3' + x_1 x_2'),$$

i.e., in coordinates, matrix multiplication in the Heisenberg group

$$G = \mathbb{H}(3, \mathbb{Z}) = \left\{ g = (x_1, x_2, x_3) = \left(\begin{smallmatrix} 1 & x_1 & x_3 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{smallmatrix}\right) : x_1, x_2, x_3 \in \mathbb{Z} \right\}.$$

For $i = 1, 2, 3$, let s_i be the triplet with a 1 in position i and 0 otherwise. For $t \geqslant 1$, let $F_{s_i}(t) = t^{\omega_i}$ with $\omega_i > 0$. Then

$$\|(x_1, x_2, x_3)\|_{\mathfrak{F}} \asymp \begin{cases} \max\{|x_1|^{1/\omega_1}, |x_2|^{1/\omega_2}, |x_3|^{1/\omega_3}\} & \text{if } \omega_3 \geqslant \omega_1 + \omega_2, \\ \max\{|x_1|^{1/\omega_1}, |x_2|^{1/\omega_2}, |x_3|^{1/(\omega_1+\omega_2)}\} & \text{if } \omega_3 \leqslant \omega_1 + \omega_2. \end{cases}$$

See [22, Examples 1.1 and 4.3].

The following proposition is technical in nature. Parts (b) and (c) will be used later in deriving the main new results of this paper.

Proposition 2.11. — *Consider a weight function system $(\mathfrak{F}, \Sigma)$ on a group G as above. Assume that for each $s \in \Sigma$, the function F_s is regularly varying of index $\omega(s) > 0$ at infinity.*

(a) *There exists a weight function system $(\mathfrak{F}_0, \Sigma)$ such that each member $F_{0,s}$ is in $\mathcal{C}^1([0,\infty))$, increasing, and of smooth variation in the sense of* [4, Section 1.8] *with $F_{0,s} \asymp F_s$, for $s \in \Sigma$.*

(b) *For any fixed $\omega^* \in (0, \infty)$ with $\omega^* > \max\{\omega(s) : s \in \Sigma\}$, there is a weight function system $(\mathfrak{F}_1, \Sigma)$ such that each member $F_{1,s}$ is in $\mathcal{C}^1([0,\infty))$, increasing, and of smooth variation of index less than 1 in the sense of* [4, Section 1.8] *with $F_{1,s} \asymp F_s \circ F^{-1}$, where $F(t) = (1+t)^{\omega^*} - 1$. In particular, there exists a positive real A such that, for all $s \in \Sigma$ and all $T \in [0, \infty)$,*

$$\sup_{[0,T]} \left\{ \frac{\mathrm{d}F_{1,s}(t)}{\mathrm{d}t} \right\} \leqslant A \frac{F_{1,s}(T)}{T} \tag{2.3}$$

and

$$\|g\|_{\mathfrak{F}} \asymp \|g\|_{\mathfrak{F}_1}^{1/\omega^*} \quad \text{over } G.$$

(c) *For any fixed ω_* with $0 < \omega_* < \min\{\omega(s) : s \in \Sigma\}$, there is a weight function system $(\mathfrak{F}_2, \Sigma)$ such that each member $F_{2,s}$ is in $\mathcal{C}^1([0,\infty))$ increasing, convex, and of smooth variation in the sense of* [4, Section 1.8] *with $F_{2,s} \asymp F_s \circ F^{-1}$, where $F(t) = (1+t)^{\omega_*} - 1$. In particular, $g \mapsto \|g\|_{\mathfrak{F}_2}$ is a norm and*

$$\|g\|_{\mathfrak{F}} \asymp \|g\|_{\mathfrak{F}_2}^{1/\omega_*} \text{ over } G.$$

Proof. — Part (a) is essentially [4, Theorem 1.8.2]. The difference is that we impose some simple additional conditions regarding the behavior of $F_{0,s}$ on $[0,a]$ for some $a>0$ (smooth regular variation is a property of $F_{0,s}$ on a neighborhood of infinity). By inspection, it is clear that these additional conditions can be achieved.

The main point of part (b) is that the functions $F_s \circ F^{-1}$, $s \in \Sigma$, are all regularly varying with positive index strictly less than one. By [4, Theorem 1.8.2], there are positive functions $\widetilde{F}_s$ (defined on a neighborhood $[a,\infty)$ of infinity), increasing, of smooth regular variation and satisfying (see the discussion on [4, Page 44]) $\widetilde{F}_s \sim F_s \circ F^{-1}$ at infinity and

$$\frac{d\widetilde{F}_s(t)}{dt} \leqslant A\frac{\widetilde{F}_s(t)}{t} \text{ on } [a,\infty).$$

We can now pick a constant $C(s)$ so that the function $F_{1,s}$ obtained by extending $C(s)+\widetilde{F}_s$ linearly on $[0,a]$, so that $F_{1,s}(0)=0$, $F_{1,s}(t)=C(s)+\widetilde{F}_s(t)$ on $[a,\infty)$, which belongs to $\mathcal{C}^1[0,\infty)$, is increasing and of smooth regular variation, and satisfies the other desired properties.

The main point of part (c) is that the functions $F_s \circ F^{-1}$, $s \in \Sigma$, are now all regularly varying with positive index strictly greater than one. In this case, [4, Theorem 1.8.2] gives positive functions $\widetilde{F}_s$ (defined on a neighborhood $[a,\infty)$ of infinity), increasing, convex, and of smooth regular variation such that $\widetilde{F}_s \sim F_s \circ F^{-1}$. Proceeding as in part (b), we can extend modified versions to $[0,\infty)$ with all the desired properties. $\square$

Remark 2.12. — In parts (b) and (c) of Proposition 2.11, we are avoiding the slightly troublesome case when the index is exactly 1. This is troublesome when the corresponding weight function is not exactly linear. For instance, in part (b), the result is still correct, but the derivative and its upper bound are not necessarily monotone. This becomes a real problem for part (c). By a further variation of this argument, one can use composition by a function F as above to avoid all integers index (indeed, there is only finitely many F_s to deal with so proper choices of ω_* and ω^* do the trick). Then one can apply [4, Theorem 1.8.3], which is more elegant than the above construction and provides similar results.

2.2. Volume counting from [22]

Given a group G equipped with a generating tuple $(s_1,\dots,s_k)$ and a weight function system $\mathfrak{F}$, it is really not clear how to compute or "understand" the map $g \mapsto \|g\|_{\mathfrak{F}}$. The article [22] considers the case of nilpotent groups and connects the results to the study of certain random walks with long range jumps.

Beyond the nilpotent case, questions such as

- What is the cardinality of $\{g \in G : \|g\|_{\mathfrak{F}} \leqslant R\}$?
- Which choice of $(\Sigma, \mathfrak{F})$ is relevant to which random walk with long range jumps?

do not seem very easy to answer.

For a better understanding of our main results, it is useful to review and emphasize the main volume counting results derived in [22] in the case of nilpotent groups. We want to apply these results in the context of Definition 2.4. We make the assumption that all the functions appearing in the system $\mathfrak{F}$ are doubling (see Remark 2.8). Recall that, by (2.1), we have a well defined total order on $\{F_s, s \in \Sigma\}$ (modulo the equivalence relation $\asymp$ on a neighborhood of infinity. The equivalence relation $\simeq$ and $\asymp$ are equal on doubling functions).

Following [22], we extend our given weight function system to the collection of all finite length abstract commutators over the alphabet $\Sigma \cup \Sigma^{-1}$ by using the rules

$$F_{s^{-1}} = F_s \text{ for } s \in \Sigma \quad \text{and} \quad F_{[c_1,c_2]} = F_{c_1} F_{c_2}.$$

In short, abstract commutators are formal entities obtained by induction via the building rule $[c_1, c_2]$ starting from $\Sigma \cup \Sigma^{-1}$ (see [22] for more details). Observe that the family of functions F_c, c running over formal commutators, have property (2.1) and thus carry a well defined total order (again, modulo the equivalence relation $\asymp$). For notational convenience, we introduce formal representatives for the linearly ordered distinct elements of $\{F_c \bmod \asymp\}$ and call these representatives

$$\overline{\omega}_1 < \overline{\omega}_2 < \overline{\omega}_3 < \dots.$$

Hence, $\overline{\omega}_1$ represents the $\asymp$ equivalence class associated with the smallest of the weight function F_c, etc. For each $\overline{\omega}_i$, let $\mathbf{F}_i$ be a representative of the $\asymp$ equivalence class of functions F_c associated with $\overline{\omega}_i$. By definition, we can pick any commutator c with $F_c \in \overline{\omega}_i$ and set

$$\mathbf{F}_i = \prod_1^{\ell} F_{\sigma_i},$$

where $\sigma_1, \dots, \sigma_\ell$ is the complete list (with repetition) of the elements of $\Sigma \cup \Sigma^{-1}$ that are used to form the formal commutator c.

Definition 2.13. — *Referring to the above notation, let $G_i^{\mathfrak{F}}$ be the subgroup of G generated by all (images in G of) formal commutators c such that*

$$F_c \in \bigcup_{j \geqslant i} \bar{\omega}_j.$$

In other words, $G_i^{\mathfrak{F}}$ is generated by all commutators such that $F_c \succeq \mathbf{F}_i$.

Obviously, these groups form a descending sequence of subgroups of G and, under the assumption that G is nilpotent, there exists a smallest integer $j_* = j_*(\mathfrak{F})$ such that $G_{j_*+1}^{\mathfrak{F}} = \{e\}$. Further, not only $G_j^{\mathfrak{F}} \supseteq G_{j+1}^{\mathfrak{F}}$ but also (see [22, Proposition 2.3])

$$[G, G_j^{\mathfrak{F}}] \subseteq G_{j+1}^{\mathfrak{F}},$$

so that $G_j^{\mathfrak{F}}/G_{j+1}^{\mathfrak{F}}$ is a finitely generated abelian group. We let

$$\mathfrak{r}_j = \operatorname{rank}(G_j^{\mathfrak{F}}/G_{j+1}^{\mathfrak{F}})$$

be the torsion free rank of this abelian group (by the finitely generated abelian group structure theorem, any such group A is isomorphic to a product of the form $K \times \mathbb{Z}^r$ where K is a finite abelian group. The integer r is the torsion free rank of the group A).

Theorem 2.14 ([22, Theorems 2.10 and 3.2]). — *Let G be a nilpotent group equipped with a finite tuple of generators Σ and a weight function system $\mathfrak{F}$ as above satisfying* (2.1)-(2.2). *From this data, extract the functions $\mathbf{F}_j$ and integers $\mathfrak{r}_j$, $1 \leqslant j \leqslant j_*$, as explained above. Then*

$$\#\{g \in G : \|g\|_{\mathfrak{F}} \leqslant R\} \asymp \prod_1^{j_*} [\mathbf{F}_i(R)]^{\mathfrak{r}_j}.$$

Furthermore, there exist an integer Q, a constant C, and a sequence $\sigma_1, \dots, \sigma_Q$ with $\sigma_j \in \Sigma$, $1 \leqslant j \leqslant Q$, such for any positive R and any element g with $\|g\|_{\mathfrak{F}} \leqslant R$, g can be written in the form

$$g = \prod_{j=1}^{Q} \sigma_j^{x_j} \text{ with } |x_j| \leqslant C F_{\sigma_j}(R).$$

Definition 2.15. — *Let G be a nilpotent group equipped with a finite tuple of generators Σ and a weight function system $\mathfrak{F}$ as above satisfying* (2.1)–(2.2). *From this data, extract the functions $\mathbf{F}_j$ and integers $\mathfrak{r}_j$, $1 \leqslant j \leqslant j_*$, as explained above. Set*

$$\mathbf{F}_{G,\mathfrak{F}} = \mathbf{F}_{\mathfrak{F}} = \prod_1^{j_*} \mathbf{F}_i^{\mathfrak{r}_j}.$$

By Theorem 2.14, for any nilpotent group and weight function system satisfying (2.1)–(2.2), we have the explicit volume estimate

$$\#\{g \in G : \|g\|_{\mathfrak{F}} \leqslant R\} \asymp \mathbf{F}_{\mathfrak{F}}(R).$$

Example 2.16. — Let us return to the Heisenberg group example $G = \mathbb{H}(3, \mathbb{Z})$, Example 2.10, with $F_{s_i}(t) = t^{\omega_i}$ with $\omega_i > 0$ for all $t \geqslant 1$ and $1 \leqslant i \leqslant 3$. Then

$$\mathbf{F}_{\mathfrak{F}}(R) \asymp \begin{cases} R^{\omega_1+\omega_2+\omega_3} & \text{if } \omega_3 \geqslant \omega_1 + \omega_2, \\ R^{2(\omega_1+\omega_2)} & \text{if } \omega_3 \leqslant \omega_1 + \omega_2. \end{cases}$$

3. Upper bounds on return probabilities

3.1. A general approach

In this section we discuss how to obtain upper bounds for the return probability of a measure $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$, a subset of $\mathcal{P}(G, \mathrm{reg})$ which is described below in Definition 3.9. The main tool is the following technical result. For the proof, we can follow the proofs of [22, Theorems 4.1 and 4.3] with minor adaptations.

Proposition 3.1. — *Let G be a countable group equipped with symmetric probability measures μ_i, $0 \leqslant i \leqslant k$. Assume that there exists a constant $C > 0$ such that for each $R > 0$ and $0 \leqslant i \leqslant k$, there is a subset $K_i(R) \subset G$ such that*

$$\sum_{x \in G} |f(xh) - f(x)|^2 \leqslant CR\mathcal{E}_{G,\mu_i}(f, f), \quad f \in L^2(G),\ h \in K_i(R). \tag{3.1}$$

Assume further that there are an integer Q and a positive monotone function $\mathbf{F}$ of regular variation of positive index with inverse $\mathbf{F}^{-1}$ such that

$$\#\left(\bigcup_{i=0}^{k} K_i(R)\right)^{Q} \geqslant \mathbf{F}(R),$$

where $\left(\bigcup_{i=0}^{k} K_i(R)\right)^{Q} = \{g = g_1 \ldots g_Q : g_i \in \bigcup_{i=0}^{k} K_i(R)\}$ is viewed as a subset of G. Set $\mu = (k+1)^{-1} \sum_{i=0}^{k} \mu_i$. Then

$$\Lambda_{2,G,\mu} \succeq 1/\mathbf{F}^{-1} \quad \text{and} \quad \mu^{(2n)}(e) \preceq 1/\mathbf{F}(n).$$

To understand how this proposition works in practice, note that the larger the set $K_i(R)$ is, the harder it is to prove (3.1), but the faster the growth from the lower-bound $F(R)$ on $\#\left(\bigcup_{i=0}^{k} K_i(R)\right)^{Q}$ one might expect.

Observe also that the inequality (3.1) is trivial if $K_i(R) = \{e\}$ but, then, $K_i(R)$ does not contribute at all to the growth of $\#\left(\bigcup_{i=0}^k K_i(R)\right)^Q$. If $K_i(R) = \{e\}$ for all $i = 0, \ldots, k$, then $F(R)$ cannot grow and the conclusion of the proposition is trivial. Formally, this case is excluded by the requirement that F is regularly varying of positive index.

The next two propositions provide a way to verify assumption (3.1) for the type of measures of interest to us here.

Proposition 3.2. — *Let G be a countable group equipped with a symmetric probability measure μ supported on a subgroup H equipped with a quasi-norm $\|\cdot\|$ such that there exist a constant $d > 0$ and a positive monotone regularly varying function $\phi : (0, \infty) \to (0, \infty)$ of positive index at infinity such that*

$$\#\{h \in H : \|h\| \leqslant r\} \asymp r^d, \quad \mu(h) \asymp \left[\phi(1 + \|h\|)(1 + \|h\|)^d\right]^{-1}.$$

Then there is a constant C such that, for all $f \in L^2(G)$, $R \geqslant 1$ and $h \in H$ with $\|h\| \leqslant R$, we have

$$\sum_{x \in G} |f(xh) - f(x)|^2 \leqslant C\phi(R)\mathcal{E}_{G,\mu}(f, f).$$

Proof. — First observe that

$$\sum_{\|h\| \geqslant r} \mu(h) \asymp 1/\phi(r). \tag{3.2}$$

This is proved by summing over A-adic annuli with A large enough so that

$$\#\{A^{n-1} \leqslant \|h\| \leqslant A^n\} \asymp \#\{\|h\| \leqslant A^n\} \asymp A^{dn}.$$

Next, note that for $h, h' \in H$, the inequality $\|h'\| \geqslant C_0\|h\|$ (with $C_0 \in (0, 1/A)$ small enough, where A is the constant in the definition of quasi-norm, see Definition 2.1) implies

$$\|h^{-1}h'\| \geqslant C^{-1}\|h'\| \geqslant C^{-1}C_0\|h\| \text{ and } \mu(h') \leqslant C\mu(h^{-1}h')$$

for some constant $C \geqslant 1$. Now, write

$$\begin{aligned}
\sum_{x\in G} |f(xh)-f(x)|^2 &\left(\sum_{|h'|\geqslant C_0|h|} \mu(h')\right) \\
&\leqslant 2 \sum_{|h'|\geqslant C_0|h|} \sum_{x\in G} (|f(xh)-f(xh')|^2 + |f(xh')-f(x)|^2)\mu(h') \\
&\leqslant C \sum_{h'\in H:|h'|\geqslant C_0|h|} \sum_{x\in G} |f(x)-f(xh^{-1}h')|^2 \mu(h^{-1}h') \\
&\qquad + 2 \sum_{h'\in H, x\in G} |f(xh')-f(x)|^2 \mu(h') \\
&\leqslant C \sum_{h'\in H:|h^{-1}h'|\geqslant C^{-1}|h'|} \sum_{x\in G} |f(x)-f(xh^{-1}h')|^2 \mu(h^{-1}h') \\
&\qquad + 2 \sum_{h'\in H, x\in G} |f(xh')-f(x)|^2 \mu(h') \\
&\leqslant (2+C)\mathcal{E}_{G,\mu}(f,f).
\end{aligned}$$

This gives the desired inequality. □

Proposition 3.3. — *Let G be a countable group equipped with symmetric probability measure μ supported on a subgroup H equipped with a finite generating set and its word length $|\cdot|$. Assume that there exist a constant $d > 0$ and a positive monotone regularly varying function $\phi : (0,\infty) \to (0,\infty)$ of positive index at infinity such that*

$$\#\{h \in H : |h| \leqslant r\} \asymp r^d, \quad \mu(h) \asymp \left[\phi(1+|h|)(1+|h|)^d\right]^{-1}.$$

Then there is a constant C such that, for all $f \in L^2(G)$, $R \geqslant 1$ and $h \in H$ with $\|h\| \leqslant R$, we have

$$\sum_{x\in G} |f(xh)-f(x)|^2 \leqslant C\Phi(R)\mathcal{E}_{G,\mu}(f,f),$$

where $\Phi(t) = t^2/\int_0^t \frac{sds}{\phi(s)}$, $t \geqslant 1$.

Proof. — Let u_r be the uniform probability measure on $\{h\in H : |h|\leqslant r\}$. Follow the proof of [23, Proposition A.4] to show that for any $f \in L^2(G)$ and any $0 < s < r < \infty$ and $h \in G$ with $|h| \leqslant r$, we have

$$\sum_{x\in G} |f(xh)-f(x)|^2 \leqslant C(r/s)^2 \sum_{x\in G}\sum_{h\in H} |f(xh)-f(x)|^2 u_s(h).$$

Observe that $\mu \asymp \sum_0^\infty \frac{1}{\phi(2^n)} u_{2^n}$ and that, for $r \geqslant 1$,

$$\sum_{n:2^n \leqslant r} \frac{2^{2n}}{\phi(2^n)} \asymp \int_0^r \frac{s\mathrm{d}s}{\phi(s)}.$$

The desired inequality follows. □

Remark 3.4. — If ϕ is regularly varying of positive index γ, then we always have that $\Phi(t) \leqslant C\phi(t)$ and $\phi(t) \asymp \Phi(t)$ if $\gamma \in (0,2)$. If $\gamma > 2$, $\Phi(t) \asymp t^2$ which is much less than ϕ on $(1,\infty)$.

Remark 3.5. — The proof of Proposition 3.3 outlined above uses the fact that, roughly speaking, an element h with $|h| = r$ can be written as a product of (r/s) elements of length at most s. This property is not necessarily true for an arbitrary quasi-norm $\|\cdot\|$ as in Proposition 3.2.

Definition 3.6. — *Given $\mu = \sum_0^k \mu_i \in \mathcal{P}(G, \mathrm{reg})$ with μ_i defined in terms of a regularly varying function ϕ_i as in* (1.3), $1 \leqslant i \leqslant k$, *set*

$$\Phi_0(t) = \max\{t, t^2\}$$

and, for $1 \leqslant i \leqslant k$,

$$\Phi_i(t) = \frac{t^2}{\int_0^t \frac{2s}{\phi_i(s)}\mathrm{d}s} \quad \textit{on} \quad [1,\infty),$$

with Φ_i extended linearly on $[0,1]$ *with* $\Phi_i(0) = 0$.

Remark 3.7. — The exact definition of Φ_i in the interval $[0,1]$ is not very important to us. For convenience, we prefer to have it vanish linearly at 0. Because ϕ_i is increasing, one can check that

$$\Phi_i \leqslant \phi_i \text{ on } [1,\infty) \text{ and } \Phi_i \text{ is increasing on } (0,\infty).$$

Further, Φ_i is regularly varying at infinity of index in $(0,2]$. More precisely, $\Phi_i \asymp \phi_i$ when the index of ϕ_i is in $(0,2)$. When the index of ϕ is at least 2, then $\Phi(t) \asymp t^2/\ell(t)$ where ℓ is increasing and slowly varying at infinity (i.e., index 0). In particular, in all cases, $\Phi_i(t) \preceq \max\{t,t^2\} = \Phi_0(t)$, $1 \leqslant i \leqslant k$.

Proposition 3.8. — *Let G be a countable group. Let $\mu \in \mathcal{P}(G, \mathrm{reg})$ and, referring to Definitions 1.5 and 3.6, set*

$$K_i(r) = \{h \in H_i : \Phi_i(|h|_i) \leqslant r\}, \quad 0 \leqslant i \leqslant k$$

and, for some fixed integer Q,

$$K(r) = \left\{ g \in G : \; g = g_1 \dots g_Q, \;\; g_i \in \bigcup_0^k K_j(r) \right\}.$$

Assume that $\mathbf{F}$ *is a positive monotone regularly varying function of positive index at infinity with the property that*

$$\forall\, r \geqslant 1, \ \ \#K(r) \geqslant \mathbf{F}(r).$$

Let $\mathbf{F}^{-1}$ *be the inverse function of* $\mathbf{F}$*. Then*

$$\Lambda_{2,G,\mu} \succeq 1/\mathbf{F}^{-1} \ \text{ and } \mu^{(n)}(e) \preceq 1/\mathbf{F}(n), \ \ n = 1, 2, \dots.$$

Proof. — This follows immediately from Propositions 3.1–3.3. □

Let us now defined the subset $\mathcal{P}_{\preceq}(G, \mathrm{reg})$ of $\mathcal{P}(G, \mathrm{reg})$ which is relevant to us because of condition (2.1) and Definition 2.4 (it requires that the functions F_s, $s \in \Sigma$, which are used to define a quasi-norm, be ordered modulo $\simeq$).

Definition 3.9 ($\mathcal{P}_{\preceq}(G, \mathrm{reg})$). — *A measure* $\mu = \sum_0^k p_i \mu_i$ *in* $\mathcal{P}(G, \mathrm{reg})$ *with associated regularly varying functions* ϕ_i, $1 \leqslant i \leqslant k$, *is in* $\mathcal{P}_{\preceq}(G, \mathrm{reg})$ *if, for each pair* i, j *of distinct indices in* $\{1, \dots, k\}$, *either* $\Phi_i \preceq \Phi_j$ *or* $\Phi_j \preceq \Phi_i$ *in a neighborhood of infinity.*

3.2. Application to nilpotent groups

The following is a corollary to Proposition 3.8 and Theorem 2.14 (i.e., [22, Theorems 2.10 and 3.2]).

Theorem 3.10. — *Assume that* G *is nilpotent. Let* $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$*. Let* S_0 *be the support of* μ_0 *and* S_i *be a symmetric generating set for the subgroup* H_i *for* $1 \leqslant i \leqslant k$*. Let*

$$\Sigma = (s_1, \dots, s_m)$$

be a tuple of distinct representatives of the set $(\bigcup_0^k S_i) \setminus \{e\}$ *under the equivalence relation* $s^{-1} \sim s$*, and set*

$$\mathfrak{F} = \left\{F_\sigma = \max\{\Phi_i^{-1} : i \in \{j : \sigma \in S_j\}\} : \sigma \in \Sigma\right\}.$$

Let $\mathbf{F} = \mathbf{F}_{G,\mathfrak{F}}$ *be as in Definition 2.15. Then, we have*

$$\Lambda_{2,G,\mu}(v) \succeq 1/\mathbf{F}(v), \ \ v \geqslant 1, \ \ \text{and} \ \ \mu^{(n)}(e) \preceq 1/\mathbf{F}(n), \ \ n = 1, 2 \dots.$$

Proof. — Theorem 2.14 (i.e., [22, Theorems 2.10 and 3.2]) shows that the hypothesis of Proposition 3.8 are satisfied with $\mathbf{F} = \mathbf{F}_{G,\mathfrak{F}}$, which is given in Definition 2.15. The assumption that μ belongs to $\mathcal{P}_{\preceq}(G, \mathrm{reg})$ (instead of $\mathcal{P}(G, \mathrm{reg})$) insures that property (2.1) is satisfied by the functions F_s, $s \in \Sigma$. □

Remark 3.11. — In this theorem, whether we choose the set S_0 to be an arbitrary finite symmetric generating set of G or the support of μ_0 (as we did in the above statement) makes no difference. The reason is that $\Phi_i \preceq \Phi_0$ for each $1 \leqslant i \leqslant k$.

Remark 3.12. — In the case each H_i is a discrete one parameter subgroup of G, Theorem 3.10 is already contained in [22]. The case when $k = 1$, $p_0 = 0$, and $H_1 = G$ is also known. The theorem provides a natural extension covering these two special cases.

Remark 3.13. — Let us emphasize here the fact that the choice of the geometry adapted to $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$ in the above theorem follows straightforwardly from the "structure" of the measure μ which is captured by the subgroups H_i and the regularly varying functions ϕ_i, $1 \leqslant i \leqslant k$. We shall see later that this is not the case when we replace the hypothesis that G is nilpotent by the hypothesis that G has polynomial volume growth.

Example 3.14. — We return again to the Heisenberg example (Example 2.10), keeping the same notation. We let $H_i = \langle s_i \rangle$ (the subgroup generated by s_i) and $\phi_i(t) = (1+t)^{\alpha_i}$ with $\alpha_i > 0$ for $1 \leqslant i \leqslant 3$. Obviously, the volume growth degree of each H_i is $d_i = 1$ for all $1 \leqslant i \leqslant 3$, and, for $t \geqslant 1$,

$$\Phi_i(t) \asymp \begin{cases} t^{\alpha_i} & \text{if } \alpha_i \in (0,2), \\ t^2/\log(1+t) & \text{if } \alpha_i = 2, \\ t^2 & \text{if } \alpha_i > 2. \end{cases}$$

If none of the α_i is equal to 2, set $\widetilde{\alpha}_i = \min\{2, \alpha_i\}$ and $\omega_i = 1/\widetilde{\alpha}_i$. In this case,

$$\mathbf{F}(t) \asymp \begin{cases} t^{\omega_1+\omega_2+\omega_3} & \text{if } \omega_3 \geqslant \omega_1 + \omega_2, \\ t^{2(\omega_1+\omega_2)} & \text{if } \omega_3 \leqslant \omega_1 + \omega_2. \end{cases}$$

The best way to treat the cases that include the possibility that $\alpha_i = 2$ is to introduce a two-coordinate weight system and set

$$\omega_i = (\omega_{i,1}, \omega_{i,2}) = \begin{cases} (1/\widetilde{\alpha}_i, 0) & \text{if } \alpha_i \neq 2, \\ (1/2, 1/2) & \text{if } \alpha_i = 2. \end{cases}$$

With this notation, the natural order over the functions $F_{s_i} = \Phi_i^{-1}$ (in a neighborhood of infinity) is the same as the lexicographical order over the weights $\omega_i = (\omega_{i,1}, \omega_{i,2})$. Furthermore, we have

$$\mathbf{F}(t) \asymp \begin{cases} t^{\omega_{1,1}+\omega_{2,1}+\omega_{3,1}}[\log(1+t)]^{\omega_{1,2}+\omega_{2,2}+\omega_{3,2}} & \text{if } \omega_3 \geqslant \omega_1 + \omega_2, \\ t^{2(\omega_{1,1}+\omega_{2,1})}[\log(1+t)]^{2(\omega_{1,2}+\omega_{2,2})} & \text{if } \omega_3 \leqslant \omega_1 + \omega_2. \end{cases}$$

The corresponding probability of return upper bound is already contained in [22]. Later in this paper we will prove the (new) matching lower bound.

Example 3.15. — We continue with the Heisenberg example (Example 2.10), keeping the same basic notation. Now, we let $H_1 = \langle s_1, s_3\rangle$ and $H_2 = \langle s_2, s_3\rangle$ (these are two abelian subgroups of the Heisenberg group with $d_1 = d_2 = 2$ as each of these subgroups is isomorphic to $\mathbb{Z}^2$). We set again $\phi_i(t) = (1+t)^{\alpha_i}$ with $\alpha_i > 0$ for $1 \leqslant i \leqslant 2$. The associated Φ_i, $i = 1, 2$, are as above. Now, we can pick $\Sigma = (s_1, s_2, s_3)$ and set $F_{s_i} = \Phi_i^{-1}$ for $i = 1, 2$ and $F_{s_3} = \max\{\Phi_1^{-1}, \Phi_2^{-1}\}$. Inspection of the construction of the weight functions on commutators shows that the function F_{s_3} will play no role because $s_3 = [s_1, s_2]$ and $F_{s_3} \preceq F_{s_1}F_{s_2}$ on a neighborhood of infinity. We introduce the two-coordinate weight system $i \in \{1, 2\}$,

$$\omega_i = (\omega_{i,1}, \omega_{i,2}) = \begin{cases} (1/\widetilde{\alpha}_i, 0) & \text{if } \alpha_i \neq 2, \\ (1/2, 1/2) & \text{if } \alpha_i = 2. \end{cases}$$

The volume function $\mathbf{F}$ is then given by

$$\mathbf{F}(t) \asymp t^{2(\omega_{1,1}+\omega_{2,1})}[\log(1+t)]^{2(\omega_{1,2}+\omega_{2,2})}.$$

3.3. Application to groups of polynomial volume growth

It may at first be surprising that the literal generalization of Theorem 3.10 to the case of groups of polynomial volume growth is actually incorrect. This is because the appropriate definition of a weight system $(\Sigma, \mathfrak{F})$ associated with the data describing a probability measure $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$ is more subtle in this case. In order to obtain an appropriate definition, we will make use of a nilpotent approximation N of the group G — that is, a normal nilpotent subgroup of G with finite index. We will build related weight systems $(\Sigma_G, \mathfrak{F}_G)$ and $(\Sigma_N, \mathfrak{F}_N)$ that are compatible in the sense that

$$\forall\, g \in H \subset G, \quad \|g\|_{\mathfrak{F}_G} \asymp \|g\|_{\mathfrak{F}_N}.$$

This construction will have the additional advantage to allow us to bring to bear on the polynomial volume growth case some of the nilpotent results of [22].

Let G be a group having polynomial volume growth, $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$, and H_i, ϕ_i, Φ_i as in Definitions 1.5 and 3.6. As G has polynomial volume growth, Gromov's theorem asserts that G has a nilpotent subgroup N with finite index. It is well known that one can choose N to be a normal subgroup (it suffices to replace N by the kernel of the homomorphism $G \mapsto \mathrm{Sym}(N\backslash G)$

defined by the action of G by right-multiplication on the right-cosets Ng, $g \in G$. This kernel is a subgroup of N because $Ng = N$ only if $g \in N$).

From now on, we assume that N is a normal nilpotent subgroup of G with finite index and we denote by $u_0, \dots, u_n$ the right-coset representatives so that G is the disjoint union of the Nu_i, $0 \leqslant i \leqslant n$, and $u_0 = e$ (since N is normal, these are also left-coset representatives). We are going to use N (a nilpotent approximation of G) to define a weighted geometry on G that is suitable to study the random walk driven by μ. Simultaneously, we will define a compatible geometry on N.

Definition 3.16 (Geometry on G). — *Let G, H_i, ϕ_i, Φ_i, $1 \leqslant i \leqslant k$, $\Phi_0(r) = \max\{r, r^2\}$, and N be as above. Set*

$$N_i = N \cap H_i.$$

Let S_0 be a fixed finite symmetric generating set of G and, for $1 \leqslant i \leqslant k$, let S_i be a symmetric generating set of N_i. Let Σ_G be a set of representatives of $(\bigcup_0^k S_i) \setminus \{e\}$ under the equivalence relation $g \sim g^{-1}$. For each $s \in \Sigma_G$, set

$$F_{G,s} = \max\left\{\Phi_i^{-1} : i \in \{0, \dots, k\}, s \in S_i\right\}.$$

We refer to this system of weight functions on G as $(\Sigma_G, \mathfrak{F}_G)$. For each $s \in \Sigma_G$, fix $i = i_G(s)$ such that

$$s \in S_i \quad \text{and} \quad F_{G,s} = \Phi_i^{-1},$$

and set $\Sigma_G(i) = \{s \in \Sigma_G : i_G(s) = i\}$.

Remark 3.17. — In this definition, it is important that the set S_i is a generating set of $H_i \cap N$, not of H_i itself. It is not hard to see that $N_i = H_i \cap N$ is normal and of finite index in H_i. Indeed, for any subgroups A, B of a group G, we have the index relation $[A : A \cap B] \leqslant [A \cup B : B]$ which we apply here with $A = H_i$ and $B = N$.

Definition 3.18 (Geometry on N). — *Referring to the setting and notation of Definition 3.16, pick a finite symmetric generating set Ξ_0 in N and, for each $1 \leqslant i \leqslant k$, set*

$$\Xi_i = \bigcup_{\ell=0}^{m} u_\ell \, S_i \, u_\ell^{-1}, \quad 1 \leqslant i \leqslant k,$$

where the elements u_ℓ are the fixed right-coset representatives of N in G. Let Σ_N be a set of representatives of $(\bigcup_0^k \Xi_i) \setminus \{e\}$ under the equivalence relation $g \sim g^{-1}$. For each $s \in \Sigma_N$, set

$$F_{N,s} = \max\left\{\Phi_i^{-1} : i \in \{0, \dots, k\}, s \in \Xi_i\right\}, \quad s \in \Sigma_N.$$

We refer to this system of weight functions as $(\Sigma_N, \mathfrak{F}_N)$. *For each* $s \in \Sigma_N$, *fix* $i = i_N(s)$ *such that*

$$s \in \Xi_i \quad \text{and} \quad F_{N,s} = \Phi_i^{-1},$$

and set $\Sigma_N(i) = \{s \in \Sigma_N : i_N(s) = i\}$.

Remark 3.19. — In this definition, it is important that the set Ξ_i is used to define $(\Sigma_N, \mathfrak{F}_N)$ instead of just the generating set S_i of $N_i = H_i \cap N$.

Remark 3.20. — We are abusing notation in denoting our two weight functions systems by $\mathfrak{F}_G$ and $\mathfrak{F}_N$. Indeed, each of them depends on the entire data including G, N, the collection of subgroups H_i, the collection of functions ϕ_i, the choice of S_i, $0 \leqslant i \leqslant k$, Ξ_0, and the choice of the coset representatives u_ℓ, $0 \leqslant \ell \leqslant m$.

One important motivation behind these two definitions is the following result.

Theorem 3.21. — *Referring to the setting and notation of Definitions 3.16 and 3.18, there are constants* $0 < c \leqslant C < \infty$ *such that*

$$\forall\, g \in N, \;\; c\|g\|_{\mathfrak{F}_G} \leqslant \|g\|_{\mathfrak{F}_N} \leqslant C\|g\|_{\mathfrak{F}_G}.$$

Proof.

(1). — We first prove that for $g \in N$, $\|g\|_{\mathfrak{F}_N} \leqslant C\|g\|_{\mathfrak{F}_G}$. For $g \in N$, let $w = \sigma_1^{\varepsilon_1} \dots \sigma_p^{\varepsilon_p}$, $\varepsilon_i \in \{\pm 1\}$, $\sigma_i \in \Sigma_G$, be a word on the alphabet $\Sigma_G \cup \Sigma_G^{-1}$ so that $g = w$ in G. Set

$$g_0 = e, \;\; g_i = \sigma_1^{\varepsilon_1} \dots \sigma_i^{\varepsilon_i}, \;\; 0 \leqslant i \leqslant p,$$

and write (using $g_0 = e$, $g_p \in N$),

$$g = \prod_{i=1}^{p} \widetilde{g_{i-1}} \sigma_i^{\varepsilon_i} (\widetilde{g_i})^{-1}.$$

where, for any $g \in G$, $\widetilde{g} \in \{u_0, \dots, u_m\}$ denotes the fixed coset representative of Ng. Observe that, by definition, $\widetilde{xy} = \widetilde{\widetilde{x}y}$ and that

$$\widetilde{g_i} = \widetilde{g_{i-1}\sigma_i^{\varepsilon_i}} = \widetilde{\widetilde{g_{i-1}}\sigma_i^{\varepsilon_i}}.$$

Setting $j_i = j$ if $\widetilde{g_i} = u_j$, we have

$$g = \prod_{i=1}^{p} u_{j_{i-1}} \sigma_i^{\varepsilon_i} (\widetilde{u_{j_{i-1}} \sigma_i^{\varepsilon_i}})^{-1}.$$

By definition, each factor of this product is in N because $x(\widetilde{x})^{-1} \in N$ for any $x \in G$. If $\sigma_i \in \Sigma_G(\ell) \subset N_\ell \subset N$, $\ell \in \{1, \dots, k\}$, then

$$\widetilde{u_{j_{i-1}}\sigma_i^{\varepsilon_i}} = u_{j_{i-1}}$$

and

$$u_{j_{i-1}}\sigma_i^{\varepsilon_i}(\widetilde{u_{j_{i-1}}\sigma_i^{\varepsilon_i}})^{-1} = u_{j_{i-1}}\sigma_i^{\varepsilon_i}u_{j_{i-1}}^{-1} \in \Xi_\ell.$$

If $\sigma_i \in \Sigma_G(0)$, then we can write

$$u_{j_{i-1}}\sigma_i^{\varepsilon_i}(\widetilde{u_{j_{i-1}}\sigma_i^{\varepsilon_i}})^{-1}$$

as a word of uniformly bounded length at most K using elements in Ξ_0.

Now, we interpret this construction as providing us with a word w' over the alphabet $\Sigma_N \cup \Sigma_N^{-1}$ that represents the given element $g \in N$. By construction, if $\xi = u\sigma u^{-1} \in \Xi_\ell$ with $\sigma \in \Sigma_G(\ell)$, $1 \leqslant \ell \leqslant k$, we have

$$\deg_\xi(w') \leqslant \deg_\sigma(w) \text{ and } F_{G,\sigma} = F_{N,\xi}.$$

Further, if $\xi \in \Xi_0$, then

$$\deg_\xi(w') \leqslant K \sum_{\theta \in \Sigma_G(0)} \deg_\theta(w).$$

This shows that

$$\forall\, g \in N, \quad \|g\|_{\mathfrak{F}_N} \leqslant K\#\Sigma_G(0)\, \|g\|_{\mathfrak{F}_G}.$$

(2). — We next prove that, for $g \in N$, $\|g\|_{\mathfrak{F}_G} \leqslant C\|g\|_{\mathfrak{F}_N}$. For this part we rely on the main result of [22]. By [22, Theorem 2.10], there exist an integer Q, a constant C and a fixed sequence $\xi_1 \dots, \xi_Q \in \Sigma_N$ such that any $g \in N$ with $\|g\|_{\mathfrak{F}_N} = R$ can be written in the form

$$g = \prod_1^Q \xi_i^{x_i} \text{ with } x_i \in \mathbb{Z}, \ \ |x_i| \leqslant CF_{N,\xi_i}(R).$$

Let w be the word over the alphabet $\Sigma_N \cup \Sigma_N^{-1}$ corresponding to this product. On the one hand, for each $i \in \{1, \dots, Q\}$ such that $\xi_i \in \Sigma_N(j)$, $j \geqslant 1$, there are $u \in \{u_1, \dots, u_m\}$ and $\sigma \in S_j$ such that $\xi_i = u\sigma u^{-1}$. Hence, for such i, we can write

$$\xi_i^{x_i} = u\sigma^{x_i}u^{-1}$$

and each u is a product of uniformly bounded length at most K over S_0. On the other hand, each $\xi_i \in \Sigma_N(0)$ can also be written as a finite product of uniformly bounded length at most K using elements in S_0.

Using these decompositions in the product corresponding to w gives us a word w' representing of g over the alphabet $\Sigma_G \cup \Sigma_G^{-1}$ with

$$\deg_\sigma(w') \leqslant \deg_{u\sigma u^{-1}}(w)$$
$$\text{if } \xi = u\sigma u^{-1} \in \Sigma_N(j),\ u \in \{u_1, \dots, u_m\},\ 1 \leqslant j \leqslant k,$$

and

$$\deg_\sigma(w') \leqslant 2KQ + K \sum_{i:\xi_i \in \Sigma_N(0)} |x_i|, \text{ if } \sigma \in S_0.$$

Hence,

$$\forall\, g \in N,\ \ \|g\|_{\mathfrak{F}_G} \leqslant 2KQ\#\Sigma_N(0)\|g\|_{\mathfrak{F}_N}.$$

The proof of Theorem 3.21 is complete. □

Definition 3.22. — *Let G be a group having polynomial volume growth, $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$ and H_i, ϕ_i, Φ_i as in Definitions 1.5 and 3.6. Referring to the notation of Definitions 3.16, 3.18 and 2.15, set*

$$\mathbf{F}_{G,\mathfrak{F}_G} = \mathbf{F}_{N,\mathfrak{F}_N},$$

where $\mathbf{F}_{N,\mathfrak{F}_N}$ is the regularly varying function associated with $(N, \Sigma_N, \mathfrak{F}_N)$ by Definition 2.15.

Corollary 3.23. — *Referring to the setting and notation of Definitions 3.16 and 3.22, we have*

$$\#\{g \in G : \|g\|_{\mathfrak{F}_G} \leqslant R) \asymp \mathbf{F}_{G,\mathfrak{F}_G}(R).$$

Furthermore, there exist a finite sequence $\theta_i \in \Sigma_G \setminus \Sigma_G(0)$, $1 \leqslant i \leqslant Q$, and a constant C such that, for any $R \geqslant 1$ and any element g of G satisfying $\|g\|_{\mathfrak{F}_G} \leqslant R$, there are elements $g_i \in G$, $0 \leqslant i \leqslant Q$, and $x_i \in \mathbb{Z}$, $1 \leqslant i \leqslant Q$, such that

$$g = g_0 \prod_{i=1}^{Q} \theta_j^{x_j} g_i$$

with

$$|x_j| \leqslant CF_{G,\theta_j}(R),\ 1 \leqslant j \leqslant Q \ \text{ and } \ |g_i|_{S_0}^2 \leqslant CR,\ 0 \leqslant i \leqslant Q.$$

Here $|a|_{S_0}$ is the word length of the element $a \in G$ over the generating set S_0.

Proof. — The volume estimate follows from Theorem 2.14 applied to $(N, \Sigma_N, \mathfrak{F}_N)$ together with Theorem 3.21. Theorem 2.14 also gives an explicit description of the function $\mathbf{F}_{N,\mathfrak{F}_N} = \mathbf{F}_{G,\mathfrak{F}_G}$.

Similarly, to prove the product decomposition of an element $g \in G$ stated in the corollary, we use the fact that any such g can be written hu with

$h \in N$ and $u \in \{u_1, \dots, u_k\}$. Obviously $\|h\|_{\mathfrak{F}_G} \leqslant C_1 \|g\|_{\mathfrak{F}_G}$ and, by Theorem 3.21, $\|h\|_{\mathfrak{F}_N} \leqslant C_2 \|h\|_{\mathfrak{F}_G}$. It then suffices to repeat the argument used in part (2) of the proof of Theorem 3.21 to obtain the desired product decomposition. □

Corollary 3.24. — *Let G be a group having polynomial volume growth, $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$ and H_i, ϕ_i, Φ_i as in Definitions 1.5 and 3.6. We have*

$$\Lambda_{2,G,\mu}(v) \succeq 1/\mathbf{F}^{-1}(v), \quad v \geqslant 1 \quad \text{and} \quad \mu^{(n)}(e) \preceq 1/\mathbf{F}(n), \quad n = 1, 2, \dots,$$

where $\mathbf{F} = \mathbf{F}_{G,\mathfrak{F}_G}$.

Proof. — Using both parts of Corollary 3.23, the stated result follows from Proposition 3.8. □

4. Lower bounds on return probabilities

In this section, we derive sharp lower bounds for the return probability in the case of random walks driven by measures $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$, when G has polynomial volume growth. According to well-known results recalled with pointers to the literature in Section 1.3, if $\mathbf{F}$ is a given regularly varying function of positive index, we have the equivalence

$$\forall\, v \geqslant 1,\ \Lambda_{2,G,\mu}(v) \preceq 1/\mathbf{F}^{-1}(v) \iff \forall\, n = 1, 2 \dots,\ \mu^{(2n)}(e) \succeq 1/\mathbf{F}(n).$$

Accordingly, one of the simplest methods to obtain a lower bound on $\mu^{(2n)}(e)$ is to find a family of test functions, ζ_R, $R \geqslant 1$, supported on a set of volume $\mathbf{F}(R)$ and such that

$$\frac{\mathcal{E}_{G,\mu}(\zeta_R, \zeta_R)}{\|\zeta_R\|_2^2} \preceq 1/R.$$

As μ is a convex combination of measures which are either finitely supported (μ_0) or supported by a subgroup H_i and associated with a regularly varying function ϕ_i, the following two lemmas will be exactly what we need. Lemma 4.1 will be used in the proof of Lemma 4.2 to control test functions constructed with the help of $\|\cdot\|_{\mathfrak{F}}$.

Lemma 4.1. — *Let G be a group of polynomial volume growth. Let $\Sigma = (s_1, \dots, s_q)$ be a generating tuple of distinct elements and let $\mathfrak{F} = \{F_s : s \in \Sigma\}$ be a weight function system so that each F_s and $F_s^{-1} \in \mathcal{C}^\infty([0,\infty))$, and F_s is an increasing function vanishing at 0 and of smooth regular variation of positive index less than 1 at infinity. Fix A_1, A_2 and $A_3 \geqslant 1$. For any g, $h \in G$, let $w \in \bigcup_{p=0}^\infty (\Sigma \cup \Sigma^{-1})^p$ and $R \geqslant 1$ with $h = w$ in G,*

$$\|g\|_{\mathfrak{F}} \leqslant A_1 R, \quad \|h\|_{\mathfrak{F}} \leqslant A_2 R \quad \text{and} \quad F_\sigma^{-1}(\deg_\sigma(w)) \leqslant A_3 R.$$

Then, we have

$$\big|\|gh\|_{\mathfrak{F}} - \|g\|_{\mathfrak{F}}\big| \leqslant C_{\mathfrak{F}}(A_1, A_2, A_3) \max_{\sigma\in\Sigma} \left\{ \frac{R}{F_\sigma(R)} \deg_\sigma(w) \right\}.$$

Proof. — Each F_s^{-1} is smooth at 0 and of smooth regular variation of degree greater than 1, so that the derivative $[F_s^{-1}]'$ of F_s^{-1} satisfies

$$\sup_{[0,T]} \left\{[F_s^{-1}]'\right\} \leqslant C_s \frac{F_s^{-1}(T)}{T}, \quad T > 0.$$

Let

$$C_{\mathfrak{F}}(A_1, A_3) = \max\left\{ C_s \frac{F_s(t)\cdot F_s^{-1}(F_s(A_1 t) + F_s(A_3 t))}{t(F_s(A_1 t) + F_s(A_3 t))} : s \in \Sigma, t > 0 \right\}.$$

If $\|gh\|_{\mathfrak{F}} > \|g\|_{\mathfrak{F}}$, let $w_g \in \bigcup_{p=0}^{\infty} (\Sigma \cup \Sigma^{-1})^p$ be such that

$$\|g\|_{\mathfrak{F}} = \max_{\sigma\in\Sigma} \{F_\sigma^{-1}(\deg_\sigma(w_g)\}.$$

Set $x_\sigma = \deg_\sigma(w_g)$ and $y_\sigma = \deg_\sigma(w)$, where w is as in the statement of the lemma. We have, by Definition 2.4,

$$\begin{aligned}
\big|\|gh\|_{\mathfrak{F}} - \|g\|_{\mathfrak{F}}\big| &= \|gh\|_{\mathfrak{F}} - \|g\|_{\mathfrak{F}} \\
&\leqslant \max_{\sigma\in\Sigma} \left\{F_\sigma^{-1}(x_\sigma + y_\sigma) - F_\sigma^{-1}(x_\sigma)\right\} \\
&\leqslant C_\sigma \frac{F_\sigma^{-1}(F_\sigma(A_1 R) + F_\sigma(A_2 R))}{F_\sigma(A_1 R) + F_\sigma(A_2 R)} y_\sigma \\
&\leqslant C_{\mathfrak{F}}(A_1, A_3) \max_{\sigma\in\Sigma} \left\{ \left(\frac{R}{F_\sigma(R)}\right) y_\sigma \right\}.
\end{aligned}$$

If, instead $\|gh\|_{\mathfrak{F}} < \|g\|_{\mathfrak{F}}$, run the same argument with $g' = gh$ and $h' = h^{-1}$ (and w replace by w^{-1}, the formal word inverse of w). This gives the same bound with A_1 replaced by $A_1 + A_2$. □

Lemma 4.2. — *Let G be a group of polynomial volume growth. Let $\Sigma = (s_1, \ldots, s_q)$ be a generating tuple of distinct elements, and let $\mathfrak{F} = \{F_s : s \in \Sigma\}$ be a weight function system satisfying* (2.1) *with F_s and $F_s^{-1} \in \mathcal{C}^1([0,\infty))$ for $s \in \Sigma$. Assume that F_s is an increasing function, vanishing at 0, and of smooth regular variation of positive index less than 1 at infinity. Let $F_\star = \min\{F_s : s \in \Sigma\}$ be the smallest of the weight functions. Let H be a subgroup of G of volume growth of index d with word length $|\cdot|$. Let ϕ be a positive increasing function on $[0,\infty)$ which is also regularly varying function of positive index, and set $\Phi(t) = t^2/\int_0^t \frac{2s}{\phi(s)}\mathrm{d}s$ for $t \geqslant 1$. Let μ_H be a probability measure supported on H and of the form $\mu_H(h) \asymp [(1+|h|)^d \phi(1+|h|)]^{-1}$. Assume that there exists A such that, for* *any* $h \in H$*, there is a word* $\theta := \theta_h \in \bigcup_{p=0}^{\infty}(\Sigma \cup \Sigma^{-1})^p$ *such that* $h = \theta$ *in* G *and, for each* $s \in \Sigma$,

- *either* $\forall\, t \geqslant 1$*, it holds that* $AF_s \circ F_\star^{-1}(\sqrt{t}) \geqslant \Phi^{-1}(t)$*, in which case* $\deg_s(\theta) \leqslant A|h|$,
- *or* $\deg_s(\theta) \leqslant A$.

Under these hypotheses the function $g \mapsto \xi_R(g) = (R - \|g\|_{\mathfrak{F}})_+$ *satisfies*

$$\frac{\mathcal{E}_{G,\mu_H}(\zeta_R, \zeta_R)}{\|\zeta_R\|_2^2} \leqslant \frac{C}{F_\star(R)^2}.$$

Also, if μ_0 *is a symmetric finitely supported measure,*

$$\frac{\mathcal{E}_{G,\mu_0}(\zeta_R, \zeta_R)}{\|\zeta_R\|_2^2} \leqslant \frac{C}{F_\star(R)^2}.$$

Proof.

(1). — We first consider μ_H. Let $W(R)$ be the cardinality of $\{g : \|g\|_{\mathfrak{F}} \leqslant R\}$. Since ζ_R is at least $R/2$ over $\{\|g\|_{\mathfrak{F}} \leqslant R/2\}$, we have $\|\zeta_R\|_2^2 \asymp R^2 W(R)$. We now bound

$$\mathcal{E}_{G,\mu_H}(\zeta_R, \zeta_R) = \frac{1}{2}\sum_{g,h} |\zeta_R(gh) - \zeta_R(g)|^2 \mu_H(h)$$

from above. Let

$$\Omega = \{(g,h) \in G \times H : \zeta_R(gh) + \zeta_R(g) > 0\}.$$

Obviously, the sum defining $\mathcal{E}_{G,\mu}(\zeta_R, \zeta_R)$ can be restricted to Ω and, for any given h,

$$\#\{g \in G : (g,h) \in \Omega\} \leqslant 2W(R).$$

We split this sum into two parts depending on whether or not $|h| \geqslant \rho$ where ρ will be chosen later, and write

$$\begin{aligned}\sum_{(g,h)\in\Omega : |h| \geqslant \rho} |\zeta_R(gh) - \zeta_R(g)|^2 \mu_H(h) &\leqslant 2R^2 W(R) \left(\sum_{|h|\geqslant\rho} \mu_H(h)\right) \\ &\leqslant CR^2 W(R)/\phi(\rho) \leqslant CR^2 W(R)/\Phi(\rho),\end{aligned}$$

where we have used the simple fact that $\phi \geqslant \Phi$ in the last inequality, see Remark 3.7. On the other hand, consider pairs $(g,h) \in \Omega$ with $|h| < \rho$ with $\rho = \Phi^{-1}\left(F_\star(R)^2\right)$. This choice of ρ ensures that, for any $\sigma \in \Sigma$ such that $AF_\sigma \circ F_\star^{-1}(\sqrt{t}) \geqslant \Phi^{-1}(t)$ for all $t > 1$, we have $\deg_\sigma(\theta) \leqslant A\rho$. Thus,

$$F_\sigma^{-1}(\deg_\sigma(\theta)) \leqslant A' F_\star^{-1}\left(\sqrt{\Phi(\rho)}\right) \leqslant A'R$$

and

$$\frac{R}{F_\sigma(R)} \leqslant A' \frac{R}{\Phi^{-1}(F_\star(R)^2)}.$$

Since $(g,h)\in\Omega$, it follows that all $\|h\|_{\mathfrak{F}}$, $\|g\|_{\mathfrak{F}}$ and $\|gh\|_{\mathfrak{F}}$ are smaller than $A''R$ for some constant $A''>0$. Combining all the estimates above with the assumptions of the lemma (in particular, $F_\star$ is the smallest of the weight functions) and Lemma 4.1 yields that

$$|\zeta_R(gh)-\zeta_R(g)| \leqslant |\|gh\|_{\mathfrak{F}}-\|g\|_{\mathfrak{F}}| \leqslant C_1\left(\frac{R}{F_\star(R)}+\frac{R}{\Phi^{-1}(F_\star(R)^2)}|h|\right).$$

It follows that

$$\begin{aligned}
&\sum_{(g,h)\in\Omega:|h|<\rho}|\zeta_R(gh)-\zeta_R(g)|^2\mu_H(h)\\
&\qquad\leqslant 4C_1^2W(R)\left(\sum_{|h|<\rho}\left(\frac{R^2}{F_\star(R)^2}+\frac{R^2}{[\Phi^{-1}(F_\star(R)^2)]^2}|h|^2\right)\mu_H(h)\right)\\
&\qquad\leqslant 4C_1^2R^2W(R)\left(\frac{1}{F_\star(R)^2}+\frac{1}{[\Phi^{-1}(F_\star(R)^2)]^2}\sum_{|h|\leqslant\rho}|h|^2\mu_H(h)\right)\\
&\qquad\leqslant C_2R^2W(R)\left(\frac{1}{F_\star(R)^2}+\frac{\rho^2}{\Phi(\rho)[\Phi^{-1}(F_\star(R)^2)]^2}\right)\\
&\qquad\leqslant \frac{C_2R^2W(R)}{F_\star(R)^2}.
\end{aligned}$$

(2). — We next consider μ_0. In the case of μ_0, for $(g,h)\in\Omega$, we have $|h|\leqslant C$ (because h is in the support of μ_0), and both $\|g\|_{\mathfrak{F}}$ and $\|gh\|_{\mathfrak{F}}$ are smaller than $(1+C)R$ (because $(g,h)\in\Omega$, h is in the support of μ_0 and $R\geqslant 1$) for some constant $C>0$. Thus, it follows from Lemma 4.1 that

$$|\zeta_R(gh)-\zeta_R(g)| \leqslant |\|gh\|_{\mathfrak{F}}-\|g\|_{\mathfrak{F}}| \leqslant C_1\left(\frac{R}{F_\star(R)}\right).$$

Using this in computing $\mathcal{E}_{G,\mu_0}(\zeta_R,\zeta_R)$ yields

$$\begin{aligned}
\sum_{(g,h)\in\Omega}|\zeta_R(gh)-\zeta_R(g)|^2\mu_0(h) &\leqslant 4C_1^2W(R)\sum_h\frac{R^2}{F_\star(R)^2}\mu_0(h)\\
&\leqslant \frac{2C_1R^2W(R)}{F_\star(R)^2}.
\end{aligned}$$

Combining all the estimates above yields the desired conclusion. ☐

THEOREM 4.3. — *Let G be a group having polynomial volume growth and μ be in $\mathcal{P}_{\preceq}(G,\mathrm{reg})$. Let H_i,ϕ_i,Φ_i as in Definitions 1.5 and 3.6. Referring to the notation of Definitions 3.16, 3.18 and 3.22, set $\mathbf{F}=\mathbf{F}_{G,\mathfrak{F}_G}$. Then, we have*

$$\Lambda_{2,G,\mu}(v)\asymp 1/\mathbf{F}^{-1}(v),\quad \mu^{(n)}(e)\asymp 1/\mathbf{F}(n).$$

Proof. — By the results mentioned in Section 1.3, it suffices to prove that

$$\Lambda_{2,G,\mu}(v) \asymp 1/\mathbf{F}^{-1}(v).$$

Corollary 3.24 provides the lower bound $\Lambda_{2,G,\mu}(v) \succeq 1/\mathbf{F}^{-1}(v)$ so it suffices to prove the upper bound. Consider the weight function system associated with $(\Sigma_G, \mathfrak{F}_G)$, where $\mathfrak{F}_G = \{F_s, s \in \Sigma_G\}$ is as in Definition 3.16. Let $\omega(s)$ be the positive index of the increasing regularly varying function F_s. By Proposition 2.11, there are functions $F_{1,s} \in \mathcal{C}^1([0,\infty))$ for all $s \in \Sigma_G$, positive, increasing and of smooth regular variation of index strictly less than 1 such that $F_{1,s}^{-1} \in \mathcal{C}^1([0,\infty))$ and $F_{1,s} \asymp F_s \circ F^{-1}$, where $F(t) = (1+t)^{\omega^*} - 1$ and ω^* is chosen so that $\omega^* > \max\{\omega(s) : s \in \Sigma_G\}$. These functions satisfy (2.3). Let $(\Sigma_G, \mathfrak{F}_{G,1})$ be the associated weight function system and recall that, by construction,

$$\|g\|_{\mathfrak{F}_G} \asymp \|g\|_{\mathfrak{F}_{G,1}}^{1/\omega^*}.$$

Let $\zeta_R(g) = (R - \|g\|_{\mathfrak{F}_{G,1}})_+$. For each H_i, μ_i, $1 \leqslant i \leqslant k$, and $s \in S_i$ (see Definition 3.16), we have

$$F_{1,s} \circ F_{1,\star}^{-1}(\sqrt{t}) \asymp F_s \circ F_\star^{-1}(\sqrt{t}) \geqslant (1/A)F_s(t) \geqslant (1/A)\Phi_i^{-1}(t).$$

Here we have used the simple fact that, because of the definition of the system $\mathfrak{F}_G$ based on the function Φ_i, $F_\star(t) \asymp \min\{t, \sqrt{t}\}$ (See Remark 3.7 and the definition of Φ_0). Further, since $N_i = H_i \cap N$ is of finite index in H_i (See Remark 3.17). Hence, for any $h \in H_i$, there is a word $\theta := \theta_h \in \bigcup_{p=0}^\infty (\Sigma_G \cup \Sigma_G^{-1})^p$ such that $h = \theta$ in G and, for each $s \in \Sigma_G \setminus S_i$, $\deg_s(\theta) \leqslant A$.

It follows that we can apply Lemma 4.2 to μ_i and $\mathfrak{F}_{1,G}$. This gives that for $R \geqslant 1$,

$$\frac{\mathcal{E}_{G,\mu_i}(\zeta_R, \zeta_R)}{\|\zeta_R\|_2^2} \leqslant \frac{C}{F_{1,\star}(R)^2}$$

with $F_{1,\star}(t) \asymp t^{1/(2\omega^*)}$ for all $t \geqslant 1$; that is,

$$\frac{\mathcal{E}_{G,\mu_i}(\zeta_R, \zeta_R)}{\|\zeta_R\|_2^2} \leqslant \frac{C}{R^{1/\omega^*}} \quad \text{for every } R \geqslant 1.$$

This holds for all $i = 1, \dots, k$, and also for μ_0. Hence it also holds for the convex combination $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$; that is,

$$\frac{\mathcal{E}_{G,\mu}(\zeta_R, \zeta_R)}{\|\zeta_R\|_2^2} \leqslant \frac{C}{R^{1/\omega^*}} \quad \text{for every } R \geqslant 1.$$

Now, ζ_R is supported in $\{\|g\|_{\mathfrak{F}_{G,1}} < R\}$ which has volume

$$\#\{\|g\|_{\mathfrak{F}_{G,1}} \leqslant R\} \asymp \#\{\|g\|_{\mathfrak{F}_G}^{\omega^*} \leqslant R\} \asymp \mathbf{F}(R^{1/\omega^*}),$$

thanks to Corollary 3.23. It follows that

$$\forall\, v \geqslant 1, \quad \Lambda_{2,G,\mu}(v) \preceq \frac{1}{\mathbf{F}^{-1}(v)}.$$

The proof is complete. □

Remark 4.4. — Even in the simplest case when G is nilpotent, each H_i is a one parameter subgroup $H_i = \langle s_i \rangle \subseteq G$ and the measure μ_i satisfies $\mu_i(h) \asymp (1+|h|)^{-1-\alpha_i}$ with $\alpha_i > 0$, $h = s_i^n$ and $n \in \mathbb{Z}$, (i.e., the basic case studied in [22]), Theorem 4.3 provides new results, since [22] gives complete two sided bounds only in the case when all α_i are in the interval $(0,2)$ or all α_i are equal to 2.

Example 4.5. — We return again to the Heisenberg example (Example 2.10), keeping the same notation. We let $H_i = \langle s_i \rangle$ (the subgroup generated by s_i) and $\phi_i(t) = (1+t)^{\alpha_i}$ with $\alpha_i > 0$ for $1 \leqslant i \leqslant 3$. Let $\mu = \sum_{i=1}^{3} \mu_i$. Obviously, $d_i = 1$ and, for $t \geqslant 1$,

$$\Phi_i(t) \asymp \begin{cases} t^{\alpha_i} & \text{if } \alpha_i \in (0,2), \\ t^2/\log(1+t) & \text{if } \alpha_i = 2, \\ t^2 & \text{if } \alpha_i > 2. \end{cases}$$

Introduce the two-coordinate weight system

$$\omega_i = (\omega_{i,1}, \omega_{i,2}) = \begin{cases} (1/\widetilde{\alpha}_i, 0) & \text{if } \alpha_i \neq 2, \\ (1/2, 1/2) & \text{if } \alpha_i = 2. \end{cases}$$

Note that the natural order over the functions $F_{s_i} = \Phi_i^{-1}$ (in a neighborhood of infinity) is the same as the lexicographical order over the weights $\omega_i = (\omega_{i,1}, \omega_{i,2})$ and that

$$\mathbf{F}(t) \asymp \begin{cases} t^{\omega_{1,1}+\omega_{2,1}+\omega_{3,1}} [\log(1+t)]^{\omega_{1,2}+\omega_{2,2}+\omega_{3,2}} & \text{if } \omega_3 \geqslant \omega_1 + \omega_2, \\ t^{2(\omega_{1,1}+\omega_{2,1})} [\log(1+t)]^{2(\omega_{1,2}+\omega_{2,2})} & \text{if } \omega_3 \leqslant \omega_1 + \omega_2. \end{cases}$$

Theorem 4.3 tells us that

$$\Lambda_{2,\mathbb{H}(3,\mathbb{Z}),\mu}(v) \asymp \frac{1}{\mathbf{F}^{-1}(v)}, \quad \mu^{(n)}(e) \asymp \frac{1}{\mathbf{F}(n)}.$$

Next consider the case when $\phi_i(t) = (1+t)^2[\log(2+t)]^{\beta_i}$ with $\beta_i \in \mathbb{R}$ for $i = 1,2$, and $\phi_3(t) = (1+t)[\log(2+t)]^{\beta_3}$ with $\beta_3 \in \mathbb{R}$. In this case, for $t \geqslant 1$ and $i = 1,2$,

$$\Phi_i(t) \asymp \begin{cases} t^2 \log(2+t)]^{\beta_i - 1} & \text{if } \beta_i < 1, \\ t^2/\log\log(4+t) & \text{if } \beta_i = 1, \\ t^2 & \text{if } \beta_i > 1, \end{cases}$$

and

$$\forall\, t \geqslant 1, \quad \Phi_3(t) \asymp t[\log(2+t)]^{\beta_3}.$$

We need to compare $\Phi_1^{-1}\Phi_2^{-1}$ to Φ_3^{-1} over $(1,\infty)$.

Assume first that $\beta_1,\beta_2 \in (-\infty,1)$ so that

$$\Phi_1^{-1}(t)\Phi_2^{-1}(t) \asymp t[\log(2+t)]^{-(\beta_1+\beta_2-2)/2}$$

and $\Phi_3^{-1}(t) \asymp t[\log(2+t)]^{-\beta_3}$. When $\frac{1}{2}(\beta_1+\beta_2-2) > \beta_3$, the function Φ_3^{-1} dominates and the volume function $\mathbf{F}$ associated with the weight function system $F_{s_i} = \Phi_i^{-1}$ is the product $\mathbf{F} = \Phi_1^{-1}\Phi_2^{-1}\Phi_3^{-1}$. If, instead, $\frac{1}{2}(\beta_1+\beta_2-2) \leqslant \beta_3$, then we have $\mathbf{F} = (\Phi_1^{-1}\Phi_2^{-1})^2$.

In the case when $\beta_1,\beta_2 \geqslant 1$, one can check that

$$\mathbf{F} \asymp \begin{cases} \Phi_1^{-1}\Phi_2^{-1}\Phi_3^{-1} & \text{if } \beta_3 < 0, \\ (\Phi_1^{-1}\Phi_2^{-1})^2 & \text{if } \beta_3 \geqslant 0. \end{cases}$$

In all cases (including those not discussed explicitly above),

$$\mathbf{F} = \Phi_1^{-1}\Phi_2^{-1}\max\{\Phi_3^{-1}, \Phi_1^{-1}\Phi_2^{-1}\}.$$

The following examples illustrate some of the subtleties related to the treatment of groups of polynomial growth that are not nilpotent.

Example 4.6 (Infinite dihedral group). — Recall that the infinite dihedral group $\mathbf{D}$ can be presented as $\langle u,v : u^2, v^2\rangle$. This means it is the quotient of the free group $\mathbf{F}(\mathbf{u},\mathbf{v})$ with two generators by the normal subgroup of $\mathbf{F}(\mathbf{u},\mathbf{v})$ generated by $\mathbf{u}^2,\mathbf{v}^2$ (and all the conjugates of these). Obviously, the images u,v of $\mathbf{u},\mathbf{v}$ in the quotient, $\mathbf{D}$, satisfy $u^2=v^2=e$. Since $\{u,v\}$ generates $\mathbf{D}$, any element g can be written uniquely as $g=(uv)^n u^{\varepsilon}$ with $n\in\mathbb{Z}$ and $\varepsilon\in\{0,1\}$ and also as $g=(uv)^m v^{\eta}$ with $m\in\mathbb{Z}$ and $\eta\in\{0,1\}$. For instance $vu=(uv)^{-1}$. The Cayley graph looks like this:

Figure 4.1. Dihedral group: u in blue, v in red.

This group is not nilpotent because the commutator of $(uv)^n$ by u is $(uv)^{2n}$. Let $z=uv$ and $Z=\langle z\rangle$, the subgroup generated by z. Since z has infinite order, this is a copy of $\mathbb{Z}$. It is also a normal subgroup with quotient the group with two elements $\{0,1\}$. From what we said before, any element g can be written uniquely $g=t^n u^{\varepsilon}$ and this gives a description of $\mathbf{D}$ as the semi-direct product

$$\mathbf{D} = Z \rtimes \langle u\rangle.$$

Obviously (from the picture) $\mathbf{D}$ has volume growth of degree 1. The subgroup Z is abelian (hence, nilpotent) with finite index. Even so this example is relatively trivial, it already illustrates significant differences with the nilpotent case.

For instance, it is easy to see that Theorem 3.9 does not apply as stated to this group. Indeed, take $\mu = (1/2)(\mu_1 + \mu_2) \in \mathcal{P}(\mathbf{D}, \text{reg})$ where μ_1, μ_2 are associated to $\phi_1(s) = s^{\alpha_1}$, $\phi_2(s) = s^{\alpha_2}$ on the subgroups $\langle u\rangle, \langle v\rangle$, respectively. Assume that $0 < \alpha_1 \leqslant \alpha_2 < 2$.

On the one hand, since u, v have order 2, it is obvious that this is just a weighted version of simple random walk on Δ with generators $\{u, v\}$. On the other hand, we can consider the norm $\|\cdot\|$ associated with the system $\{u, v\}$, $F_u(s) = (1+s)^{1/\alpha_1} - 1$, $F_u(v) = (1+s)^{1/\alpha_2} - 1$. That would be the norm used in Theorem 3.9 if Δ was nilpotent. Obviously, the element $g = (uv)^n u^{\varepsilon}$ as norm $\|g\| \asymp (|\varepsilon| + |n|)^{\alpha_2}$. If they where applicable, Theorems 2.10 and 3.9 would say $\mu^{(2n)}(e) \asymp n^{-1/\alpha_2}$ whereas, obviously, $\mu^{(2n)}(e) \asymp n^{-1/2}$.

Why is it the case that such situation does not appear in the nilpotent case treated in Theorems 2.10 and 3.9? The reason is algebraic, namely, in any finitely generated nilpotent group, the torsion elements form a *finite* normal subgroup and torsion elements cannot play a significant role in computing the type of length considered in Theorem 2.10. Obviously, the dihedral group $\mathbf{D}$ is very different from this view point.

Example 4.7. — The goal of this example is to construct a multi-dimensional version of the dihedral group above which will illustrate how our results apply to groups of polynomial growth that are not nilpotent. Let Δ be generated by s, s', t, t' which are all of order two (involutions) and which satisfy the following commutation relations

$$tt' = t't,\ ss' = s's,\ s't' = t's',\ stst' = st'st,\ sts't = s'tst,\ st's't = s'tst'.$$

Technically, this means that Δ is the quotient of the free group generated by four letters s, s', t, t' by the normal subgroup generated by the indicated relations which are all commutation relations between two elements: t and t' commute ($[t, t'] = 1$), s and s' commute ($[s, s'] = 1$), s' and t' commute ($[s', t'] = 1$), st and st' commutes ($[st, st'] = 1$), st and $s't$ commute ($[st, s't] = 1$), and st' and $s't$ commutes ($[st', s't] = 1$). The next lemma gives a concrete description of this group. But in the present case it is possible to obtain a good picture of what happens.

There are three homomorphisms onto the Dihedral group $\mathbf{D}$, ψ_1, ψ_2, ψ_3, obtained by sending the pair (s', t) (resp. (s, t), (s, t')) to (u, v) and all other generators to the identity. For instance ψ_1 is obtained by looking at the homomorphism ψ from the free group $\mathbf{F}_4$ (with four generators $\mathbf{t}, \mathbf{t}, \mathbf{s}, \mathbf{s}'$) onto $\mathbf{D}$ which send $\mathbf{s}'$ to u, $\mathbf{t}$ to v, and the other generators to the identity. By construction, the kernel of this homomorphism contains the defining kernel N_Δ of the group $\Delta = \mathbf{F}_4/N_\Delta$. It follows that ψ descends to an surjective homomorphism $\psi_1 : \Delta \longrightarrow \mathbf{D}$. We get one such homomorphism for each pair (s', t), (s, t) and (s, t'). This proves that the commuting elements st, st' and $s't$ each have infinite order. There are also three copies of the dihedral group $\langle s', t\rangle$, $\langle s, t\rangle$ and $\langle s, t'\rangle$ in Δ.

Lemma 4.8. — *Any element g of Δ can be represented uniquely in the form*

$$g = (s't)^{n_1}(st)^{n_2}(st')^{n_3}s^{\varepsilon}, \;\; \varepsilon \in \{0, 1\}, \;\; n_1, n_2, n_3 \in \mathbb{Z}$$

and the commuting elements $s't, st, st'$ generate a copy of $\mathbb{Z}^3$ as a subgroup of Δ.

Proof. — Existence of $(n_1, n_2, n_3, \varepsilon)$ can be proved by induction on the length of words on the generators s, s', t, t'. For this, it suffices to check that if g has the desired form, then we can also write gs, gs', gt, gt' in the desired form. This is easy for s (!) and t'. If $\varepsilon = 1$, it is also easy to move t through since st commutes with st'. If $\varepsilon = 0$, move t by writing $t = (ts)s$. Finally, s' commutes with s and t' so that

$$\begin{aligned} gs' &= (s't)^{n_1}(st)^{n_2}s'(st')^{n_3}s^{\varepsilon} = (s't)^{n_1}(st)^{n_2}(s't)(ts)s(st')^{n_3}s^{\varepsilon} \\ &= (s't)^{n_1+1}(st)^{n_2-1}(st')^{n_3'}s^{\varepsilon'}, \end{aligned}$$

where, on the right, we have used the fact that $s(st')^{n_3}s^{\varepsilon} = (st')^{n_3'}s^{\varepsilon'}$ since any element in $< s, t' >$ can be written in that form (of course, one can also compute explicitly (n_3', ε') as a function of (n_3, ε)).

To prove uniqueness, assume an element g in Δ can be written in two different ways $(n_1, n_2, n_3, \varepsilon)$ and $(n_1', n_2', n_3', \varepsilon')$ or, equivalently,

$$s^{\varepsilon'}(s't)^{n_1-n_1'}(st)^{n_2-n_2'}(st')^{n_3-n_3'}s^{\varepsilon} = 1 \text{ in } \Delta.$$

We consider the images of this by ψ_1, ψ_2, ψ_3. It will be useful to note that in $\mathbf{D} = \langle u, v\rangle$, $u^{\eta}(uv)^k u^{\eta'} = 1$ implies $k = 0$ and also $v^{\eta}(uv)^k v^{\eta'} = 1$ implies $k = 0$. Taking the image by ψ_1 gives

$$v^{n_2'}(uv)^{n_1-n_1'}v^{n_2} = 1 \text{ in the dihedral group.}$$

This implies $n_1 = n_1'$. Taking the image by ψ_2 gives

$$u^{\varepsilon'}(uv)^{n_2-n_2'}u^{n_3-n_3'+\varepsilon} = 1$$

in the dihedral group which implies $n_2 = n_2'$. Finally, taking the image by ψ_3 gives

$$u^{\varepsilon' + n_2 - n_2'}(uv)^{n_3 - n_3'}u^{\varepsilon} = 1$$

in the dihedral group which implies $n_3 = n_3'$. Obviously, we also must have $\varepsilon = \varepsilon'$. □

Next we show that this writing is (almost) minimizing the degree of each of the generator s, t, s', t' in a word representing g. For this, we first need to observe that for each element w in the dihedral group $\langle u, v : u^2, v^2\rangle$, there exists a smallest $n \geqslant 0$ such that $w = w_w$ with $w_w = (uv)^n$ or $w_w = (uv)^{-n}$ or $w_w = (uv)^n u$ or $w_w = (uv)^{-n} v$. Moreover, if w is any word over $\{u, v\}$ such that $w = w$ then we have

$$\deg_x(w) \geqslant \deg_x(w_w), \quad x = u, v. \tag{4.1}$$

This fact is nothing difficult. Each element of $\mathbf{D}$ is, uniquely and minimally, a product iterating between u and v. The four cases mentioned above are exactly the cases when w starts with u and finishes with v, starts with v and finishes with u, starts with u and finishes with u, and starts with v and finishes with v.

Now, for any word w over $\{s, t, s', t'\}$ representing $g = (n_1, n_2, n_3, \varepsilon)$ and $i = 1, 2, 3$, let $\overline{\psi}_i(w)$ be the word obtained in an obvious way by cancelling those letters that are sent to the identity by ψ_i. Note that the word $\overline{\psi}_i(w)$ represents the element $\psi_i(g)$ in the dihedral group. For $x = s, t, s', t'$, we have

$$\deg_x(w) \geqslant \deg_{\psi_i(x)}(\overline{\psi}_i(w)) \geqslant |n_i| - 1$$

because, for instance, $\psi_2(g) = t^{n_1}(st)^{n_2}s^{n_3+\varepsilon} = t^{\eta_1}(st)^{n_2}s^{\eta_2}$ with $\eta_1, \eta_2 \in \{0, 1\}$. Compare with (4.1) to obtain the desired result. Note that the previous inequality is optimal since in order to write t in the form $(n_1, n_2, n_3, \varepsilon)$ we have to write $t = tss = (0, -1, 0, 1)$.

Let

$$\theta_1 = s't, \ \theta_2 = st, \ \theta_3 = st'$$

be the generators of the subgroup $\mathbb{Z}^3 = \langle \theta_1, \theta_2, \theta_3\rangle$ in Δ. Let

$$H_1 = \langle s', t\rangle, \ H_2 = \langle s, t\rangle, \ H_3 = \langle s, t'\rangle$$

be the three copies of the dihedral group. Pick

$$\alpha_1, \alpha_2, \alpha_3 \in (0, 2) \text{ and } \phi_i(t) = (1+t)^{\alpha_i}.$$

Let μ_i be a measure supported on H_i with $\mu_i(h) \asymp [(1+|h|_i)\phi_i(1+|h|_i)]^{-1}$ and

$$\mu = (1/3)\sum_1^2 \mu_i \in \mathcal{P}(\Delta, \mathrm{reg}).$$

If Δ were nilpotent (it is not!), we would obtain an appropriate geometry by setting $\Sigma = \{s, s', t, t'\}$, $\omega(\sigma) = \max\{1/\alpha_i : i \in \{j : \sigma \in H_i\}\}$, and $F_\sigma(t) = (1+t)^{\omega(\sigma)}-1$. (See Definition 2.4). Here we have picked a generating set for each H_i, assigned the weight $1/\alpha_i$ to the generators coming from H_i, reduced to Σ to avoid repetition and picked the highest available weight for each $\sigma \in \Sigma$. We can compute the associated geometry. We have

$$\omega(s) = \max\{1/\alpha_2, 1/\alpha_3\}, \;\; \omega(t) = \max\{1/\alpha_1, 1/\alpha_2\},$$

and $\omega(s') = 1/\alpha_1$, $\omega(t') = 1/\alpha_3$. Set

$$\begin{aligned}
\omega_1 &= \min\{\omega(s'), \omega(t)\} = 1/\alpha_1,\\
\omega_2 &= \min\{\omega(s), \omega(t)\} = \min\{\max\{1/\alpha_2, 1/\alpha_3\}, \max\{1/\alpha_1, 1/\alpha_2\}\},\\
\omega_3 &= \min\{\omega(s), \omega(t')\} = 1/\alpha_3.
\end{aligned}$$

Note that $\omega_2 = 1/\alpha_2$ if and only if $\alpha_2 \leqslant \max\{\alpha_1, \alpha_3\}$. By inspection, using the lower bound on degrees derived above in terms of the $|n_i|$, we find that if g is represented by $(n_1, n_2, n_3, \varepsilon)$, we have

$$\|g\|_{\mathfrak{F}} \asymp \max\{|n_1|^{1/\omega_1}, |n_2|^{1/\omega_2}, |n_3|^{1/\omega_3}, |\varepsilon|\}.$$

Accordingly,

$$\#\{\|g\|_{\mathfrak{F}} \leqslant R\} \asymp R^{\omega_1+\omega_2+\omega_3}.$$

However, this quasi-norm is not appropriate to study the measure μ.

Now, let us follow carefully Definition 3.16 in order to obtain a quasi-norm that is adapted to the study of the random walk driven by μ. As a normal nilpotent subgroup N with finite index in Δ, we take

$$N = \mathbb{Z}^3 = \langle s't, st, st'\rangle = \langle \theta_1, \theta_2, \theta_3\rangle.$$

We obtain $N_i = N \cap H_i = \langle \theta_i \rangle$ with generating set $S_i = \{\theta_i\}$. We pick a generating set $S_0 = \{s, t, s', t'\}$ of Δ. We form $\Sigma_\Delta = \{\theta_1, \theta_2, \theta_3, , s, t, s', t'\}$ equipped with the weight functions system $\mathfrak{F}_\Delta$,

$$F_{\theta_i}(r) = (1+r)^{1/\alpha_i} - 1, \;\; i = 1, 2, 3,$$

and

$$F_s(r) = F_{s'}(r) = F_t(r) = F_{t'}(r) = \min\{r, r^{1/2}\}.$$

In this system, we can compute (this takes a bit of inspection) that if g is represented by $(n_1, n_2, n_3, \varepsilon)$ then

$$\|g\|_{\mathfrak{F}_D} \asymp \{|n_1|^{1/\alpha_1}, |n_2|^{1/\alpha_2}, |n_3|^{1/\alpha_3}, |\varepsilon|\}$$

and $\#\{\|g\|_{\mathfrak{F}_\Delta} \leqslant R\} \asymp R^d$ with $d = \sum_1^3 1/\alpha_i$. Regarding the random walk driven by μ, we obtain

$$\Lambda_{2,\Delta,\mu}(v) \asymp v^{-1/d}, \;\; \mu^{(n)}(e) \asymp n^{-d}.$$

Example 4.9. — Consider semi-direct product $G = \mathbb{Z} \ltimes_\rho \mathbb{Z}^2$ where $k \in \mathbb{Z}$ acts on $\mathbb{Z}^2$ by ρ^k where $\rho(n_1, n_2) = (-n_2, n_1)$ (i.e., ρ is the counter-clockwise rotation of 90°). Concretely, the group elements are elements (k, n_1, n_2) of $\mathbb{Z}^3$ and multiplication is given by

$$(k, n_1, n_2) \cdot (k', n_1', n_2') = (k + k', m_1, m_2),$$

where

$$m = (m_1, m_2) = n + \rho^k(n'), \; n = (n_1, n_2), \; n' = (n_1', n_2').$$

This group is not nilpotent, but it is of polynomial volume growth of degree 3 with normal abelian subgroup $N = 4\mathbb{Z} \times \mathbb{Z}^2$ because $\rho^4 = \mathrm{id}$. Let s, v_1, v_2 be canonical generators of $\mathbb{Z}$ and $\mathbb{Z}^2$. Consider the subgroups

$$H_1 = \langle 4s, v_1\rangle, \quad H_2 = \langle 4s, v_2\rangle,$$

and set $\phi(t) = (1+t)^{\alpha_i}$, $\alpha_i \in (0,2)$, $i = 1, 2$. Let

$$\mu = (1/3)\sum_0^3 \mu_i$$

with $\mu_0 = \frac{1}{2}\mathbf{1}_{\{s,s^{-1}\}}$ and

$$\mu_i((k, n_1, n_2) \asymp (1 + |k| + |n_i|)^{-2-\alpha_i}\mathbf{1}_{H_i}, \quad i = 1, 2.$$

We now describe the geometries $\mathfrak{F}_G$ and $\mathfrak{F}_N$ which are such that

$$\|g\|_{\mathfrak{F}_G} \asymp \|g\|_{\mathfrak{F}_N} \text{ for all } g \in N$$

and are compatible with the random walk driven by μ.

For S_i (a generating set of $N \cap H_i = H_i$), we take $S_i = \{4s, v_i\}$. For S_0, a generating set of G, we take $S_0 = \{s, v_1, v_2\}$. We obtain $\Sigma_G = \{s, 4s, v_1, v_2\}$ with associated F_σ functions $F_s(t) = \min\{t, t^{1/2}\} \asymp (1+t)^{1/2} - 1$ and

$$F_{4s}(t) = (1+t)^{\max\{1/\alpha_1, 1/\alpha_2\}} - 1, \; f_{v_i}(t) = (1+t)^{1/\alpha_i} - 1.$$

Set $\alpha = \min\{\alpha_1, \alpha_2\}$. By inspection, for $g = (k, n_1, n_2)$,

$$\|(k, n_1, n_2)\|_{\mathfrak{F}_G} \asymp \max\{|k|^\alpha, |n_1|^\alpha, |n_2|^\alpha\}.$$

The same power α appears for both n_1, n_2 because of the use of the rotation ρ.

Regarding the associated geometry on N, we take $\Xi_0 = \{4s, v_1, v_2\}$, and

$$\Xi_i = \{4s, v_i, \rho(v_i) = \pm v_j, \rho^2(v_i) = -v_i, \rho^3(v_i) = \mp v_j\},\ i \neq j,\ i, j \in \{1, 2\}.$$

This gives $\Sigma_N = \{4s, v_1, v_2\}$ and

$$F_\sigma(t) = (1+t)^{1/\alpha}, \quad \sigma = 4s, v_1, v_2.$$

For $g \in N$ with $g = (4k, n_1, n_2)$,

$$\|(4k, n_1, n_2)\|_{\mathfrak{F}_N} \asymp \max\{|k|^\alpha, |n_1|^\alpha, |n_2|^\alpha\}.$$

This confirms the fact that $\|g\|_{\mathfrak{F}_G} \asymp \|g\|_{\mathfrak{F}_N}$ when $g \in N$. Regarding μ, we have

$$\Lambda_{2,G,\mu}(v) \asymp v^{-\alpha/3}, \;\; \mu^{(n)}(e) \asymp n^{-3/\alpha}.$$

5. Pseudo-Poincaré inequality and control

We now turn to the extension of some results in [24] to random walk driven by measures $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$, when G has polynomial volume growth. Some of these applications are new even in the case of nilpotent groups.

We recall two definitions from [24] and a general result involving these definitions. Note that these statements involve the more restrictive notion of norm rather than the one of quasi-norm.

Definition 5.1. — *Let μ be a symmetric probability measure on a group G. Let $\|\cdot\|$ be a norm with volume function V. Let $r : (0,\infty) \to (0,\infty)$ be a continuous and increasing function with inverse ρ. Let $(X_n)_0^\infty$ be the random walk on G driven by μ.*

- *We say that μ is $(\|\cdot\|, r)$-controlled if the following properties are satisfied:*
  1. *For all n, $\mu^{(2n)}(e) \asymp V(r(n))^{-1}$.*
  2. *For all $\varepsilon > 0$ there exists $\gamma \in (0,\infty)$ such that for all $n \geqslant 1$,*
  $$\mathbf{P}_e\left(\sup_{0\leqslant k\leqslant n}\{\|X_k\|\} \geqslant \gamma r(n)\right) \leqslant \varepsilon.$$
- *We say that μ is strongly $(\|\cdot\|, r)$-controlled if the following properties are satisfied:*
  1. *There exists $C \in (0,\infty)$ and, for any $\kappa > 0$, there exists $c(\kappa) > 0$ such that, for all $n \geqslant 1$ and $g \in G$ with $\|g\| \leqslant \kappa r(n)$,*
  $$c(\kappa)V(r(n)))^{-1} \leqslant \mu^{(2n)}(g) \leqslant CV(r(n))^{-1}.$$

(2) *There exist* $\varepsilon, \gamma_1 \in (0,\infty)$ *and* $\gamma_2 \geqslant 1$, *such that for all* n, τ *with* $\frac{1}{2}\rho(\tau/\gamma_1) \leqslant n \leqslant \rho(\tau/\gamma_1)$,

$$\inf_{x:\|x\|\leqslant\tau} \mathbf{P}_x\left(\sup_{0\leqslant k\leqslant n}\{\|X_k\|\} \leqslant \gamma_2\tau; \|X_n\| \leqslant \tau\right) \geqslant \varepsilon.$$

Proposition 5.2 ([24, Propositin 1.4]). — *Let* $\|\cdot\|$ *be a norm on* G. *Assume that* $r : (0,\infty) \to (0,\infty)$ *is continuous and increasing with inverse* ρ *and that the symmetric probability measure* μ *is strongly* $(\|\cdot\|, r)$ *-controlled. Then, for any* n *and* τ *such that* $\gamma_1 r(2n) \geqslant \tau$, *we have*

$$\inf_{x:\|x\|\leqslant\tau} \mathbf{P}_x\left(\sup_{0\leqslant k\leqslant n}\|X_k\| \leqslant \gamma_2\tau;\ \|X_n\| \leqslant \tau\right) \geqslant \varepsilon^{1+2n/\rho(\tau/\gamma_1)}.$$

Another key notion is that of pseudo-Poincaré inequality (see [9]).

Definition 5.3. — *Let* G *be a discrete group equipped with a symmetric probability measure* ν, *a quasi-norm* $\|\cdot\|$ *and a continuous increasing function* $r : (0,\infty) \to (0,\infty)$ *with inverse* ρ. *We say that* ν *satisfies a pointwise* $(\|\cdot\|, r)$*-pseudo-Poincaré inequality if, for any* f *with finite support on* G,

$$\forall\, g \in G,\ \sum_{x\in G} |f(xg) - f(x)|^2 \leqslant C\rho(\|g\|)\mathcal{E}_\nu(f, f).$$

Proposition 5.4. — *Let* G *be a group of polynomial volume growth and* $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$, *see Definition 3.9. Let* $(\Sigma_G, \mathfrak{F}_G)$ *be a geometry adapted to* μ *as in Definition 3.16. Then the symmetric probability measure* μ *satisfies a pointwise linear (i.e.,* $r(t) = t$*)* $\|\cdot\|_{\mathfrak{F}_G}$*-pseudo-Poincaré inequality.*

Proof. — This follows from Proposition 3.3 and Corollary 3.23 via a simple telescopic sum argument and the Cauchy–Schwarz inequality for finite sums. ☐

Theorem 5.5. — *Let* G *be a group of polynomial volume growth and* μ *be a probability in* $\mathcal{P}_{\preceq}(G, \mathrm{reg})$. *Let* $(\Sigma_G, \mathfrak{F}_G)$ *be a geometry adapted to* μ *as in Definition 3.16. Let*

$$\mathbf{F} = \mathbf{F}_{G,\mathfrak{F}_G}$$

be as in Definition 3.22. Let $\mathfrak{F}_{G,2}$ *be a system of functions related to the system* $\mathfrak{F}_G$ *as in Proposition 2.11 (c) so that, for each* $\sigma \in \Sigma_G$,

$$F_{2,\sigma} \text{ is convex},\ \ F_{2,\sigma} \asymp F_\sigma \circ F^{-1} \ \text{ and }\ F(t) = (1+t)^{\omega_*} - 1$$

with $\omega_* > 0$ *as in Proposition 2.11. Then* $\|\cdot\|_{\mathfrak{F}_{G,2}}$ *is a norm on* G, *and it satisfies* $\|\cdot\|_{\mathfrak{F}_{G,2}} \asymp \|\cdot\|_{\mathfrak{F}_G}^{\omega_*}$ *over* G. *Set*

$$V(R) = \#\{\|g\|_{\mathfrak{F}_{G,2}} \leqslant R\}.$$

With this notation, the following properties are satisfied:

(1) *For all* $R \geqslant 1$, $V(R) \asymp \mathbf{F}(R^{1/\omega_*})$ *and, for all* n,
$$\mu^{(n)}(e) \asymp 1/\mathbf{F}(n) \asymp 1/V(n^{\omega_*}).$$

(2) *There exists a constant* C_1 *such that, for all* $g \in G$, $f \in L^2(G)$,
$$\sum_{x\in G} |f(xg) - f(x)|^2 \leqslant C_1 \|g\|_{\mathfrak{F}_{G,2}}^{1/\omega_*} \mathcal{E}_\mu(f,f).$$

(3) *There exists a constant* C_2 *such that, for all* $n, m \in \mathbb{N}$ *and* $x, y \in G$,
$$|\mu^{(n+m)}(xy) - \mu^{(n)}(x)| \leqslant C_2 \left(\frac{m}{n} + \frac{\|y\|_{\mathfrak{F}_{G,2}}^{1/2\omega_*}}{\sqrt{n}} \right) \mu^{(n)}(e).$$

(4) *There exists* $\eta \in (0,1]$ *such that, for all* $n \geqslant 1$ *and* $g \in G$ *with* $\|g\|_{\mathfrak{F}_{G,2}} \leqslant \eta n^{\omega_*}$, *we have*
$$\mu^{(n)}(g) \asymp 1/V(n^{\omega_*}).$$

(5) *The symmetric probability measure* μ *is* $(\|\cdot\|_{\mathfrak{F}_{G,2}}, r)$*-controlled with* $r(t) = t^{\omega_*}$.

(6) *The symmetric probability measure* μ *is strongly* $(\|\cdot\|_{\mathfrak{F}_{G,2}}, r)$*-controlled with* $r(t) = t^{\omega_*}$.

Proof. — The first two items are results that have already been proved above and which are now stated in terms of $\mathfrak{F}_{G,2}$ instead of $\mathfrak{F}_G$. The point in doing this is that $\|\cdot\|_{\mathfrak{F}_{G,2}}$ is a norm whereas $\|\cdot\|_{\mathfrak{F}_G}$ might only be a quasi-norm. The third item (regularity) follows from the first two and Appendix A, see Proposition A.3. Item (4) immediately follows from (1) and (3).

Given the relation between $\|\cdot\|_{\mathfrak{F}_G}$ and $\|\cdot\|_{\mathfrak{F}_{G,2}}$, item (5) follows from (1), the estimate on $\Lambda_{2,G,\mu}$ stated in Theorem 4.3 and [19, Lemma 4.1].

To prove item (6), observe that the norm $\|\cdot\|_{\mathfrak{F}_{G,2}}$ is well-connected in the sense of [24, Definition 3.3] (this is a rather weak property satisfied by any such norm), and that the pointwise pseudo-Poincaré inequality stated as item (2) holds. Because of these properties and [24, Proposition 3.5], $(\|\cdot\|_{\mathfrak{F}_{G,2}}, r)$-control (that is, item (5)) implies strong $(\|\cdot\|_{\mathfrak{F}_{G,2}}, r)$-control (item (6)). □

Note that, by the symmetry of μ, $\|\mu^{(2n)}\|_\infty = \mu^{(2n)}(e)$. Since $n \mapsto \|\mu^{(n)}\|_\infty$ is obviously non-increasing, the on-diagonal estimate in Theorem 5.5 (1) implies

$$\|\mu^{(n)}\|_\infty \asymp 1/\mathbf{F}(n). \tag{5.1}$$

6. Hölder continuity of caloric functions

In this section, we place ourselves in the context of Theorem 5.5. Namely, we consider a group G, finitely generated and of polynomial volume growth, equipped with a probability measure $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$. Having chosen a normal nilpotent subgroup N of finite index in G, we construct the quasi-norm $\|\cdot\|_{\mathfrak{F}_G}$ as in Definitions 3.16 and the associated function $\mathbf{F} = \mathbf{F}_{G,\mathfrak{F}_G}$ introduced in Definition 3.22 so that $\#\{g \in : \|g\|_{\mathfrak{F}_G} \leqslant R\} \asymp \mathbf{F}(R)$, see Corollary 3.23. Consider further the norm $\|\cdot\|_{\mathfrak{F}_{G,2}}$ and positive real ω_* as in Theorem 5.5 so that $\|\cdot\|_{\mathfrak{F}_{G,2}} \asymp \|\cdot\|_{\mathfrak{F}_G}^{\omega_*}$. For any $x \in G$ and $r > 0$, let

$$B(x, r) = \{z \in G : \|x^{-1}z\|_{\mathfrak{F}_{G,2}} < r\}.$$

Also, for any $x \in G$ and $A \subset G$, let

$$\|x^{-1}A\|_{\mathfrak{F}_{G,2}} = \inf_{y \in A} \|x^{-1}y\|_{\mathfrak{F}_{G,2}}$$

denote the distance between x and A with respect to $\|\cdot\|_{\mathfrak{F}_{G,2}}$.

The goal of this section is to prove the Hölder regularity of bounded solutions of the discrete parabolic equation (1.1). More precisely, given a (discrete) time interval $I = [0, N]$ and a subset A of G, we say that a real valued bounded function q defined on $I \times G$ is a solution of (1.1) in $I \times A$ if

$$q(n+1, x) - q(n, x) = [q(n, \cdot\,) * (\mu - \delta_e)](x), \quad n, n+1 \in I,\ x \in A,$$

or, equivalently

$$q(n+1, x) = q(n, \cdot\,) * \mu(x), \quad n, n+1 \in I, x \in A.$$

for all n with $n, n+1 \in I$ and all $x \in A$. Note that this definition requires q to be defined over all of G at all times in I. This is natural in the present context since, typically, the measure μ has infinite support. Because such a solution q is bounded and defined for each $k \in I$ on all of G, it is always possible to extend it forward in time.

Let $(X_n)_0^\infty$ denote the random walk driven by μ started at X_0. For any subset $A \subset G$, define

$$\tau_A = \inf\{n \geqslant 0 : X_n \notin A\}, \quad \sigma_A = \inf\{n \geqslant 1 : X_n \in A\}.$$

The argument developed below is mainly based on that of [2, Theorem 4.9], which in turn is the discrete analog of that for [6, Theorem 4.14], and makes use of the following notion of *caloric function* (we followed the definition from [2, 6], though co-caloric might be more appropriate, see [11, p. 263]).

Let $\Upsilon = \mathbf{Z}_+ \times G$. We will make use of the Υ-valued Markov chain $Z_k := (V_k, X_k)$ where $V_k = V_0 + k$. Write $\mathbf{P}_{(i,x)}$ for the law of Z_k started at (i, x)

and let $\mathscr{F}_i = \sigma\{Z_k : k \leqslant i\}$. A bounded function $u(k,x)$ on Υ is said to be *caloric* on $D \subset \Upsilon$, if $k \mapsto u(V_{k\wedge\tau_D}, X_{k\wedge\tau_D})$ is a martingale, where

$$\tau_D = \inf\{k \geqslant 0 : (V_k, X_k) \notin D\}.$$

In the case $V_0 = 0$ and $D = I \times A = [k_1, k_2] \times A$, $k_1 < k_2$, the condition that $k \mapsto u(V_{k\wedge\tau_D}, X_{k\wedge\tau_D})$ is a martingale is equivalent to

$$u(k-1, x) = [u(k,\cdot) * \mu](x), \quad k, k-1 \in I, \ x \in A.$$

This is the "backward version" of the parabolic discrete equation (1.1) in the sense that, for any N, $q(k,x) = u(N-k,x)$ is a solution of (1.1) in $J \times A$, $J = [N-k_2, N-k_1]$ (and vice versa). In particular, for any fixed integer N and bounded function f on G, the functions

$$(k,x) \longmapsto \mu^{(N-k)}(x) \quad \text{and} \quad (k,x) \longmapsto f * \mu^{(N-k)}(x)$$

are caloric functions on $[0,N] \times G$.

Lemma 6.1. — *Given any $\delta > 0$, there exists a constant $\kappa := \kappa(\delta) > 0$ such that for any $x, y \in G$, $A \subset G$ with* $\min\left\{\|x^{-1}A\|_{\mathfrak{F}_{G,2}}, \|y^{-1}A\|_{\mathfrak{F}_{G,2}}\right\} \geqslant \kappa n^{\omega_*}$, *and any $n \geqslant 1$,*

$$\mathbf{P}_x(X_n = y, \sigma_A \leqslant n) \leqslant \delta/\mathbf{F}(n).$$

Proof. — By the strong Markov property of X_n and (5.1), we have

$$\begin{aligned}\mathbf{P}_x(X_n = y, \sigma_A \leqslant [n/2]) &= \mathbf{E}_x\left[\mathbf{1}_{\{\sigma_A \leqslant [n/2]\}} \mathbf{P}_{X_{\sigma_A}}(X_{n-\sigma_A} = y)\right] \\ &\leqslant \frac{c_1}{\mathbf{F}(n-[n/2])} \mathbf{P}_x(\sigma_A \leqslant [n/2]) \\ &\leqslant \frac{c_2}{\mathbf{F}(n)} \mathbf{P}_x(\sigma_A \leqslant [n/2]).\end{aligned}$$

Furthermore, due to Theorem 5.5 (5), by choosing $\kappa := \kappa(\delta) > 0$ large enough, we have

$$\mathbf{P}_x(\sigma_A \leqslant [n/2]) \leqslant \mathbf{P}_x\left(\sup_{0\leqslant k\leqslant[n/2]} \|x^{-1}X_k\|_{\mathfrak{F}_{G,2}} \geqslant \kappa n^{\omega_*}\right) \leqslant \delta/(2c_2),$$

which yields that

$$\mathbf{P}_x(X_n = y, \sigma_A \leqslant [n/2]) \leqslant \delta/(2\mathbf{F}(n)). \tag{6.1}$$

We now consider $\mathbf{P}_x(X_n = y, [n/2] \leqslant \sigma_A \leqslant n)$. If the first hitting time of A is between time $[n/2]$ and time n, then the last hitting time of A before time n is larger than $[n/2]$. So, setting $\widehat{\sigma}_A = \sup\{k \leqslant n : X_k \in A\}$, we have

$$\mathbf{P}_x(X_n = y, [n/2] \leqslant \sigma_A \leqslant n) \leqslant \mathbf{P}_x(X_n = y, [n/2] \leqslant \widehat{\sigma}_A \leqslant n).$$

We claim that by time reversal,

$$\mathbf{P}_x(X_n = y, [n/2] \leqslant \widehat{\sigma}_A \leqslant n) = \mathbf{P}_y(X_n = x, \sigma_A \leqslant n - [n/2]).$$

To see this, observe that the symmetry of the transition probability $p(x,y) := \mu(x^{-1}y)$ implies that we have, for any $[n/2] \leqslant k \leqslant n$,

$$\begin{aligned}\mathbf{P}_x(X_k = z_k, X_{k+1} = z_{k+1}, \ldots, X_n = y)& \\ &= p_k(x, z_k)p(z_k, z_{k+1}) \ldots p(z_{n-1}, y) \\ &= \mathbf{P}_y(X_1 = z_{n-1}, \ldots, X_{n-k} = z_k, X_n = x),\end{aligned}$$

where $p_k(x,z) := \mathbf{P}_x(X_k = z) = \mu^{(k)}(x^{-1}y)$. Summing over all $z_k \in A$ and $z_{k+1}, \ldots, z_{n-1} \notin A$, we have

$$\begin{aligned}\mathbf{P}_x(X_k \in A, X_{k+1} \notin A, \ldots, X_{n-1} \notin A, X_n = y)& \\ &= \mathbf{P}_y(X_1 \notin A, \ldots, X_{n-k-1} \notin A, X_{n-k} \in A, X_n = x).\end{aligned}$$

Further, summing over $[n/2] \leqslant k \leqslant n$, this yields that

$$\mathbf{P}_x([n/2] \leqslant \widehat{\sigma}_A \leqslant n, X_n = y) = \mathbf{P}_y(0 \leqslant \sigma_A \leqslant n - [n/2], X_n = x).$$

This proves the desired assertion. Arguing as in the first part of the proof, we find that

$$\begin{aligned}\mathbf{P}_x([n/2] \leqslant \widehat{\sigma}_A \leqslant n, X_n = y)& \\ &= \mathbf{P}_y(0 \leqslant \sigma_A \leqslant n - [n/2], X_n = x) \leqslant \delta/(2\mathbf{F}(n)).\end{aligned}$$

Therefore, the lemma follows from the estimate above and (6.1). □

Proposition 6.2. — *Let $\eta \in (0,1]$ be the constant in Theorem 5.5 (4). For all $n \geqslant 1$, there exist constants $c_1 \in (0,\infty)$ and $\theta \in (0,1)$ such that for any $x, y, z \in G$ with* $\max\left\{\|x^{-1}z\|_{\mathfrak{F}_{G,2}}, \|y^{-1}z\|_{\mathfrak{F}_{G,2}}\right\} \leqslant \eta n^{\omega_*}/2$ *and* $r \geqslant (n/\theta)^{\omega_*}$,

$$\mathbf{P}_x\left(X_n = y, \sup_{0 \leqslant k \leqslant n} \|z^{-1}X_k\|_{\mathfrak{F}_{G,2}} \leqslant r\right) \geqslant \frac{c_1}{\mathbf{F}(n)}.$$

Proof. — Note that since $\|x^{-1}y\|_{\mathfrak{F}_{G,2}} \leqslant \|x^{-1}z\|_{\mathfrak{F}_{G,2}} + \|y^{-1}z\|_{\mathfrak{F}_{G,2}} \leqslant \eta n^{\omega_*}$, we have by Theorem 5.5 (4) that

$$\mathbf{P}_x(X_n = y) \geqslant \frac{c_0}{\mathbf{F}(n)}.$$

Let $\delta = c_0/2$ in Lemma 6.1. Then for all $r > (\kappa+1)n^{\omega_*}$ (where κ is the constant in Lemma 6.1 associated with this δ), we have by Lemma 6.1 that

$$\begin{aligned}\mathbf{P}_x(X_n = y, \sup_{0\leqslant k\leqslant n} \|z^{-1}X_k\|_{\mathfrak{F}_{G,2}} \leqslant r) \\ &= \mathbf{P}_x(X_n = y) - \mathbf{P}_x(X_n = y, \sup_{0\leqslant k\leqslant n} \|z^{-1}X_k\|_{\mathfrak{F}_{G,2}} > r) \\ &\geqslant \mathbf{P}_x(X_n = y) - \mathbf{P}_x(X_n = y, \sigma_{B(z,r)^c} \leqslant n) \\ &\geqslant c_0/(2\mathbf{F}(n)),\end{aligned}$$

where in the last inequality we used the facts that

$$\|x^{-1}B(z,r)^c\|_{\mathfrak{F}_{G,2}} \geqslant r - \|x^{-1}z\|_{\mathfrak{F}_{G,2}} \geqslant (\kappa+\eta)n^{\omega_*} - \eta n^{\omega_*}/2 \geqslant \kappa n^{\omega_*}$$

and, similarly, $\|y^{-1}B(z,r)^c\|_{\mathfrak{F}_{G,2}} \geqslant \kappa n^{\omega_*}$. This proves the desired assertion with $\theta = (\kappa+1)^{-1/\omega_*}$. □

The following is an immediate consequence of the last proposition.

Corollary 6.3. — *Let $\eta \in (0,1]$ be the constant in Theorem 5.5(4). There exist constants $\theta \in (0,1)$ and $c_1 > 0$ such that for every $z \in G$, $n \geqslant 1$, and $A \subset B(z, \eta n^{\omega_*}/2)$,*

$$\mathbf{P}_x\left(X_n \in A, \tau_{B(z,(n/\theta)^{\omega_*})} > n\right) \geqslant c_1 \frac{\#A}{\mathbf{F}(n)} \quad \text{for every } x \in B(z, \eta n^{\omega_*}/2).$$

The following is a key proposition concerning the space-time process $Z_k = (V_k, X_k) = (V_0 + k, X_k)$ on $\Upsilon = \mathbf{Z}_+ \times G$ discussed before Lemma 6.1, which will be used in the proof of Theorem 6.8.

Proposition 6.4. — *Let $\eta \in (0,1]$ be the constant in Theorem 5.5(4), and $\theta \in (0,1)$ be the constant in Corollary 6.3. Let m be the counting measure on Υ, and set*

$$C_0 = 2^{3+4\omega_*}/(\eta(\theta \wedge \theta^{\omega_*})) \quad \text{and} \quad C_1 = \eta\theta^{\omega_*}/2^{1+4\omega_*}. \tag{6.2}$$

For every $\delta \in (0,1)$ and $\gamma \geqslant C_0$, there is a constant $c_0 = c_0(\delta, C_0) > 0$ (independent of γ) such that for any $x_0 \in G$, $n_0 \geqslant 0$, $R > 1$ with $\theta R^{1/\omega_*} \geqslant 1$ and

$$A \subset \left[n_0 + \left[\frac{1}{2}\theta(\gamma R)^{1/\omega_*}\right], n_0 + \left[\frac{1}{2}\theta(\gamma R)^{1/\omega_*}\right] + [\theta R^{1/\omega_*}]\right] \times B(x_0, C_1\gamma R)$$

satisfying

$$\frac{m(A)}{[\theta R^{1/\omega_*}] \cdot \#B(x_0, C_1\gamma R)} \geqslant \delta,$$

we have

$$\mathbf{P}_z(\sigma_A < \tau_{Q(n_0,x_0,\gamma R)}) \geqslant c_0 \quad \text{for every } z \in \left[n_0, n_0 + [\theta R^{1/\omega_*}]\right] \times B(x_0, R).$$

Here $Q(n_0,x_0,r):=\big[n_0,n_0+[\theta r^{1/\omega_*}]\big]\times B(x_0,r)$*, and by abusing the notation,* $\sigma_A:=\inf\{k\geqslant 1: Z_k\in A\}$*.*

Proof. — First note that by the definition (6.2) of C_0, $C_0\geqslant 8^{\omega_*+1}/(\eta(\theta\wedge\theta^{\omega_*})$, and so for any $\gamma\geqslant C_0$,

$$[\theta(\gamma R)^{1/\omega_*}]>8[\theta R^{1/\omega_*}].$$

Let A be the subset of $\Upsilon=\mathbf{Z}_+\times G$ in the proposition. For $j\in\mathbb{N}$, let $A_j=\{x\in G:(j,x)\in A\}$. There exists some

$$k\in\left[n_0+\left[\frac{1}{2}\theta(\gamma R)^{1/\omega_*}\right],n_0+\left[\frac{1}{2}\theta(\gamma R)^{1/\omega_*}\right]+[\theta R^{1/\omega_*}]\right]$$

so that

$$\#A_k\geqslant\frac{m(A)}{[\theta R^{1/\omega_*}]}\geqslant\delta\cdot\#B(x_0,C_1\gamma R).\tag{6.3}$$

Note that for any $z=(n,x)\in\big[n_0,n_0+[\theta R^{1/\omega_*}]\big]\times B(x_0,R)$,

$$\frac{5}{8}[\theta(\gamma R)^{1/\omega_*}]\geqslant k-n\geqslant\frac{1}{8}[\theta(\gamma R)^{1/\omega_*}].\tag{6.4}$$

In particular,

$$\eta(k-n)^{\omega_*}/2\geqslant C_1\gamma R\geqslant 4R\quad\text{ and }\quad((k-n)/\theta)^{\omega_*}<\gamma R.\tag{6.5}$$

It follows from Theorem 5.5 (1) and (6.4) that $\#B(x_0,\gamma R)\asymp\mathbf{F}(k-n)$. Hence we have by (6.3), (6.5) and Corollary 6.3 that

$$\mathbf{P}_{(n,x)}\left(\sigma_A<\tau_{Q(n_0,x_0,\gamma R)}\right)\geqslant\mathbf{P}_x(X_{k-n}\in A_k,\tau_{B(x_0,\gamma R)}>k-n)\geqslant c_0.$$

This completes the proof of the proposition. □

The following is a special case of the Lévy system formula for Markov chains. For any $(k,x)\in\Upsilon$ and $A\subset\Upsilon$, define $N_A(k,x)=\mathbf{P}_{(k,x)}(X_1\in A(k+1))$ if $(k,x)\notin A$ and 0 otherwise.

Lemma 6.5. — *For the Υ-valued Markov chain (V_k,X_k), let $A\subset\Upsilon$ and*

$$J_n=\mathbf{1}_A(V_n,X_n)-\mathbf{1}_A(V_0,X_0)-\sum_{k=0}^{n-1}N_A(V_k,X_k).$$

Then $\{J_{n\wedge\sigma_A};n\in\mathbb{N}\}$ is a martingale.

Proof. — See [3, Lemma 3.2] for the proof. □

We also need the following lemma.

Lemma 6.6. — *For $r>0$, there exists a constant $c_1>0$ such that*

$$\mathbf{E}_x(\tau_{B(x,r)})\leqslant c_1 r^{1/\omega_*}.$$

Proof. — By (5.1) and Theorem 5.5 (1), for any $x, y \in G$ and $n \geqslant 1$, we have

$$\mathbf{P}_y(\tau_{B(x,r)} > n) \leqslant \mathbf{P}_y(X_n \in B(x,r)) \leqslant \frac{c_1}{\mathbf{F}(n)}\mathbf{F}(r^{1/\omega_*}).$$

By taking $n = c_2 r^{1/\omega_*}$ for a proper constant $c_2 > 0$, $\mathbf{P}_y(\tau_{B(x,r)} > n) \leqslant 1/2$. Using the Markov property at time kn for $k = 1, 2, \dots,$

$$\begin{aligned}\mathbf{P}_x(\tau_{B(x,r)} > (k+1)n) &\leqslant \mathbf{E}_x\left(\mathbf{P}_{X_{kn}}(\tau_{B(x,r)} > n); \tau_{B(x,r)} > kn\right)\\ &\leqslant \frac{1}{2}\mathbf{P}_x(\tau_{B(x,r)} > kn).\end{aligned}$$

By induction,

$$\mathbf{P}_x(\tau_{B(x,r)} > kn) \leqslant 2^{-k}$$

for $k = 1, 2, \dots$. With this choice of n, we obtain that

$$\mathbf{E}_x\tau_{B(x,r)} \leqslant \sum_{k=1}^{\infty} kn\mathbf{P}_x((k-1)n < \tau_{B(x,r)} \leqslant kn) \leqslant n\sum_{k=1}^{\infty} k2^{k-1} =: cn,$$

which proves our result. □

Proposition 6.7. — *There is a constant $c_0 > 0$ such that for any $s \geqslant 2r$ and $x \in G$,*

$$\mathbf{P}_x(X_{\tau_{B(x,r)}} \notin B(x,s)) \leqslant c_0\left(\frac{r}{s}\right)^{1/\omega_*}.$$

Proof. — Applying Lemma 6.5 with $A = \mathbf{Z}_+ \times B(x,s)^c$ at stopping times $n \wedge \tau_{B(x,r)}$ and then letting $n \to \infty$, we see that

$$\begin{aligned}\mathbf{P}_x(X_{\tau_{B(x,r)}} \notin B(x,s)) &= \mathbf{E}_x\left(\sum_{k=0}^{\tau_{B(x,r)}-1} N_{B(x,s)^c}(X_k)\right)\\ &= \mathbf{E}_x\left(\sum_{k=0}^{\tau_{B(x,r)}-1}\sum_{y\in B(x,s)^c}\mu(X_k^{-1}y)\right)\\ &\leqslant \mathbf{E}_x\tau_{B(x,r)}\sup_{z\in B(x,r)}\sum_{y\in B(x,s)^c}\mu(z^{-1}y)\\ &\leqslant c_1 r^{1/\omega_*}s^{-1/\omega_*},\end{aligned}$$

where in the last inequality we used Lemma 6.6 and the fact that

$$\sum_{\|h\|_{\mathfrak{F}_{G,2}}\geqslant r}\mu(h) \leqslant c_2 r^{-1/\omega_*}, \quad r > 0. \tag{6.6}$$

Note that, (6.6) can be obtained following the same line of reasoning as for (3.2), by using Theorem 5.5 (1), $\|\cdot\|_{\mathfrak{F}_{G,2}} \asymp \|\cdot\|_{\mathfrak{F}_G}^{\omega_*}$, $F_{N,s} = \Phi_i^{-1}$ (see Definition 3.18) and Remark 3.7. □

From now, we take $\theta \in (0,1)$ and C_0, C_1 be the constants in Corollary 6.3 and (6.2), respectively. For $n \in \mathbb{N}$, $x \in G$ and $r > 1$ with $\theta r^{1/\omega_*} \geqslant 1$, define

$$Q(n,x,r) = [n, n + [\theta r^{1/\omega_*}]] \times B(x,r).$$

The proof of the following theorem is similar to the proofs of [6, Theorem 4.14] and [2, Theorem 4.9]. We remark that we also took this opportunity to correct an error in selecting subsets A and A' in the proofs of [6, Theorem 4.14] and [2, Theorem 4.9]. Such a correction was previously made in the proof of [7, Theorem 6.3] on Hölder regularity of caloric functions for certain diffusion processes. Our proof is based on Propositions 6.4 and 6.7.

Theorem 6.8. — *Let $C_0 > 0$ be the constant in* (6.2). *There are constants $C > 0$ and $\beta > 0$ such that for any $R > 1$ with $\theta R^{1/\omega_*} \geqslant 1$ (where θ is the constant in Corollary 6.3) and bounded caloric function q in $Q(0, x_0, C_0 R) = [0, [\theta (C_0 R)^{1/\omega_*}]] \times B(x_0, C_0 R)$, we have*

$$|q(m_1, x) - q(m_2, y)| \leqslant C \|q\|_\infty \left(\frac{|m_1 - m_2|^{\omega_*} + \|x^{-1}y\|_{\mathfrak{F}_{G,2}}}{R} \right)^\beta$$

for all (m_1, x), $(m_2, y) \in Q(0, x_0, R)$, and

$$\|q\|_\infty = \sup_{(i,x) \in [0, [\theta (C_0 R)^{1/\omega_*}]] \times G} q(i,x).$$

In particular, we have

$$\sum_{x \in G} |\mu^{(m_1)}(x) - \mu^{(m_2)}(xy)| \leqslant C \left(\frac{|m_1 - m_2|^{\omega_*} + \|y\|_{\mathfrak{F}_{G,2}}}{n_0^{\omega_*}} \right)^\beta$$

for all $y \in G$ and $m_1, m_2 \geqslant n_0 \geqslant 1$.

Proof. — Recall that $Z_k = (V_k, X_k)$ is the space-time process of X, where $V_k = V_0 + k$. Without loss of generality, assume that $0 \leqslant q(z) \leqslant \|q\|_\infty = 1$ for all $z \in [0, [\theta (C_0 R)^{1/\omega_*}]] \times G$.

By Proposition 6.4, there exists a constant $c_1 \in (0,1)$ such that if $x_0 \in G$, $n_0 \geqslant 0$, $r > 1$ with $\theta r^{1/\omega_*} \geqslant 1$, $\gamma \geqslant C_0$, and

$$A \subset \left[n_0 + \left[\frac{1}{2} \theta (\gamma r)^{1/\omega_*} \right], n_0 + \left[\frac{1}{2} \theta (\gamma r)^{1/\omega_*} \right] + [\theta r^{1/\omega_*}] \right] \times B(x_0, C_1 \gamma r)$$

satisfying

$$\frac{m(A)}{[\theta r^{1/\omega_*}] \cdot \# B(x_0, C_1 \gamma r)} \geqslant 1/3,$$

then

$$\mathbf{P}_z(\sigma_A < \tau_{Q(n_0, x_0, \gamma r)}) \geqslant c_1 \quad \text{for } z \in \left[n_0, n_0 + [\theta r^{1/\omega_*}] \right] \times B(x_0, r). \tag{6.7}$$

Here $C_1 = \eta\theta^{\omega_*}/2^{1+4\omega_*} \in (0,1)$ as defined in (6.2), where $\eta \in (0,1]$ is the constant in Theorem 5.5 (4). On the other hand, according to Proposition 6.7, there is a constant $c_2 > 0$ such that if any $x \in G$, and $s \geqslant 2r$, then

$$\mathbf{P}_x(X_{\tau_{B(x,r)}} \notin B(x,s)) \leqslant c_2(r/s)^{1/\omega_*}. \tag{6.8}$$

Let $\eta_0 = 1 - (c_1/4)$ and

$$\rho = C_0^{-1} \wedge (\eta_0/2)^{\omega_*} \wedge (c_1\eta_0/(8c_2))^{\omega_*} < 1.$$

Note that for every $(n_0, x) \in Q(0, x_0, R)$, q is caloric in $Q(n_0, x, R) \subset Q(0, x_0, C_0 R)$. We will show that

$$\sup_{Q(n_0,x,\rho^{2k}R)} q - \inf_{Q(n_0,x,\rho^{2k}R)} q \leqslant \eta_0^k \tag{6.9}$$

for all $k \leqslant K_0$, where K_0 is the largest integer k so that $\theta(\rho^{2k}R)^{1/\omega_*} \geqslant 1$. For notational convenience, we write Q_i for $Q(n_0, x, \rho^i R)$ and $\tau_i = \tau_{Q(n_0,x,\rho^i R)}$. Define

$$a_i = \sup_{Q_{2i}} q, \qquad b_i = \inf_{Q_{2i}} q.$$

Clearly $b_i - a_i \leqslant 1 \leqslant \eta_0^i$ for all $i \leqslant 0$. Now suppose that $b_i - a_i \leqslant \eta_0^i$ for all $i \leqslant k$ and we are going to show that $b_{k+1} - a_{k+1} \leqslant \eta_0^{k+1}$ as long as $k+1 \leqslant K_0$. Observe that $Q_{2k+2} \subset Q_{2k+1} \subset Q_{2k}$ and $a_k \leqslant q \leqslant b_k$ on Q_{2k}. Define

$$Q'_{2k+1} = \left[n_0 + \left[\frac{1}{2}\theta(\rho^{2k+1}R)^{1/\omega_*}\right], n_0 + \left[\frac{1}{2}\theta(\rho^{2k+1}R)^{1/\omega_*}\right] + [\theta(\rho^{2k+2}R)^{1/\omega_*}]\right] \times B(x_0, C_1\rho^{2k+1}R)$$

and

$$A' = \left\{z \in Q'_{2k+1} : q(z) \leqslant (a_k + b_k)/2\right\}.$$

It is clear that $Q'_{2k+1} \subset Q_{2k+1}$. We may suppose that

$$\frac{m(A')}{[\theta(\rho^{2k+2}R)^{1/\omega_*}] \cdot \#B(x_0, C_1\rho^{2k+1}R)} \geqslant 1/2.$$

Otherwise, we use $1-q$ instead of q. Let A be a compact subset of A' such that

$$\frac{m(A)}{[\theta(\rho^{2k+2}R)^{1/\omega_*}] \cdot \#B(x_0, C_1\rho^{2k+1}R)} \geqslant 1/3.$$

For any given $\eta_0 > 0$, pick $z_1, z_2 \in Q_{2(k+1)}$ so that $q(z_1) \geqslant b_{k+1} - \varepsilon$ and $q(z_2) \leqslant a_{k+1} + \varepsilon$. Then, according to (6.7), (6.8) and (6.9),

$$\begin{aligned}
&b_{k+1} - a_{k+1} - 2\varepsilon \\
&\leqslant q(z_1) - q(z_2) \\
&= \mathbf{E}_{z_1}[q(Z_{\sigma_A \wedge \tau_{2k+1}}) - q(z_2)] \\
&= \mathbf{E}_{z_1}[q(Z_{\sigma_A}) - q(z_2); \sigma_A < \tau_{2k+1}] \\
&\quad + \mathbf{E}_{z_1}[q(Z_{\tau_{2k+1}}) - q(z_2); \sigma_A > \tau_{2k+1}, Z_{\tau_{2k+1}} \in Q_{2k}] \\
&\quad + \sum_{i=1}^{\infty} \mathbf{E}_{z_1}[q(Z_{\tau_{2k+1}}) - q(z_2); \sigma_A \geqslant \tau_{2k+1}, Z_{\tau_{2k+1}} \in Q_{2(k-i)} \setminus Q_{2(k+1-i)}] \\
&\leqslant \left(\frac{a_k + b_k}{2} - a_k\right) \mathbf{P}_{z_1}(\sigma_A < \tau_{2k+1}) + (b_k - a_k)\mathbf{P}_{z_1}(\sigma_A \geqslant \tau_{2k+1}) \\
&\quad + \sum_{i=1}^{\infty} (b_{k-i} - a_{k-i}) \mathbf{P}_{z_1}(Z_{\tau_{2k+1}} \notin Q_{2(k+1-i)}) \\
&\leqslant (b_k - a_k)\left(1 - \frac{\mathbf{P}_{z_1}(\sigma_A < \tau_{2k+1})}{2}\right) + \sum_{i=1}^{\infty} c_2 \eta_0^k (\rho^{1/\omega_*}/\eta_0)^i \\
&\leqslant (1 - c_1/2)\eta_0^k + 2c_2 \eta_0^{k-1} \rho^{1/\omega_*} \\
&\leqslant (1 - c_1/2)\eta_0^k + c_1 \eta_0^k/4 = \eta_0^{k+1}.
\end{aligned}$$

Since ε is arbitrary, we have $b_{k+1} - a_{k+1} \leqslant \eta_0^{k+1}$, and this proves the claim (6.9).

For $z = (i, x)$ and $w = (j, y)$ in $Q(0, x_0, R)$ with $i \leqslant j$, let k be the largest integer such that

$$\|z - w\| := (|j - i|/\theta)^{\omega_*} + \|x^{-1}y\|_{\mathfrak{F}_{G,2}} \leqslant \rho^{2k} R.$$

Then $\log(\|z - w\|/R) \geqslant 2(k+1)\log\rho$, $w \in Q(n, x, \rho^{2k}R)$ and

$$|q(z) - q(w)| \leqslant \eta_0^k = e^{k \log \eta_0} \leqslant c_3 \left(\frac{\|z - w\|}{R}\right)^{\log \eta_0/(2\log\rho)}.$$

This proves the first desired assertion.

Fix $n_0 \geqslant 1$, $N_0 \geqslant 2$ and a bounded function u on G with $\|u\|_\infty = 1$. Set $q(n, x) = \sum_z u(z) p_{N_0 - n}(z, x) = u * \mu^{(N_0 - n)}(x)$. This function q is a caloric function on $[0, N_0 - n_0]$ (for example see [6, Lemma 4.5]), and is bounded above by $\|u\|_\infty = 1$. Take $R > 1$ such that $\theta^{\omega_*} R = n_0^{\omega_*}$. (Note that in particular $\theta R^{1/\omega_*} \geqslant 1$ so the first assertion will apply.) Let $m_1, m_2 \in [n_0, N_0]$ with $m_1 > m_2$ and $x_1, x_2 \in G$. Assume first that

$$|m_1 - m_2|^{\omega_*} + \|x_1^{-1} x_2\|_{\mathfrak{F}_{G,2}} < \theta^{\omega_*} R = n_0^{\omega_*} \tag{6.10}$$

and so $(N_0 - m_2, x_2) \in Q(N_0 - m_1, x_1, R) \subset [0, N_0 - n_0] \times G$. Applying the first assertion to this caloric function $q(n,x)$ with $(N_0 - m_1, x_1)$, $(N_0 - m_2, x_2)$ and $Q(N_0 - m_1, x_1, R)$ in place of (m_1, x), (m_2, y) and $Q(0, x_0, R)$ respectively, we have

$$|u*\mu^{(m_1)}(x^{-1}z) - u*\mu^{(m_2)}(y^{-1}z)| \leqslant \frac{c}{n_0^{\beta\omega_*}} \left(|m_1 - m_2|^{\omega_*} + \|x_1^{-1}x_2\|_{\mathfrak{F}_{G,2}}\right)^{\beta}.$$

This inequality is also trivially true when (6.10) does not hold. So the inequality above holds for every $m_1, m_2 \in [n_0, N_0]$ and $x_1, x_2 \in G$ for all $n_0 \geqslant 1$ and $N_0 \geqslant 2$. This proves the second assertion after taking the supremum over all u with $\|u\|_\infty = 1$. □

Remark 6.9. — From (6.8) we can get that

$$|\mu^{(m_1)}(x) - \mu^{(m_2)}(y)| \leqslant \frac{c}{n_0^{\beta\omega_*}\mathbf{F}(n_0)} \left(|m_1 - m_2|^{\omega_*} + \|x^{-1}y\|_{\mathfrak{F}_{G,2}}\right)^{\beta}$$

for any $x, y \in G$ and $m_1, m_2 \geqslant n_0 \geqslant 1$. However, this assertion is weaker than that in Theorem 5.5(3).

As an easy application of Theorem 6.8, we have the following. Recall that G is a finitely generated group of polynomial volume growth, equipped with a probability measure $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$.

Corollary 6.10. — *The pair (G, μ) has the (weak) Liouville property, namely, all bounded μ-harmonic functions on G are constant. Moreover there are constants $C, \beta > 0$ such that, for any $x_0 \in G, R > 1$ and any bounded function u defined on G and μ-harmonic in $B(x_0, R) = \{z \in G : \|x_0^{-1}z\|_{\mathfrak{F}_{G,2}} \leqslant R\}$, we have*

$$\forall\, x, y \in B(x_0, R/2), \;\; |u(x) - u(y)| \leqslant C \left(\frac{\|x^{-1}y\|_{\mathfrak{F}_{G,2}}}{R}\right)^{\beta} \|u\|_\infty. \tag{6.11}$$

Remark 6.11. — For finitely generated nilpotent groups and any probability measure μ, the weak Liouville property was proved in [12]. In the case of symmetric probability measures on finitely generated nilpotent groups, the strong Liouville property (i.e., all non-negative μ-harmonic functions are constant) follows from [18]. Because any finitely generated group of polynomial growth contains a nilpotent subgroup of finite index, these two results extend to groups of polynomial growth (given a symmetric measure on a group of polynomial volume growth G, one constructs a symmetric measure on a nilpotent subgroup of finite index N so that the restriction to N of any harmonic function on G is harmonic on N). In addition, for finitely generated group of polynomial volume growth and measures $\mu \in \mathcal{P}_{\preceq}(G, \mathrm{reg})$ as treated here, the weak Liouville property follows also from a more general and direct argument given in [13, Corollary 2.3]. The different methods used in these papers do not provide estimates such as (6.11).

Appendix A. Space-time regularity for $\mu^{(n)}$

This section provides details concerning the intrinsic regularity afforded to the iterated convolutions of a symmetric measure, providing a straightforward extension and complement to [15, Theorem 4.2].

LEMMA A.1. — *Fix $\varepsilon \in (0,1]$ and $\alpha > 0$. If μ is symmetric and the spectrum of $f \mapsto f * \mu$ on $L^2(G)$ is contained in $[-1+\varepsilon, 1]$, then there exists a constant $C_{\varepsilon,\alpha}$ such that*

$$\forall\, f \in L^2(G),\ \forall\, n = 1, 2, \ldots,\ \ \|f * (\delta_e - \mu)^\alpha * \mu^{(n)}\|_2 \leqslant C_{\varepsilon,\alpha} n^{-\alpha} \|f\|_2.$$

Proof. — This is a simple consequence of spectral theory and calculus. Indeed, (using the spectral resolution E_λ of the (self-adjoint) operator of convolution by μ), spectral theory shows that $\|f * (\delta_e - \mu)^\alpha * \mu^{(n)}\|_2 \leqslant M\|f\|_2$, where

$$M = \sup_{\lambda \in J_\varepsilon} \{|1-\lambda|^\alpha |\lambda|^n\},\ \ J_\varepsilon = [-1+\varepsilon, 1].$$

The local maxima of the function $\lambda \mapsto |1-\lambda|^\alpha|\lambda|^n$ on J_ε are at $-1+\varepsilon$ or $n/(\alpha+n)$ so that

$$M \leqslant \max\left\{(1-\varepsilon)^n(2-\varepsilon)^\alpha, \left(\frac{n}{\alpha+n}\right)^n \left(\frac{\alpha}{\alpha+n}\right)^\alpha\right\} \leqslant C_{\varepsilon,\alpha} n^{-\alpha}. \quad \square$$

LEMMA A.2. — *Fix $\varepsilon \in (0,1]$. If μ is symmetric and the spectrum of $f \mapsto f * \mu$ on $L^2(G)$ is contained in $[-1+\varepsilon, 1]$, then there exists a constant C_ε such that, for all pairs of positive integers u, v such that $n \geqslant u + 2v$, we have*

$$\|\mu^{(n+m)} - \mu^{(n)}\|_\infty \leqslant C_\varepsilon \frac{m}{u} \mu^{(2v)}(e).$$

Proof. — It suffices to prove this with $m = 1$. To that end, observe that

$$\begin{aligned} \|\mu^{(n+1)} - \mu^{(n)}\|_\infty &\leqslant \|\mu^{(u+v+1)} - \mu^{(u+v)}\|_2 \|\mu^{(v)}\|_2 \\ &= \|\mu^{(v)} * (\mu^{(u+1)} - \mu^{(u)})\|_2 \|\mu^{(v)}\|_2 \\ &\leqslant C_{\varepsilon,1} u^{-1} \|\mu^{(v)}\|_2^2. \end{aligned}$$

The result follows because $\|\mu^{(v)}\|_2^2 = \mu^{(2v)}(e)$. $\square$

Proposition A.3. — *Fix $\varepsilon \in (0,1]$. Let G be a discrete group equipped with a symmetric probability measure μ, a quasi-norm $\|\cdot\|$ and a continuous increasing function $r : (0,\infty) \to (0,\infty)$ with inverse ρ. Assume μ satisfies a pointwise $(\|\cdot\|, r)$-pseudo-Poincaré inequality with constant C and that the spectrum of $f \mapsto f * \mu$ on $L^2(G)$ is contained in $[-1+\varepsilon, 1]$. Then there exists a constant C_ε such that for all positive integers n, m, u, v such that $n = u + 2v$ and all $x, y \in G$, we have*

$$|\mu^{(n+m)}(xy) - \mu^{(n)}(x)| \leqslant C_\varepsilon \left(\frac{m}{u} + \sqrt{\frac{C\rho(\|y\|)}{u}} \right) \mu^{(2v)}(e).$$

Proof. — It suffices to prove the case $m = 0$ (for $m > 0$, use the previous lemma to reduce to the case $m = 0$). By the Cauchy–Schwarz inequality and the assumed Poincaré inequality,

$$\begin{aligned} |\mu^{(n)}(xy) - \mu^{(n)}(x)| &\leqslant \|\mu^{(u+v)}(\cdot y) - \mu^{(u+v)}(\cdot)\|_2 \|\mu^{(v)}\|_2 \\ &\leqslant [C\rho(\|y\|)\mathcal{E}_\mu(\mu^{(u+v)}, \mu^{(u+v)})]^{1/2} \|\mu^{(v)}\|_2 \\ &\leqslant [C\rho(\|y\|)]^{1/2} \|\mu^{(v)} * (\delta_e - \mu)^{1/2} * \mu^{(u)}\|_2] \|\mu^{(v)}\|_2 \\ &\leqslant C_{\varepsilon,1/2} \sqrt{\frac{C\rho(\|y\|)}{u}} \|\mu^{(v)}\|_2^2 \\ &= C_{\varepsilon,1/2} \sqrt{\frac{C\rho(\|y\|)}{u}} \mu^{(2v)}(e). \end{aligned}$$

□

Acknowledgements

We are grateful to the referee for helpful comments.

BIBLIOGRAPHY

[1] R. F. Bass & Z.-Q. Chen, "Regularity of harmonic functions for a class of singular stable-like processes", *Math. Z.* **266** (2010), no. 3, p. 489-503.

[2] R. F. Bass & T. Kumagai, "Symmetric Markov chains on $\mathbb{Z}^d$ with unbounded range", *Trans. Am. Math. Soc.* **360** (2008), no. 4, p. 2041-2075.

[3] R. F. Bass & D. A. Levin, "Transition probabilities for symmetric jump processes", *Trans. Am. Math. Soc.* **354** (2002), no. 7, p. 2933-2953.

[4] N. H. Bingham, C. M. Goldie & J. L. Teugels, *Regular variation*, Encyclopedia of Mathematics and Its Applications, vol. 27, Cambridge University Press, 1987, xx+491 pages.

[5] J. Chaker & M. Kassmann, "Nonlocal operators with singular anisotropic kernels", *Commun. Partial Differ. Equations* **45** (2020), no. 1, p. 1-31.

[6] Z.-Q. Chen & T. Kumagai, "Heat kernel estimates for stable-like processes on d-sets", *Stochastic Processes Appl.* **108** (2003), no. 1, p. 27-62.

[7] Z.-Q. Chen & S. Lou, “Brownian motion on spaces with varying dimension”, *Ann. Probab.* **47** (2019), no. 1, p. 213-269.

[8] T. Coulhon, “Ultracontractivity and Nash type inequalities”, *J. Funct. Anal.* **141** (1996), no. 2, p. 510-539.

[9] T. Coulhon & L. Saloff-Coste, “Isopérimétrie pour les groupes et les variétés”, *Rev. Mat. Iberoam.* **9** (1993), no. 2, p. 293-314.

[10] T. Delmotte, “Parabolic Harnack inequality and estimates of Markov chains on graphs”, *Rev. Mat. Iberoam.* **15** (1999), no. 1, p. 181-232.

[11] J. L. Doob, *Classical potential theory and its probabilistic counterpart*, Classics in Mathematics, Springer, 2001, Reprint of the 1984 edition, xxvi+846 pages.

[12] E. B. Dynkin & M. B. Malyutov, “Random walk on groups with a finite number of generators”, *Dokl. Akad. Nauk SSSR* **137** (1961), p. 1042-1045.

[13] A. Erschler & T. Zheng, “Growth of periodic Grigorchuck groups”, *Invent. Math.* **219** (2020), no. 3, p. 1069-1155.

[14] R. Grigorchuk, “Milnor’s problem on the growth of groups and its consequences”, in *Frontiers in complex dynamics*, Princeton Mathematical Series, vol. 51, Princeton University Press, 2014, p. 705-773.

[15] W. Hebisch & L. Saloff-Coste, “Gaussian estimates for Markov chains and random walks on groups”, *Ann. Probab.* **21** (1993), no. 2, p. 673-709.

[16] V. A. Kaimanovich & A. M. Vershik, “Random walks on discrete groups: boundary and entropy”, *Ann. Probab.* **11** (1983), no. 3, p. 457-490.

[17] H. Kesten, “Symmetric random walks on groups”, *Trans. Am. Math. Soc.* **92** (1959), no. 2, p. 336-354.

[18] G. A. Margulis, “Positive harmonic functions on nilpotent groups”, *Sov. Math., Dokl.* **7** (1966), p. 241-244.

[19] Y. Peres & T. Zheng, “On groups, slow heat kernel decay yields Liouville property and sharp entropy bounds”, *Int. Math. Res. Not.* (2020), no. 3, p. 722-750.

[20] C. Pittet & L. Saloff-Coste, “On the stability of the behavior of random walks on groups”, *J. Geom. Anal.* **10** (2000), no. 4, p. 713-737.

[21] L. Saloff-Coste, “Probability on groups: random walks and invariant diffusions”, *Notices Am. Math. Soc.* **48** (2001), no. 9, p. 968-977.

[22] L. Saloff-Coste & T. Zheng, “Random walks on nilpotent groups driven by measures supported on powers of generators”, *Groups Geom. Dyn.* **9** (2015), no. 4, p. 1047-1129.

[23] ———, “Random walks and isoperimetric profiles under moment conditions”, *Ann. Probab.* **44** (2016), no. 6, p. 4133-4183.

[24] ———, “On some random walks o driven by spread-out measures”, in *Groups, graphs and random walks*, London Mathematical Society Lecture Note Series, vol. 436, Cambridge University Press, 2017, p. 444-474.

[25] N. T. Varopoulos, “Théorie du potentiel sur les groupes nilpotents”, *C. R. Math. Acad. Sci. Paris* **301** (1985), no. 5, p. 143-144.

[26] N. T. Varopoulos, L. Saloff-Coste & T. Coulhon, *Analysis and geometry on groups*, Cambridge Tracts in Mathematics, vol. 100, Cambridge University Press, 1992, xii+156 pages.

[27] W. Woess, *Random walks on infinite graphs and groups*, Cambridge Tracts in Mathematics, vol. 138, Cambridge University Press, 2000, xii+334 pages.

Manuscrit reçu le 1er juillet 2018,
révisé le 12 août 2020,
accepté le 13 novembre 2020.

Zhen-Qing CHEN
University of Washington
Department of Mathematics
Seattle, WA 98195 (USA)
zqchen@uw.edu

Takashi KUMAGAI
Kyoto University
Research Institute for Mathematical Sciences
Kyoto University, Kyoto 606-8502 (Japan)
kumagai@kurims.kyoto-u.ac.jp

Laurent SALOFF-COSTE
Cornell University
Department of Mathematics
Ithaca, NY 14853 (USA)
lsc@math.cornell.edu

Jian WANG
Fujian Normal University
School of Mathematics and Statistics & Fujian Key Laboratory of Mathematical Analysis and Applications (FJKLMAA) & Center for Applied Mathematics of Fujian Province (FJNU)
Fuzhou 350007 (P.R. China)
jianwang@fjnu.edu.cn

Tianyi ZHENG
UC San Diego
Department of Mathematics
San Diego, CA 92093-0112 (USA)
tzheng2@math.ucsd.edu