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

Limit theorems for some long range random walks on torsion free nilpotent groups

Preprint, July 30, 2022

Springer Nature

Preface

The celebrated limit theorem of DeMoivre and Laplace concerns the convergence of the law of simple random walk on the integers, properly rescaled, to Gauss law. It serves as the starting point of many of probability theory's most important developments. In 1963, Ulf Grenander published a little book [31] titled *Probabilities on Algebraic Structures*, which, among other things, spelled out the natural problem of extending basic limit theorems to the case when addition of numbers is replaced by a more general group law. When taken literally, such extensions face several major difficulties that are easy to explain.

The most natural extensions of simple random on the integers are random walks on countable groups (in particular, finitely generated groups). On the one hand, the classical limit theorems of probability theory are based on the the fact that proper rescaling allows us to approximate the real axis (or Euclidean space of dimension d) by finer and finer embeddings of the integers (or the square lattice of dimension d). On the other hand, it is relatively rare that a finitely generated group embeds into a Lie group, and even rarer that such an embedding can be done at smaller and smaller scales. Indeed, limits obtained though "rescaling" typically inherit an invariance property under the considered rescaling and this applies to both the underlying limit space and the limit stochastic process. Very few connected Lie groups admit rescaling structures of any sort as only certain nilpotent groups do (see, e.g., Theorem 2.1.2 in [36]).

Triangular arrays provide an ingenious way to state results that contain classical limit theorems on abelian groups as special cases and circumvent the difficulties just explained. The tread-off is that such results are not directly applicable to the study of random walks on finitely generated groups unless one finds a way to "rescale" those random walks into a proper triangular array, which bring us back to the previous difficulties.

The most basic example of a non-abelian discrete random walk for which limit theorems through rescaling have been obtained is simple random walk on the Heisenberg group $\mathbb{H}_3(\mathbb{Z})$ of 3 by 3 upper triangular matrices with diagonal entries equal to 1. This group is generated by the four matrices

$$s_1^{\pm 1} = \begin{pmatrix} 1 & \pm 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad s_2^{\pm 1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \pm 1 \\ 0 & 0 & 1 \end{pmatrix}$$

because

$$s_3 = s_1^{-1} s_2^{-1} s_1 s_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The (lazy) simple random walk on this group (associated with the symmetric generating set $S = \{e, s_1^{\pm 1}, s_2^{\pm 1}\}$ where e stands for the identity matrix) is driven by the probability measure

$$\mu = \frac{1}{5} \mathbb{1}_S.$$

If $(\xi_i)_1^\infty$ is an i.i.d. sequence of matrices distributed according to μ, then, at time n, the position of this random walk started at the identity is the matrix $\xi_1 \xi_2 \dots \xi_n$.

What makes it easy to state limit theorems in this case is the combination of the following two facts:

1. The discrete Heisenberg group $\mathbb{H}_3(\mathbb{Z})$ embeds as a subgroup of the real Heisenberg group $\mathbb{H}_3(\mathbb{R})$;
2. The maps

$$\delta_t : \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & tx & t^2 z \\ 0 & 1 & ty \\ 0 & 0 & 1 \end{pmatrix}, \quad t > 0,$$

form a group of group automorphisms of $\mathbb{H}_3((R)$.

Sophisticated versions of the classical limit theorems for this example follow (functional limit theorem, local limit theorem, Edgeworth expansions), see [3, 12, 23, 38, 50, 52, 53] and the references therein. Some of these works treat random walks of finite range or having finite moments of high order on finitely generated nilpotent groups in much greater generality and involve the consideration of more complicated scaling mechanisms.

This monograph is concerned with the extensions of these ideas in the context of stable-like random walks. The simplest family of examples of such walks on the Heisenberg group $\mathbb{H}_3(\mathbb{Z})$ is obtained by considering the measures

$$\mu_{\boldsymbol{\alpha}} = \frac{1}{3} \sum_{1=1}^{3} \sum_{k \in \mathbb{Z}} \frac{c_{\alpha_i}}{(1 + |k|)^{1+\alpha_i}} \mathbb{1}_{s_i^k} \quad \text{with } \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3) \in (0,2)^3.$$

In words, the walk associated with one of these measures on $\mathbb{H}_3(\mathbb{Z})$, takes random long-range steps along each of the one dimensional subgroups of $\mathbb{H}_3(\mathbb{Z})$ associated with the matrices s_1, s_2, s_3. In the direction of s_i, these random long-range steps are stable-like with index $\alpha_i \in (0, 2)$. Obviously, the rescaling mechanism used to study such a walk must be properly adapted to its structural parameters (i.e., to $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$). One interesting phenomenon is that the limit group structure supporting the corresponding limit process also depends on these parameters. Namely, in this case, it is $\mathbb{H}_3(\mathbb{R})$ when $1/\alpha_1 + 1/\alpha_2 \leq 1/\alpha_3$ and it is $\mathbb{R}^3$ otherwise.

The aim of the authors is to develop limit theorems for stable-like random walks in the context of torsion free finitely generated nilpotent groups, theorems that naturally cover these examples and many others.

Seattle, Tokyo, Ithaca, Fuzhou, and San Diago
July, 2022

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

Acknowledgements

The authors thank Mark Meerschaert for references [47, Corollary 8.2.12] and [48, Theorem 4.1].

The research of ZC is partially supported by Simons Foundation Grant 520542, TK by the Grant-in-Aid for Scientific Research (A) 17H01093 and 22H00099, Japan, LSC by NSF grants DMS-1707589 and DMS-2054593, and JW by NNSFC grant 11831014 and 12071076.

Contents

Chapter 1
Setting the stage

Abstract In this opening chapter, we set the stage for later chapters by recalling the very general statements of some basic limit theorems in $\mathbb{R}^n$. Considering the simplest non abelian nilpotent group, the Heisenberg group, in its discrete and Lie versions, and using the abelian results, we provide a first glimpse at the form limit theorems have to take in this context when they involve stable-like long range random walks.

The aim of this work is to present limit theorems (of both functional and local types) for certain long jump random walks on nilpotent groups. Recall that a nilpotent group is a group G with identity element e that has a central series of finite length, that is, there is a finite sequence of normal subgroups so that

$$\{e\} = K_0 \lhd K_1 \lhd \cdots \lhd K_n = G$$

with K_{i+1}/K_i contained in the center of G/K_i for $0 \le i \le n-1$. See the Appendix for a very brief introduction to nilpotent groups.

Before we explain our particular setup and the tools and techniques that we will use, we attempt to put this research in perspective by discussing a small selection of related results concerning random walks and limit theorems in finite dimensional vector spaces (i.e., the torsion free abelian case), and applications of these classical results to the simplest example of non-abelian nilpotent groups, the (discrete and continuous) Heisenberg groups $\mathbb{H}_3(\mathbb{Z}) \subseteq \mathbb{H}_3(\mathbb{R})$.

The "limit theorems" that concern us always have three key ingredients: The first ingredient is a discrete random walk $S = \{S_n; n = 1, 2, \dots\}$ on a group G with independent identically distributed (i.i.d. in short) increments of probability distribution μ. The second ingredient is a method of renormalization via some sort of "dilations" acting on the underlying space G. We remain vague here on purpose. The third ingredient is a continuous time process that appears in the limit, call it Z. Hopefully, Z has properties that make it relatively easy to study although this entire story can also be viewed as a way to understand Z in terms of the more elementary process S. The following fundamental questions arise:

1. What is the nature of those limiting processes Z that may appear through such a scheme?
2. Given a limit process Z, what are all the one-step increment probability distributions μ whose associated random walk converges to Z after renormalization?
3. Given a one-step increment probability distribution, how to find a renormalization procedure that leads to an interesting non-trivial limit process Z, if any exists?

1.1 Review of some abelian results

In this chapter, we discuss some aspects of these vaguely stated questions in the context of finite dimensional vector spaces where detailed answers to the first two questions are well-known and understood. The answer to the first question involves the notions of infinitely divisible probability distribution and Lévy process, and the additional notion of operator stability which relates directly to the "normalization procedure" that allows us to pass from S to Z. See, e.g., [26, Chapter 6], [36, Section 1.6] and [47, Chapter 8]. The second question concerns the "domain of operator-attraction" of the limit Z and falls outside the scope of our interest. The third question is not easily answered in general (see [32]) but it plays an important role in the results we develop in this work for nilpotent groups. Indeed, for the particular class of examples we treat on nilpotent groups, a key step consists in identifying appropriate renormalization procedures.

Recall that an $\mathbb{R}^d$-valued random variable Y (or its probability distribution) is said to be *infinitely divisible* if, for each integer $n \geq 1$, there are i.i.d $\mathbb{R}^d$-valued random variables $\{X_1, \ldots, X_n\}$ such that $\sum_{k=1}^n X_k$ has the same distribution as Y. It is well known (see, e.g., [10, 39, 47, 56]) that the distribution of Y is infinitely divisible if and only if it is the distribution at time 1 of a Lévy process $Z = \{Z_t; t \geq 0\}$ with $Z_0 = 0$. An infinitely divisible probability is uniquely characterized by the Lévy exponent ϕ of its characteristic function $\phi(\lambda) := -\log \mathbb{E}\left[e^{i\lambda\cdot Y}\right]$, which takes the following form. There are a symmetric non-negative definite constant matrix $A = (a_{ij})_{1\leq i,j\leq d}$, a constant vector $b = (b_1, \ldots, b_d)$, and a non-negative Borel measure ν on $\mathbb{R}^d \setminus \{0\}$ satisfying $\int_{\mathbb{R}^d}(1 \wedge |z|^2)\nu(dz) < \infty$ so that

$$\phi(\lambda) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\lambda_i\lambda_j + \sum_{i=1}^{d} b_i\lambda_i + \int_{\mathbb{R}^d}\left(1 - e^{i\lambda\cdot z} + i\lambda\cdot z\mathbb{1}_{\{|z|\leq 1\}}\right)\nu(dz) \quad (1.1)$$

for any $\lambda = (\lambda_1, \ldots, \lambda_d) \in \mathbb{R}^d$. The triplet (A, b, ν) and the measure ν are called the Lévy triplet and the Lévy measure of the infinitely divisible distribution of Y, respectively. They are uniquely determined by Y, and vice-versa. See, e.g., [36, 1.3.2]. The expression (1.1) is called the Lévy-Khintchine formula for the infinitely divisible distribution of Y. We say the random variable Y has no Gaussian part if $A = 0$. Clearly, if the distribution of Y is symmetric, that is, $-Y$ has the same distribution as Y, then $b = 0$ and the Lévy measure ν is symmetric. We say that an $\mathbb{R}^d$-valued random variable X is full if there is no non-zero $\lambda \in \mathbb{R}^d$ so that $\lambda \cdot X$ is a constant, that is, if the distribution of X is not supported on any $(d-1)$-dimensional affine subspace of $\mathbb{R}^d$. An infinite divisible probability distribution having no Gaussian part is full if and only if its Lévy measure ν is not supported on any $d-1$ dimensional linear subspace of $\mathbb{R}^d$; see [47, Proposition 3.1.20]. In this work, we are interested in results involving limits that are symmetric with no Gaussian part (symmetric random walks with jumps having heavy tails).

We start with the following elegant result. Let $S = \{S_n; n \geq 0\}$ be a random walk in $\mathbb{Z}^d$ with i.i.d. steps $\{\xi_k; k \geq 1\}$ having distribution μ. That is,

$$\mathbb{P}(\xi_k = (j_1, \dots, j_d)) = \mu((j_1, \dots, j_d)) \quad \text{for } (j_1, \dots, j_d) \in \mathbb{Z}^d, \tag{1.2}$$

and $S_n = \xi_1 + \cdots + \xi_n$.

Proposition 1.1 ([47, Corollary 8.2.12]) Let η be a full infinitely divisible probability distribution on $\mathbb{R}^d$ with no Gaussian part and Lévy measure ν. Let $\{S_n; n \geq 0\}$ be a random walk in $\mathbb{R}^d$ driven by a probability measure μ as above.

There are linear operators $A_n : \mathbb{R}^d \to \mathbb{R}^d$ and vectors $b_n \in \mathbb{R}^d$ such that $A_n S_n + b_n$ converges in distribution to η if and only if

$$n\, \mu \circ A_n^{-1} \text{ converges vaguely to } \nu \text{ on } \mathbb{R}^d \setminus \{0\}. \tag{1.3}$$

In this case, $\lim_{n\to\infty} \|A_n\| = 0$. □

Here, $\mu \circ A_n^{-1}$ is the probability measure on $\mathbb{R}^d$ defined by

$$\mu \circ A_n^{-1}(B) = \mu(\{x \in \mathbb{R}^d : A_n x \in B\}) \quad \text{for every } B \in \mathcal{B}(\mathbb{R}^d).$$

Denote by $C_c(\mathbb{R}^d \setminus \{0\})$ the space of continuous functions on $\mathbb{R}^d \setminus \{0\}$ with compact support. Then (1.3) means that

$$\lim_{n\to\infty} n \int_{\mathbb{R}^d} f(A_n x)\mu(dx) = \int_{\mathbb{R}^d} f(x)\nu(dx) \quad \text{for any } f \in C_c(\mathbb{R}^d \setminus \{0\}), \tag{1.4}$$

Note that from the Lévy-Khintchine formula (1.1), two infinitely divisible random variables without Gaussian components and having the same Lévy measure ν can only differ by a constant vector.

If (1.3) holds, we say the Lévy measure ν is operator-stable (see below) and the measure μ (or equivalently ξ_1) belongs to the generalized domain of attraction of η (or ν, by abuse of language). The matrix A_n is automatically invertible for all large n. See [47, Lemma 3.3.25].

Remark 1.2 Suppose (1.3) holds with the Lévy measure ν not supported in a $(d-1)$-dimensional vector subspace and μ being symmetric (that is, $\mu(A) = \mu(-A)$).

(i) The vector b_n in Proposition 1.1 can be taken to be the zero vector in $\mathbb{R}^d$ and the limiting distribution η is symmetric. This is because in this case, $\{S_n; n \in \mathbb{N}\}$ has the same distribution as $\{-S_n; n \in \mathbb{N}\}$, and so, $\{A_n S_n - b_n; n \in \mathbb{N}\}$ has the same distribution as $\{-(A_n S_n + b_n); n \in \mathbb{N}\}$. Consequently, $\{A_n S_n - b_n; n \in \mathbb{N}\}$ also converges weakly. It then follows from the characterization of weak convergence that $\{A_n S_n; n \in \mathbb{N}\}$ converges weakly to a symmetric random variable η.

(ii) By [47, Theorem 8.1.5] and its proof, there are a sequence of invertible $d \times d$ matrices $(M_n)_{n\geq 1}$ that keeps the distribution of η invariant (that is, $M_n \eta$ has the same distribution as η for each $n \geq 1$) and a $d \times d$-matrix E with real entries such that $\widetilde{A}_n := M_n A_n$ satisfies

$$\lim_{n\to\infty} \widetilde{A}_{[\lambda n]} \widetilde{A}_n^{-1} = \lambda^E \quad \text{for all } \lambda > 0, \tag{1.5}$$

and $\widetilde{A}_n S_n$ converges in distribution to η as $n \to \infty$. Here, $[a]$ stands for the largest integer not exceeding the real a.
Using the independent stationary increments property of random walks, we can easily deduce from Proposition 1.1that both $\{A_n S_{[nt]}; t \geq 0\}$ and $\{\widetilde{A}_n S_{[nt]}; t \geq 0\}$ converge in finite dimensional distributions to the symmetric Lévy process $Z = \{Z_t; t \geq 0\}$ with Z_1 having the same distribution as η; see the proof of Proposition 1.3. Further, Z has the following scaling property by (1.5): for any $\lambda > 0$,

$$\{Z_{\lambda t}; t \geq 0\} \text{ has the same distribution as } \lambda^E Z = \{\lambda^E Z_t; t \geq 0\}.$$

See [47, Example 11.2.18] and [48, p.625]. For this reason, the Lévy process Z is called an operator-stable process (or operator-Lévy motion) and its Lévy measure, ν, is also said to be operator-stable in the literature. If $E = \alpha^{-1} I_{d\times d}$, where $I_{d\times d}$ denotes the $d \times d$ identity matrix, $\lambda^E = \lambda^{1/\alpha} I_{d\times d}$. In this case, Z is an α-stable Lévy process on $\mathbb{R}^d$.

(iii) The matrices $\{A_n; n \in \mathbb{N}\}$ and the limiting Lévy measure ν in (1.3) are not unique. Suppose (1.3) holds. Then for any non-degenerate matrix M, we clearly have that $n\,\mu \circ (MA_n)^{-1}$ converges vaguely to $\nu \circ M^{-1}$ on $\mathbb{R}^d \setminus \{0\}$. Thus ν depends not only on μ but also on the "dilation structure" A_n. □

Denote by $\mathbb{D}([0,\infty); \mathbb{R}^d)$ the space of right continuous $\mathbb{R}^d$-valued functions on $[0,\infty)$ having left limits. We refer the reader to [28] for the definition of $\mathcal{J}_1$-topology on the Skorohod space $\mathbb{D}([0,\infty); \mathbb{R}^d)$.

Proposition 1.3 Suppose that the one step distribution μ of the random walk $\{S_n; n = 0, 1, 2, \ldots\}$ is symmetric and satisfies condition (1.3). Let η be an infinitely divisible symmetric probability distribution with no Gaussian component and Lévy measure ν. Let $Z = \{Z_t; t \geq 0\}$ be the symmetric Lévy process on $\mathbb{R}^d$ so that Z_1 has distribution η. Then $\{A_n S_{[nt]}; t \geq 0\}$ converges weakly in the Skorohod space $\mathbb{D}([0,\infty); \mathbb{R}^d)$ equipped with $\mathcal{J}_1$-topology to the Lévy process Z as $n \to \infty$. □

Proof Let $\widetilde{A}_n = M_n A_n$ be defined as in Remark 1.2(ii), where $(M_n)_{n\geq 1}$ is a sequence of invertible matrices that keeps the distribution of η invariant. We know from [48, Theorem 4.1] that $\{\widetilde{A}_n S_{[nt]}; t \geq 0\}$ converges weakly in the Skorohod space $\mathbb{D}([0,\infty); \mathbb{R}^d)$ equipped with $\mathcal{J}_1$-topology to Z as $n \to \infty$. By [47, Theorem 3.2.10], $(M_n)_{n\geq 1}$ is relatively compact in the spaces of invertible $d \times d$-matrices. Thus for any subsequence of $(n)_{n\geq 1}$, there is a sub-subsequence $(n')_{n'\geq 1}$ so that $M_{n'}$ converges to a non-degenerate $d \times d$-matrix M that also keeps the distribution of η and hence its Lévy measure ν invariant. Note that $A_n = M_n^{-1} \widetilde{A}_n$ and the Lévy process $M^{-1} Z$ is of the same distribution as that of Z. It follows that $\{A_{n'} S_{[n't]}; t \geq 0\}$ converges weakly in the Skorohod space $\mathbb{D}([0,\infty); \mathbb{R}^d)$ equipped with $\mathcal{J}_1$-topology to Z as $n' \to \infty$. Since this holds for any subsequence of n, we conclude that $\{A_n S_{[nt]}; t \geq 0\}$ converges weakly to Z as $n \to \infty$. □

The two propositions above and the accompanying remarks tell us that, if we expect that a given symmetric measure μ on $\mathbb{Z}^d$ drives a random walk whose functional limit process $Z = \{Z_t; t \geq 0\}$ has no Gaussian part and Lévy measure ν, we should concentrate on finding the sequence of invertible matrices A_n such that (1.3) holds. Indeed, that property is necessary and sufficient for the desired limit theorems to hold.

1.2 Illustrative examples on nilpotent matrix groups

In this section, we describe some illustrative examples, let us emphasize that, although one can easily formulate versions of Proposition 1.1 in the context of certain nilpotent groups, it is not known if such generalizations hold true. In a similar vein, in $\mathbb{R}^d$, a full operator-stable Lévy process always admits a smooth density whereas in the context of nilpotent group, it is not known if a full operator-stable Lévy process always has a density. For details on how to formulate these questions more precisely on nilpotent groups, see [36, Chapter 2].

Example 1.4 In this example, we consider a random walk on $\mathbb{Z}^3$ with i.i.d. steps $\{\xi_k; k \geq 1\}$ distributed according to the probability measure μ concentrated along the coordinate axes of $\mathbb{Z}^3$ given by

$$\mu((i_1,i_2,i_3)) = \frac{\kappa_1}{(1+|i_1|)^{1+\alpha_1}}\mathbb{1}_{\{i_2=i_3=0\}} + \frac{\kappa_2}{(1+|i_2|)^{1+\alpha_2}}\mathbb{1}_{\{i_1=i_3=0\}} + \frac{\kappa_3}{(1+|i_3|)^{1+\alpha_3}}\mathbb{1}_{\{i_1=i_2=0\}} \quad \text{for } (i_1,i_2,i_3) \in \mathbb{Z}^3 \setminus \{0\}.$$

We assume $\alpha_i \in (0,2)$, $i = 1,2,3$. Let $A_n = \begin{pmatrix} n^{-1/\alpha_1} & 0 & 0 \\ 0 & n^{-1/\alpha_2} & 0 \\ 0 & 0 & n^{-1/\alpha_3} \end{pmatrix}$. It is easy to check that

$$n\mu \circ A_n^{-1} \text{ converges vaguely to } \nu \text{ on } \mathbb{R}^3 \setminus \{0\},$$

where

$$\nu(dx) = \sum_{i=1}^{3} \frac{\kappa_i}{|x_i|^{1+\alpha_i}} dx_i \otimes_{j\in\{1,2,3\}\setminus\{i\}} \delta_{\{0\}}(dx_j).$$

Here $\delta_{\{0\}}$ is the Dirac measure concentrated at 0. Since μ is symmetric, by Proposition 1.1and Remark 1.2(i), $A_n S_n$ converges weakly to a random vector whose distribution is symmetric infinite divisible with no Gaussian part and Lévy measure ν. By Proposition 1.3, $\{A_n S_{[nt]}; t \geq 0\}$ converges weakly in the Skorohod space $\mathbb{D}([0,\infty);\mathbb{R}^3)$ equipped with $\mathcal{J}_1$-topology to the purely discontinuous symmetric Lévy process $Z = \{Z_t; t \geq 0\}$ having ν as its Lévy measure. Note that the coordinate processes of $Z = (Z^{(i)})_1^3$ are independent with $Z^{(i)}$ being a one-dimensional symmetric α_i-stable process, $1 \leq i \leq 3$.

As pointed out earlier in Remark 1.2(iii), it is worth noting that the choice of A_n above is not unique even though it seems most natural in this example. To simplify the discussion, assume that $\alpha_1 < \alpha_2 < \alpha_3$. Let (e_1, e_2, e_3) be the canonical basis of $\mathbb{R}^3$ used implicitly above. Construct a linear operator B_n as follows. First, set

$B_n e_1 = n^{-1/\alpha_1} e_1$. Second, pick an arbitrary vector e_2' which is linearly independent from e_1 and belongs to the plane spanned by e_1 and e_2, and set $B_n e_2' = n^{-1/\alpha_2} e_2'$. Finally, pick an arbitrary non-zero vector e_3' that does not belong to the plane spanned by e_1 and e_2, and set $B_n e_3' = n^{-1/\alpha_3} e_3'$. Then $n\mu \circ B_n^{-1}$ converges vaguely to a Lévy measure ν' having essentially the same form as ν but carried by the axes associated with e_1, e_2', e_3'. More precisely,

$$\nu'(dx) = \sum_{i=1}^{3} \frac{c_i}{|x_i'|^{1+\alpha_i}} dx_i' \otimes_{j\in\{1,2,3\}\setminus\{i\}} \delta_{\{0\}}(dx_j'),$$

where (x_1', x_2', x_3') is the coordinate of $x \in \mathbb{R}^3$ under the coordinate system (e_1, e_2', e_3'). Note that the Lévy measure ν' is thus a linear transformation of ν.

Example 1.5 (Random walk on the Heisenberg group $\mathbb{H}_3(\mathbb{Z})$) Recall that the discrete Heisenberg group $\mathbb{H}_3(\mathbb{Z})$ is the family of upper triangle matrices of the form $\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$, with $x, y, z \in \mathbb{Z}$, equipped with matrix multiplication; that is

$$\begin{pmatrix} 1 & x_1 & z_1 \\ 0 & 1 & y_1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & x_2 & z_2 \\ 0 & 1 & y_2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x_1 + x_2 & z_1 + z_2 + x_1 y_2 \\ 0 & 1 & y_1 + y_2 \\ 0 & 0 & 1 \end{pmatrix}.$$

For $a = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$, its inverse a^{-1} is $\begin{pmatrix} 1 & -x & xy - z \\ 0 & 1 & -y \\ 0 & 0 & 1 \end{pmatrix}$. If we identify matrix a with (x, y, z), then the discrete Heisenberg group $\mathbb{H}_3(\mathbb{Z})$ can be identified with $\mathbb{Z}^3$ equipped with the group multiplication

$$(x_1, y_1, z_1) \cdot (x_2, y_2, z_2) := (x_1 + x_2, y_1 + y_2, z_1 + z_2 + x_1 y_2). \tag{1.6}$$

We will use this realization of $\mathbb{H}_3(\mathbb{Z})$. This is one of the simplest example of a non-abelian nilpotent group.

Let $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ and $e_3 = (0, 0, 1)$, which are generators of $\mathbb{H}_3(\mathbb{Z})$. Note that for $k \in \mathbb{Z} \setminus \{0\}$, $e_1^k = (k, 0, 0)$, $e_2 = (0, k, 0)$ and $e_3^k = (0, 0, k)$. Let $\alpha_k \in (0, 2)$ be a constant, $1 \le k \le 3$, and write $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$. Consider the following probability measure on $\mathbb{H}_3(\mathbb{Z}) = \mathbb{Z}^3$:

$$\mu_{\boldsymbol{\alpha}}(g) = \sum_{i=1}^{3} \sum_{n\in\mathbb{Z}} \frac{\kappa_i}{(1 + |n|)^{1+\alpha_i}} \mathbb{1}_{\{e_i^n\}}(g), \quad g \in \mathbb{H}_3(\mathbb{Z}),$$

where κ_j, $1 \le j \le 3$, are appropriate positive constants. Let $\left(\xi_k = (\xi_k^{(1)}, \xi_k^{(2)}, \xi_k^{(3)})\right)_{k\ge 1}$ be an i.i.d sequence of random variables taking values in $\mathbb{H}_3(\mathbb{Z})$ of distribution $\mu_{\boldsymbol{\alpha}}$. Then $S_n = S_0 \cdot \xi_1 \cdot \ldots \cdot \xi_n$, $\quad n \ge 1$, defines a random walk on the Heisenberg group

$\mathbb{H}_3(\mathbb{Z})$. Write S_n as (X_n, Y_n, Z_n). By (1.6),

$$X_{n+1} = X_n + \xi^{(1)}_{n+1}, \quad Y_{n+1} = Y_n + \xi^{(2)}_{n+1}, \quad Z_{n+1} = Z_n + \xi^{(3)}_{n+1} + X_n \xi^{(2)}_{n+1}. \tag{1.7}$$

If we define $\widehat{Z}_n = Z_0 + \sum_{k=1}^n \xi^{(3)}_k$, then

$$Z_n = \widehat{Z}_n + \sum_{k=1}^n X_{k-1}\xi^{(2)}_k = \widehat{Z}_n + \sum_{k=1}^n X_{k-1}(Y_k - Y_{k-1}), \quad n \geq 1. \tag{1.8}$$

We know from Example 1.4 that

$$\left\{ \left(n^{-1/\alpha_1} X_{[nt]},\, n^{-1/\alpha_2} Y_{[nt]},\, n^{-1/\alpha_3} \widehat{Z}_{[nt]}, \right); t \geq 0 \right\} \Longrightarrow \{(\bar{X}_t, \bar{Y}_t, \bar{Z}_t),\, t \geq 0\} \tag{1.9}$$

weakly in the Skorohod space $\mathbb{D}([0,\infty), \mathbb{R}^3)$ equipped with $\mathcal{J}_1$-topology as $n \to \infty$, where $\bar{X}$, $\bar{Y}$, $\bar{Z}$ are symmetric α_1-, α_2- and α_3-stable processes on $\mathbb{R}$, respectively, and they are independent. For simplicity, let

$$\widetilde{X}^n_t := n^{-1/\alpha_1} X_{[nt]}, \quad \widetilde{Y}^n_t := n^{-1/\alpha_2} Y_{[nt]} \quad \text{and} \quad \widetilde{Z}^n_t := n^{-1/\alpha_2} \widehat{Z}_{[nt]}.$$

Now, we can use the following key facts. Lévy processes are semimartingales so stochastic integrals such as Lévy area $\int_0^t \bar{X}_{s-} d\bar{Y}_s$ are well-defined. Furthermore, [44, Theorem 7.10] shows that $\left\{ \left(\widetilde{X}^n_t, \widetilde{Y}^n_t, \widetilde{Z}^n_t, \int_0^t \widetilde{X}^n_{s-} d\widetilde{Y}^n_s \right);\ t \geq 0 \right\}$ converges weakly in the Skorohod space $\mathbb{D}([0,\infty); \mathbb{R}^4)$ equipped with $\mathcal{J}_1$-topology as $n \to \infty$ to

$$\left\{ \left(\bar{X}_t, \bar{Y}_t, \bar{Z}_t, \int_0^t \bar{X}_{s-} d\bar{Y}_s \right);\ t \geq 0 \right\}. \tag{1.10}$$

Indeed, to prove (1.10), for any $\delta > 0$, let $h_\delta(r) = (1 - \delta/r)^+$. Define

$$\widetilde{X}^{n,\delta}_t = \widetilde{X}^n_t - \sum_{0<s\leq t} h_\delta(|\Delta \widetilde{X}^n_s|)\Delta \widetilde{X}^n_s \quad \text{and} \quad \widetilde{Y}^{n,\delta}_t = \widetilde{Y}^n_t - \sum_{0<s\leq t} h_\delta(|\Delta \widetilde{Y}^n_s|)\Delta \widetilde{Y}^n_s.$$

One can define $\widetilde{Z}^{n,\delta}$ in a similar way. Observe that $\widetilde{X}^{n,\delta}$, $\widetilde{Y}^{n,\delta}$ and $\widetilde{Z}^{n,\delta}$ are again symmetric random walks but with i.i.d step sizes

$$\left\{ \left(1 - h_\delta(n^{-1/\alpha_j} n^{-1/\alpha_j} \xi^{(j)}_k) \right) n^{-1/\alpha_j} \xi^{(j)}_k;\, k \geq 1 \right\} \quad \text{for } j = 1, 2, 3,$$

respectively. Let $[\widetilde{X}^{n,\delta}]$, $[\widetilde{Y}^{n,\delta}]$ and $[\widetilde{Z}^{n,\delta}]$ denote the quadratic variation processes of the square integrable martingales $\widetilde{X}^{n,\delta}$, $\widetilde{Y}^{n,\delta}$ and $\widetilde{Z}^{n,\delta}$, respectively. Note that

$$\mathbb{E}\left(\left[\widetilde{X}^{n,\delta} \right]_t \right) = [nt]\, \mathbb{E}\left[\left(1 - h_\delta(n^{-1/\alpha_1} \xi^{(1)}_1) \right)^2 \left(n^{-1/\alpha_1} \xi^{(1)}_1 \right)^2 \right]$$

$$
\begin{aligned}
&\leq c_1\kappa_1 n^{-2/\alpha_1}[nt]\left(\sum_{k=1}^{[n^{1/\alpha_1}\delta]}\frac{k^2}{(1+k)^{1+\alpha_1}}+\sum_{[n^{1/\alpha_1}\delta]+1}^{\infty}\frac{n^{2/\alpha_1}\delta^2}{k^2}\frac{k^2}{(1+k)^{1+\alpha_1}}\right)\\
&\leq c_1\kappa_1 n^{-2/\alpha_1}[nt]\left(\frac{[n^{1/\alpha_1}\delta]^{2-\alpha_1}}{2-\alpha_1}+\frac{n^{2/\alpha_1}\delta^2}{\alpha_1(1+[n^{1/\alpha_1}\delta])^{\alpha_1}}\right)\\
&\leq c_1\kappa_1 n^{1-2/\alpha_1}t\left(\frac{(n^{1/\alpha_1}\delta)^{2-\alpha_1}}{2-\alpha_1}+\frac{n^{2/\alpha_1}\delta^2}{\alpha_1(n^{1/\alpha_1}\delta)^{\alpha_1}}\right)\\
&=\frac{2c_1\kappa_1\delta^{2-\alpha_1}}{\alpha_1(2-\alpha_1)}\,t,
\end{aligned}
$$

where $c_1 > 0$ is a constant independent of n and δ. In the same way, there is a constant $c_k > 0$, $k = 2,3$, independent of n and δ so that

$$
\mathbb{E}\left(\left[\widetilde{Y}^{n,\delta}\right]_t\right)\leq\frac{2c_2\kappa_2\delta^{2-\alpha_2}}{\alpha_2(2-\alpha_2)}\,t\quad\text{and}\quad\mathbb{E}\left(\left[\widetilde{Z}^{n,\delta}\right]_t\right)\leq\frac{2c_3\kappa_3\delta^{2-\alpha_3}}{\alpha_3(2-\alpha_3)}\,t
$$

for all $n \geq 1$ and $t > 0$. So these three sequences of square integrable martingales $\{\widetilde{X}^n; n \geq 1\}$, $\{\widetilde{Y}^n; n \geq 1\}$ and $\{\widetilde{Z}^n; n \geq 1\}$ have uniformly controlled variations in the sense of [44, Definition 7.5]. Thus by taking $(\widetilde{X}^n_t, 0)$ and $(\widetilde{Y}^n, \widetilde{Z}_n)$ for the vector-valued process H^n and X^n in [44, Theorem 7.10], we conclude $\{(\widetilde{X}^n_t, \widetilde{Y}^n_t, \widetilde{Z}^n_t, \int_0^t \widetilde{X}^n_{s-}d\widetilde{Y}^n_s); t \geq 0\}$ converges weakly in the Skorohod space $\mathbb{D}([0,\infty);\mathbb{R}^4)$ equipped with $\mathcal{J}_1$-topology to $\{(\bar{X}_t, \bar{Y}_t, \bar{Z}_t, \int_0^t \bar{X}_{s-}d\bar{Y}_s); t \geq 0\}$. This proves the claim (1.10).

Using the almost sure Skorohod representation theorem, we can assume without loss of generality that $\left\{\left(\widetilde{X}^n_t, \widetilde{Y}^n_t, \widetilde{Z}^n, \int_0^t \widetilde{X}^n_{s-}d\widetilde{Y}^n_s\right);\ t \geq 0\right\}$ converges a.s. in the Skorohod space $\mathbb{D}([0,\infty);\mathbb{R}^4)$ as $n \to \infty$ to $\left\{\left(\bar{X}_t, \bar{Y}_t, \bar{Z}_t, \int_0^t \bar{X}_{s-}d\bar{Y}_s\right);\ t \geq 0\right\}$. Consequently, we have the following conclusions. The weak convergence below (denoted by $\Longrightarrow$) is in the Skorohod space $\mathbb{D}([0,\infty);\mathbb{R}^3)$ equipped with $\mathcal{J}_1$-topology.

(i) If $1/\alpha_3 < 1/\alpha_1 + 1/\alpha_2$,

$$
\begin{aligned}
&\left\{\left(n^{-1/\alpha_1}X_{[nt]},\, n^{-1/\alpha_2}Y_{[nt]},\, n^{-1/\alpha_1-1/\alpha_2}Z_{[nt]}\right); t\geq 0\right\}\\
&\qquad\Longrightarrow\left\{\left(\bar{X}_t,\bar{Y}_t,\int_0^t\bar{X}_{s-}d\bar{Y}_s\right);\ t\geq 0\right\}\quad\text{as } n\to\infty;
\end{aligned}
$$

(ii) If $1/\alpha_3 = 1/\alpha_1 + 1/\alpha_2$,

$$
\begin{aligned}
&\left\{\left(n^{-1/\alpha_1}X_{[nt]},\, n^{-1/\alpha_2}Y_{[nt]},\, n^{-1/\alpha_3}Z_{[nt]}\right); t\geq 0\right\}\\
&\qquad\Longrightarrow\left\{\left(\bar{X}_t,\bar{Y}_t,\bar{Z}_t+\int_0^t\bar{X}_{s-}d\bar{Y}_s\right);\ t\geq 0\right\}\quad\text{as } n\to\infty;
\end{aligned}
$$

(iii) If $1/\alpha_3 > 1/\alpha_1 + 1/\alpha_2$,

$$\left\{ \left(n^{-1/\alpha_1} X_{[nt]},\, n^{-1/\alpha_2} Y_{[nt]},\, n^{-1/\alpha_3} Z_{[nt]}\right); t \geq 0\right\}$$
$$\Longrightarrow \left\{\left(\bar{X}_t, \bar{Y}_t, \bar{Z}_t\right);\ t \geq 0\right\} \quad \text{as } n \to \infty. \qquad \square$$

Let us interpret the results above in group theoretical terms. In the treatment above, we have taken the coordinate components of the measure μ and considered the one-dimensional random walks, X, Y, Z, independently of each other. We have then reconstructed the group law effect of the random walk on $\mathbb{H}_3(\mathbb{Z})$ by considering the Lévy area generated by the X and Y components. This is easy to do in this case because the Z component commutes with anything else (it is in the center of the group). Now, the renormalization process involved is to make particular somewhat ad-hoc choices of scalings.

In the first two cases, (i)-(ii), we used the anisotropic dilations

$$\delta_t((x,y,z)) = (t^{-1/\alpha_1} x, t^{-1/\alpha_2} y, t^{-(1/\alpha_1+1/\alpha_2)} z), \quad t > 0.$$

This one parameter group of diffeomorphisms has the very special property of being a one parameter group of automophisms of $\mathbb{H}_3(\mathbb{R})$. That is,

$$\delta_t((x,y,z)\cdot(x',y',z')) = \delta_t((x,y,z))\cdot\delta_t((x',y',z')).$$

The consequence of this property is that the limit processes obtained above, $\left\{\left(\bar{X}_t, \bar{Y}_t, \int_0^t \bar{X}_{s-} d\bar{Y}_s\right); t \geq 0\right\}$ in case (i), $\left\{\left(\bar{X}_t, \bar{Y}_t, \bar{Z}_t + \int_0^t \bar{X}_{s-} d\bar{Y}_s\right); t \geq 0\right\}$ in case (ii), are symmetric Lévy processes on the real nilpotent group $\mathbb{H}_3(\mathbb{R})$ which are operator-stable with respect to the one parameter group of automorphisms $\{\delta_t : t > 0\}$. See [36, Chapter 2, Definition 2.3.13].

In the third case when $1/\alpha + 1/\alpha_2 < 1/\alpha_3$, we used

$$\delta_t((x,y,z)) = (t^{-1/\alpha_1} x, t^{-1/\alpha_2} y, t^{-1/\alpha_3} z), \quad t > 0.$$

These diffeomorphisms are not automorphisms of $\mathbb{H}_3(\mathbb{R})$ and it follows that using them in rescaling the random walk driven by μ on $\mathbb{H}_3(\mathbb{Z}) \subset \mathbb{H}_3(\mathbb{R})$ produces a non-trivial change in the underlying group structure. This is visible in the nature of the limiting process, $\left\{\left(\bar{X}_t, \bar{Y}_t, \bar{Z}_t\right);\ t \geq 0\right\}$, which is not a Lévy process on $\mathbb{H}_3(\mathbb{R})$ but a Lévy process on the abelian group $\mathbb{R}^3$.

Although it is certainly possible to push this approach further in specific examples, there are serious difficulties in treating large classes of examples in this way. For this reason, the approach presented in this monograph is quite different. It does not involve explicitly the stochastic calculus involved in studying the Lévy area and higher degree functionals of the same type that are known to appear when expressing random walks on nilpotent groups in coordinates. The interested reader might try the following two informal exercises before reading further.

Exercise 1.6 Pick a tuple of 10 elements $(s_1, \ldots, s_{10})$ in either $\mathbb{Z}^3$ or in $\mathbb{H}_3(\mathbb{Z})$, $s_i = (x_i, y_i, z_i)$, and a tuple of ten reals $\alpha_i \in (0,2)$, $1 \leq i \leq 10$. Consider the probability measure

$$\mu(g) = \sum_{i=1}^{10} \sum_{n\in\mathbb{Z}} \frac{\kappa_i}{(1+|n|)^{1+\alpha_i}} \mathbb{1}_{\{s_i^n\}}(g).$$

What to do to formulate a limit theorem? in $\mathbb{Z}^3$? in $\mathbb{H}_3(\mathbb{Z})$? □

Exercise 1.7 Repeat Exercise 1.6 with $\mathbb{H}_3(\mathbb{Z})$ replaced by the group of four by four upper-triangular matrices with diagonal entries equal to 1 (this group is nilpotent).□

As the reader will see, the approach developed in this work is amenable to detailed computation in concrete cases. Using the theory developed in this monograph, we will revisit Example 1.5 in Section 7.3.

We close this preliminary chapter by describing the organization of this monograph. The next chapter provides an introduction to our main results while avoiding most technical details. In particular, Section 2.3 describes special cases which we hope the reader will find both interesting and informative, and Section 2.5 discusses prior results. Section 3.1 introduces polynomial coordinate systems and the key notions of group dilation and approximate group dilation relative to such a coordinate system. Approximate group dilations lead to the definition of "limit group structures" and we present some basic properties of these limit group structures that are important for our purpose. Chapter 4 introduces the vague convergence of a probability measures under rescaling by an approximate group dilation and how the vague limit and the limit group structure interact (see Proposition 4.7). Chapter 5 describes our main technical results concerning functional limit theorem. It identifies a list of strong hypothesss that allows us to state such a theorem. See Theorem 5.10. Chapter 6 presents the corresponding local limit theorem, Theorem 6.2. Chapter 7 describes how to identify in concrete terms (in coordinates), the limiting Lévy process (on the associated limit group). They are then used together with the main results of this monograph to give several examples on the weak convergence of long range random walks on various nilpotent groups. Chapter 8 describes the main class of probability measures, $\mathcal{SM}$, to which we want to apply the results obtained in previous chapters. Chapter 9 shows how to choose appropriate coordinate systems and dilations for measures in $\mathcal{SM}$ whereas Chapter 10 demonstrates that the hypotheses needed in Chapters 5 -6 are essentially satisfied by measures in $\mathcal{SM}$.

Notation

We use := as a way of definition. For $a, b \in \mathbb{R}$, $a \wedge b := \min\{a, b\}$. We use $\delta_{\{x_0\}}$ to denote the Dirac measure concentrated at $x_0 \in \mathbb{R}^d$, and $\mathbb{1}_A$ for the indicator function of a Borel measurable set $A \subset \mathbb{R}^d$. For an open subset $D \subset \mathbb{R}^d$, the space of bounded continuous functions on D and the space of continuous functions on D with compact support will be denoted by $C_b(D)$ and $C_c(D)$, respectively.

Chapter 2
Introduction

Abstract This chapter introduces the particular problems studied in this book. A set of compelling special cases are presented in order to describe the ingredients used in our study and to illustrate the results we obtained. The chapter ends with a quick review of what is known and how it differs from what is presented in the book.

2.1 Basic question

The aim of this work is to prove limit theorems for a class of random walks on nilpotent groups driven by probability measures allowing for long jumps in certain directions. The class of probability measures we study can be described roughly as follows. Let Γ be a finitely generated nilpotent group with neutral element e. Assume that we are given a finite family of subgroups of Γ, $H_1, \dots, H_k$, each equipped with a finite symmetric generating set S_i and the associated word-length $|\cdot|_{H_i,S_i} = |\cdot|_i$. For each $i \in \{1, \dots, k\}$, fix $\alpha_i \in (0, 2)$. On each H_i, set $V_i(r) = \#\{g \in H_i : |g|_i \le r\}$ and consider the probability measure on H_i:

$$\mu_i(g) = \frac{c_i}{(1 + |g|_i)^{\alpha_i} V_i(|g|_i)}, \quad g \in H_i.$$

Now, on Γ, consider the symmetric probability measure

$$\mu = \sum_{i=1}^{k} \lambda_i \mu_i,$$

where λ_i's are positive constants with $\sum_{i=1}^k \lambda_i = 1$. The class of measures we will treat is slightly larger than what we just described. Two special cases of this construction are particularly compelling. The first is the case when $k = 1$, $H_1 = \Gamma$, $S_1 = S$ is a finite symmetric generating set for Γ, and $\mu(g) = \frac{c_\Gamma}{(1+|g|_S)^\alpha V_\Gamma(|g|_S)}$. This is reminiscent of radially symmetric α-stable process. The second is the case when each H_i is an infinite cyclic subgroup in Γ, a case reminiscent of more singular symmetric operator-stable process whose coordinate processes are independent to each other. See [15, 39, 47, 55].

In an earlier work [20], we proved that there are a positive constant $\gamma_0 = \gamma_0(\mu)$ (which can be computed relatively easily from the data) and positive constants $c = c(\mu), C = C(\mu)$ such that

$$cn^{-\gamma_0} \le \mu^{(n)}(e) \le Cn^{-\gamma_0}.$$

Here $\mu^{(n)}$ is the n-fold convolution power of the measure μ. One motivation for the present work is to provide the more precise asymptotic

$$\lim_{n\to\infty} n^{\gamma_0}\mu^{(n)}(e) = a(\mu)$$

with, hopefully, a description of the constant $a(\mu)$. One classical approach to such problems is to find a way to rescale the random walk on Γ so as to obtain some sort of limit theorem proving convergence of the law of the rescaled random walk towards the law of a limit process on an appropriate limit space. Typically, the limit space and the limit process will have some self-similarity properties with respect to some scaling structure. In the most classical cases, e.g., when μ is a symmetric probability measure on $\mathbb{Z}^d$ which drives a symmetric random walk converging towards some symmetric stable process on $\mathbb{R}^d$, the limit space supporting the limit process, and its group law, are always the same, $(\mathbb{R}^d, +)$, independently of μ. In the present context, one interesting new phenomenon is that the group structure of the limit space supporting the limit process depends not only of the discrete group Γ but also on the measure μ.

2.2 Description of the basic ingredients and results

For simplicity, in this work, we restrict ourselves to random walks on torsion-free finitely generated nilpotent groups, that is, finitely generated nilpotent groups whose only element of finite order is the identity element. These countable groups are both similar to and more complicated than the square lattice $\mathbb{Z}^d$ in $\mathbb{R}^d$. Let Γ be such a group. By a celebrated theorem of Malcev [46], the countable group Γ can be realized as a co-compact discrete subgroup of a simply connected nilpotent Lie group G. Moreover, any simply connected nilpotent Lie group G can be identified with the d-dimensional coordinate space $\mathbb{R}^d$ equipped with an appropriate group structure whose (multiplication) law is given, in coordinates, by polynomial functions. This accounts for the similarity with the square lattice in dimension d. Note however that the description of G as $\mathbb{R}^d$ equipped with a polynomial product is very far from being unique (and it may sometimes be difficult to recognize that two such descriptions give the same group G up to isomorphism). One way to understand the complexity of such structures is to attempt to give a list of all non-isomorphic simply connected nilpotent groups in a fix dimension d. No such lists exist for relatively large d (we are not aware of such lists when d is greater than 8). See [11] and the references therein.

Once Γ is represented as a subset of $\mathbb{R}^d$, a probability measure μ on Γ can be viewed as a weighted series of Dirac masses on $\mathbb{R}^d$. For such a measure μ in a certain relatively large class of “stable-like” probability measures on Γ, we are going to find an adapted dilation structure $(\delta_t^\mu)_{t>0}$, expressed in coordinates over $G = (\mathbb{R}^d, \cdot)$ by

$\delta_t^\mu(u) = (t^{1/\alpha_i^\mu} u_i)_1^d$, with carefully chosen exponents $\alpha_i^\mu \in (0,2)$, so that the measure

$$\mu_t = t\delta_{1/t}^\mu(\mu) : \ \phi \mapsto t \int_{\mathbb{R}^d} \phi(\delta_{1/t}^\mu(u))\mu(du)$$

has a vague limit $\mu_\bullet$ (a non-negative Radon measure) on $\mathbb{R}^d \setminus \{0\}$ as t tends to ∞. By construction, the limit $\mu_\bullet$ will satisfy the self-similar property

$$(\mu_\bullet)_t = \mu_\bullet \quad \text{for any } t > 0.$$

At the same time, the rescaled group laws

$$x \cdot_t y = \delta_{1/t}^\mu \left(\delta_t^\mu(x) \cdot \delta_t^\mu(y)\right), \quad x, y \in \mathbb{R}^d, \quad t > 0,$$

will have a limit as t tends to infinity

$$\lim_{t\to\infty} x \cdot_t y = x \bullet^\mu y,$$

which defines a group law $\bullet^\mu$ on $\mathbb{R}^d$. Most of the time, we will drop the reference to μ and write $\bullet^\mu = \bullet$ but it is an essential feature of this work that this limit law actually depends on μ via the choice of a proper dilation structure. It will automatically have the self-similar property

$$x \bullet^\mu y = \delta_{1/t}^\mu \left(\delta_t^\mu(x) \bullet^\mu \delta_t^\mu(y)\right) \quad \text{for every } x, y \in \mathbb{R}^d \text{ and } t > 0.$$

Of course, we are most interested in cases when this can be done in such a way that the symmetric measure $\mu_\bullet$ is not supported on a proper closed connected subgroup of $G_\bullet = (\mathbb{R}^d, \bullet^\mu)$. In general, the limit measure $\mu_\bullet$ defines a left-invariant jump process on the group $G_\bullet$ and the key results of this monograph are:

1. A "stable-like" limit theorem expressing the convergence of the rescaled long jump random walk on Γ associated with μ to the left-invariant Lévy process on the group $(\mathbb{R}^d, \bullet^\mu)$ associated with $\mu_\bullet$.
2. A characterization of the left-invariant Lévy process on the nilpotent group $(\mathbb{R}^d, \bullet^\mu)$ associated with $\mu_\bullet$.
3. A companion local limit theorem providing a proper statement of convergence relating the densities of the distributions of these processes.

The reader should be warned that, given μ, the choice of the appropriate dilation structure $(\delta_t^\mu)_{t>0}$ is not unique and that, consequently, we have made various abuse of notation in the explanations given above.

The simplest instances of these results are the well-known convergence theorems relating the "stable-like" random walk on $\mathbb{Z}$ associated with the probability measure $\mu(x) = c_\alpha(1 + |x|)^{-1-\alpha}, x \in \mathbb{Z}$, $\alpha \in (0,2)$, to the symmetric α-stable process on $\mathbb{R}$, and its rather rich and complex extension to higher dimensions which includes both rotationally symmetric stable processes and some more singular operator stable processes as illustrated in Chapter 1. See also [39, 47, 48]. We note that, in so far as this monograph focusses on a particular class of probability measures, it only offers a limited extension of these classical abelian theories to nilpotent groups.

2.3 Detailed description of some special cases

In this section, we spell out in an informal way how our results of this monograph apply to a series of specific examples that are of particular interest. These cases all illustrate our main result, Theorem 10.1, which follows from Theorems 5.10 and 6.2, and the discussions in Sections 10.2-10.4.

Word length radial stable walks

On a finitely generated group Γ equipped with a symmetric finite generating set S, the word length $|g|_S$ is the minimal length k of a string $(g_1, \ldots, g_k)$ of elements of S such that g is equal to the product of that string, $g = g_1 \ldots g_k$. By Gromov's polynomial volume growth theorem [34], to say that Γ has polynomial volume growth is equivalent to the fact that there are an integer D (independent of S) and constants $0 < c_S \leq C_S < \infty$ such that

$$c_S r^D \leq \#\{g \in \Gamma : |g|_S \leq r\} \leq C_S r^D \quad \text{for all } r \geq 1.$$

This is known to hold for any finitely generated nilpotent group; see Section A.4. In this context, we call *word length radial stable probability measure of index* $\alpha \in (0, 2)$ the probability measure

$$\mu_{S,\alpha}(g) = \frac{c(\Gamma, S, \alpha)}{(1 + |g|_S)^{\alpha + D}}, \quad g \in \Gamma.$$

It is known (see [55, Section 5.1] and [49, Theorem 1,1] as well as the references given therein) that there are constants $0 < a = a(\Gamma, S) \leq A = A(\Gamma, S) < \infty$ such that the iterated convolutions of this measure satisfy

$$\frac{an}{(n + |g|_S^{\alpha})^{1+D/\alpha}} \leq \mu_{S,\alpha}^{(n)}(g) \leq \frac{An}{(n + |g|_S^{\alpha})^{1+D/\alpha}}, \quad g \in \Gamma, n \in \mathbb{N}.$$

So, one has a remarkably good control of the behavior of the associated random walk. However, there are no existing limit theorems in the literature for such walks, even if we assume that Γ is a torsion free nilpotent group (such groups are basic examples of groups with polynomial volume growth). Our results provide limit theorems (functional, and also local) for any random walk driven by a word length radial stable probability measure $\mu_{S,\alpha}$, $\alpha \in (0, 2)$, on a torsion free finitely generated nilpotent group. We now briefly describe these results.

First, because we assume that Γ is a finitely generated torsion free nilpotent group, there is a simply connected nilpotent Lie group $G = (\mathbb{R}^d, \cdot)$ which contains Γ as a co-compact discrete subgroup. The Lie algebra, $\mathfrak{g}$, of this Lie group is equipped with its central descending series

$$\mathfrak{g}_1 = \mathfrak{g} \supseteq \mathfrak{g}_2 = [\mathfrak{g}, \mathfrak{g}] \supseteq \cdots \supseteq \mathfrak{g}_j = [\mathfrak{g}_{j-1}, \mathfrak{g}] \supseteq \cdots \supseteq \{0\},$$

and this series become trivial (i.e., constant equal to $\{0\}$) after finitely many steps. Let j^* be the smallest j such that $\mathfrak{g}_{j+1} = \{0\}$. One can choose a direct sum decomposition by vector subspaces, $\mathfrak{n}_i$, $1 \le i \le j^*$, compatible with the central descending series above, so that

$$\mathbb{R}^d = \mathfrak{g} = \oplus_{i=1}^{j^*} \mathfrak{n}_i \quad \text{and} \quad \mathfrak{g}_j = \sum_{i \ge j} \mathfrak{n}_i, \ j \in \{1, \dots, j^*\}.$$

The linear invertible maps

$$\delta_t(x) = t^i x \quad \text{ if } x \in \mathfrak{n}_i, \ 1 \le i \le j^*, \quad t > 0$$

form an approximate Lie dilation structure in the sense that

$$[x, y]_\bullet = \lim_{t \to \infty} \delta_t^{-1}([\delta_t(x), \delta_t(y)])$$

is a Lie bracket on $\mathbb{R}^d$ with the property that $\delta_t([x, y]_\bullet) = [\delta_t(x), \delta_t(y)]_\bullet$. Using exponential coordinate (of the first type) to represent G as $(\mathbb{R}^d, \cdot)$, the approximate Lie dilations $\delta_t, t > 0$, define approximate group dilations on G for which we use the same notation. The limit group $G_\bullet$ is the simply connected Lie group associated with the Lie algebra $(\mathbb{R}^d, [\cdot, \cdot]_\bullet)$ defined above. It follows from (A.1) of the Appendix that the volume growth exponent D of the original group Γ is given by $D = \sum_{i=1}^{j^*} i \dim(\mathfrak{n}_i)$. Thus we have $\det(\delta_t) = t^D$ for every $t > 0$ In [51], Pansu proves the fundamental results that there is a norm $\| \cdot \|_\bullet$ on $(\mathbb{R}^d, \bullet)$, homogeneous with respect to $(\delta_t)_{t>0}$, such that the geometry of $(\Gamma, |\cdot|_S)$ is well approximated at large scale by that of $(\mathbb{R}^d, \| \cdot \|_\bullet)$ in the sense that

$$\lim_{g \in \Gamma, g \to \infty} \frac{|g|_S}{\|g\|_\bullet} = 1.$$

Further, one has

$$\lim_{r \to \infty} \frac{\#\{g \in \Gamma : |g|_S \le r\}}{|\{x \in \mathbb{R}^d : \|x\|_\bullet \le r\}|} = 1,$$

where $|\Omega|$ is the Haar volume of $\Omega \subset G_\bullet$. See also, [13]. Haar measures on G and $G_\bullet$ are both Lebesgue measure dx on $\mathbb{R}^d$. When considering densities on these groups, we mean densities with respect to dx.

The importance of these results for us is that they enable us to establish the convergence of the measure $t\delta_{1/t^{1/\alpha}}(\mu_{S,\alpha})$, vaguely on $\mathbb{R}^d \setminus \{0\}$, to the radial stable jump measure $\mu_{\bullet,\alpha}$ with density

$$\phi_{\bullet,\alpha}(x) = \frac{c(\Gamma, S, \alpha)}{\|x\|_{\bullet}^{\alpha+D}}.$$

This measure is the jumping measure of a left-invariant (strong) Markov process $(X_t^{\bullet})_{t>0}$ on $(\mathbb{R}^d, \bullet)$ which is self-similar in the sense that $(X_s^{\bullet})_{s>0}$ equals $(\delta_{1/t^{1/\alpha}}(X_{ts}^{\bullet}))_{s>0}$ in distribution. In a proper global coordinate system, the coordinates of this process can be expressed in terms of suitable stable processes and their (possibly iterated) Lévy areas. The Lévy process $X^{\bullet}$ admits a continuous convolution density with respect to the Lebesgue measure on $\mathbb{R}^d$:

$$p_{\bullet,\alpha}(t,x), \quad (t,x) \in (0,\infty) \times \mathbb{R}^d.$$

This density satisfies

$$p_{\bullet,\alpha}(t,x) = t^{-D/\alpha} p_{\bullet,\alpha}(1, \delta_{1/t^{1/\alpha}}(x)), \quad (t,x) \in (0,\infty) \times \mathbb{R}^d,$$

and

$$\frac{at}{(t + \|x\|_{\bullet}^{\alpha})^{1+D/\alpha}} \le p_{\bullet,\alpha}(t,x) \le \frac{At}{(t + \|x\|_{\bullet}^{\alpha})^{1+D/\alpha}}, \quad (t,x) \in (0,\infty) \times \mathbb{R}^d.$$

In this context, the results developed in this work establish two limit theorems for the random walk $(X_n)_{n\ge 0}$ on Γ driven by $\mu_{S,\alpha}$. These limit theorems capture the fact that, after proper rescaling in time and space, the limit of the random walk $(X_n)_{n\ge 0}$ is the Markov process $(X_s^{\bullet})_{s>0}$. Namely, the functional limit theorem establishes the convergence of $(\delta_{1/t^{1/\alpha}}(X_{[st]}))_{s>0}$ to $(X_s^{\bullet})_{s>0}$ as t tends to infinity. In particular, for any continuous function ϕ with compact support on $\mathbb{R}^d$,

$$\sum_{g\in\Gamma} \phi(\delta_{t^{1/\alpha}}(g)) \mu_{S,\alpha}^{[(ts])}(g) \to \int_{\mathbb{R}^d} \phi(x) p_{\bullet,\alpha}(s,x)dx$$

as t tends to infinity. For any compact set $K \subset \mathbb{R}^d$ and any functions $g_n : K \to \Gamma$, $n = 1, 2, \ldots$, such that the sequence of functions $\delta_{1/n^{1/\alpha}} \circ g_n : K \to \mathbb{R}^d$, $n = 1, 2, \ldots$, converges uniformly over K to the identity function, the local limit theorem of this monograph establishes the uniform convergence to zero over K of

$$n^{D/\alpha} \mu_{S,\alpha}^{(n)}(g_n(x)) - p_{\bullet,\alpha}(1,x)$$

when n tends to infinity. In particular, this shows that, for any fixed $g \in \Gamma$ (e.g., $g = e$),

$$\lim_{n\to\infty} n^{D/\alpha} \mu^{(n)}(g) = p_{\bullet,\alpha}(1,e).$$

Walks taking stable-like steps along one parameter subgroups

Let Γ be a torsion free finitely generated nilpotent subgroup of a simply connected nilpotent group G. One of the cases that motivates our study can be described as follows: We are given a tuple $S = (s_1, \ldots, s_k)$ of elements of Γ, which, together with their inverses, generates Γ. We are also given a tuple of reals $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k) \in (0,2)^k$. Note that the letter S is used here in a slightly different way than in the previous case. Now, set

$$\mu_{S,\boldsymbol{\alpha}}(g) = \frac{1}{k} \sum_{i=1}^{k} \sum_{m \in \mathbb{Z}} \frac{c_{\alpha_i}}{(1+|m|)^{1+\alpha_i}} \mathbb{1}_{\{s_i^m\}}(g).$$

It was proved in [55] that, for any such probability measure, there exist $0 < a = a(\Gamma, S, \boldsymbol{\alpha}) \le A = A(\Gamma, S, \boldsymbol{\alpha}) < \infty$ and $\gamma_0 = \gamma_0(\Gamma, S, \boldsymbol{\alpha})$ such that

$$a n^{-\gamma_0} \le \mu^{(n)}_{S,\boldsymbol{\alpha}}(e) \le A n^{-\gamma_0}. \tag{2.1}$$

In Chapter 8 , we introduce the space of probability measures $\mathcal{SM}_1(\Gamma)$, see Definition 8.2, which contains all such measures. We then explain how to choose a coordinate system of polynomial type, $G = (\mathbb{R}^d, \cdot)$, and an approximate dilation structure $(\delta_t)_{t>0}$ with limit group $G_\bullet = (\mathbb{R}^d, \bullet)$, which are adapted to the pair $(S, \boldsymbol{\alpha})$, and such that, with $\mu_t := t\delta_{1/t}(\mu_{S,\boldsymbol{\alpha}})$, the family of measures $(\|z\|_2^2 \wedge 1)\mu_t(dz)$ converges weakly on $\mathbb{R}^d \setminus \{0\}$ to a measure $(\|z\|_2^2 \wedge 1)\mu_\bullet(dz)$ as $t \to \infty$, that is,

$$\lim_{t\to\infty} \int_{\mathbb{R}^d \setminus \{0\}} f(z)(\|z\|_2^2 \wedge 1)\mu_t(dz) = \int_{\mathbb{R}^d \setminus \{0\}} f(z)(\|z\|_2^2 \wedge 1)\mu_\bullet(dz)$$

for any $f \in C_b(\mathbb{R}^d \setminus \{0\})$. Here $C_b(\mathbb{R}^d \setminus \{0\})$ denotes the space of bounded continuous functions on $\mathbb{R}^d \setminus \{0\}$. The measure $\mu_\bullet$ is supported on the union of a finite number of one parameters subgroups of $G_\bullet$ and its support generates $G_\bullet$. It can be interpreted as the Lévy measure of a convolution semigroup of probability measures, associated with a left-invariant Lévy process on $G_\bullet$. The convolution transition kernel of this semigroup admits a continuous density, $p_\bullet(t,x)$, with respect to the Lebesgue measure on $(\mathbb{R}^d, \bullet)$ and satisfies

$$p_\bullet(t,x) = t^{-\gamma_0} p_\bullet(1, \delta_{1/t}(x)) \quad \text{for } (t,x) \in (0,\infty) \times \mathbb{R}^d.$$

Note that the limit objects introduced here, e.g., $G_\bullet$ and $p_\bullet$, all depend on S and $\boldsymbol{\alpha}$, even so we did not capture that dependence in the notation used above. A notable difference with the earlier description of the radial stable-like case is that, in general, there are no particular canonical choices of the approximate dilation structure $(\delta_t)_{t>0}$ and we have not made a canonical choice of coordinates either. To a certain extent, the entire results and the associated limit objects depend on the choices of coordinates and adapted dilation structure while, of course, there are great commonalities shared by all the limit objects obtained based on these different choices. This, however, will not be deeply investigated here.

As in the case of radial stable walks, the results of this monograph establish the convergence of the discrete time random walk driven by $\mu_{S,\alpha}$, properly rescaled in time and space, to the left-invariant Lévy process $(X^{\bullet}_s)_{s>0}$ with convolution density $p_{\bullet}(t,x)$ mentioned above. More precisely, the functional limit theorem establishes the weak convergence of $(\delta_{1/t}(X_{[st]}))_{s>0}$ to $(X^{\bullet}_s)_{s>0}$ as t tends to infinity. In particular, for any continuous function ϕ with compact support on $\mathbb{R}^d$,

$$\sum_{g\in\Gamma}\phi(\delta_t(g))\mu_{S,\alpha}^{[(ts])}(g)\to\int_{\mathbb{R}^d}\phi(x)p_{\bullet}(s,x)dx$$

as t tends to infinity. The local limit theorem asserts that

$$\lim_{n\to\infty}\sup_{x\in K}\left|n^{\gamma_0}\mu_{S,\alpha}^{(n)}(g_n(x))-p_{\bullet}(1,x)\right|=0, \tag{2.2}$$

where K is a compact in $\mathbb{R}^d$ and $g_n : K\to\Gamma$ is a sequence of functions such that $\delta_{1/n}\circ g_n : K\to\mathbb{R}^d$ converges uniformly over K to the identity function. Of course, the non-negative real γ_0 appearing in (2.1) and in (2.2) is the same in both equations. It is also given by $\det(\delta_t)=t^{\gamma_0}$.

Walks associated with measure in $\mathcal{SM}(\Gamma)$

In Chapter 8 , we introduce a particular set of "stable-like" measures on Γ, $\mathcal{SM}(\Gamma)$, which interpolates between the radially symmetric measures considered above and the convex combinations of one dimensional measures described in the last section. These measures were studied in our earlier work [20]. With any measure in $\mathcal{SM}(\Gamma)$ we can associate in a natural way a (non-unique) polynomial coordinate system $G=\mathbb{R}^d$ and a family of dilations $(\delta_t)_{t>0}$ which define a limit group structure $G_{\bullet}=(\mathbb{R}^d,\bullet)$. The approximate dilation structure $(\delta_t)_{t>0}$ is built so that the family $\mu_t=t\delta_{1/t}(\mu)$, $t>0$, has well defined limit points which are all Lévy measures of $(\delta_t)_{t>0}$-stable symmetric convolution semigroups of probability measures on $G_{\bullet}$ with continuous positive densities on $G_{\bullet}$. One of the key contributions of this work is to describe explicitly how one can construct such an approximate dilation structure based on a proper description of μ on Γ. If it is the case that, with $\mu_t := t\delta_{1/t}(\mu)$, the measure $(\|z\|_2^2\wedge 1)\mu_t(dz)$ converges weakly to a finite measure $(\|z\|_2^2\wedge 1)\mu_{\bullet}(dz)$ on $\mathbb{R}^d\setminus\{0\}$ as $t\to\infty$, then we obtain both a functional theorem and local limit theorem. The results described in the previous two paragraphs are, in fact, special cases of these more general theorems. The structure of the Lévy measures of the limit Lévy processes on $G_{\bullet}$ appearing in these limit theorem is described at the end of the next section.

2.4 Symmetric continuous convolution semigroup of probability measure and Lévy processes

For this very minimal vocabulary review, we follow [36]. Let G be a connected Lie group. Recall that there is a one-to-one correspondence between symmetric continuous convolution semigroups of probability measures on G and symmetric Lévy processes on G. Here $(\mu_t)_{t>0}$ is a symmetric continuous convolution semigroup of probability measures on G if the map $t \mapsto \mu_t$ is continuous, $\mu_t * \mu_s = \mu_{t+s}, s, t > 0$, $\mu_0 = \delta_e$, and $\mu_t(\phi) = \mu_t(\check{\phi})$ for any continuous function ϕ on G with compact support, where $\check{\phi}(y) := \phi(y^{-1})$ for $y \in G$. A symmetric Lévy process X on G is a G-valued time-homogeneous càdlàg Markov process $(X_t)_{t\geq 0}$ with stationary independent increments, started at e and such that $X_t^{-1} = X_t$ in distribution for every $t > 0$. In this setting, the notion of infinitesimal generator of X can be captured in a more elementary way via the so-called generating functional (defined on smooth compactly supported functions): if the infinitesimal generator of the symmetric Lévy process X is $\mathcal{L}$, the associated generating functional is simply $\phi \mapsto \mathcal{L}\phi(e)$. The Lévy-Khinchin-Hunt formula provides a description of the generating functional of a Lévy process. Under the symmetry condition, the generating functional has two parts, a diffusion part, and a jump part described by a symmetric measure ν on $G \setminus \{e\}$ in the form $\phi \mapsto \text{p.v.} \int_{G\setminus\{e\}} (\phi(y) - \phi(e))\nu(dy)$ with

$$\int_{G\setminus\{e\}} \min\{1, \|y\|_2^2\}\nu(dy) < \infty. \tag{2.3}$$

Here, $\|y\|_2$ is the Riemannian distance between e and y in some fixed left-invariant metric on G, and

$$\text{p.v.} \int_{G\setminus\{e\}} (\phi(y) - \phi(e))\nu(dy) := \frac{1}{2} \int_{G\setminus\{e\}} (\phi(y) + \phi(y^{-1}) - 2\phi(e))\nu(dy).$$

In this work, we are only interested in pure-jump symmetric Lévy processes, that is, generating functional of the form

$$\phi \mapsto \mathcal{L}\phi(e) = \text{p.v.} \int_{G\setminus\{e\}} (\phi(y) - \phi(e))\nu(dy)$$

where ν is a symmetric measure on $G \setminus \{e\}$ satisfying (2.3). Equivalently, the infinitesimal generator is given on smooth compactly supported functions by

$$\langle -\mathcal{L}u, v\rangle = \frac{1}{2} \int_G \int_{G\setminus\{e\}} (u(xy) - u(x))(v(xy) - v(x))\nu(dy)dx.$$

This, of course, is also a description of the associated Dirichlet form (on a dense subspace of its domain).

In this work, these objects come about through a limit procedure which implies that they have additional properties. First, the underlying Lie group is a simply connected nilpotent Lie group which we call $G_\bullet$. Second, by construction, $G_\bullet$ carries a group of dilations, $(\delta_t)_{t>0}$, $\delta_t : G \to G$, $\delta_1 = \mathrm{Id}$, $\delta_{ts} = \delta_t \circ \delta_s = \delta_s \circ \delta_t$, $s,t > 0$, where δ_t is also a group isomorphism for every $t > 0$, and $\lim_{t\to 0} \delta_t(x) = e$ for all $x \in G$. In addition, the convolution semigroups and associated Lévy processes of interest to us are self-similar with respect to such a dilation structure, that is, $(X_s)_{s>0}$ equals $(\delta_{1/t}(X_{ts}))_{s>0}$, in distribution, for any $t > 0$. Moreover, there is a linear basis $\varepsilon = (\varepsilon_1, \dots, \varepsilon_d)$ of the Lie algebra of $G_\bullet$ in which the dilation δ_t has the form $\delta_t(\varepsilon_i) = t^{1/\beta_i}\varepsilon_i$, $\beta_i \in (0,2)$, $1 \le i \le d$. This last condition, $\beta_i \in (0,2)$, $1 \le i \le d$, is related to the fact the processes in question are pure-jump operator-stable Lévy processes. See, e.g., [36, Theorem 2.3.17]. Finally, when the original random walk is driven by a probability measure μ in $\mathcal{SM}(\Gamma)$ (a class of stable-like measures on Γ described in Chapter 8), the Lévy measure

$$\mu_\bullet = \lim_{t\to\infty} t\delta_{1/t}(\mu) = \lim_{t\to\infty} t\mu \circ \delta_t$$

of our limit process has a particular structure that it inherits from the facts that $\mu \in \mathcal{SM}(\Gamma)$. Namely, there is a finite family of closed Lie subgroups of $G_\bullet$, call them $H_{\bullet,i}$, $1 \le i \le k$, which are each invariant under $(\delta_t)_{t>0}$, and functions $\psi_i : H_{\bullet,i} \to (0,\infty)$ satisfying $t\psi_i(\delta_t(x)) = \psi_i(x)$, and $\psi_i(x^{-1}) = \psi_i(x)$, $x \in H_{\bullet,i}$, such that

$$\mu_\bullet(dx) = \sum_1^m \nu_i(dx), \quad \nu_i(\phi) = \int_{H_{\bullet,i}} \phi(x)\psi_i(x) d_{H_{\bullet,i}}x \text{ for } 1 \le i \le m,$$

where $d_{H_{\bullet,i}}x$ is the Haar measure on $H_{\bullet,i}$; see Proposition 10.8. Each ν_i satisfies (2.3). The group generated by the union $\cup_1^m H_{\bullet,i}$ of the subgroups $\{H_{\bullet,i}, 1 \le i \le m\}$ is $G_\bullet$.

2.5 Prior results

To put our results in perspective, we briefly review limit theorems (functional and/or local) relating random walks on discrete groups to Lévy processes on a related Lie group. Very few results of this type exist outside the setting of nilpotent groups (and closely related groups such as groups of polynomial volume growth). The classical (functional) limit theorems can be interpreted in two distinct ways:

(i) As providing approximation of a (continuous time) Lévy process by a discrete time process. This can be motivated by the desire to actually construct the limiting process, or to simulate it, or to understand it in more concrete terms. In this case, one should read the limit theorem as follows: at each stage, we take a greater number of smaller steps to approximate the behavior of a continuous time process on a fix bounded time interval. A natural setup for this interpretation is the triangular array setup.

(ii) As a result illuminating the long term behavior of a discrete time process by providing a continuous time scaling limit. In this case, at each stage we take a greater number of identically distributed steps and approximate the probability of larger and larger scale events for the discrete time process by the probability of the same large scale events for the limiting continuous time process, at a large time. Whenever that limiting process is self-similar, the limit computation can be rephrased as a computation within a fixed bounded time interval. There is more rigidity in this viewpoint than in the first as we cannot choose the different individual steps taken as one possibly can in a triangular array formulation of the first viewpoint.

For random walks in $\mathbb{R}^n$, it is somewhat difficult to see the differences between these two interpretations. The reason is that we have a relatively obvious way to turn the identically distributed steps appearing in the second interpretations into smaller and smaller steps appearing in the first interpretation. Indeed, we typically assume that the limiting process is self-similar with respect to a dilation structure that commute with addition and this dilation structure can be used to turn the fixed-size steps of (ii) into the small-size steps of (i).

Both viewpoints are present in this work. Our main focus is on using (ii) to study long term behavior of a class of discrete long range random walks on a finitely generated torsion free nilpotent group Γ. One can then use (i) to better understand the limiting self-similar Lévy processes on the limit nilpotent group $G_\bullet$. See Chapter 7 .

2.5.1 Functional type limit theorems

On a general Lie group, there are results stated in terms of triangular arrays that go back to Wehn [60, 61]. Later, Stroock and Varadhan [58] rediscovered Wehn's results. These works concern the case when the limit Lévy process is a diffusion. These triangular array results have been extended to cover the case when the limit Lévy process may have jumps. An exposition of such results is found in [36] which contains a very long list of references. They are also found in work by Kunita [40, 41, 42, 43]. These results must be understood as an extension of the first interpretation of the classical limit theorem discussed above. From this viewpoint, the title of the Stroock-Varadhan paper, Limit Theorems for Random Walks on Lie Groups, is somewhat misleading. What the results of Wehn and Stroock-Varadhan do is to provide discrete time steps approximations of diffusions on Lie groups. They do not, in general, help us understand the behavior of random walks on Lie groups. That is because, on a general Lie group, there is no clear way to turn identically distributed steps into small-size steps. There are, however, many ways to create arrays of smaller and smaller size steps, not related to any identically distributed model. The theorems described by Wehn, Stroock-Varadhan, Hazod and Siebert, Kunita, and others thus provide functional limit theorems along the line of the first interpretation. See the excellent discussion in [12].

There is one setting in which these triangular array limit theorems provide an understanding of random walk (in the sense of a process taking repeated identically distributed steps). This is, informally, when the limiting continuous time process is self-similar with respect to a dilation structure that preserve the multiplication law of the underlying group. Unfortunately, this is a rather rare occurrence as the only Lie groups admitting such dilation structures are simply connected nilpotent Lie groups of a very special kind. Moreover, outside the case of diffusion limit, whether or not a dilation structure exists that is suitable for a given random walk on a given group depends, to a large extend, on the particular random walk in question. In fact, given a driving measure μ, constructing a proper dilation structure for μ (deciding if such exists!) is a major problem, one that is completely ignored by the triangular array formulation of limit theorems. This is illustrated by the results of the present work.

Somewhat independently of the above circle of ideas, Crepel, Raugi, and others, obtained rather satisfying random walk limit theorems for general nilpotent groups in the case the limit is a diffusion [23, 53, 54]. The proofs in these works can be viewed as using two steps: the first step proves the result in the presence of a canonical adapted dilation structure (that is, in the case of stratified nilpotent groups). The second step is closely related to one of the key ingredients we will use here and involves the idea behind our definition of an approximate group dilation structure, a dilation structure that does not preserve the group structure of the original underlying Lie group. In general, because of the second step, the original group carrying the random walk has a group structure that is different from that of the Lie group carrying the limit diffusion. This, clearly, takes us outside the realm of Wehn-type results.

One key point in the results by Crépel and Raugi is that the structure of the group carrying the limit diffusion depend only on the original group, not of the particular (diffusive) random walk one wants to study. In general, this cannot be the case when the random walk to be studied calls for a limit process that has jumps as we do here. As we shall see, in this case, the limit structure depends on both the original group and the particular probability measure that drive the given random walk. One thus has to discover what this proper limit structure is for each studied random walk.

2.5.2 Local limit theorems

The first local limit theorem in the context of general nilpotent groups and groups of polynomial volume growth is due to G. Alexopoulos [2, 3, 4]. See also the discussion in [12]. It concerns centered random walks driven by a finitely supported measure. For nilpotent groups, following a very different approach, and covering random walks driven by measures that have a high enough finite moment (much higher than 2, in general), the best known results are due to R. Hough [38] which provides an informative review of earlier results. We do not know of references treating cases when the limit is not a diffusion process.

Chapter 3
Polynomial coordinates and approximate dilations

Abstract This chapter introduces the notions of polynomial coordinate systems and approximate group dilations relative to such coordinate systems. Rescaling via suitable dilation structures is key to the formulation of limit theorems for random walks on groups. One of the main tools used in this book is the notion of approximate group dilations. The limit group structures that appear when one use rescaling associated with approximate group dilations are discussed.

3.1 Polynomial coordinate systems

Even though some related results can be stated in an intrinsic manner, in practice, limit theorems are coordinate dependent. This applies to the results of this monograph and, consequently, we discuss in some details the notion of global coordinate system for simply connected nilpotent Lie groups. A number of different choices are possible for this purpose. In this chapter, we outline basic characteristics of the coordinate systems we will use. A given group G can be described via many different such global polynomial coordinate charts and it is often desirable to allow for such a choice to be made by circumstances. This is discussed further in Chapter 9 .

A simply connected nilpotent Lie group G can always be described by a global coordinate chart $\mathbb{R}^d \to G$, $0 \to e$, in which the group multiplication and inverse map are given by polynomials

$$x \cdot y = P(x, y) = (p_1(x, y), \dots, p_d(x, y)), \quad x^{-1} = Q(x) = (q_1(x), \dots, q_d(x)).$$

As $P(x, 0) = x$ and $P(0, y) = y$ for any $x, y \in \mathbb{R}^d$, we have

$$p_i(x, y) = x_i + y_i + \bar{p}_i(x, y), \quad 1 \le i \le d, \tag{3.1}$$

where $\bar{p}_i(x, y)$'s are polynomials having no constant nor first order terms. Moreover, for any compact $K \subset \mathbb{R}^d$, there is a constant C_K such that

$$\|x^{-1} \cdot y\|_2 := \|P(Q(x), y)\|_2 \le C_K \|x - y\|_2 \quad \text{for every } x, y \in K, \tag{3.2}$$

because $P(Q(x), x) = 0$. Here $\|\cdot\|_2$ is the canonical Euclidean norm in $\mathbb{R}^d$. Similarly, for any compact $K \subset \mathbb{R}^d$, there is a constant C'_K such that

$$\|x - y\|_2 \le C'_K \|x^{-1} \cdot y\|_2 \quad \text{for every } x, y \in K. \tag{3.3}$$

Indeed, this is the same as $\|x - x \cdot z\|_2 \le C'_K \|z\|_2$ and the polynomial $x - P(x,z)$ vanishes at $z = 0$.

We assume throughout that the Jacobian of the maps $y \mapsto x \cdot y$, $x \in G$, is 1 so that the Lebesgue measure on $\mathbb{R}^d$ is a Haar measure for our group G. This assumption follows from the much more demanding assumption that (3.1) has the additional property that

$$\bar{p}_1(x,y) = 0 \quad \text{and} \quad \bar{p}_i(x,y) = \bar{p}_i((x_j)_1^{i-1}, (y_j)_1^{i-1}) \text{ for } 2 \le i \le d. \tag{3.4}$$

In other word, the polynomial

$$\bar{p}_i(x,y) = p_i(x,y) - x_i - y_i$$

depends only on the first $i-1$ coordinates of x and y and has no constant nor first order terms. Clearly, this triangular structure implies that the Jacobian of the map $y \mapsto x \cdot y$ is 1. Morever, for $x^{-1} = Q(x) = (q_1(x), \dots, q_d(x))$, we deduce from $P(Q(x), x) = 0$ that

$$q_1(x) = -x_1 \quad \text{and} \quad q_i(x) = -x_i + \bar{q}_i(x_1, \dots, x_{i-1}) \text{ for } 2 \le i \le d, \tag{3.5}$$

where $\bar{q}_i(x_1, \dots, x_{i-1})$, $2 \le i \le d$, are polynomials having no constant nor first order terms.

Example 3.1 (Matrix coordinates) The most commonly used coordinate system (as Molière's Mr. Jourdain with *prose*, we may use it without realizing we do!) comes from matrix groups. Indeed, the group G is often given as a subgroup of a group of invertible matrices of a certain dimension, say N. In particular, a nilpotent group is often given as a subgroup of the group of unipotent upper-triangular matrices. The most obvious example is when G is the group of unipotent upper-triangular matrices itself

$$\mathbf{U}_N = \left\{ \begin{pmatrix} 1 & x_{12} & x_{13} & \dots & x_{1N} \\ 0 & 1 & x_{23} & \dots & x_{2N} \\ 0 & 0 & 1 & \dots & x_{3N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} : x_{ij} \in \mathbb{R}, 1 \le i < j \le d \right\}.$$

This group has dimension $d = \binom{N}{2}$. In the case $N = 3$, this is the Heisenberg group $\mathbb{H}_3(\mathbb{R})$ in its matrix form with

$$P(x,y) = (x_1 + y_1, x_2 + y_2, x_3 + y_3 + x_1 y_2), \quad Q(x_1, x_2, x_3) = (-x_1, -x_2, -x_3 + x_1 x_2)$$

and

$$P(x^{-1}, y) = (y_1 - x_1, y_2 - x_2, y_3 - x_3 - x_1(y_2 - x_2)).$$

Example 3.2 (Exponential coordinates of the first type) The second most commonly encountered coordinate system is given by the canonical exponential map

$$\exp : \mathfrak{g} \to G$$

between the Lie algebra $\mathfrak{g}$ of the group G and G itself. We can think of $\mathfrak{g} = \mathbb{R}^d$ as the tangent space at e. Given a tangent vector $x \in \mathbb{R}^d$, we first consider the (unique) left invariant vector field X on G such that $X(e) = x$ and the solution $\gamma_x : [0, 1] \to G$ of $\frac{d}{dt}\gamma_x(t) = X(\gamma_x(t))$ with initial condition $\gamma_x(0) = e$, and set

$$\exp(x) = \gamma_x(1).$$

Using the fact that, for any two left-invariant vector fields X, Y, the well defined differential operator $XY - YX$ is a left-invariant vector field, we obtain the Lie bracket $(x, y) \mapsto [x, y] = (XY - YX)(e)$. Moreover,

$$[x, y] = \partial_s \partial_t \left(\exp(tx) \cdot \exp(sy) \cdot \exp(-tx)\right)\big|_{s=t=0}.$$

In the case of simply connected Lie group, the exponential map is a global invertible diffeomorphism and the multiplication is given in universal form by the famous Campbell-Hausdorff formula

$$\exp(x) \cdot \exp(y) = \exp\left(P_{\mathrm{CH}}(x, y)\right),$$

where

$$P_{\mathrm{CH}}(x, y) = x + y + \frac{1}{2}[x, y] + \frac{1}{12}([x, [x, y]] + [y, [y, x]]) + \cdots. \tag{3.6}$$

In other words, in the exponential coordinate system, the group law is

$$x \cdot y = P_{\mathrm{CH}}(x, y) = x + y + \frac{1}{2}[x, y] + \frac{1}{12}([x, [x, y]] + [y, [y, x]]) + \cdots.$$

This has the desirable polynomial form because iterated Lie brackets with more than r entries are equal to 0 if r is the nilpotency class of G. In these coordinates, it is always the case that

$$x^{-1} = -x.$$

Applying this to the Heisenberg group we obtain the often-used description of $\mathbb{H}_3(\mathbb{R})$ as $\mathbb{R}^3$ equipped with the product

$$x \cdot y = P_{\mathrm{CH}}(x, y) = \left(x_1 + y_1, x_2 + y_2, y_3 + x_3 + \tfrac{1}{2}(x_1 y_2 - x_2 y_1)\right)$$

and

$$P_{\mathrm{CH}}(x^{-1}, y) = \left(y_1 - x_1, y_2 - x_2, y_3 - x_3 + \tfrac{1}{2}(x_2(y_1 - x_1) - x_1(y_2 - x_2))\right).$$

Example 3.3 (Exponential coordinates of the second kind) For a simply connected nilpotent Lie group G, exponential coordinate systems of the second kind are typically associated with a filtration of the Lie algebra $\mathfrak{g}$ by subalgebras (resp. ideals)

$$\mathfrak{g} = \mathfrak{g}_1 \supset \mathfrak{g}_2 \supset \cdots \supset \mathfrak{g}_\ell \supset \{0\}$$

with $\mathfrak{g}_j$ of dimension m_j, and a linear basis $(\varepsilon_i)_1^d$ such that the linear span of $(\varepsilon_i)_{i\geq j}$ is a subalgebra (resp. ideal) for all $1 \leq j \leq d := m_1$, and $(\varepsilon_i)_{d-m_j+1}^d$ is a basis of $\mathfrak{g}_j$. In such a situation, the maps from $\mathbb{R}^d$ to G defined by

$$\Phi(x_1, \ldots, x_d) = \exp(x_1\varepsilon_1) \cdot \cdots \cdot \exp(x_d\varepsilon_d)$$

and

$$\Psi(x_1, \cdots, x_d) = \exp(x_d\varepsilon_d) \cdot \cdots \cdot \exp(x_1\varepsilon_1)$$

give two distinct global polynomial coordinate systems for G.

For example, the matrix coordinate system of the group of $n \times n$ upper-triangular matrices with entries equal to 1 on the diagonal, is an exponential coordinate system of the second kind associated with the lower central series

$$\mathfrak{g} = \mathfrak{g}_1, \quad \mathfrak{g}_{i+1} = [\mathfrak{g}_i, \mathfrak{g}], \ 1 \leq j \leq n,$$

which, in this case, has last non-trivial member $\mathfrak{g}_{n-1}$ corresponding to the upper-right corner entry. Here, we can realize $\mathfrak{g}$ as the algebra of the strictly upper-triangular matrix. We then enumerate the entries $(x_i)_1^d$, $d = n(n-1)/2$, going down along each upper-diagonal in order so that x_d is the entry in the upper-right corner, and consider the corresponding map Ψ. For instance, in the 4×4 case,

$$\begin{pmatrix} 1 & x_1 & x_4 & x_6 \\ 0 & 1 & x_2 & x_5 \\ 0 & 0 & 1 & x_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & x_6 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & x_5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & x_1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Each of the matrices on the right is the matrix exponential of the corresponding strictly triangular matrix. Note that Φ defined above leads to a different coordinate system. □

These classical constructions concerning exponential coordinates of the first and second kinds are explained in more details in [22, Section 1.2]. See also [21, 30, 46].

3.2 Dilations, approximate dilations, and $G_\bullet$

Straight dilations

Let G be a nilpotent simply connected Lie group given in a global polynomial coordinate system $G = (\mathbb{R}^d, \cdot)$. Call *straight dilations with exponents* $\mathbf{a} = (a_1, \ldots, a_d) \in \mathbb{R}_+^d$, the group of diffeomorphisms

$$\phi_t(x) = (t^{a_1}x_1, \ldots, t^{a_d}x_d), \quad t > 0.$$

Note that $\phi_s \circ \phi_t = \phi_{st}$, $s, t > 0$, and $\phi_1 = \mathrm{Id}$.

Definition 3.4 We say that $(\phi_t)_{t>0}$ as above is a *straight group dilation structure* if

$$\phi_t(x \cdot y) = \phi_t(x) \cdot \phi_t(y), \qquad t > 0, x, y \in G. \tag{3.7}$$

This, of course, is a very restrictive property and not every simply connected nilpotent Lie group G admits such a structure. In the case of the Heisenberg group in matrix form, for given $a, b, c \geq 0$, set

$$\phi_t\left(\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}\right) = \begin{pmatrix} 1 & t^a x & t^c z \\ 0 & 1 & t^b y \\ 0 & 0 & 1 \end{pmatrix}, \quad t > 0, \; x, y, z \in \mathbb{R}.$$

These straight dilations structures are group dilation structures if and only if $a+b = c$.

Remark 3.5 More generally, without reference to any coordinate system, a group of diffeomorphisms $(\phi_t)_{t>0}$, $\phi_t : G \to G$, $\phi_1 = \mathrm{Id}$, satisfying (3.7) and such that $\lim_{t\to 0} \phi_t(g) = e$ is called an expanding group dilation structure. See [36, 45]. By a theorem of Siebert [57], a connected locally compact group carrying such a structure must be a simply connected nilpotent Lie group (and not every simply connected nilpotent groups admit such a structure). □

Definition 3.6 Let $\mathbb{R}^d$ be equipped with a straight dilation structure

$$(\phi_t)_{t>0}, \quad \phi_t(x) = (t^{a_i} x_i)_1^d, \quad a_i > 0, \quad 1 \leq i \leq d.$$

A positive function N on $\mathbb{R}^d$ is called homogeneous with respect to $(\phi_t)_{t>0}$ if $N(\phi_t(x)) = tN(x)$. □

Example 3.7 The function $x \mapsto N(x) = \max_{1\leq i\leq d}\{|x_i|^{1/a_i}\}$ is homogenous with respect to $(\phi_t)_{t>0}$, $\phi_t(x) = (t^{a_i} x_i)_1^d$, $a_i > 0$, $1 \leq i \leq d$. It is a norm on $(\mathbb{R}^d, +)$ (i.e., satisfies the triangle inequality) if $a_i \geq 1$ for all $1 \leq i \leq d$. If M is another homogeneous function with respect to $(\phi_t)_{t>0}$, such that the set $x : M(x) \leq 1$ is compact, then there are constants $0 < c \leq C < \infty$ such that $cN \leq M \leq CN$. □

Approximate group dilations and $G_\bullet$

Let G be a simply connected nilpotent Lie group given in a global polynomial chart $G = (\mathbb{R}^d, \cdot)$ and equipped with a straight dilation (not necessarily a group dilation structure) $(\phi_t)_{t>0}$. For each $t > 0$, we obtain a new group structure $\cdot_t$ on $\mathbb{R}^d$ by setting

$$x \cdot_t y = \phi_{1/t}(\phi_t(x) \cdot \phi_t(y)), \quad x, y \in \mathbb{R}^d.$$

Moreover,

$$\phi_{1/t} : (\mathbb{R}^d, \cdot) \to (\mathbb{R}^d, \cdot_t)$$

is a group isomorphism between $G = (\mathbb{R}^d, \cdot)$ and $G_t = (\mathbb{R}^d, \cdot_t)$. Additionally, $(\phi_t)_{t>0}$ is a group dilation structure if and only if $\cdot_t = \cdot$ for all $t > 0$.

Definition 3.8 (Approximate group dilation structure) Let G be a simply connected nilpotent Lie group described by a global polynomial chart $(\mathbb{R}^d, \cdot)$. Let $(\phi_t)_{t>0}$ be a straight dilation structure. We say that this dilation structure is an *approximate group dilation structure* if, for any $x, y \in \mathbb{R}^d$, the limits

$$\lim_{t\to\infty} \phi_{1/t}(\phi_t(x)^{-1}) = x_\bullet^{-1} \quad \text{and} \quad \lim_{t\to\infty} \phi_{1/t}(\phi_t(x) \cdot \phi_t(y)) = x \bullet y$$

exist. □

Lemma 3.9 The pairing $(x, y) \mapsto x \bullet y$ yields a nilpotent Lie group $G_\bullet = (\mathbb{R}^d, \bullet)$ and $x_\bullet^{-1}$ is the inverse of x for the group law $\bullet$, that is $x_\bullet^{-1} \bullet x = x \bullet x_\bullet^{-1} = e_\bullet$. For the group $(\mathbb{R}^d, \bullet)$, the straight dilations $\{\phi_t;\ t > 0\}$ form a *group dilation structure*, i.e., satisfy (3.7). □

Proof By construction, the maps $P_t(x, y) = \phi_{1/t}(\phi_t(x) \cdot \phi_t(y))$ and $I_t(x) = \phi_{1/t}(\phi_t(x)^{-1})$ are polynomial maps in x, y with coefficients equal to linear combinations of power functions of t with exponents in $\mathbb{R}$. If the limits $\lim_{t\to\infty} P_t(x, y)$ and $\lim_{t\to\infty} I_t(x)$ exist for all x, y, it means that only non-positive powers of t occur and this implies that the families P_t, I_t are uniformly equicontinuous on compact sets. A sequence of simple considerations then yields that

$$x \bullet (y \bullet z) = \lim_{t\to\infty} \phi_{1/t}(\phi_t(x) \cdot \phi_t(y) \cdot \phi_t(z)) = (x \bullet y) \bullet z$$

and

$$x_\bullet^{-1} \bullet x = x \bullet x_\bullet^{-1} = e_\bullet = 0.$$

Note that this also implies

$$\lim_{t\to\infty} \phi_{1/t}\left(\phi_t(x)^{-1} \cdot \phi_t(y)\right) = x_\bullet^{-1} \bullet y. \tag{3.8}$$

Lemma 3.10 Let $(\phi_t)_{t>0}$ be a straight approximate group dilation structure on $(\mathbb{R}^d, \cdot)$. For any compact $K \subset \mathbb{R}^d$ there is a constant C_K such that, for any $x, y \in K$ and $t \geq 1$,

$$\|\phi_{1/t}\left(\phi_t(x)^{-1} \cdot \phi_t(y)\right)\|_2 \leq C_K \|y - x\|_2$$

and

$$\|\phi_{1/t}\left(\phi_t(x)^{-1} \cdot \phi_t(y)\right)\|_2 \leq C_K \|x_\bullet^{-1} \bullet y\|_2.$$

Proof The function $(t, x, y) \mapsto \phi_{1/t}\left(\phi_t(x)^{-1}\phi_t(y)\right)$ is a polynomial in

$$(x, y) = (x_1, \ldots, x_d, y_1, \ldots, y_d)$$

with coefficients equal to linear combinations of powers of t with exponents in $\mathbb{R}$. By (3.8), only non-positive powers of t appear. The desired inequality follows because this polynomial function equals 0 when $x = y$. The second inequality follows from the first and (3.3) applied to $(\mathbb{R}^d, \bullet)$. □

Remark 3.11 When working in exponential coordinates, we have the extra structure of the Lie bracket $[\cdot,\cdot]$ at our disposal and we can replace the conditions in Definition 3.8 by the condition that

$$\lim_{t\to\infty} \phi_{1/t}([\phi_t(x), \phi_t(y)]) = [x, y]_\bullet$$

exists. Call this an approximate Lie dilation structure. Note that, in this case, $x^{-1} = x_\bullet^{-1} = -x$ and $\phi_t(x)^{-1} = \phi_t(x^{-1})$ so that the inverse map condition is automatically satisfied. □

Remark 3.12 If $(\phi_t)_{t>0}$ is a group dilation structure (resp. an approximate group dilation structure) then so is $(\phi_{t^a})_{t>0}$, for any $a > 0$. Moreover, in the case of an approximate group dilation structure, this change does not affect the limit structure. □

Remark 3.13 The basic idea behind Definition 3.8 is well-known in two different related contexts. It appears in the study of the large scale geometry of groups of polynomial volume growth, see, e.g., [13, Section 2.2], and in the work of Alexopoulos on local limit theorems in the context of groups of polynomial volume growth, see [4, Section 5.2]. In these works, there is a unique relevant structure at infinity and it follows that the "dilation structures" considered there are very special examples of those defined here. Various forms of the same idea play an important role in the local study of sub-elliptic second order operators but in that context the limit is taken when the parameter t goes to 0. See for instance [59, Chapter V]. □

The following lemma is not used explicitly but serves as an exercise in manipulating the notion introduced above. See also Section 10.3.

Lemma 3.14 Let H be a subgroup of $G = (\mathbb{R}^d, \cdot)$ and $(\phi_t)_{t>0}$ be an approximate group dilation structure with limit law $\bullet$. Set

$$H_\bullet = \left\{ x \in \mathbb{R}^d : \text{ there exists } (x_k)_1^\infty \subset H \text{ so that } \lim_{k\to\infty} \phi_{1/k}(x_k) = x \right\}.$$

Then $H_\bullet$ is a subgroup of $G_\bullet = (\mathbb{R}^d, \bullet)$.

Proof Let $x, y \in H_\bullet$ with witness sequences $(x_k)_1^\infty$, $(y_k)_1^\infty$ in H. Fix $\varepsilon > 0$. By the continuity of $\bullet$, there exists $\delta > 0$ such that $\|x - x'\|_2 < \delta$ and $\|y - y'\|_2 \le \delta$ imply $\|x \bullet y - x' \bullet y'\|_2 < \varepsilon/2$. By the efinition of $H_\bullet$, there exists $N > 0$ such that $\|x - \phi_{1/k}(x_k)\|_2 < \delta$ and $\|y - \phi_{1/k}(y_k)\|_2 < \delta$ for all $k \ge N$. By the uniform convergence of $\phi_{1/t}(\phi_t(u) \cdot \phi_t(v))$ to $u \bullet v$ on compact sets, there exists N' such that, for all $k \ge N$ and $k' \ge N'$,

$$\|\phi_{1/k}(x_k) \bullet \phi_{1/k}(y_k) - \phi_{1/k'}(\phi_{k'}(\phi_{1/k}(x_k)) \cdot \phi_{k'}(\phi_{1/k}(y_k)))\|_2 < \varepsilon/2.$$

Hence, for $k \ge \max\{N, N'\}$,

$$\|x \bullet y - \phi_{1/k}(\phi_k(\phi_{1/k}(x_k)) \cdot \phi_k(\phi_{1/k}(y)))\|_2 < \varepsilon,$$

and thus, $\|x \bullet y - \phi_{1/k}(x_k \cdot y_k)\|_2 < \varepsilon$. Because $x_k \cdot y_k \in H$, this proves that $x \bullet y = \lim_{k\to\infty} \phi_{1/k}(x_k \cdot y_k) \in H_\bullet$. A similar proof applies to show that $x_\bullet^{-1} \in H_\bullet$ for $x \in H_\bullet$. □

Example 3.15 Consider the Heisenberg group viewed as the group of matrices

$$\mathbb{H}_3(\mathbb{R}) = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} : (x, y, z) \in \mathbb{R}^3 \right\}.$$

Here, the product of the matrices associated with (x, y, z) and (x', y', z') is associated with the triplet

$$(x + x', y + y', z + z' + xy').$$

The inverse of (x, y, z) is

$$(x, y, z)^{-1} = (-x, -y, -z + xy). \tag{3.9}$$

This is isomorphic but different from the "exponential coordinate description" discussed earlier where

$$(u, v, w) \cdot (u', v', w') = (u + u', v + v', w + w' + \tfrac{1}{2}(uv' - u'v)).$$

The map

$$q : (x, y, z) \to q(x, y, z) = (u, v, w) = (x, y, z - \tfrac{1}{2}xy) \tag{3.10}$$

provides the group isomorphism between these two descriptions.

Now, consider the group of diffeomorphisms $(\phi_t)_{t>0}$ (straight dilations in that system) given in the (x, y, z) matrix-coordinates by

$$\phi_t(x, y, z) = (t^a x, t^b y, t^c z) \quad \text{for some fixed } a, b, c > 0.$$

These are group diffeomorphisms for all $t > 0$ if and only if $c = a + b$. They form an approximate group dilation structure at infinity if and only if $c \geq a + b$. When $c > a + b$,

$$(x, y, z) \bullet (x', y', z') = (x + x', y + y', z + z')$$

and

$$(x, y, z)_\bullet^{-1} = (-x, -y, -z) \neq (x, y, z)^{-1}.$$

If we write down these same diffeomorphisms in the "exponential coordinate" description (u, v, w) they are given by the maps

$$\psi_t(u, v, w) = q^{-1} \circ \phi_t \circ q(u, v, w) = (t^a u, t^b v, t^c w + \tfrac{1}{2}(t^c - t^{a+b})uv).$$

In the (u, v, w) global coordinate chart exp = log = id, and if we assume $c \geq a + b$, the straight dilations

$$\delta_t(u,v,w) = (t^a u, t^b v, t^c w), \quad t > 0,$$

give both an approximate Lie dilation structure and an associated approximate group dilation structure which are distinct from the ϕ_t/ψ_t approximate group dilation structure even so they share the same differential at the identity. They lead to isomorphic limit group structures. □

Example 3.16 Consider the group

$$G = \left\{ \begin{pmatrix} 1 & x_{12} & x_{13} & x_{14} \\ 0 & 1 & x_{23} & x_{24} \\ 0 & 0 & 1 & x_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} : x_{ij} \in \mathbb{R} \right\}$$

and the straight dilation structures associated with any tuple

$$1/\alpha_{ij},\ ij = (i,j) \in \{12, 13, 14, 23, 24, 34\}$$

so that

$$\delta_t \left(\begin{pmatrix} 1 & x_{12} & x_{13} & x_{14} \\ 0 & 1 & x_{23} & x_{24} \\ 0 & 0 & 1 & x_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \right) = \begin{pmatrix} 1 & y_{12} & y_{13} & y_{14} \\ 0 & 1 & y_{23} & y_{24} \\ 0 & 0 & 1 & y_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad y_{ij} = t^{1/\alpha_{ij}} x_{ij}.$$

Such a $(\delta_t)_{t>0}$ is a group dilation structure if and only if

$$1/\alpha_{k\ell} = 1/\alpha_{kj} + 1/\alpha_{j\ell} \quad \text{for all } 1 \le k < j < \ell \le 4,$$

that is,

$$(1):\ 1/\alpha_{13} = 1/\alpha_{12} + 1/\alpha_{23}, \qquad (2):\ 1/\alpha_{24} = 1/\alpha_{23} + 1/\alpha_{34}$$

and

$$(3):\ 1/\alpha_{14} = 1/\alpha_{12} + 1/\alpha_{24}, \qquad (4):\ 1/\alpha_{14} = 1/\alpha_{13} + 1/\alpha_{34}.$$

The group $(\phi_t)_{t>0}$ is an approximate group dilation structure at infinity if and only if

$$1/\alpha_{k\ell} \ge 1/\alpha_{kj} + 1/\alpha_{j\ell} \quad \text{for all } 1 \le k < j < \ell \le 4. \tag{3.11}$$

We now list all the possible Lie structures that appear as a limit of such an approximate group dilation structure on G.

1. When equality holds in all of the inequalities (3.11), we have $G_\bullet = G$.
2. When strict inequality holds in all of the inequalities (3.11), we have $G_\bullet = \mathbb{R}^6$ (abelian).
3. When equations (1) and (2) are equalities then equations (3) and (4) become equivalent. Assume a strict inequality holds in (3) and (4). Then the limit $G_\bullet$ is

$$\left\{ \left(\begin{pmatrix} 1 & x_{12} & x_{13} \\ 0 & 1 & x_{23} \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & x_{23} & x_{24} \\ 0 & 1 & x_{34} \\ 0 & 0 & 1 \end{pmatrix}, \left(x_{14} \right) \right) : x_{ij} \in \mathbb{R} \right\}.$$

Here multiplication for these triplets of matrices is matrix-coordinate by matrix-coordinate. Note how the same x_{23} appears in the first and second matrix-coordinates.

4. When strict inequality holds in both equations (1) and (2) and equality holds in both (3) and (4), then the limit $G_\bullet$ is

$$\left\{ \left(\begin{pmatrix} 1 & x_{12} & x_{13} & x_{14} \\ 0 & 1 & 0 & x_{24} \\ 0 & 0 & 1 & x_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix}, \left(x_{23} \right) \right) : x_{ij} \in \mathbb{R} \right\}$$

(this is the direct product of the 5 dimensional Heisenberg group $\mathbb{H}_5(\mathbb{R})$ and a copy of $\mathbb{R}$).

5. When strict inequality holds in both equations (1) and (2) and equality holds in (3) but not in (4) (resp. (4) but not in (3)), then the limit $G_\bullet$ is

$$\left\{ \left(\begin{pmatrix} 1 & x_{12} & x_{14} \\ 0 & 1 & x_{24} \\ 0 & 0 & 1 \end{pmatrix}, \left(x_{13} \right), \left(x_{23} \right), \left(x_{34} \right) \right) : x_{ij} \in \mathbb{R} \right\}$$

(resp. exchange the roles of pairs x_{12}, x_{24} and x_{13}, x_{34}). This is the direct product of a copy of $\mathbb{H}_3(\mathbb{R})$ and $\mathbb{R}^3$).

6. When strict inequality holds in (1) (resp. (2)) and equality holds in (2) (resp. (1)) then strict inequality must hold in (3) (resp. (4)). If equality holds in (4) (resp. (3)), the limit group is isomorphic to

$$\left\{ \left(\left(x_{12} \right), \begin{pmatrix} 1 & 0 & x_{13} & x_{14} \\ 0 & 1 & x_{23} & x_{24} \\ 0 & 0 & 1 & x_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \right) : x_{ij} \in \mathbb{R} \right\}$$

(resp. exchange the roles of x_{12} and x_{34}, the limit groups in both cases are isomorphic).

7. When strict inequality holds in (1) (resp. (2)) and equality holds in (2) (resp. (1)) and strict inequality holds in each of (3) and (4), the limit group is isomorphic to

$$\left\{ \left(\begin{pmatrix} 1 & x_{23} & x_{24} \\ 0 & 1 & x_{34} \\ 0 & 0 & 1 \end{pmatrix}, \left(x_{12} \right), \left(x_{13} \right), \left(x_{14} \right) \right) : x_{ij} \in \mathbb{R} \right\}$$

(resp. replace the triplet (x_{23}, x_{24}, x_{34}) with (x_{12}, x_{13}, x_{23}) and the triplet (x_{12}, x_{13}, x_{14}) with (x_{34}, x_{24}, x_{14})). This is the direct product of a copy of $\mathbb{H}_3(\mathbb{R})$ and $\mathbb{R}^3$). □

Chapter 4
Vague convergence and change of group law

Abstract This short chapter is devoted to a key technical result which consists in passing from the vague convergence of the family of rescaled measures associated with the driving probability measure of a long-range random walk to the vague convergence of the associated jump kernels. This involves taking into account the change of group law induced by the rescaling of space through an approximate group dilation.

4.1 Vague convergence under rescaling

We consider a rather general situation pertaining to the problem we want to study. We are given the following data:

(a) A finitely generated torsion free nilpotent group Γ given as a co-compact closed subgroup of a simply connected nilpotent Lie group G. It is useful for our purpose to be more explicit and write $G = (\mathbb{R}^d, \cdot)$ where this coordinate system is a polynomial coordinate system as explained earlier.
(b) A probability measure μ on Γ.
(c) An approximate group dilation structure $(\delta_t)_{t>0}$ on G with Lie group limit $G_\bullet = (\mathbb{R}^d, \bullet)$.

Definition 4.1 We say that the approximate group dilation structure $(\delta_t)_{t>0}$ is admissible for μ if the family of measures

$$\mu_t = t\delta_{1/t}(\mu) \quad \text{defined by } \mu_t(\phi) := t\int_{\mathbb{R}^d} \phi(\delta_{1/t}(u))\mu(du) \tag{4.1}$$

converge vaguely to a Radon measure $\mu_\bullet$ on $\mathbb{R}^d \setminus \{0\}$ as $t \to \infty$. Recall that, by definition, this means that, for any continuous function ϕ with compact support in $\mathbb{R}^d \setminus \{0\}$,

$$\lim_{t\to\infty} \int \phi(x)d\mu_t(x) = \int \phi(x)d\mu_\bullet(x).$$

Remark 4.2 Note the following identities:

$$\mu_t(A) = t\mu(\delta_t(A)) = t\sum_{y\in\Gamma} \mathbb{1}_{\delta_t A}(y)\mu(y) = t\sum_{x\in\delta_t^{-1}\Gamma} \mathbb{1}_A(x)\mu(\delta_t x)$$

and

$$\int \phi(x)d\mu_t(x) = t\int \phi(\delta_t^{-1}y)d\mu(y) = t\sum_{x\in\delta_t^{-1}\Gamma}\phi(x)\mu(\delta_t x).$$

Remark 4.3 The normalization by a factor of t in $\mu_t = t\delta_{1/t}(\mu)$ is less restrictive than it may first appear because of Remark 3.12. If there is an approximate Lie dilation structure $(\delta_t)_{t>0}$ (with limit law $\bullet$) such that the measure $\mu_t = t^a\delta_t^{-1}(\mu)$ converges vaguely to $\mu_\bullet$ on $\mathbb{R}^d\setminus\{0\}$ then the modified approximate Lie dilation structure $(\delta_{t^{1/a}})_{t>0}$ gives the same limit law $\bullet$ and is admissible for μ. In this sense, the choice of the linear t factor in the definition of μ_t amounts, more or less, to a scaling normalization. □

Example 4.4 Fix $\alpha\in(0,2)$ and let μ be the probability measure on $\mathbb{Z}\subset\mathbb{R}$ with

$$\mu(k) = c_\alpha(1+|k|)^{-\alpha-1}.$$

Let $\delta_t(x) = t^{1/\alpha}x$. Then $t\delta_t^{-1}(\mu)$ converges vaguely on $\mathbb{R}\setminus\{0\}$ as $t\to\infty$ to the measure $\mu_\bullet$ with density $c_\alpha|x|^{-\alpha-1}$ with respect to the Lebesgue measure on $\mathbb{R}$. □

Example 4.5 Fix $\alpha\in(0,2)$ and $\beta\in(0,\alpha)$. Let μ be the probability measure on $\mathbb{Z}^2$ given by

$$\mu((x,y)) = c(1+|x|+|y|)^{-\alpha-2},\quad (x,y)\in\mathbb{Z}^2\subset\mathbb{R}^2.$$

Let $\delta_t((x,y)) = (t^{1/\alpha}x, t^{1/\beta}y)$. Then $t\delta_t^{-1}(\mu)$ converges vaguely on $\mathbb{R}^2\setminus\{(0,0)\}$ as $t\to\infty$ to the measure $\mu_\bullet(dxdy) = f_\bullet(x)dx\otimes\delta_0(dy)$ supported on the x-axis with $f_\bullet(x) = c'|x|^{-\alpha-1}$, where $c' = c\int_{\mathbb{R}}(1+u)^{-\alpha-2}du$. □

Example 4.6 On the Heisenberg group $\mathbb{H}_3(\mathbb{Z})$ viewed as the group of matrix

$$\left\{\begin{pmatrix}1 & x_1 & x_3\\ 0 & 1 & x_2\\ 0 & 0 & 1\end{pmatrix} : x_1,x_2,x_3\in\mathbb{Z}\right\}, \tag{4.2}$$

consider the measure

$$\mu((x_1,x_2,x_3)) = \frac{c_\alpha}{\left(1+\sqrt{x_1^2+x_2^2+|x_3-x_1x_2/2|}\right)^{\alpha+4}}, \tag{4.3}$$

(note that this is a symmetric measure). Consider an approximate Lie dilation structure $(\delta_t)_{t>0}$ of the form $\delta_t((x_i)_1^3) = (t^{1/\gamma_i}x_i)_1^3$. For this to be an approximate Lie dilation structure, it must be that $1/\gamma_3\geq 1/\gamma_1+1/\gamma_2$ which we assume. For the measure $t\delta_t^{-1}(\mu_t)$ to have a vague limit, it is necessary that $1/\gamma_1\geq 1/\alpha, 1/\gamma_2\geq 1/\alpha$ and $1/\gamma_3\geq 2/\alpha$. Note that the roles of x and y are the same so that we can assume for the sake of the computations described below that $1/\gamma_1\leq 1/\gamma_2$.

1. Assume that $1/\gamma_2\geq 1/\gamma_1 > 1/\alpha$. Then $1/\gamma_3 > 2/\alpha$ and it is not hard to see that $t\delta_t^{-1}(\mu)$ converges vaguely to 0 as $t\to\infty$.

2. Assume $\gamma_1 = \gamma_2 = \gamma_3/2 = \alpha$. Then $(\delta_t)_{t>0}$ is a group dilation structure and $t\delta_t^{-1}(\mu)$ converges vaguely as $t \to \infty$ to

$$\mu_\bullet((dx_1, dx_2, dx_3)) = \frac{c_\alpha dx_1 dx_2 dx_3}{\left(\sqrt{x_1^2 + x_2^2 + |x_3 - x_1 x_2/2|}\right)^{\alpha+4}}.$$

3. Assume that $1/\gamma_1 = 1/\gamma_2 = 1/\alpha$ and $1/\gamma_3 > 2/\alpha$. Then $t\delta_t^{-1}(\mu) \Rightarrow \mu_\bullet$ as $t \to \infty$, where

$$\mu_\bullet(dx_1 dx_2 dx_3) = \frac{c'}{\left(\sqrt{x_1^2 + x_2^2}\right)^{\alpha+2}} dx_1 dx_2 \otimes \delta_0(dx_3)$$

with $c' = 2c \int_0^\infty (1+s)^{-(2+\alpha/2)} ds$.

4. Assume that $1/\gamma_2 > 1/\gamma_1 = 1/\alpha$. It follows that $1/\gamma_3 \geq 1/\gamma_1 + 1/\gamma_2 > 2/\alpha$. In this case $t\delta_t^{-1}(\mu) \Rightarrow \mu_\bullet$ as $t \to \infty$, where

$$\mu_\bullet(dx_1 dx_2 dx_3) = c'|x_1|^{-\alpha+1} dx_1 \otimes \delta_0(dx_2) \otimes \delta_0(dx_3)$$

with

$$c' = 2c \int_{-\infty}^{\infty} \left(\int_0^\infty \left(\sqrt{1 + u^2 + v}\right)^{-(\alpha+4)} dv \right) du.$$

We provide details for the third case (the fourth case is similar). Let f be a continuous function with compact support in $\mathbb{R}^3 \setminus \{0\}$. We want to show that

$$\begin{aligned}
\lim_{t\to\infty} \int f(x) d\mu_t(x) &= \lim_{t\to\infty} t \sum_{x \in \mathbb{Z}^3} \frac{c_\alpha f(\delta_t^{-1}(x))}{\left(1 + \sqrt{x_1^2 + x_2^2 + |x_3 - x_1 x_2/2|}\right)^{\alpha+4}} \\
&= \int_{\mathbb{R}^3} \frac{c' f(x)}{\left(\sqrt{x_1^2 + x_2^2}\right)^{\alpha+2}} dx_1 dx_2 \otimes \delta_0(dx_3) \\
&= \int_{\mathbb{R}^2} \frac{c' f(x_1, x_2, 0)}{\left(\sqrt{x_1^2 + x_2^2}\right)^{\alpha+2}} dx_1 dx_2.
\end{aligned}$$

Recall that $\gamma_1 = \gamma_2 = \alpha$ and $2\gamma_3 < \alpha$ and write

$$\begin{aligned}
\int f(x) t\delta_t^{-1}(\mu(dx)) &= t \sum_{x \in \mathbb{Z}^3} f(\delta_t^{-1}(x)) \frac{c_\alpha}{\left(1 + \sqrt{x_1^2 + x_2^2 + |x_3 - x_1 x_2/2|}\right)^{\alpha+4}} \\
= t \sum_{z \in \delta_t^{-1}(\mathbb{Z}^3)} f(z) &\frac{c_\alpha}{\left(1 + \sqrt{t^{2/\gamma_1} z_1^2 + t^{2/\gamma_2} z_2^2 + |t^{1/\gamma_3} z_3 - t^{1/\gamma_1 + 1/\gamma_2} z_1 z_2/2|}\right)^{\alpha+4}}
\end{aligned}$$

$$= t^{-4/\alpha} \sum_{z \in \delta_t^{-1}(\mathbb{Z}^3)} f(z) \frac{c_\alpha}{\left(t^{-1/\alpha} + \sqrt{z_1^2 + z_2^2 + |t^{1/\gamma_3 - 2/\alpha} z_3 - z_1 z_2/2|}\right)^{\alpha+4}}$$

$$= t^{-4/\alpha} \sum_{z \in (t^{-1/\alpha}\mathbb{Z})^2 \times t^{-2/\alpha}\mathbb{Z}} \frac{c_\alpha f((z_1, z_2, 0))}{\left(t^{-1/\alpha} + \sqrt{z_1^2 + z_2^2 + |z_3 - z_1 z_2/2|}\right)^{\alpha+4}}$$

$$+ t^{-4/\alpha} \sum_{z \in (t^{-1/\alpha}\mathbb{Z})^2 \times t^{-2/\alpha}\mathbb{Z}} \frac{c_\alpha (f((z_1, z_2, t^{-1/\gamma_3 + 2/\alpha} z_3)) - f((z_1, z_2, 0)))}{\left(t^{-1/\alpha} + \sqrt{z_1^2 + z_2^2 + |z_3 - z_1 z_2/2|}\right)^{\alpha+4}}.$$

The first term is, essentially, a (multivariate, generalized) Riemann sum of a uniformly continuous integrable function on $\mathbb{R}^3$ over the lattice $(t^{-1/\alpha}\mathbb{Z})^2 \times t^{-2/\alpha}\mathbb{Z}$ and, consequently, it converges when t tends to infinity to

$$\int_{\mathbb{R}^3} \frac{c_\alpha f(x_1, x_2, 0)}{\left(\sqrt{x_1^2 + x_2^2 + |x_3 - x_1 x_2/2|}\right)^{\alpha+4}} dx_1 dx_2 dx_3 = \int_{\mathbb{R}^2} \frac{c'_\alpha f(z_1, z_2, 0)}{(z_1^2 + z_2^2)^{(\alpha+2)/2}} dx_1 dx_2,$$

where $c'_\alpha = 2c_\alpha \int_0^\infty \frac{du}{(1+u)^{(\alpha+4)/2}}$.

The second term goes to 0 when t tends to ∞ because f is uniformly continuous and $1/\gamma_3 - 2/\alpha > 0$: for any $\varepsilon > 0$ there is a T_ε such that for all $t > T_\varepsilon$,

$$|((z_1, z_2, t^{-1/\gamma_3 + 2/\alpha} z_3)) - f((z_1, z_2, 0))| < \varepsilon.$$

This gives

$$t^{-4/\alpha} \sum_{(t^{-1/\alpha}\mathbb{Z})^2 \times t^{-2/\alpha}\mathbb{Z}} \frac{c_\alpha |f((z_1, z_2, t^{-1/\gamma_3 + 2/\alpha} z_3)) - f((z_1, z_2, 0))|}{\left(\sqrt{z_1^2 + z_2^2 + |z_3 - z_1 z_2/2|}\right)^{\alpha+4}}$$

$$\leq \varepsilon t^{-4/\alpha} \sum_{(t^{-1/\alpha}\mathbb{Z})^2 \times t^{-2/\alpha}\mathbb{Z}} \frac{c_\alpha}{\left(\sqrt{z_1^2 + z_2^2 + |z_3 - z_1 z_2/2|}\right)^{\alpha+4}}.$$

When t tends to infinity, the limit of the right-hand side is

$$\varepsilon \int_{\mathbb{R}^3} \frac{c_\alpha}{\left(\sqrt{x_1^2 + x_2^2 + |x_3 - x_1 x_2/2|}\right)^{\alpha+4}} dx_1 dx_2 dx_3.$$

As $\varepsilon > 0$ is arbitrary, this proves that

$$\lim_{t \to \infty} t^{-4/\alpha} \sum_{(t^{-1/\alpha}\mathbb{Z})^2 \times t^{-2/\alpha}\mathbb{Z}} \frac{c_\alpha (f((z_1, z_2, t^{1/\gamma_3} z_3)) - f((z_1, z_2, 0)))}{\left(t^{-1/\alpha} + \sqrt{z_1^2 + z_2^2 + |z_3 - z_1 z_2/2|}\right)^{\alpha+4}} = 0$$

as desired. □

4.2 Vague convergence of jump measures and kernels

Next, we relate the vague convergence of μ_t to $\mu_\bullet$ to the vague convergence of jump kernels.

Proposition 4.7 Let $\Gamma \subset G$ be a discrete co-compact subgroup of the simply connected nilpotent Lie group $G = (\mathbb{R}^d, \cdot)$. Let $c(\Gamma, G)$ be the Haar volume of G/Γ (i.e., of a fundamental domain for Γ in G). Let μ be a probability measure on Γ, $(\delta_t)_{t>0}$ be an approximate group dilation structure on G which is admissible for μ, and let $\mu_t := t\delta_{1/t}(\mu)$. Suppose that μ_t converges vaguely on $\mathbb{R}^d \setminus \{0\}$ to a Radon measure $\mu_\bullet$ as t tends to infinity. Then, for any continuous and compactly supported function ϕ in $\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta$, the positive Radon measure $J_t(dxdy)$ on $\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta$ defined by

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta} \phi(x,y) J_t(dxdy) = \\ c(\Gamma, G) t \det(\delta_{1/t}) \sum_{x,y \in \delta_{1/t}(\Gamma),\ x \neq y} \phi(x,y) \mu(\delta_t(x)^{-1} \cdot \delta_t(y))$$

converges vaguely as t tends to infinity to the positive Radon measure $J_\bullet$ defined on $(\mathbb{R}^d \times \mathbb{R}^d) \setminus \Delta$ by

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta} \phi(x,y) J_\bullet(dxdy) = \iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta} \phi(x, x \bullet y) dx \mu_\bullet(dy),$$

where $\bullet$ is the limit law $x \bullet y = \lim_{t \to \infty} \delta_t^{-1}(\delta_t(x) \cdot \delta_t(y))$ for the approximate Lie dilation structure $(\delta_t)_{t>0}$. □

Remark 4.8 Of course, in the group $G_\bullet = (\mathbb{R}^d, \bullet)$, we can write

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta} \phi(x, x \bullet y) dx \mu_\bullet(dy) = \iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta} \phi(x,y) dx \mu_\bullet(x_\bullet^{-1} \bullet dy)$$

(the inverse operation is in $(G, \bullet)$) so that

$$J_\bullet(dxdy) = dx \mu_\bullet(x_\bullet^{-1} \bullet dy).$$

Remark 4.9 Note that the measure $J_t^1(dx)$ defined by

$$\int \phi(x) J_t^1(dx) = c(\Gamma, G) \det(\delta_{1/t}) \sum_{x \in \delta_{1/t}(\Gamma)} \phi(x)$$

obviously converges to $\int \phi(x) dx$ as t tends to infinity. That is, the vague limit of $J_t^1(dx)$ is the Lebesgue (=Haar) measure on $\mathbb{R}^d$. □

Proof Observe that

$$x \cdot_t y = \delta_t^{-1}(\delta_t(x) \cdot \delta_t(y)) \tag{4.4}$$

is a group law which turns $\mathbb{R}^d$ into a Lie group $G_t = (\mathbb{R}^d, \cdot_t)$ (this group is actually isomorphic to G). For any fixed $x \in \delta_{1/t}(\Gamma)$, consider

$$\begin{aligned} t \sum_{y \in \delta_{1/t}(\Gamma) \setminus \{x\}} \phi(x,y)\mu(\delta_t(x)^{-1} \cdot \delta_t(y)) &= t \sum_{y \in \delta_{1/t}(\Gamma) \setminus \{x\}} \phi(x,y)\mu(\delta_t(x^{-1} \cdot_t y)) \\ &= t \sum_{y \in \delta_{1/t}(\Gamma) \setminus \{x\}} \phi(x, x \cdot_t y)\mu(\delta_t(y)). \end{aligned}$$

Now, write

$$\begin{aligned} &\iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta} \phi(x,y) J_t(dxdy) - \iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta} \phi(x,y) J_\bullet(dxdy) \\ &= \iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta} \phi(x,y) J_t(dxdy) - \sum_{y \in \delta_{1/t}(\Gamma) \setminus \{e\}} \int_{\mathbb{R}^d} \phi(x, x \bullet y) t\mu(\delta_t(y)) dx \\ &\quad + \sum_{y \in \delta_{1/t}(\Gamma) \setminus \{e\}} \int_{\mathbb{R}^d} \phi(x, x \bullet y) t\mu(\delta_t(y)) dx - \iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta} \phi(x,y) J_\bullet(dxdy) \\ &= I_1(t) + I_2(t). \end{aligned}$$

To bound $|I_1|$, write $c_t = c(\Gamma, G)\det(\delta_{1/t})$ and

$$|I_1(t)| = \left| \sum_{y \in \delta_{1/t}(\Gamma) \setminus \{e\}} t\mu(\delta_t(y)) \left(c_t \sum_{x \in \delta_{1/t}(\Gamma)} \phi(x, x \cdot_t y) - \int \phi(x, x \bullet y) dx \right) \right|.$$

Note that ϕ is continuous and compactly supported in $\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta$ and $x \cdot_t y$ converges (uniformly on compact sets) to $x \bullet y$. It follows that there is a compact set $K = K_\phi$ in $\mathbb{R}^d \setminus \{0\}$ with the property that, for any $\varepsilon > 0$, there is T such that, for all $y \in \mathbb{R}^d$ and all $t > T$,

$$\left| c_t \sum_{x \in \delta_t^{-1}\Gamma} \phi(x, x \cdot_t y) - \int \phi(x, x \bullet y) dx \right| \le \varepsilon \mathbb{1}_K(y).$$

Also, there exist C_K and T' such that for all $t > T'$, $t\mu(\delta_t(K)) \le C_K$. It follows that $|I_1(t)| \le \varepsilon C_K$. As for $|I_2(t)|$, the fact that it converges to 0 is a consequence of the vague convergence of $t\delta_t^{-1}(\mu)$ to $\mu_\bullet$ on $\mathbb{R}^d \setminus \{0\}$. □

The jump kernel J_t introduced above is defined on $\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta$ and acts on functions of $x, y \in \mathbb{R}^d \times \mathbb{R}^d \setminus \Delta$. It is useful to consider also a related discrete jump kernel supported on

$$\Gamma_t \times \Gamma_t \setminus \Delta,$$

where $\Gamma_t = \delta_t^{-1}(\Gamma)$ (by abuse of notation, we use the letter Δ to demote the diagonal on $R \times R$ for any space R, e.g., $R = \mathbb{R}$ or $R = \Gamma_t$). Note that Γ_t is a co-compact subgroup of the group $G_t = (\mathbb{R}^d, \cdot_t)$ defined at (4.4) and that δ_t provides a group isomorphism from Γ_t onto Γ. We equipped Γ_t with the rescaled counting measure

$$m_t(A) = c(\Gamma, G)\det(\delta_t^{-1})|A|, \quad \text{where } |A| = \#A \tag{4.5}$$

for any finite subset $A \subset \Gamma_t$. On Γ_t, we consider the jump kernel measure j_t defined by

$$j_t(x, y) = c(\Gamma, G)t\det(\delta_{1/t})\mu(\delta_t(x)^{-1} \cdot \delta_t(y)), \quad (x, y) \in \Gamma_t \times \Gamma_t \setminus \Delta. \tag{4.6}$$

We now assume that the probability measure μ on Γ is symmetric. Then $j_t(x, y)$ is symmetric in (x, y) and it gives arise to an associated symmetric Dirichlet form in $L^2(\Gamma_t, m_t)$ with domain $\mathcal{F}^{(t)} := L^2(\Gamma_t, m_t)$ defined by

$$\mathcal{E}^{(t)}(u, v) = \frac{1}{2}\sum_{x,y\in\Gamma_t}(u(x) - u(y))(v(x) - v(y))j_t(x, y), \quad u, v \in \mathcal{F}^{(t)}. \tag{4.7}$$

The infinitesimal generator of this Dirichlet form on $L^2(\Gamma_t, m_t)$ is

$$f \mapsto -t(f - f *_{\Gamma_t} \delta_t^{-1}(\mu)) \tag{4.8}$$

on Γ_t.

Recall that $\Gamma_t \subset \mathbb{R}^d$. For each $x \in \mathbb{R}^d$, let $[x]_t \in \Gamma_t$ be the point closest to x in the $\|\cdot\|$-norm (if there are more than two such points, we choose one arbitrary and fix it). When needed, extend a function f on Γ_t to a function $\tilde{f}$ on $\mathbb{R}^d$ by setting $f(x) = f([x]_t)$ for each $x \in \mathbb{R}^d$. We say a family of functions $\{f_t : \Gamma_t \to \mathbb{R}\}_{t\ge 1}$ converges uniformly to a function f on $\mathbb{R}^d$ if $\tilde{f}$ converges uniformly to f.

The following is an easy consequence of Proposition 4.7 that relates to j_t. It is stated for continuous limit but it obviously holds as well for sequential limits based on an arbitrary sequence t_k tending to infinity.

Lemma 4.10 Let $\{f_t : \Gamma_t \to \mathbb{R}\}_{t>0}$ (resp. $\{g_t : \Gamma_t \to \mathbb{R}\}_t$) be a family of continuous functions that converges uniformly to a continuous function f (resp. g) on $\mathbb{R}^d$. Then, under the assumptions of Proposition 4.7, for any open set $U \subset \mathbb{R}^d \times \mathbb{R}^d \setminus \{(x, y) : \|x_\bullet^{-1} \bullet y\|_2 \le \eta\}$ with $\eta > 0$ whose closure is compact, it holds that

$$\begin{aligned}&\lim_{t\to\infty}\sum_{(x,y)\in(\Gamma_t\times\Gamma_t)\cap U}(f_t(x) - f_t(y))(g_t(x) - g_t(y))j_t(x, y)\\ &= \iint_U (f(x) - f(y))(g(x) - g(y))J_\bullet(dxdy).\end{aligned}$$

□

Proof Set $\psi_t(x, y) := (f_t(x) - f_t(y))(g_t(x) - g_t(y))$ and $\psi(x, y) := (f(x) - f(y))(g(x) - g(y))$. Then

$$
\begin{aligned}
&\Big|\sum_{(x,y)\in(\Gamma_t\times\Gamma_t)\cap U}\psi_t(x,y)j_t(x,y)-\iint_U\psi(x,y)J_\bullet(dxdy)\Big|\\
&\le\Big|\sum_{(x,y)\in(\Gamma_t\times\Gamma_t)\cap U}(\psi_t(x,y)-\psi(x,y))j_t(x,y)\Big|\\
&\quad+\Big|\iint_U(\psi(x,y)J_t(dxdy)-\psi(x,y)J_\bullet(dxdy))\Big|=:I_1+I_2.
\end{aligned}
$$

By Proposition 4.7, $\sup_{t\geq 1}\sum_{(x,y)\in(\Gamma_t\times\Gamma_t)\cap U} j_t(x,y)<\infty$. It follows that $\lim_{t\to\infty} I_1 = 0$ because ψ_t converges uniformly to ψ. By the proof of Proposition 4.7 (and the fact that U is compact in $\mathbb{R}^d\times\mathbb{R}^d\setminus\Delta$), $\lim_{k\to\infty} I_2 = 0$. □

Chapter 5
Weak convergence of the processes

Abstract This chapter is devoted to our main functional limit theorem for symmetric long range random walks on finitely generated torsion free nilpotent groups. A set of technical conditions are identified as sufficient conditions to establish a functional limit theorem by adapting established techniques in the present setting. These conditions are phrased mostly in terms of the given random walk but also involve the existence of an appropriate approximate dilation. All together, they are rather strong conditions and finding ways to work under less stringent hypotheses is an interesting open problem.

5.1 Assumption (A)

In this chapter and next, we prove limit theorems involving

(a) A finitely generated torsion free nilpotent group Γ embedded as a co-compact lattice in a simply connected nilpotent Lie group G;
(b) A symmetric probability measure μ on Γ;
(c) A polynomial coordinate system for $G = (\mathbb{R}^d, \cdot)$ and straight dilation structure

$$\delta_t,\ t > 0, \quad \delta_t((u_i)_1^d) = (t^{1/\beta_i} u_i)_1^d, \quad \beta_i \in (0,2), \quad i = 1, \dots, d, \tag{5.1}$$

which is an approximate group dilation structure for G with limit group $G_\bullet = (\mathbb{R}^d, \bullet)$.

The key hypothesis we will make that links together the probability measure μ, the dilation structure $(\delta_t)_{t>0}$ and the limit group $G_\bullet$ is that

(A) The straight dilation structure $(\delta_t)_{t>0}$ is admissible for the probability measure μ, that is, the (positive) measure $\mu_t = t\delta_{1/t}(\mu)$, $t \geq 1$, defined at (4.1) converges vaguely to a non-trivial Radon measure $\mu_\bullet$ on $\mathbb{R}^d \setminus \{0\}$ as t tends to infinity.

Remark 5.1 (i) The Radon measure $\mu_\bullet$ appeared in **(A)** is on $\mathbb{R}^d \setminus \{0\}$ and is expressed under the global coordinate system we use for the nilpotent group G and hence for $G_\bullet$. It induces a Radon measure of $G_\bullet$ through this global coordinate system. By abusing the notations, we use the same notation $\mu_\bullet$ for the induced measure on $G_\bullet$.
(ii) Under assumption **(A)**, it follows from the definition of μ_t that $\mu_\bullet$ is a symmetric measure on $G_\bullet \setminus \{e\}$ and has the following scaling property

$$\delta_r(\mu_\bullet) = r\mu_\bullet \quad \text{for every } r > 0; \tag{5.2}$$

that is, for any Borel measurable set $A \subset \mathbb{R}^d \setminus \{0\}$, $\mu_\bullet(A) = \mu_\bullet(A_\bullet^{-1})$, where $A_\bullet^{-1} := \{x \in \mathbb{R}^d : x_\bullet^{-1} \in A\}$, and

$$\mu_\bullet(\delta_r^{-1}(A)) = \mu_\bullet(\delta_{1/r}(A)) = r\mu_\bullet(A) \quad \text{for every } r > 0.$$

We are most interested in the case the limit measure $\mu_\bullet$ is not supported on a proper closed connected subgroup of $G_\bullet$. In that case, the condition that the exponents $\{\beta_i, 1 \le i \le d\}$ for the straight dilation structure $\{\delta_t; t \ge 0\}$ of (5.1) are in $(0,2)$ means that the original measure μ must have some sort of heavy tail characteristics, i.e., μ has to be "stable-like".

Geometries on $\mathbb{R}^d$ and $G_\bullet$

Fix $\beta \ge \max_{1\le i\le d}\{\beta_i\}$. By [37], there is a norm $\|\cdot\|$ on $G_\bullet = (\mathbb{R}^d, \bullet)$ (this means that $\|x \bullet y\| \le \|x\| + \|y\|$ for all $x, y \in \mathbb{R}^d$, $\|x_\bullet^{-1}\| = \|x\|$ and $\|x\| = 0$ if and only if $x = 0$) such that

$$\|\delta_t(u)\| = t^{1/\beta}\|u\| \quad \text{for every } t > 0 \text{ and } u = (u_i)_1^d \in \mathbb{R}^d. \tag{5.3}$$

This implies, of course, that there are constants $c, C \in (0,\infty)$ such that

$$c \max_{1\le i\le d}\{|u_i|^{\beta_i/\beta}\} \le \|u\| \le C \max_{1\le i\le d}\{|u_i|^{\beta_i/\beta}\} \quad \text{for } u = (u_i)_1^d \in \mathbb{R}^d. \tag{5.4}$$

Note that $\max_{1\le i\le d}\{|u_i|^{\beta_i/\beta}\}$ itself is a norm on $(\mathbb{R}^d, +)$ but not necessarily on $G_\bullet = (\mathbb{R}^d, \bullet)$ (it may not be symmetric on $G_\bullet$ and only satisfies the triangle inequality up to a multiplicative constant in general). Set

$$B(r) = \left\{x \in \mathbb{R}^d : \|x\| < r\right\}.$$

Obviously, we have

$$\|\delta_t(u)\| = t^{1/\beta}\|u\| \quad \text{and} \quad \delta_t(B(r)) = B(rt^{1/\beta}).$$

This means that the volume (the Lebesgue measure) of $B(r)$ is

$$m(B(r)) = m(\delta_{r^\beta}(B(1))) = m(B(1))\det(\delta_{r^\beta}) = m(B(1))r^{\beta(\sum_{i=1}^d 1/\beta_i)}.$$

Recall that $\mathbb{R}^d$ is also equipped with the Euclidean norm $\|u\|_2 = \sqrt{\sum_1^d |u_i|^2}$. Let

$$\beta_- = \min_{1\le i\le d}\beta_i \quad \text{and} \quad \beta_+ = \max_{1\le i\le d}\beta_i.$$

From the definition, it is clear that

$$c\min\{\|u\|^{\beta/\beta_-}, \|u\|^{\beta/\beta_+}\} \le \|u\|_2 \le C\max\{\|u\|^{\beta/\beta_-}, \|u\|^{\beta/\beta_+}\}. \tag{5.5}$$

Similarly, for any $u \in \mathbb{R}^d$ with $\|u\| \le C_1 r^{1/\beta}$, we have

$$c_2\left(\frac{\|u\|_2}{r}\right)^{\beta_+/\beta} \le \frac{\|u\|}{r^{1/\beta}} \le C_2\left(\frac{\|u\|_2}{r}\right)^{\beta_-/\beta}. \tag{5.6}$$

We will need the following version of Lemma 3.10 with respect to the norm $\|\cdot\|$.

Lemma 5.2 For any compact $K \subset \mathbb{R}^d$ there is a constant C_K such that, for any $x, y \in K$ and $t \ge 1$,

$$\|\delta_{1/t}\left(\delta_t(x)^{-1}\cdot\delta_t(y)\right)\| \le C_K\|y-x\|^{\beta_-/\beta_+}$$

and

$$\|\delta_{1/t}\left(\delta_t(x)^{-1}\cdot\delta_t(y)\right)\| \le C_K\|x_\bullet^{-1}\bullet y\|^{\beta_-/\beta_+}.$$

Proof In view of (3.1) and (3.5), the function $(t,x,y)\mapsto \delta_{1/t}\left(\delta_t(x)^{-1}\delta_t(y)\right)$ is a polynomial in

$$(x,y) = (x_1,\dots,x_d,y_1,\dots,y_d)$$

with coefficients equal to linear combination of powers of t with exponents in $\mathbb{R}$. By (3.8), only non-positive powers of t appear. The desired inequality follows from (5.5) because this polynomial function equals 0 when $x = y$. For the second inequality, we first note from (5.5) again that for $x, y \in K$,

$$\|\delta_{1/t}\left(\delta_t(x)^{-1}\cdot\delta_t(y)\right)\| \le C_K\|x-y\|_2^{\beta_-/\beta}$$

and then observe that $\|x-y\|_2 \le C'_K\|x_\bullet^{-1}\bullet y\|^{\beta/\beta_+}$. □

5.2 Further hypotheses

Under the general circumstances described above, in order to obtain limit theorems relating the random walk on Γ driven by μ to the continuous time left-invariant jump process on $G_\bullet$ associated with the jump measure $J_\bullet$ of Proposition 4.7, we need several additional hypotheses which we now spell out in details. One important feature of the various hypotheses described in this section is that they do not involve the precise limit behavior of μ_t as t tends to infinity. In a non-technical sense, they are of a coarser, more robust nature. In Chapter 10, we will exhibit a large class of “stable-like” measures on Γ, all of which satisfy these hypotheses thanks to the results of [55, 20].

The random walk on Γ (regularity)

A bounded function u on Γ is called μ-harmonic in a subset U if it satisfies

$$u * \mu = u \quad \text{ in } U.$$

Consider the following basic regularity assumption regarding μ-harmonic functions. Note that we consider that Γ as a subgroup of $G = (\mathbb{R}^d, \cdot)$ and use the $G_\bullet$-norm $\|\cdot\|$ to state this property.

(R1) There are constants C_1 and κ such that, for any bounded function u defined on Γ and μ-harmonic in $B(r) = \{x \in \mathbb{R}^d : \|x\| < r\}$, $r > 0$, and all $x, y \in \Gamma \cap B(r/2)$, we have

$$|u(y) - u(x)| \le C_1 \|u\|_\infty \left(\frac{\|x^{-1} \cdot y\|}{r} \right)^\kappa . \tag{5.7}$$

Remark 5.3 For any fixed $a > 0$, changing $\|\cdot\|$ to $\|\cdot\|^a$ (including in the definition of balls) amounts to changing κ to $\kappa/a > 0$. □

Exit time estimates

We consider the following exit time hypotheses formulated in terms of the norm $\|\cdot\|$ and the scaling exponent $\beta > 0$ associated with it in (5.3). In particular, the balls appearing in the definition below are the balls $B(r) = \{x \in \mathbb{R}^d : \|x\| < r\}$, $r \ge 0$, even so the exit probability estimates below concern the random walk on Γ.

(E1) There exists $A > 1$ such that the following holds: for any $\varepsilon \in (0, 1)$, there exists $\gamma = \gamma(A, \varepsilon) > 0$ such that for any $r > 0$, we have

$$\mathbb{P}^x \left(\tau_{B(Ar)} \le \gamma r^\beta \right) \le \varepsilon \quad \text{ for all } x \in \Gamma \cap B(r).$$

(E2) There exists $0 < C < \infty$ such that for any $r > 0$, we have

$$\mathbb{E}^x \left[\tau_{B(r)} \right] \le C r^\beta \quad \text{ for all } x \in \Gamma \cap B(r).$$

Here, $\mathbb{P}^x$ and $\mathbb{E}^x$ refer to the random walk on Γ starting from x driven by the probability measure μ.

Remark 5.4 One can esily check that Assumptions **(E1)-(E2)**, together, are equivalent to $\mathbb{E}^x[\tau_{B(r)}] \asymp r^\beta$ for any $x \in \Gamma \cap B(r/2)$. □

Remark 5.5 For our limit theorems to hold, the exponent $\beta > 0$ in **(E1)** and **(E2)** needs to be the same exponent β in (5.3). Thus, in this context, conditions **(E1)**

and **(E2)** as well as condition **(R1)** are not only a condition on the measure μ (which determines the random walk X_n on Γ and hence its harmonic functions) but also a condition on its comparability with the dilation structure $(\delta_t)_{t>0}$, scaled measure $\mu_t = t\delta_{1/t}(\mu)$, and norm $\|\cdot\|$ on $\mathbb{R}^d$. We expect the rescaled random walks $(\delta_{1/k}(X_{[kt]}))_{t>0}$ to converge when k tends to infinity to a self-similar process $(Z_t)_{t>0}$ satisfying $\delta_{1/s}(Z_{st}) = Z_t$ for all $s,t > 0$. From the definition of $\|\cdot\|$ at (5.3), the expected exit time out of a ball of radius r for this process should scale as r^β. Moreover, the random walk exit time of the ball of radius r is

$$\tau_{B(r)} = \inf\{n : X_n \notin B(r)\} = r^\beta \inf\left\{n/r^\beta : \delta_{1/r^\beta}(X_{r^\beta(n/r^\beta)}) \notin B(1)\right\}$$

and we expect that, as r tends to infinity,

$$\inf\left\{n/r^\beta : \delta_{1/r^\beta}(X_{r^\beta(n/r^\beta)}) \notin B(1)\right\} \to \inf\{s : Z_s \in B(1)\}$$

so that $\mathbb{E}^e\left[\tau_{B(r)}\right]$ should indeed behave as r^β. □

Tails properties for J_t and $J_\bullet$

We now discuss two related sets of hypotheses that are more technical but essential to obtain the desired results. They concern the limit jump measure $J_\bullet$ and the rescaled jump measures J_t for large $t > 0$. Theses hypotheses will have a natural flavor to anyone familiar with Lévy processes and Dirichlet forms. They complement the vague convergence of J_t to J on $(\mathbb{R}^d \times \mathbb{R}^d) \setminus \Delta$.

Set

$$B_\bullet(x,r) = x \bullet B(r) = \left\{y \in \mathbb{R}^d : \|x_\bullet^{-1} \bullet y\| < r\right\}.$$

Concerning the limit Radon measure

$$J_\bullet(dxdy) = dx\mu_\bullet(x_\bullet^{-1} \bullet dy)$$

on $(\mathbb{R}^d \times \mathbb{R}^d) \setminus \Delta$ from Proposition 4.7, which is symmetric by Remark 5.1, consider the hypothesis that

(T•) For any fixed compact set $K \subset \mathbb{R}^d$,

$$\lim_{\eta\to 0} \iint_{\{(x,y)\in K\times K : \|x_\bullet^{-1}\bullet y\|_2 \le \eta\}} \|x_\bullet^{-1} \bullet y\|_2^2 \, J_\bullet(dx,dy) = 0, \tag{5.8}$$

$$\lim_{R\to\infty} \int_{x\in K} \int_{y\in B_\bullet(x,R)^c} J_\bullet(dx,dy) = 0. \tag{5.9}$$

Note that, because $J_\bullet(dxdy) = dx\mu_\bullet(x_\bullet^{-1}dy)$, where $\mu_\bullet$ is a Radon measure on $G_\bullet \setminus \{e\} = \mathbb{R}^d \setminus \{0\}$, condition **(T•)** is equivalent to

$$\int_{G_\bullet} \min\{1, \|z\|_2^2\}\, \mu_\bullet(dz) < \infty, \tag{5.10}$$

which is (2.3) for $\nu = \mu_\bullet$.

Under this hypothesis, $J_\bullet$ is the jump measure of a symmetric bilinear form

$$\mathcal{E}_\bullet(u,v) := \frac{1}{2} \iint_{\mathbb{R}^d\times\mathbb{R}^d\setminus\Delta} (u(x)-u(y))(v(x)-v(y))J_\bullet(dx,dy) \tag{5.11}$$

on $\mathrm{Lip}_c(\mathbb{R}^d)$, which is the space of Lipschitz functions with compact support. Moreover, this form is closable in $L^2(G_\bullet; dx)$ and its closure is a regular conservative Dirichlet form $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$ – see, e.g., [27, Example 1.2.4] and [33, Theorem 1.3]. Hence by [18, Corollary 6.6.6],

$$\mathcal{F}_\bullet = \left\{u \in (\mathcal{F}_\bullet)_{\mathrm{loc}} \cap L^2(G_\bullet; dx) : \mathcal{E}_\bullet(u,u) < \infty\right\}. \tag{5.12}$$

Recall that by Lemma 3.9, the straight dilation $\{\delta_t, t > 0\}$ is a group dilation structure for the group $(G_\bullet, \bullet)$. Denote by $(\mathcal{L}_\bullet, \mathrm{Dom}(\mathcal{L}_\bullet))$ the infinitesimal generator of $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$ on $L^2(G_\bullet; dx)$. Under the hypothesis **(T•)**, we have $C_b^2(\mathbb{R}^d)\cap L^2(\mathbb{R}^d; m) \subset \mathrm{Dom}(\mathcal{L}_\bullet)$ and

$$\mathcal{L}_\bullet f(x) = \lim_{\varepsilon\to 0} \int_{\{z\in G_\bullet : \|z\|\ge\varepsilon\}} (f(x\bullet z) - f(x))\mu_\bullet(dz) \quad \text{for } f \in \mathrm{Dom}(\mathcal{L}_\bullet). \tag{5.13}$$

For $f \in \mathrm{Dom}(\mathcal{L}_\bullet)$ and $r > 0$, we have by (5.2) and (5.3) that for $x \in G_\bullet$,

$$\begin{aligned}
\mathcal{L}_\bullet(f\circ\delta_r)(x) &= \lim_{\varepsilon\to 0} \int_{\{z\in G_\bullet : \|z\|\ge\varepsilon\}} (f(\delta_r(x\bullet z)) - f(\delta_r(x)))\mu_\bullet(dz) \\
&= \lim_{\varepsilon\to 0} \int_{\{z\in G_\bullet : \|w\|\ge\varepsilon\}} (f(\delta_r(x)\bullet\delta_r(z)) - f(\delta_r(x)))\mu_\bullet(dz) \\
&= \lim_{\varepsilon\to 0} \int_{\{z\in G_\bullet : \|w\|\ge\varepsilon\}} (f(\delta_r(x)\bullet w) - f(\delta_r(x)))(\delta_r\mu_\bullet)(dw) \\
&= r\lim_{\varepsilon\to 0} \int_{\{z\in G_\bullet : \|w\|\ge\varepsilon\}} (f(\delta_r(x)\bullet w) - f(\delta_r(x)))\mu_\bullet(dw) \\
&= r\mathcal{L}_\bullet f(\delta_r(x)).
\end{aligned} \tag{5.14}$$

In particular, we have for $f \in C_c^2(\mathbb{R}^d)$ and $r > 0$,

$$\mathcal{L}_\bullet(f\circ\delta_r)(e) = r\mathcal{L}_\bullet f(e). \tag{5.15}$$

Remark 5.6 Under the hypothesis **(T•)**, let $X^\bullet$ be the symmetric Hunt process on $G_\bullet$ associated with the regular Dirichlet form $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$ on $L^2(G_\bullet, dx)$; see [18, 27]. In view of (5.13), $X^\bullet$ has stationary independent increment property; that is, for any $t > s \ge 0$, $(X_s^\bullet)_\bullet^{-1} \bullet X_t^\bullet$ is independent of $\sigma(X_r^\bullet; r \le s)$ and has the same distribution as $(X_0^\bullet)_\bullet^{-1} \bullet X_{t-s}^\bullet$. Thus $X^\bullet$ can be refined to start from every point in $G_\bullet$. Moreover, it follows from (5.14) that if $X_0^\bullet = e$, then

$$\{\delta_r(X_t^\bullet); t \geq 0\} \text{ has the same distribution as } \{X_{rt}^\bullet; t \geq 0\}. \tag{5.16}$$

We know from Lemma 3.9 the straight dilations $\{\phi_t,\, t > 0\}$ form a group dilation structure for the nilpotent group $(G_\bullet, \bullet)$. Thus in the terminology of [41, p.170], the Lévy process $X^\bullet$ is stable with respect to the dilations $\{\phi_t,\, t > 0\}$. □

Remark 5.7 In the terminology of [29, p.31], the scaling property (5.15) says that the generating functional $f \to \mathcal{L}_\bullet f(e)$ is a kernel of order $\beta_+ = \max_{1\leq i\leq d} \beta_i$. Observe that in [29], the exponents $\{d_j, 1 \leq j \leq d\}$ for the straight dilation structure $\{\delta_t, t > 0\}$ are our $\{1/\beta_j, 1 \leq j \leq d\}$ and the smallest d_j there (which corresponds our β_+, the largest of β_j) is normalized to 1; see Remark 3.12 for the procedure of doing such a normalization. Note also that the norm $|\cdot|$ defined on [29, (1.1)] is comparable to our norm $\|\cdot\|$. □

Regarding the scaled jump kernel J_t, consider the property

(TΓ) For any fixed compact set $K \subset \mathbb{R}^d$,

$$\lim_{\eta\to 0} \limsup_{t\to\infty} \iint_{\{(x,y)\in K^2 : \|x_\bullet^{-1}\bullet y\|_2 \leq \eta\}} \|x_\bullet^{-1} \bullet y\|_2^2 J_t(dx, dy) = 0, \tag{5.17}$$

$$\lim_{R\to\infty} \limsup_{t\to\infty} \int_{x\in K} \int_{y\in B_\bullet(x,R)^c} J_t(dx, dy) = 0. \tag{5.18}$$

Remark 5.8 In the estimates (5.8) and (5.17), it is crucial to use the norm $\|\cdot\|_2$ in the integrant in order to measure the strength of small jumps allowed by theses jump kernels in a classical fashion. In the estimates (5.9) and (5.18), it is natural to use the norm $\|\cdot\|$ due to the scaling property of $\{\delta_t\}$, but we may also use $\|\cdot\|_2$ if desired because of (5.5). □

Remark 5.9 Recalling (5.10), one can check that conditions **(A)**, **(T•)** and **(TΓ)** combined are equivalent to the following condition:

(A′) The straight dilation structure $(\delta_t)_{t>0}$ is admissible for the probability measure μ in the sense that the finite positive measure $(\|z\|_2^2 \wedge 1)\, \mu_t(dz)$ converges weakly on $\mathbb{R}^d \setminus \{0\}$ to a finite measure $(\|z\|_2^2 \wedge 1)\, \mu_\bullet$ as t tends to infinity, where μ_t is the measure defined at (4.1). □

Note that Examples 4.4-4.6 satisfy any of these conditions **(A)**, **(T•)**, **(TΓ)**, **(R1)-(R2)** and **(E1)-(E2)**.

5.3 Weak convergence

Throughout this section, we generally assume that **(A)-(R1)-(E1)-(E2)** and **(T•)-(TΓ)** are all satisfied even so we will list exactly which properties are used for different results stated in this section.

Because of assumptions **(A)** and **(T•)**, we can consider the (continuous time) Markov semigroup of operators

$$\{P_{\bullet,s}\}_{s\geq 0}$$

corresponding to $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$ at (5.11). Let $\{U_\bullet^\lambda; \lambda > 0\}$, and $\{\mathbb{P}_\bullet^x; x \in G_\bullet\}$ be the resolvent, and probabilities corresponding to the regular Dirichlet form $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$ on $L^2(G_\bullet; dx)$. Our goal is to prove that the continuous time conservative Markov process associated with this regular Dirichlet form is the limit of the properly rescaled discrete time random walk on Γ driven by the probability measure μ. We let $(X_n)_{n\geq 0}$ denote this random walk. Assumptions **(R1)** and **(E1)-(E2)** are assumptions regarding the behavior of this discrete time random walk on Γ.

Fix an arbitrary sequence of positive reals $\{T_k\}$ that goes to ∞. We write $(\mathcal{E}^{(k)}, \mathcal{F}^{(k)})$ for $(\mathcal{E}^{(T_k)}, \mathcal{F}^{(T_k)})$ defined by (4.7) with T_k in place of t there, which corresponds to the rescaled discrete time process by

$$(X_n^{(k)} := \delta_{T_k}^{-1}(X_{T_k n}))_{n\in(1/T_k)\mathbb{N}\cup\{0\}}.$$

Note that this is just discrete time random walk on Γ_{T_k} where time has been rescaled linearly according to the scaling sequence T_k. Let $\{P_n^k; n \in (1/T_k)\mathbb{N}\cup\{0\}\}$, $\{U_k^\lambda; \lambda > 0\}$, and $\{\mathbb{P}_k^x; x \in \Gamma_{T_k}\}$ be the associated semigroup, resolvent, and probabilities. For $t \geq 0$, we write

$$\hat{X}_t^{(k)} := \delta_{T_k}^{-1}(X_{[T_k t]}) = X_{[T_k t]/T_k}^{(k)}, \quad \hat{P}_t^k := P_{[T_k t]/T_k}^k \tag{5.19}$$

and denote the corresponding probabilities by $\{\hat{\mathbb{P}}_k^x; x \in \Gamma_{T_k}\}$. So for $x, y \in \Gamma_{T_k}$, and $n = m/T_k$, $m \in \mathbb{N}\cup\{0\}$,

$$\mathbb{P}_k^x(X_n^{(k)} = y) = \mu^{(m)}(\delta_{T_k}(x)^{-1}\cdot\delta_{T_k}(y)),$$

and for $x, y \in \Gamma_{T_k}$, and $t > 0$,

$$\hat{\mathbb{P}}_k^x(\hat{X}_t^{(k)} = y) = \mu^{([tT_k])}(\delta_{T_k}(x)^{-1}\cdot\delta_{T_k}(y)). \tag{5.20}$$

For a constant $M_0 > 0$, let $\mathbb{D}([0, M_0], \mathbb{R}^d)$ be the space of right continuous functions on $[0, M_0]$ having left limits and taking values in $\mathbb{R}^d$ that is equipped with the Skorohod $\mathcal{J}_1$ topology. Our goal is to prove the following theorem. Recall that $\beta_i \in (0,2)$, $1 \leq i \leq d$, are the parameters in (5.1) for the straight dilation structure $\{\delta_t; t > 0\}$ and $\beta_+ = \max\{\beta_i : 1 \leq i \leq d\}$.

Theorem 5.10 Referring to the setup and notation introduced above, assume that (5.3), **(A)-(R1)-(E1)-(E2)** and **(T•)-(TΓ)** are all satisfied with the same exponent $\beta > 0$. Then

(i) The symmetric Hunt process $X^\bullet$ associated with the regular Dirichlet form $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$ on $L^2(G_\bullet; dx)$ is a Lévy process on $G_\bullet$. The Lévy process $X_t^\bullet$ has a bounded, strictly positive, jointly continuous transition density function $p(t, x, y) = p(t, x_\bullet^{-1} \bullet y)$ with respect to dy that has the following properties.

(a) Let $\gamma_0 := \sum_{i=1}^d 1/\beta_i$. For every $(t, x) \in (0, \infty) \times G_\bullet$,

$$p(t, x) = t^{-\gamma_0} p(1, \delta_{1/t}(x)). \tag{5.21}$$

In particular, there is a constant $C_1 > 0$ so that $p(t, x) \le C_1 t^{-\gamma_0}$ for every $(t, x) \in (0, \infty) \times G_\bullet$.

(b) For every $\gamma \in (0, \beta_+ \wedge 1)$, there is a constant $C_2 > 0$ so that

$$|p(1, x) - p(1, y)| \le C_2 \|x_\bullet^{-1} \bullet y\|^\gamma \quad \text{for } x, y \in G_\bullet. \tag{5.22}$$

(c) For every $\alpha \in (0, \beta_+)$, there is a constant $C_3 > 0$ so that for every $(t, x) \in (0, \infty) \times G_\bullet$,

$$p(t, x) \le \min \left\{ C_1 t^{-\gamma_0}, \, C_3 \frac{t^{\alpha/\beta_+}}{\|x\|^{d+\alpha}} \right\}. \tag{5.23}$$

(ii) For any bounded continuous function f on $\mathbb{R}^d$, $\hat{P}_s^k f$ converges uniformly on compacts to $P_{\bullet, s} f$. Furthermore, for each $M_0 > 0$, for every $x \in \mathbb{R}^d$, $\hat{\mathbb{P}}_k^{[x]_k}$ converges weakly to $\mathbb{P}_\bullet^x$ on the space $\mathbb{D}([0, M_0], \mathbb{R}^d)$. □

Remark 5.11 Note that under its conditions, Theorem 5.10 in particular implies that the Lévy process $X^\bullet$ is always non-degenerate in the sense that it has a strictly positive convolution density kernel $p(t, x)$ with respect to the Haar measure dx on $G_\bullet$. Consequently, the support of its Lévy measure $\mu_\bullet$ generates the whole group $G_\bullet$. □

5.4 Proof of Theorem 5.10

In this section, we prove Theorem 5.10. The main part of the argument is based on Section 4 of [7]. Similar arguments for discrete setting (including a diffusion term in the limit) are given in [9, Theorem 5.5].

Recall that $X_n^{(k)} = \delta_{T_k}^{-1}(X_{T_k n})$, $n \in T_k^{-1}\mathbb{N} \cup \{0\}$. We first state a lemma that is an easy consequence of rescaling, and assumptions **(R1)-(E1)-(E2)**, and Lemma 5.2 (with $\phi_t = \delta_t$). For $x_0 \in \Gamma_{T_k}$, let

$$B_{T_k}(x_0, r) = x_0 \cdot_{T_k} B(r).$$

Note that this is different from $B_\bullet(x_0, r) = x_0 \bullet B(r)$ which we have used earlier. Also, $y \in B_{T_k}(x_0, r)$ if and only if $\delta_{1/T_k}(\delta_{T_k}(x_0)^{-1}) \cdot_{T_k} y \in B(r)$ (i.e., we have to take the inverse of x_0 in $(\Gamma_t, \cdot_t)$ with $t = T_k$).

Lemma 5.12 (i) Assume **(E1)**. Then, there exists $A > 1$ such that the following holds: for any $\varepsilon \in (0,1)$, there exists $\gamma = \gamma(A,\varepsilon) > 0$ such that for all $k \geq 1$, $x_0 \in \Gamma_{T_k}$, $r \in (0,1)$ and $x \in B_{T_k}(x_0,r) \cap \Gamma_{T_k}$,

$$\mathbb{P}_k^x \left(\tau_{B_{T_k}(x_0,Ar)}(X^{(k)}) \leq \gamma r^{\beta} \right) \leq \varepsilon.$$

(ii) Under **(E2)**, there exists $c_1 > 0$ such that the following hold for all $k \geq 1$, $x_0 \in \Gamma_{T_k}$, $r \in (0,1)$, and all $x \in B_{T_k}(x_0,r) \cap \Gamma_{T_k}$,

$$\mathbb{E}_k^x \left[\tau_{B_{T_k}(x_0,r)}(X^{(k)}) \right] \leq c_1 r^{\beta}.$$

(iii) Under **(R1)**, there exists $\kappa \in (0,\infty)$ such that, for any compact set $K \subset \mathbb{R}^n$, there is $c_{2,K} > 0$ for which, for any $k \geq 1$, $x_0 \in K \cap \Gamma_{T_k}$ and $r \in (0,1)$, if h_k is bounded in Γ_{T_k} and harmonic with respect to $X^{(k)}$ in a ball $B_{T_k}(x_0,r) \cap \Gamma_{T_k}$ then, for $x, y \in B_{T_k}(x_0, r/2) \cap K \cap \Gamma_{T_k}$,

$$|h_k(x) - h_k(y)| \leq c_{2,K} \left(\frac{\|x_\bullet^{-1} \bullet y\|^{\beta_-/\beta_+}}{r} \right)^{\kappa} \|h_k\|_\infty.$$

Proof In view of the scaling property (5.3) of the norm $\|\cdot\|$ on $G_\bullet$, properties (i) and (ii) are just reformulation of conditions **(E1)** and **(E2)**, respectively, under the approximate dilation δ_{T_k}.

(iii) follows from condition **(R1)** under the approximate dilation δ_{T_k} and Lemma 5.2. □

Recall that for $\lambda > 0$, the resolvent U_k^λ is given by

$$\begin{aligned} U_k^\lambda f(x) &= (\lambda I - T_k(P-I))^{-1} f(\delta_{T_k}^{-1}(x)) \\ &= (T_k + \lambda)^{-1} \sum_{n=0}^{\infty} \left(\frac{1}{1 + \lambda T_k^{-1}} \right)^n P^n f(\delta_{T_k}^{-1}(x)) \quad \text{for } x \in \Gamma_{T_k} = \delta_{T_k}(\Gamma), \end{aligned}$$

where P is the transition matrix for the random walk $\{X_n\}_n$ on Γ.

The following proposition is based on [17, Proposition 2.4] (see also [7, Proposition 3.3]). We outline the proof for the reader's convenience.

Proposition 5.13 Under **(R1)** and **(E2)**, for any compact set K, there exist $C_{\lambda,K} \in (0,\infty)$ and $\gamma \in (0, (\beta \wedge \kappa)/2]$ such that the following holds for any bounded function f on Γ_{T_k}, for any $k \geq 1$ and any $x, y \in K \cap \Gamma_{T_k}$ with $\|x_\bullet^{-1} \bullet y\| \leq 1$,

$$|U_k^\lambda f(x) - U_k^\lambda f(y)| \leq C_{\lambda,K} \|x_\bullet^{-1} \bullet y\|^{\gamma} \|f\|_\infty. \tag{5.24}$$

In particular, we have

$$\lim_{\delta \to 0} \sup_{k \geq 1} \sup_{\substack{x,y \in K \cap \Gamma_{T_k}: \\ \|x_\bullet^{-1} \bullet y\| < \delta}} |U_k^\lambda f(x) - U_k^\lambda f(y)| = 0. \tag{5.25}$$

Proof Recall the notation $B_{T_k}(z,r) = z\cdot_{T_k} B(r)$. Let $x, y \in K\cap\Gamma_{T_k}$ and let $r\in(0,1]$ be such that $\|x_\bullet^{-1}\bullet y\| \le r$. By Lemma 5.2, $y\in B_{T_k}(x,\rho)$, $\rho = C_K r^{\beta_-/\beta_+}$. Set $\tau_r^k := \tau_{B_{T_k}(x,2\rho)}(X^{(k)})$. In what follows the constant C_K depend only on K and can change from line to line. By the strong Markov property,

$$\begin{aligned}
U_k^\lambda f(x) &= (T_k+\lambda)^{-1}\mathbb{E}_k^x\Big[\sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ n\in[0,\tau_r^k T_k]}}\Big(\frac{1}{1+\lambda T_k^{-1}}\Big)^n P^n f(\delta_{T_k}^{-1}(x))\Big]\\
&\quad+\mathbb{E}_k^x\Big[\Big(\frac{1}{1+\lambda T_k^{-1}}\Big)^{\tau_r^k T_k} U_k^\lambda f(X^{(k)}_{\tau_r^k})\Big]\\
&= (T_k+\lambda)^{-1}\mathbb{E}_k^x\Big[\sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ n\in[0,\tau_r^k T_k]}}\Big(\frac{1}{1+\lambda T_k^{-1}}\Big)^n P^n f(\delta_{T_k}^{-1}(x))\Big]\\
&\quad+\mathbb{E}_k^x\left[\Big(\Big(\frac{1}{1+\lambda T_k^{-1}}\Big)^{\tau_r^k T_k}-1\Big)U_k^\lambda f(X^{(k)}_{\tau_r^k})\right]+\mathbb{E}_k^x\left[U_k^\lambda f(X^{(k)}_{\tau_r^k})\right]\\
&=: I_1+I_2+I_3,
\end{aligned}$$

and similarly when x is replaced by y. Because of Lemma 5.12(ii) and the fact that $\|Pf\|_\infty \le \|f\|_\infty$, we have

$$|I_1| \le \frac{T_k}{T_k+\lambda}\mathbb{E}_k^x\left[\tau_r^k\right]\|f\|_\infty \le c_1 r^\zeta\|f\|_\infty, \quad \text{where } \zeta := \beta\beta_-/\beta_+.$$

Note that

$$\|U_k^\lambda f\|_\infty \le (T_k+\lambda)^{-1}\frac{1}{1-\frac{1}{1+\lambda T_k^{-1}}}\|f\|_\infty = \lambda^{-1}\|f\|_\infty.$$

Using this and applying $1-e^{-s}\le s$, $s\ge 0$, with $s = \tau_r^k T_k\log(1+\lambda T_k^{-1})$, we have

$$|I_2| \le \mathbb{E}_k^x\left[\tau_r^k\right]T_k\log(1+\lambda T_k^{-1})\|U_k^\lambda f\|_\infty \le \mathbb{E}_k^x\left[\tau_r^k\right]T_k\lambda T_k^{-1}\lambda^{-1}\|f\|_\infty \le c_1 r^\zeta\|f\|_\infty.$$

Similar statements also hold when x is replaced by y. So,

$$\begin{aligned}
&\left|U_k^\lambda f(x) - U_k^\lambda f(y)\right| \le\\
&\quad c_1 r^\zeta\|f\|_\infty + \left|\mathbb{E}_k^x\left[U_k^\lambda f(X^{(k)}_{\tau_r^k})\right] - \mathbb{E}_k^y\left[U_k^\lambda f(X^{(k)}_{\tau_r^k})\right]\right|.
\end{aligned}\tag{5.26}$$

But $z\to\mathbb{E}_k^z\left[U_k^\lambda f(X^{(k)}_{\tau_r^k})\right]$ is bounded in Γ_{T_k} and harmonic in $B_{T_k}(x,2\rho)\cap\Gamma_{T_k}$. By Lemma 5.12(iii), for $y\in B_{T_k}(x,\rho)$, the second term in (5.26) is bounded by

$$C_K(\|x_\bullet^{-1}\bullet y\|^{\beta_-/\beta_+}/r^{\beta_-/\beta_+})^\kappa\|U_k^\lambda f\|_\infty.$$

So using $\|U_k^\lambda f\|_\infty \le \lambda^{-1}\|f\|_\infty$ again, for $y\in B_\bullet(x,\rho)\cap\Gamma_{T_k}$ we have

$$\left|U_k^\lambda f(x) - U_k^\lambda f(y)\right| \le C_K \left(r^{\beta\beta_-/\beta_+} + \lambda^{-1} \left(\frac{\|x_\bullet^{-1} \bullet y\|^{\beta_-/\beta_+}}{r^{\beta_-/\beta_+}} \right)^\kappa \right) \|f\|_\infty. \qquad (5.27)$$

Now choose r such that $r = \|x_\bullet^{-1} \bullet y\|^{1/2}$ (then $\|x_\bullet^{-1} \bullet y\| = r^2 \le r \le 1$). For this choice of r, we obtain

$$\begin{aligned} &\left|U_k^\lambda f(x) - U_k^\lambda f(y)\right| \\ &\le C_K \left(\|x_\bullet^{-1} \bullet y\|^{\beta\beta_-/(2\beta_+)} + \lambda^{-1} \|x_\bullet^{-1} \bullet y\|^{\kappa\beta_-/(2\beta_+)} \right) \|f\|_\infty \\ &\le C_K (1 + \lambda^{-1}) \|x_\bullet^{-1} \bullet y\|^\gamma \|f\|_\infty, \end{aligned}$$

where $\gamma = \min\left\{ \frac{\beta\beta_-}{2\beta_+}, \frac{\kappa\beta_-}{2\beta_+} \right\} \in (0, (\beta \wedge \kappa)/2]$. □

The first part of the next proposition is based on [17, Proposition 2.8] (see also [8, Proposition 6.2] and [5, Section 6]). In the following, m denotes the Lebesgue measure on $\mathbb{R}^d$.

Proposition 5.14 Assume **(A)**-**(R1)**-**(E1)**-**(E2)**. For every subsequence $\{k_j\}$, there exist a sub-subsequence $\{k_{j(l)}\}$ and a conservative m-symmetric Hunt process $(\widetilde{X}, \widetilde{\mathbb{P}}^x, x \in \mathbb{R}^d)$, which is a Lévy process on $(G_\bullet, \bullet)$, such that for every $x_{k_{j(l)}} \to x$, $\hat{\mathbb{P}}_{k_{j(l)}}^{x_{k_{j(l)}}}$ converges weakly in $\mathbb{D}([0,\infty), \mathbb{R}^d)$ to $\widetilde{\mathbb{P}}^x$. Moreover, the resolvents of the conservative Hunt process $\widetilde{X}$ map bounded functions on $\mathbb{R}^d$ into bounded local Hölder continuous functions on $\mathbb{R}^d$ and so for each $t > 0$, $\widetilde{X}_t$ has a transition density function $p(t,x,y) = p(t, e, x_\bullet^{-1} \bullet y)$ with respect to dy. □

Proof For simplicity, denote by the subsequence $\{k_j\}$ by $\{k\}$. Let $T_0 > 0$ an arbitrary constant and $x_k \in \Gamma_k$. For any stopping time η_k of $X^{(k)}$ that is bounded by T_0 and any positive constant $\delta_k \to 0$, it follows from Proposition 5.12(i) and the strong Markov property of $X^{(k)}$ that for any $\varepsilon > 0$,

$$\begin{aligned} &\limsup_{k\to\infty} \mathbb{P}_k^{x_k} \left(\|\delta_{1/k}(\delta_k((X_{\eta_k}^{(k)})^{-1})) \cdot_k X_{\eta_k+\delta_k}^{(k)}\| > \varepsilon \right) \\ &\le \limsup_{k\to\infty} \mathbb{E}_k^{x_k} \left[\mathbb{P}_k^{X_{\eta_k}^{(k)}} (\tau_{B_k(X_0^{(k)},\varepsilon)} < \delta_k) \right] = 0. \end{aligned}$$

Thus by [1], the probability laws $\{\hat{\mathbb{P}}_k^{x_k}; k \ge 1\}$ are tight on $\mathbb{D}([0,T_0), \mathbb{R}^d)$. Under conditions **(R1)**-**(E1)**-**(E2)**, the proof of the first part of this proposition (on weak convergence) is then similar to that of [17, Proposition 2.8], modulo modifying the arguments for continuous time processes there to discrete time processes, so we omit this part of the proof. Since $X^{(k)}$ has stationary independent increments on Γ_k, so does $\widetilde{X}$ on $(G_\bullet, \bullet)$.

That the resolvents of $\widetilde{X}$ maps bounded functions on $\mathbb{R}^d$ into bounded Hölder continuous functions on $\mathbb{R}^d$ follows readily from Proposition 5.13. For $\lambda > 0$, denote by $\widetilde{U}^\lambda$ the λ-resolvent of $\widetilde{X}$. For any Borel measurable set $A \subset \mathbb{R}^d$ having $m(A) = 0$, by the m-symmetry and conservativeness of $\widetilde{X}$, we have $\int_{\mathbb{R}^d} \widetilde{U}^\lambda(x,A) m(dx) =$

$\lambda^{-1}m(A) = 0$ for every $\lambda > 0$. As $\widetilde{U}^\lambda(x, A)$ is continuous in $x \in \mathbb{R}^d$, we conclude that $\widetilde{U}^\lambda(x, A) = 0$ for every $x \in \mathbb{R}^d$. By [27, Theorem 4.2.4], this implies that the law of X_t under $\widetilde{\mathbb{P}}^x$ is absolutely continuous with respect to m for each $t > 0$ and $x \in \mathbb{R}^d$. Denote its density by $p(t, x, y)$. By the Lévy property of $X^\bullet$, we have $p(t, x, y) = p(t, e, x_\bullet^{-1} \bullet y)$. □

Proof of Theorem 5.10. In view of Proposition 5.14, it suffices to show that the Dirichlet form in $L^2(G_\bullet; dx)$ of the conservative m-symmetric process $\widetilde{X}$ in Proposition 5.14 is $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$ and establish (i). As in the proof of Proposition 5.13, we know that any subsequence $\{k_j\}$ has a further subsequence $\{k_{j_l}\}$ such that $U^\lambda_{k_{j_l}} f$ converges uniformly on compacts whenever e $\lambda > 0$ and f is bounded and continuous on $\mathbb{R}^d$.

Now suppose we have a subsequence $\{k'\}$ such that the $U^\lambda_{k'} f$ on $\Gamma_{T_{k'}}$ are equicontinuous and converge uniformly on compacts whenever $\lambda > 0$ and f is bounded and continuous with compact support on $\mathbb{R}^d$. Fix $\lambda > 0$ and such an f, and let $H \in C_b(\mathbb{R}^d)$ be the limit of $U^\lambda_{k'} f$. We will show that $H \in \mathcal{F}_\bullet$ and

$$\mathcal{E}_\bullet(H, g) = \langle f, g \rangle - \lambda \langle H, g \rangle \tag{5.28}$$

whenever g is a Lipschitz function on $\mathbb{R}^d$ with compact support, where $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$ is the Dirichlet form of (5.11) and $\langle \cdot, \cdot \rangle$ is the L^2-inner product with respect to the Lebesgue measure m on $\mathbb{R}^d$. This will prove that H is the λ-resolvent of f with respect to $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$ in $L^2(\mathbb{R}^d; dx)$, that is, $H = U^\lambda f$. We can then conclude that the full sequence $U^\lambda_k f$ converges to $U^\lambda f$ whenever f is bounded and continuous with compact support. The assertions about the convergence of P^k_t and $\mathbb{P}^x_k$ then follow by Proposition 5.14.

So we need to prove H satisfies (5.28). We drop the primes for legibility. We know

$$\mathcal{E}^{(k)}(U^\lambda_k f, U^\lambda_k f) = \langle f, U^\lambda_k f \rangle_{L^2(\Gamma_{T_k}, m_{T_k})} - \lambda \langle U^\lambda_k f, U^\lambda_k f \rangle_{L^2(\Gamma_{T_k}, m_{T_k})}, \tag{5.29}$$

where for $t > 0$, m_t is the measure on Γ_t defined by (4.5). Since

$$\|U^\lambda_k f\|_{L^2(\Gamma_{T_k}, m_{T_k})} \le (1/\lambda) \|f\|_{L^2(\Gamma_{T_k}, m_{T_k})},$$

we have by the Cauchy-Schwarz inequality that

$$\sup_k \mathcal{E}^{(k)}(U^\lambda_k f, U^\lambda_k f) \le \sup_k \lambda^{-1} \|f\|^2_{L^2(\Gamma_{T_k}, m_{T_k})} \le c < \infty.$$

Set $B_2(r) = \{x : \|x\|_2 < r\}$. Since $U^\lambda_k f$ converge uniformly to H on $\overline{B_2(1/\eta)}$ for every $\eta \in (0, 1)$, it follows from Lemma 4.10 that

$$\begin{aligned} &\iint_{D_\eta} (H(y) - H(x))^2 J_\bullet(dx, dy) \\ &\le \limsup_{k \to \infty} \sum_{(x,y) \in (\Gamma_{T_k} \times \Gamma_{T_k}) \cap D_\eta} (U^\lambda_k f(x) - U^\lambda_k f(y))^2 j_k(x, y) \end{aligned}$$

$$\leq \limsup_{k\to\infty} \mathcal{E}^{(k)}(U_k^\lambda f, U_k^\lambda f) \leq c < \infty,$$

where $D_\eta := \{(x,y) \in B_2(e,\eta^{-1}) \times B_2(e,\eta^{-1}) : \eta < \|x_\bullet^{-1} \bullet y\|_2 \leq \eta^{-1}\}$. Letting $\eta \to 0$, we have

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \Delta} (H(y) - H(x))^2 J_\bullet(dx, dy) \leq c < \infty.$$

Since $H \in C_b(\mathbb{R}^d)$, the above in particular implies that $H \in (\mathcal{F}_\bullet)_{\text{loc}}$. Note that by Fatou's lemma, $H \in L^2(\mathbb{R}^d; dx)$ as it is the pointwise limit of $U_k^\lambda f$. Thus we conclude from (5.12) that

$$H \in \mathcal{F}_\bullet \quad \text{with} \quad \mathcal{E}_\bullet(H,H) < \infty. \tag{5.30}$$

Fix a Lipschitz function g on $\mathbb{R}^d$ with compact support, and choose $r_0 > 0$ large enough so that the support of g is contained in the L^2-ball $B_2(e,r_0)$. Then, setting $H_{\geq\eta^{-1}} := \{\|x_\bullet^{-1} \bullet y\|_2 \geq \eta^{-1}\}$,

$$\begin{aligned}
&\Big| \sum_{(x,y)\in(\Gamma_{T_k}\times\Gamma_{T_k})\cap H_{\geq\eta^{-1}}} (U_k^\lambda f(y) - U_k^\lambda f(x))(g(y) - g(x)) j_k(x,y) \Big| \\
&\qquad\qquad \leq \Big(\sum_{(x,y)} (U_k^\lambda f(y) - U_k^\lambda f(x))^2 j_k(x,y) \Big)^{1/2} \\
&\qquad\qquad\quad \times \Big(\sum_{(x,y)\in(\Gamma_{T_k}\times\Gamma_{T_k})\cap H_{\geq\eta^{-1}}} (g(y) - g(x))^2 j_k(x,y) \Big)^{1/2}.
\end{aligned}$$

The first factor is $(\mathcal{E}^{(k)}(U_k^\lambda f, U_k^\lambda f))^{1/2}$, while the second factor is bounded by

$$\sqrt{2}\, \|g\|_\infty \Big(\int_{B_2(e,r_0)} \int_{\|x_\bullet^{-1}\bullet y\|_2 \geq \eta^{-1}} J_k(dx, dy) \Big)^{1/2},$$

which, in view of (5.18) in **(TΓ)**, will be small if η is small. Similarly, setting $H_{\leq\eta} := \{\|x_\bullet^{-1} \bullet y\|_2 \leq \eta\}$, it holds that

$$\begin{aligned}
&\Big| \sum_{(x,y)\in(\Gamma_{T_k}\times\Gamma_{T_k})\cap H_{\leq\eta}} (U_k^\lambda f(y) - U_k^\lambda f(x))(g(y) - g(x)) j_k(x,y) \Big| \\
&\qquad\qquad \leq \Big(\sum_{(x,y)} (U_k^\lambda f(y) - U_k^\lambda f(x))^2 j_k(x,y) \Big)^{1/2} \\
&\qquad\qquad\quad \times \Big(\sum_{(x,y)\in(\Gamma_{T_k}\times\Gamma_{T_k})\cap H_{\leq\eta}} (g(y) - g(x))^2 j_k(x,y) \Big)^{1/2}.
\end{aligned}$$

The first factor is as before, while the second is bounded by

$$\|g\|_{\text{Lip}} \Big(\int_{B_2(e,r_0)} \int_{\|x_\bullet^{-1}\bullet y\|_2 \leq \eta} \|x_\bullet^{-1} \bullet y\|_2^2 J_k(dx, dy) \Big)^{1/2},$$

where

$$\|g\|_{\mathrm{Lip}} := \sup_{x,y\in\mathbb{R}^d} \frac{|g(x)-g(y)|}{\|x_\bullet^{-1}\bullet y\|_2} < \infty.$$

In view of (5.17) in **(TΓ)**, the second factor will be small if η is small. Similarly, using (5.30), we have

$$\Big|\iint_{\|x_\bullet^{-1}\bullet y\|_2\notin(\eta,\eta^{-1})} (H(y)-H(x))(g(y)-g(x))J_\bullet(dx,dy)\Big|$$

will be small if η is taken small enough, due to Remark 5.8.

Note that $U_k^\lambda f$ are equi-continuous and converge to H uniformly on compacts, and g is a compactly supported function. For $\eta > 0$ we have by Lemma 4.10,

$$\begin{aligned}&\sum_{(x,y)\in(\Gamma_{T_k}\times\Gamma_{T_k})\cap\{\|x_\bullet^{-1}\bullet y\|_2\in(\eta,\eta^{-1})\}} (U_k^\lambda f(y)-U_k^\lambda f(x))(g(y)-g(x))j_k(x,y)\\ &\to \iint_{\|x_\bullet^{-1}\bullet y\|_2\in(\eta,\eta^{-1})} (H(y)-H(x))(g(y)-g(x))J_\bullet(dx,dy).\end{aligned}$$

It follows that

$$\lim_{k\to\infty}\mathcal{E}^{(k)}(U_k^\lambda f,g)=\mathcal{E}_\bullet(H,g). \tag{5.31}$$

But as $k\to\infty$,

$$\mathcal{E}^{(k)}(U_k^\lambda f,g)=\langle f,g\rangle_{L^2(\Gamma_{T_k},m_{T_k})}-\lambda\langle U_k^\lambda f,g\rangle_{L^2(\Gamma_{T_k},m_{T_k})}\to\langle f,g\rangle-\lambda\langle H,g\rangle.$$

Combining this with (5.31) proves (5.28). This proves that $\widetilde{X}$ has the same distribution as the Lévy process $X^\bullet$ associated with the regular Dirichlet form $(\mathcal{E}_\bullet,\mathcal{F}_\bullet)$ on $L^2(G_\bullet;m)$, which in particular establishes part (ii) of the theorem by Proposition 5.14.

We next show part (i) of the theorem. By Proposition 5.14, $X_t^\bullet$ has transition density function $p(t,x_\bullet^{-1}\bullet y)$ with respect to the Lebesgue measure dy on $G_\bullet$. By Remark 5.7, the generating functional $f\mapsto\mathcal{L}f(e)$ is a kernel of order β_+. Thus by [29, Theorem 2.2], $p(t,x)$ is square-integrable for every $t>0$ and so $c_t := p(t,e)=\int_{\mathbb{R}^d}p(t/2,x)^2dx<\infty$. By the Cauchy-Schwarz inequality, for any $x\in G_\bullet$,

$$\begin{aligned}p(t,x)=p(t,e,x)&=\int_{\mathbb{R}^d}p(t/2,e,z)p(t/2,z,x)dz\\ &\le\|p(t/2,e,\cdot)\|_2\,\|p(t/2,\cdot,x)\|_2\le c_t.\end{aligned}$$

That is, $p(t,x)$ is bounded on $G_\bullet$ for every $t>0$. Property (5.21) follows from the self-similarity property (5.16) of $X^\bullet$, and Hölder regularity (5.22) follows from [29, Corollary 3.12]. The joint continuity of $p(t,x,y)=p(t,x_\bullet^{-1}\bullet y)$ in (t,x,y) follows from the scaling property (5.21) and the Hölder continuity of $p(t,x)$ in x. Note that $p(t,e)=\int_{G_\bullet}p(t/2,y)^2dy>0$ and $\lim_{t\to\infty}\delta_{1/t}(x)=e$ uniformly on every compact subset of $G_\bullet$. Thus by (5.21), for any $n \geq 1$, there is some $t_n > 0$ so that $p(t, x) > 0$ for every $(t, x) \in (0, t_n] \times B(0, n)$. It then follows from the Chapman-Kolmogorov equation that $p(t, x) > 0$ for every $(t, x) \in (0, \infty) \times G_\bullet$.

For any $\alpha \in (0, \beta_+)$, by [29, Theorem 5.1], there is a constant C_3 so that $p(t, x) \leq C_3 t^{\alpha/\beta_+}/\|x\|^{d+\alpha}$ for every $t > 0$ and $x \in G_\bullet$. Together with (5.21), it gives the estimate (5.23). This establishes part (i) of the theorem, and thus completes the proof of the theorem. □

Chapter 6
Local limit theorem

Abstract Continuing in the same vein as in the previous chapter where the functional limit theorems are treated, sufficient conditions on the original long range random walk are provided in order to apply and adapt existing local limit theorems to the problems considered here. The local theorem presented in this chapter, Theorem 6.2, is one of the central results of this monograph.

6.1 Assumption (R2)

In this chapter, we discuss the local limit theorem for $(X^{(k)}_{nT_k^{-1}})_{n\in\mathbb{N}\cup\{0\}}$ based on [24, Theorem 1] and [16, Theorem 4.5] (c.f. [6, Section 4] for the case the limit heat kernel is Gaussian).

For this purpose, we introduce an additional hypothesis, **(R2)**, which reads as follows. Let $\mu^{(n)}$ be the n-th convolution power of the probability measure μ on Γ. This is the law at time n of the random walk driven by μ, started at the identity element on Γ.

(R2) There are positive constants $C_2 > 0$ and $\beta > 0$ such that, for all $n, m \in \mathbb{N}$ and $x, y \in \Gamma$,

$$|\mu^{(n+m)}(xy) - \mu^{(n)}(x)| \leq \frac{C_2}{V(n^{1/\beta})}\left(\frac{m}{n+1} + \sqrt{\frac{\|y\|^\beta}{n+1}}\right), \tag{6.1}$$

where $V(r) := \sharp\{g \in \Gamma : \|g\| < r\}$.

For our local limit theorem to hold, the exponent $\beta > 0$ in **(R2)** should be the same as those in (5.3) and in **(E1)**-**(E2)**. We start with verifying the needed convergence of the volume of appropriate balls.

6.2 Convergence of volume

Recall that $\Gamma_t := \delta_t^{-1}(\Gamma) = \delta_{1/t}(\Gamma)$,

$$B(r) = \{x : \|x\| < r\} \quad \text{and} \quad B_\bullet(x, r) = x \bullet B(r) = \left\{y \in \mathbb{R}^d : \|x_\bullet^{-1} \bullet y\| < r\right\}.$$

Recall also that m is the Lebesgue measure on $\mathbb{R}^d$ and

$$m_t(A) = c(\Gamma, G)\det(\delta_t^{-1})\#A$$

for any finite subset A of Γ_t, where $c(\Gamma, G)$ is given in Proposition 4.7 (see also below). We need the following lemma.

Lemma 6.1 For all $x \in \mathbb{R}^d, r \geq 0$,

$$\lim_{t\to\infty} m_t(B_\bullet(x,r) \cap \Gamma_t) = m(B_\bullet(x,r)). \tag{6.2}$$

Proof Fix $x \in \mathbb{R}^d$ and $r \geq 0$. Recall that $B_\bullet(x,r) = x \bullet B(r)$ and

$$\delta_t(x \bullet B(r)) = \delta_t(x) \bullet B(rt^{1/\beta}) = B_\bullet(\delta_t(x), rt^{1/\beta}),$$

so that $B_\bullet(x,r) \cap \Gamma_t$ is the finite set of all points $y \in \mathbb{R}^d$ such that

$$z = \delta_t(y) \in B_\bullet(\delta_t(x), rt^{1/\beta}) \cap \Gamma.$$

Let $\mathrm{dist}_\bullet$ be a left-invariant Riemannian metric on the Lie group $G = (\mathbb{R}^d, \cdot)$ and take the Voronoi cell for the discrete subgroup Γ:

$$U = \left\{x \in \mathbb{R}^d : \mathrm{dist}_\bullet(x,e) = \min_{\gamma\in\Gamma} \mathrm{dist}_\bullet(x,\gamma)\right\}$$

so that

$$\mathbb{R}^d = \bigcup_{\gamma\in\Gamma} \gamma \cdot U$$

and $(\gamma \cdot U) \cap (\gamma' \cdot U) \subset \partial U$ and thus $m((\gamma \cdot U) \cap (\gamma' \cdot U)) = 0$ for any $\gamma \neq \gamma' \in \Gamma$. Note that this definition is based on the law $\cdot$ of the Lie group $G = (\mathbb{R}^d, \cdot)$ and its closed subgroup Γ, not on the rescaled limit law $\bullet$. Since $m(\partial U) = 0$, by definition, $c(\Gamma, G) = m(U)$. For any $S \subset \mathbb{R}^d$, we have

$$c(\Gamma, G)\#\{z \in \Gamma \cap S\} \leq m(S \cdot U).$$

In particular, for $S = B_\bullet(\delta_t(x), rt^{1/\beta}) = \delta_t(B_\bullet(x,r))$,

$$c(\Gamma, G)\#\{z \in \Gamma \cap B_\bullet(\delta(x), rt^{1/\beta})\} \leq m(\delta_t[\delta_{1/t}(\delta_t(B_\bullet(x,r)) \cdot \delta_t(\delta_{1/t}(U)))]).$$

Note that since Γ is a co-compact closed subgroup of G, U is bounded and closed and hence compact. Consequently, $\delta_{1/t}(U)$ converges uniformly to $\{e\}$ as $t \to \infty$. By the uniform convergence of the product $\cdot_t$ to $\bullet$ on compact sets (e.g., see Lemma 5.2), for any fixed $\varepsilon > 0$, there exists a constant $T > 0$ large enough such that, for all $t > T$, the set $[\delta_{1/t}(\delta_t(B_\bullet(x,r)) \cdot \delta_t(\delta_{1/t}(U)))]$ is contained in an ε neighborhood for the norm $\|\cdot\|$ in the group $(G, \bullet)$ of the set $B_\bullet(x,r) \bullet \delta_{1/t}(U)$. This means that, for t large enough,

$$\delta_{1/t}(\delta_t(B_\bullet(x,r)) \cdot \delta_t(\delta_{1/t}(U))) \subset B_\bullet(x, r+2\varepsilon).$$

Hence,

$$\begin{aligned}\det(\delta_{1/t})c(\Gamma, G)\#\{z \in \Gamma \cap B_\bullet(\delta(x), rt^{1/\beta})\} &\leq \det(\delta_{1/t})m(\delta_t(B_\bullet(x, r+2\varepsilon))) \\ &= m(B_\bullet(x, r+2\varepsilon)).\end{aligned}$$

Take the limsup in $t \to \infty$, note that $m(B_\bullet(x, r+2\varepsilon)) = c(r+2\varepsilon)^{\sum_1^d \beta/\beta_i}$, and let ε tend to 0, to obtain

$$\limsup_{t\to\infty} m_t(B_\bullet(x,r) \cap \Gamma_t) \leq m(B_\bullet(x,r)).$$

To prove the complementing inequality, namely,

$$\liminf_{t\to\infty} m_t(B_\bullet(x,r) \cap \Gamma_t) \geq m(B_\bullet(x,r)),$$

we use the same line of reasoning as above to see that, for any fixed $\varepsilon > 0$ and all t large enough,

$$B_\bullet(\delta_t(x), (r-2\varepsilon)t^{1/\beta}) \subset \bigcup_{\gamma \in B_\bullet(\delta_t(x), rt^{1/\beta})} \gamma \cdot U.$$

From this, it follows that, for all t large enough,

$$m(B_\bullet(x, r-2\varepsilon)) \leq \det(\delta_{1/t})c(\Gamma, G)\#\{z \in \Gamma \cap B_\bullet(\delta(x), rt^{1/\beta})\}.$$

The desired lower bound follows. □

6.3 Statement and proof of the LLT

Given an arbitrary sequence of positive reals T_k tending to infinity and $t > 0$, let $\hat{\mu}_k^{(t)}$ be the probability distribution of $(\hat{X}_t^{(k)})_{t>0}$, i.e.,

$$\hat{\mu}_k^{(t)}(x) = \mathbb{P}^e(\hat{X}_t^{(k)} = x) = \mu^{[tT_k]}(\delta_{T_k}(x)), \quad x \in \Gamma_{T_k}.$$

Recall that for each $x \in \mathbb{R}^d$, $[x]_k \in \Gamma_{T_k}$ is the point closest to x in the $\|\cdot\|$-norm.

We know from Theorem 5.10 that the Lévy process $X^\bullet$ s corresponding to $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$ has a jointly continuous convolution kernel

$$(t,x) \mapsto p_\bullet(t,x) = t^{-\gamma_0} p_\bullet(1, \delta_{1/t}(x))$$

with $t > 0$, $x \in \mathbb{R}^d$.

Theorem 6.2 **(Local limit theorem)** Assume (5.3), **(A)-(R1)-(R2)-(E1)-(E2)** and **(T•)-(TΓ)** with the same exponent $\beta > 0$. Then, for any $U_2 > U_1 > 0$ and $r > 1$,

$$\lim_{k\to\infty} \sup_{x\in\mathbb{R}^d:\|x\|\le r} \sup_{t\in[U_1,U_2]} \left|\det(\delta_{T_k})\mu_k^{([tT_k])}(\delta_{T_k}([x]_k)) - p_\bullet(t,x)\right| = 0.$$

Proof We adopt the notations in [24]. Let $E = \mathbb{R}^d$ with $d_E(x,y) = \|x_\bullet^{-1} \bullet y\|$, and $G^k = \delta_{T_k}^{-1}(\Gamma) \subset \mathbb{R}^d = E$ with the same distance $d_{G^k}(x,y) = \|x_\bullet^{-1} \bullet y\|$. (Note that $d_{G^k}(\cdot,\cdot)$ is a graph distance on G^k in [24]. However, the proof of [24, Theorem 1] works for any distance on G^k.) Then, conditions (a) and (b) in [24, Assumption 1] hold with $\alpha(k) = 1$. Let $\nu = m$ and $\nu^k = m_{T_k}$. Then by (6.2), (c) in [24, Assumption 1] holds with $\beta(k) = \det(\delta_{T_k})$. Set

$$q_t^k(x) = \hat{\mu}_k^{(t)}(x) \quad \text{and} \quad q_t(\cdot) = p_\bullet(t,\cdot).$$

It suffices to prove that the conclusion of [24, Theorem 1] holds for $q_t^k(x)$. We now check that (d) in [24, Assumption 1] holds. Let $U_0 > 0$ be a fixed constant. By Theorem 5.10, for every bounded and continuous function f on $\mathbb{R}^d$, $t \in (0,U_0]$ and $x \in \mathbb{R}^d$, we have

$$\lim_{k\to\infty} \left|\hat{\mathbb{E}}_k^{[x]_k}\left[f(\hat{X}_t^{(k)})\right] - \int_{\mathbb{R}^d} f(z)q_t(x_\bullet^{-1} \bullet z)\,dz\right| = 0. \tag{6.3}$$

We need to prove that this convergence is uniform in t over any compact time interval in $(0,\infty)$. This would easily follow if we could prove the equi-uniform continuity of the function $t \mapsto \hat{\mathbb{E}}_k^{[x]_k}\left[f(\hat{X}_t^{(k)})\right]$ on compact time intervals. However, because we are dealing with what is essentially a discrete time process, these functions are not even continuous. Nevertheless, condition **(R2)** says that, for all non-negative integers n,m and all $x,z \in \Gamma$ (inverse and multiplication are in Γ), we have

$$|\mu^{(n+m)}(z) - \mu^{(n)}(x)| \le C_2\left(\frac{m}{n+1} + \frac{\|x^{-1}\cdot z\|^{\beta/2}}{\sqrt{n+1}}\right)\cdot\frac{1}{V(n^{1/\beta})}. \tag{6.4}$$

It follows that, for $0 < s < t$,

$$\begin{aligned}&\left|\mu^{([tT_k])}([\delta_{T_k}([x]_k)]^{-1}\cdot\delta_{T_k}(y)) - \mu^{([sT_k])}([\delta_{T_k}([x]_k)]^{-1}\cdot\delta_{T_k}(y))\right|\\ &\le C_2\frac{[T_k(t-s)]+1}{[T_k s]+1}\frac{1}{V([T_k s]^{1/\beta})} \le C_2\frac{t-s+T_k^{-1}}{s}\frac{1}{V([T_k s]^{1/\beta})}.\end{aligned}$$

For any fixed time interval $[U_1,U_2]$, $0 < U_1 < U_2$, this is a version of "equi-uniform continuity," call it "equi-uniform continuity modulo T_k^{-1}." Together with the fact that $t \mapsto \int_{\mathbb{R}^d} f(z)q_t(x_\bullet^{-1} \bullet z)\,dz$ is uniformly continuous for $t \in [U_1,U_2]$, and (6.3), this equi-uniform continuity modulo T_k^{-1} yields

$$\lim_{k\to\infty} \sup_{t\in[U_1,U_2]} \left|\hat{\mathbb{E}}_k^{[x]_k}\left[f(\hat{X}_t^{(k)})\right] - \int_{\mathbb{R}^d} f(z)q_t(x_\bullet^{-1} \bullet z)\,dz\right| = 0. \tag{6.5}$$

By the joint continuity of $q_t(x)$, we have $\int_{\partial B(x_0,r)} q_t(x_\bullet^{-1} \bullet z)\,dz = 0$ for every $x, x_0 \in E$ and $r > 0$. Hence, (6.5) yields that

$$\lim_{k\to\infty} \sup_{t\in[U_1,U_2]} \left| \mathbb{P}_k^{[x]_k}\left(X^{(k)}_{[T_k t]/T_k} \in B(x_0,r)\right) - \int_{B(x_0,r)} q_t(x_\bullet^{-1} \bullet z)\,dz \right| = 0,$$

and (d) in [24, Assumption 1] is satisfied with $\gamma(k) = T_k$.

On the other hand, by (6.4) again, we have for $x, z \in B(2r) \cap \delta_{T_k}^{-1}(\Gamma)$,

$$\begin{aligned}
\det(\delta_{T_k})|q^k_{[T_k t]}(z) - q^k_{[T_k t]}(x)| &= \det(\delta_{T_k})|\mu^{([T_k t])}(\delta_{T_k}(z)) - \mu^{([T_k t])}(\delta_{T_k}(x))| \\
&\le C_2 \frac{\|\delta_{T_k}(x)^{-1} \cdot \delta_{T_k}(z)\|^{\beta/2}}{\sqrt{T_k t}} \frac{\det(\delta_{T_k})}{V((T_k t)^{1/\beta})} \\
&\le C_3 \|x_\bullet^{-1} \bullet z\|^{\beta_- \beta/(2\beta_+)} t^{-1/2-\gamma} T_k^{-1/2}.
\end{aligned}$$

For the last inequality, we have used Lemma 5.2, and the fact that $V(t^{1/\beta}) \asymp \det(\delta_t) \asymp t^\gamma$ for $\gamma = \sum_1^d 1/\beta_i > 0$. Hence it holds that for any $0 < U_1 < U_2$, $r > 0$, $\delta \in (0, r]$ and $k \ge 1$,

$$\sup_{\substack{x,y\in B_{G^k}(0,r),\\ d_{G^k}(z,x)\le\delta}} \sup_{t\in[U_1,U_2]} \det(\delta_{T_k})|q^k_{[T_k t]}(z) - q^k_{[T_k t]}(x)| \le C_4 \frac{\delta^{\beta_- \beta/(2\beta_+)}}{U_1^{1/2+\gamma}}.$$

Taking $\lim_{\delta\to 0} \limsup_{k\to\infty}$, we obtain [24, Assumption 2]. Therefore, the desired assertion follows from [24, Theorem 1]. □

Remark 6.3 It is well known (see, e.g., [14, 19]) that if a Nash's inequality holds for $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$, then $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$ has transition density function $p(t,x,y)$ that is bounded for each $t > 0$. It follows from Theorem 5.10, $p(t,x,y)$ is jointly locally Hölder continuous in (x,y) for each fixed $t > 0$. Since $X^\bullet$ is a Lévy process on $(G_\bullet, \bullet)$ and $\{\delta_t; t > 0\}$ is a group dilation structure for $(G_\bullet, \bullet)$, we have

$$p(t,x,y) = p(t,e,x_\bullet^{-1} \bullet y) = \det(\delta_t^{-1}) p(1, e, \delta_t^{-1}(x_\bullet^{-1} \bullet y)).$$

Define $p_\bullet(t,x) = \det(\delta_t^{-1}) p(1,e,\delta_t^{-1}(x_\bullet^{-1} \bullet y))$. Then $p_\bullet(t,x)$ is jointly continuous in (t,x), symmetric in $x \in G_\bullet$ (that is, $p_\bullet(t,x) = p_\bullet(t, x_\bullet^{-1})$) and is the convolution kernel for $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$. Moreover, for any $t > 0$ and $x \in G_\bullet$,

$$p_\bullet(t,x) = \det(\delta_t^{-1}) p_\bullet(1, \delta_t^{-1}(x)) = \det(\delta_{1/t}) p_\bullet(1, \delta_{1/t}(x)).$$

□

Chapter 7
Symmetric Lévy processes on nilpotent groups

Abstract This chapter anticipates on later results that show how the limit theorems of Chapters 5 and 6 apply to certain long range random walks. It focuses on the problem of identifying the limit Lévy process that is obtained through these limit theorems when applied to a given explicit long-range random walk. This is done by drawing interesting links between variations on established results regarding approximations of Lévy processes on Lie groups on one hand, and the limit theorems obtained in Chapters 5 and 6 on the other hand. Several examples are given to illustrate the limit theorems explicitly using this approach. Examples discussed in Chapter 1 are also revisited in this new light.

7.1 The problem of identifying the limit process

Theorem 5.10 gives the functional central limit theorem for a class of random walks on simply connected nilpotent groups driven by probability measures μ, which are the distributions of the one-step increments of the random walks. However, the limit symmetric Lévy process $X^\bullet$ is characterized in an abstract way by a non-local pure jump Dirichlet form $(\mathcal{E}_\bullet, \mathcal{F}_\bullet)$ on $L^2(G_\bullet; dx)$ of the form (5.11) with $J_\bullet(dx, dy) = dx\mu_\bullet(x_\bullet^{-1} \bullet dy)$. A natural question is whether we can use Theorem 5.10 to give explicit limit theorems in concrete examples as those studied in Examples 1.4 and 1.5. This amounts to ask whether we can explicitly identity or describe the Lévy process $X^\bullet$ in concrete cases. These are the questions we are going to address in this chapter and the answer is affirmative. In fact, we will do this in a more general context for any symmetric Lévy measure $\mu_\bullet$ on any simply connected nilpotent group $G_\bullet$; that is, $(G_\bullet, \bullet)$ does not need to be the limit group obtained from a simply connected nilpotent group G through an approximate group dilation structure $\{\phi_t; t > 0\}$ on G, and $\mu_\bullet$ does not need to be the weak limit of $\mu_t = t\delta_{1/t}(\mu)$ of some symmetric probability measure μ on a discrete subgroup Γ of G as in condition **(A)**. This is achieved in Theorem 7.3. Then we use this concrete description to illustrate our convergence theorem, Theorem 5.10, by revisiting Example 1.5 and presenting several more examples through this approach, without using the limit results for operator stable processes on $\mathbb{R}^d$, Propositions 1.1 and 1.3, from the literature.

7.2 Symmetric Lévy processes and their approximations

Let N be any simply connected nilpotent group. As mentioned in Section 3.1, there is a global polynomial coordinate system on N satisfying (3.4)-(3.5). Unless mentioned otherwise, this is the default coordinate system we use on N in this section. Through this global system $\Phi : \mathbb{R}^d \to N$ with $\Phi(0) = e$, N can be identified with $\mathbb{R}^d$ and dx is a Haar measure for N. The coordinate system also induces a function on N: $\|\sigma\|_2 := \|\Phi^{-1}(\sigma)\|_2$ for $\sigma \in N$, where $\|\Phi^{-1}(\sigma)\|_2$ is the Euclidean norm of $\Phi^{-1}(\sigma) \in \mathbb{R}^d$. As we already see from Section 3.1, there are many choices of the global coordinate systems for N. One of the commonly used coordinate system is the exponential map. However, sometimes it is more convenient or more natural to use other coordinate systems, for example, matrix coordinates in the Heisenberg group case. Thus with this in mind, we do not fix a particular choice of the polynomial coordinate systems, except for the assumption that (3.4)-(3.5).

Let ν be any non-zero symmetric Lévy measure on N; that is, ν is a non-negative Borel measure on N satisfying $0 < \int_N (1 \wedge \|x\|_2^2)\nu(dx) < \infty$ and $\nu(A) = \nu(A^{-1})$ for any $A \subset N \setminus \{e\}$, where $A^{-1} = \{x \in N : x^{-1} \in A\}$. Note that we do not impose any additional conditions on ν. Define

$$\mathcal{E}(u,v) := \frac{1}{2}\iint_{N\times N\setminus\Delta} (u(xz) - u(x))(v(xz) - v(x))dx\nu(dz), \tag{7.1}$$

and $\mathcal{F}$ is the closure of $\mathrm{Lip}_c(N)$, which is the space of Lipschitz functions on N with compact support, with respect to the norm $\sqrt{\mathcal{E}(u,u) + \int_N u(x)^2 dx}$. Here xz is the group multiplication of two elements $x, z \in N$. Let X be the symmetric Hunt process associated with the regular Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(N; dx)$; cf. [18, 27]. Note that in this section, as mentioned above, we do not assume the Lévy measure ν on N generates N.

Lemma 7.1 The Hunt process X is a Lévy process on N. □

Proof For each fixed $\sigma \in N\backslash\{e\}$, the process $Y^\sigma = \{Y_t^\sigma, t \geq 0\}$, with $Y_t^\sigma = \sigma X_t$ for any $t > 0$, is a symmetric Hunt process on N as dx is a left Haar measure on N and its transition semigroup

$$P_t^\sigma f(x) = \mathbb{E}\left[f(Y_t^\sigma)|Y_0^\sigma = x\right] = \mathbb{E}\left[f(\sigma X_t)|X_0 = \sigma^{-1}x\right] = (P_t f_\sigma)(\sigma^{-1}x),$$

where $f_\sigma(\eta) := f(\sigma\eta)$. Thus

$$\begin{aligned}
&\lim_{t\to 0}\frac{1}{t}(f - P_t^\sigma f, f)_{L^2(N;dx)} \\
&= \lim_{t\to 0}\frac{1}{t}\int_N (f_\sigma(\sigma^{-1}x) - (P_t f_\sigma)(\sigma^{-1}x))f_\sigma(\sigma^{-1}x)dx \\
&= \lim_{t\to 0}\frac{1}{t}\int_N (f_\sigma(y) - (P_t f_\sigma)(y))f_\sigma(y)dy
\end{aligned}$$

$$= \frac{1}{2} \iint_{N\times N\setminus\Delta} (f_\sigma(xz) - f_\sigma(x))^2 \nu(dz)dx$$
$$= \frac{1}{2} \iint_{N\times N\setminus\Delta} (f(xz) - f(x))^2 \nu(dz)dx = \mathcal{E}(f,f).$$

This shows that Y^σ is a symmetric Hunt process associated with the Dirichlet form $(\mathcal{E},\mathcal{F})$ on $L^2(N;dx)$ and so it has the same distribution as X. In other words, $\{\sigma X_t; t \geq 0\}$ with $X_0 = x \in N$ has the same distribution as $\{X_t; t \geq 0\}$ starting from σx. This combined with the Markov property of X shows that X is a symmetric Lévy process on N. □

We next investigate how the Lévy process X is determined by ν in a more explicit way; that is, given a symmetric Lévy measure ν on N, how to construct or approximate its corresponding symmetric Lévy process X in a concrete way. We will show in Theorem 7.3 that X can be approximated by a sequence of random walks on N whose one-step increments are from the small increments of a common Lévy process Z on $\mathbb{R}^d$ through the identification of the global coordinate system Φ. The key is to identify the Lévy measure and the drift of the Lévy process Z on $\mathbb{R}^d$. Our approach uses Hunt's characterization for Lévy processes on Lie groups and Kunita's triangular array type limit result for random walks on Lie groups, which we recall in Theorem 7.2.

We identify each element $\sigma \in N$ with its global coordinate

$$\Phi^{-1}(\sigma) =: x = (x_1, \ldots, x_d) \in \mathbb{R}^d.$$

For each $1 \leq j \leq d$, let $\mathcal{X}_j$ be the left-invariant vector field in the Lie algebra $\mathfrak{g}$ of the group N at e determined by the coordinate function $x_j \mapsto \mathcal{X}_j$; that is, for any C^2 function $f(x)$ on $N = \mathbb{R}^d$,

$$(\mathcal{X}_j f)(e) = \frac{\partial f(x)}{\partial x_j}\Big|_{x=0}.$$

These vector fields $(\mathcal{X}_1, \ldots, \mathcal{X}_d)$ form a natural base of $\mathfrak{g}$ at e. On the other hand, it is well known that the simply connected nilpotent group N admits an exponential map of first type from its Lie algebra $\mathfrak{g} = \mathbb{R}^d$ to N which is surjective. Under its exponential coordinates exp: $\mathfrak{g} \to N$ (of first type), $x^{-1} = -x$. Let $\{x^1, \ldots, x^d\}$ be the exponential coordinate of $\sigma \in N$ with respect to the base $\{\mathcal{X}_1, \ldots, \mathcal{X}_d\}$; that is, $\exp(\sum_{j=1}^d x^j \mathcal{X}_j) = \sigma$. Note that $x^j(\sigma^{-1}) = -x^j(\sigma)$ and $\mathcal{X}_i x^j = \delta_{ij}$. Let $\|\cdot\|$ be the norm on N defined by

$$\|\sigma\| := \Big(\sum_{j=1}^d (x^j(\sigma))^2 \Big)^{1/2} \tag{7.2}$$

in terms of the exponential coordinates of $\sigma \in N$. Note that the norm $\|\cdot\|$ is symmetric on the group N in the sense that $\|\sigma^{-1}\| = \|\sigma\|$ for any $\sigma \in N$.

Denote by C the space of real-valued functions on N that are continuous and have limit at infinity, and C^2 be the space of C^2 functions f on N so that $f, \mathcal{X}_k f$ and

$\mathcal{X}_k\mathcal{X}_j f$ are all in C. Let $\psi \in C^2$ be such that $\psi > 0$ on $G \setminus \{e\}$, $\psi(\eta) \asymp \sum_{j=1}^d x^j(\eta)^2$ near e, and $\lim_{\eta\to\infty} \psi(\eta) > 0$. Note that in view of (3.2)-(3.3),

$$\psi(\eta) \asymp 1 \wedge \|\eta\|^2 \asymp 1 \wedge \|\eta\|_2^2 \quad \text{for } \eta \in N.$$

We recall the following triangular array type limit result on N from [42], which in fact holds for any Lie group.

Theorem 7.2 (Theorem 3 of [42]) In the above setting, suppose the following hold.

(i) For each $n \geq 1$, $k \mapsto S_k^{(n)} = \xi_{n,1} \cdots \xi_{n,k}$ is a discrete time random walk on the Lie group N, where $\{\xi_{n,k}; k \geq 1\}$ are i.i.d N-valued random variables having distribution ν_n.

(ii) As $n \to \infty$, the measure $n\nu_n$ converges vaguely to a measure ν on $N \setminus \{e\}$ satisfying $\int_{N\setminus\{e\}} \psi(x)\nu(dx) < \infty$.

(iii) For $\varepsilon > 0$, let

$$U_\varepsilon := \{\eta \in N : \|\eta\| < \varepsilon\} = \left\{\eta \in N : \sum_{j=1}^d x^j(\eta)^2 < \varepsilon^2\right\}$$

be an ε-neighborhood of e in N. For each $\varepsilon > 0$,

$$\lim_{n\to\infty} n \int_{U_\varepsilon} x^i(\eta) x^j(\eta) \nu_n(d\eta) =: a_{ij}^{(\varepsilon)} \text{ exists.}$$

Clearly, $(a_{ij}^{(\varepsilon)})$ is symmetric and non-negative definite, which decreases to (a_{ij}) as $\varepsilon \to 0$.

(iv) For each $\varepsilon > 0$, $\lim_{n\to\infty} n \int_{U_\varepsilon} x(\eta)\nu_n(d\eta) =: b^{(\varepsilon)} \in \mathbb{R}^d$ exists. Here $x(\eta) = (x^1(\eta), \ldots, x^d(\eta))$.

Take $\varepsilon > 0$ so that ∂U_ε has zero ν-measure. Define

$$b = b_\varepsilon + \int_{U_\varepsilon^c} x(\eta)\nu(d\eta),$$

whose value is independent of the choice of ε. Then for each $T > 0$, $\{Z_t^{(n)} := S_{[nt]}^{(n)}; t \in [0,T]\}$ converges weakly in the Skorokhod space $\mathbb{D}([0,T]; N)$ as $n \to \infty$ to a Lévy process $Z = \{Z_t; t \in [0,T]\}$ on N, whose generator is characterized by

$$\begin{aligned}\mathcal{L}f(\eta) = \frac{1}{2}\sum_{i,j=1}^d a_{ij}\mathcal{X}_i\mathcal{X}_j f(\eta) + \sum_{i=1}^d b_i \mathcal{X}_i f(\eta) \\ + \int_{N\setminus\{e\}} \Big(f(\eta\sigma) - f(\eta) - \sum_{i=1}^d x^i(\sigma)\mathcal{X}_i f(\tau)\Big)\nu(d\sigma) \qquad (7.3)\end{aligned}$$

for any $f \in C^2$. □

Let $\{\varphi_1(\sigma), \cdots, \varphi_d(\sigma)\}$ be C^2 functions on N such that under exponential coordinates for $\sigma = \exp(\sum_{j=1}^d x^j \mathcal{X}_j) \in N$, $\varphi_j(\sigma)$ is an odd increasing function of x^j with $\varphi_j(\sigma) = x^j$ for $x^j \in (-1,1)$. Since N is also identified with $\mathbb{R}^d$ through the global coordinate system Φ satisfying (3.5) mentioned above, sometimes we also write $\varphi_j(x)$ for $\varphi_j(\sigma)$ through this global coordinate system Φ. Since ν is a symmetric measure on N and $\varphi_j(\sigma) = -\varphi_j(\sigma^{-1})$ for any $\sigma \in N$, we have for every $1 \le j \le d$ and every $r > 0$,

$$\int_{\{\sigma \in N: \|\sigma\| > r\}} \varphi_j(\sigma)\nu(d\sigma) = 0. \tag{7.4}$$

Through the identification of N with $\mathbb{R}^d$ under the global coordinate system Φ, the Lévy measure ν can also be viewed as a Lévy measure on the Euclidean space $\mathbb{R}^d$. More precisely, let $\bar{\nu}$ be the Radon measure on $\mathbb{R}^d \setminus \{0\}$ defined by

$$\bar{\nu}(A) := \nu(\Phi(A)) \quad \text{for any } A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\}). \tag{7.5}$$

Note that $\bar{\nu}$ satisfies $\int_{\mathbb{R}^d}(1 \wedge \|z\|_2^2)\bar{\nu}(dz) < \infty$ and thus is a Lévy measure on $\mathbb{R}^d$. However, we point out that even though ν is a symmetric Lévy measure on N, $\bar{\nu}$ may not be a symmetric measure on $\mathbb{R}^d$; see Examples 7.8 and 7.9(i). It is not hard to see or guess that the Lévy process Z on $\mathbb{R}^d$ that will be used to approximate the Lévy process X on N should have Lévy measure $\bar{\nu}$, however in general it also need a proper drift correction term. For this, define for $1 \le j \le d$,

$$\bar{b}_j = \int_{\{z \in N: \|z\|_2 \le 1\}} (z_j - \varphi_j(z))\nu(dz) - \int_{\{z \in N: \|z\|_2 > 1\}} \varphi_j(z)\nu(dz), \tag{7.6}$$

where $(z_1, \dots, z_d) = \Phi^{-1}(z)$ is the coordinates of $z \in N$ under the global coordinate system Φ. Recall that the coordinates of $z \in N$ under the exponential coordinate system is denoted by $(z^1, \dots, z^d)$. Observe that the integral in (7.6) is well-defined and is finite, because

$$\frac{\partial \varphi_j}{\partial z_i}\Big|_{z=0} = \mathcal{X}_i \varphi_j(z)\Big|_{z=0} = \mathcal{X}_i z^j\Big|_{z=0} = \delta_{ij}$$

and so

$$|z_j - \varphi_j(z)| = |z_j - z^j(z)| \le c\|z\|_2^2 \quad \text{for } \|z\|_2 \le 1.$$

In view of (7.4), we can rewrite (7.6) as

$$\bar{b}_j = \lim_{r \to 0} \int_{\{z \in N: \|z\|_2 \le 1 \text{ and } \|z\| \ge r\}} z_j \, \nu(dz). \tag{7.7}$$

Since the Lévy measure ν is symmetric on N, we have from (7.7) that

$$\bar{b}_j = \frac{1}{2} \int_{\{z \in N: \, \|z\|_2 \le 1 \text{ and } \|z^{-1}\|_2 \le 1\}} (z_j + (z^{-1})_j)\nu(dz)$$

$$+\int_{\{z\in N:\|z\|_2\le 1 \text{ and } \|z^{-1}\|_2>1\}} z_j \nu(dz). \tag{7.8}$$

Here z^{-1} denotes the group inverse of $z \in N$, and $(z^{-1})_j$ is the j-th coordinate of the element $z^{-1} \in N$ under the original global coordinate system Φ. Note that both integrals in (7.8) are absolutely convergent. This is because under the global coordinate system Φ, we know from (3.5) that

$$z^{-1} = -z + (0, \bar{q}_2(z_1), \ldots, \bar{q}_d(z_1, \ldots, z_{d-1})),$$

where for $2 \le j \le d$, $\bar{q}_j(z_1, \cdots, z_{j-1})$ is polynomial having no constant and first order terms. Thus on any compact set $K \subset N$, there is a constant $C_K > 0$ so that

$$\|z + z^{-1}\|_2 \le C_K \|z\|_2^2 \quad \text{for every } z \in K, \tag{7.9}$$

and $\{z \in N : \|z\|_2 < 1 \text{ and } \|z^{-1}\|_2 < 1\}$ is an open neighborhood of $e \in N$. Since $\int_{N\setminus\{e\}}(1 \wedge \|z\|_2^2)\nu(dz) < \infty$, both integrals in (7.8) are absolutely convergent. In general, the constant vector $\bar{b} := (b_1, \ldots, \bar{b}_d)$ may not be zero. However, if Φ is the exponential coordinate system, then $\bar{b}_j = 0$ for every $1 \le j \le d$ as $z^{-1} = -z$ for every $z \in N$ and $\|z\|_2 = \|z\|$.

Let $Z := \{Z_t : t \ge 0\}$ be the Lévy process on the Euclidean space $\mathbb{R}^d$ with Lévy triplet $(0, \bar{b}, \bar{\nu})$, see (1.1), where $\bar{b} = (\bar{b}_1, \cdots, \bar{b}_d)$. In other words,

$$Z_t = \bar{b}t + \int_0^t \mathbb{1}_{\{\|z\|_2\le 1\}} z\,(N(ds,dz) - ds\bar{\nu}(dz)) + \int_0^t \mathbb{1}_{\{\|z\|_2>1\}} zN(ds,dz), \tag{7.10}$$

where $N(ds, dz)$ is the Poisson random measure on $[0,\infty)\times\mathbb{R}^d$ with intensity measure $ds\bar{\nu}(dz)$.

Recall that $\Phi : \mathbb{R}^d \to N$ is the global polynomial coordinate system for the simply connected nilpotent group N. Most of the time, we identify $x \in \mathbb{R}^d$ with $\sigma := \Phi(x) \in N$ and use the notations interchangeably. In the next theorem and its proof, to be absolutely clear, we explicitly use the notation $\Phi(x)$ for emphasis when $x \in \mathbb{R}^d$ is viewed as an element in the group N.

Theorem 7.3 Let Z be the Lévy process on $\mathbb{R}^d$ with $Z_0 = 0$, Lévy measure $\bar{\nu}$ of (7.5) and drift $\bar{b}$ of (7.8). For each $T > 0$, the random walk

$$Z_t^{(n)} := \Phi(Z_{1/n})\,\Phi(Z_{2/n} - Z_{1/n}) \cdots \Phi(Z_{[nt]/n} - Z_{([nt]-1)/n}) \tag{7.11}$$

on N converges weakly in the Skorokhod space $\mathbb{D}([0,T]; N)$ as $n \to \infty$ to the left-invariant Hunt process $\{(Y_0)^{-1}Y_t;\ t \in [0,T]\}$ on N. The Hunt process Y has the same distribution as the symmetric Lévy process X on N having Lévy measure ν determined by the Dirichlet form $(\mathcal{E}, \mathcal{F})$ of (7.1) on $L^2(N; dx)$. □

Proof By Ito's formula, for any $f \in C_b^2(\mathbb{R}^d)$,

$$f(Z_t) - f(Z_0)$$

$$= \int_0^t \bar{b} \cdot \nabla f(Z_s) ds + \int_0^t \int_{\{\|z\|_2 \le 1\}} (f(Z_{s-} + z) - f(Z_{s-}))\,(N(ds, dz) - ds\bar{\nu}(dz))$$
$$+ \int_0^t \int_{\{\|z\|_2 > 1\}} (f(Z_{s-} + z) - f(Z_{s-}))N(ds, dz)$$
$$+ \int_0^t \int_{\{\|z\|_2 \le 1\}} (f(Z_s + z) - f(Z_s) - \nabla f(Z_s) \cdot z)\,\bar{\nu}(dz)ds.$$

Thus

$$\begin{aligned}
&\mathbb{E}\, f(Z_t) - f(0) \\
&= \mathbb{E} \int_0^t \bar{b} \cdot \nabla f(Z_s) ds + \mathbb{E} \int_0^t \int_{\{\|z\|_2 > 1\}} (f(Z_{s-} + z) - f(Z_{s-})\, N(ds, dz) \\
&\quad + \mathbb{E} \int_0^t \int_{\{\|z\|_2 \le 1\}} (f(Z_s + z) - f(Z_s) - \nabla f(Z_s) \cdot z)\,\bar{\nu}(dz)ds \\
&= \mathbb{E} \int_0^t \bar{b} \cdot \nabla f(Z_s) ds + \mathbb{E} \int_0^t \int_{\{\|z\|_2 > 1\}} (f(Z_{s-} + z) - f(Z_{s-})\,\bar{\nu}(dz)ds \\
&\quad + \mathbb{E} \int_0^t \int_{\{\|z\|_2 \le 1\}} (f(Z_s + z) - f(Z_s) - \nabla f(Z_s) \cdot z)\,\bar{\nu}(dz)ds \\
&= \mathbb{E} \int_0^t \int_{\mathbb{R}^d} (f(Z_{s-} + z) - f(Z_{s-}) - \nabla f(Z_s) \cdot \varphi(\Phi(z)))\,\bar{\nu}(dz)ds, \qquad (7.12)
\end{aligned}$$

where $\varphi(\sigma) := (\varphi_1(\sigma), \dots, \varphi_d(\sigma))$ for $\sigma \in N$, and the last equality is due to the definition of $\bar{b}$.

For $f_0 \in C^2$ on N with $f_0(e) = 0$ and $X_j f_0(e) = 0$ for $1 \le j \le d$, the function $f := f_0 \circ \Phi$ is C_b^2 on $\mathbb{R}^d$ with $f(0) = 0$ and $\nabla f(0) = 0$. Applying (7.4) and (7.12) to this f, we have by the dominated convergence theorem that

$$\lim_{t\to 0} \frac{1}{t} \mathbb{E}\, f_0(\Phi(Z_t)) = \lim_{t\to 0} \frac{1}{t} \mathbb{E}\, f(Z_t) = \int_{\mathbb{R}^d \setminus \{0\}} f(z)\bar{\nu}(z) = \int_{N\setminus\{e\}} f_0(\sigma)\nu(d\sigma). \qquad (7.13)$$

If we denote the law of $\Phi(Z_t)$ on N by $\widetilde{\nu}_t$, then the above in particular implies that $t^{-1}\widetilde{\nu}_t$ converges vaguely to ν on $N \setminus \{e\}$ as $t \to 0$.

Since φ_j is an odd function on N, taking $f_0 = \varphi_j$ in (7.13) in particular yields that

$$\lim_{t\to 0} \frac{1}{t} \mathbb{E}\, \varphi_j(\Phi(Z_t)) = 0 \quad \text{for every } 1 \le j \le d. \qquad (7.14)$$

On the other hand, since ν is a symmetric measure on N and φ is an odd $\mathbb{R}^d$-valued function on N, we have from (7.12) that for any $f \in C_b^2(\mathbb{R}^d)$,

$$\mathbb{E}\, f(Z_t) - f(0) = \mathbb{E} \int_0^t \int_{\mathbb{R}^d} \left(f(Z_{s-} + z^{-1}) - f(Z_{s-}) + \nabla f(Z_s) \cdot \varphi(\Phi(z)) \right) \bar{\nu}(dz)ds$$

and so

$$\mathbb{E} f(Z_t) - f(0) = \frac{1}{2}\mathbb{E} \int_0^t \int_{\mathbb{R}^d} \Big(f(Z_s + z) + f(Z_{s-} + z^{-1}) - 2f(Z_{s-}) \Big) \, \bar{\nu}(dz) ds. \tag{7.15}$$

Note that by (7.9),

$$\int_{\mathbb{R}^d} |f(z) + f(z^{-1}) - 2f(0)| \, \bar{\nu}(dz) < \infty.$$

It follows from (7.15) and the dominated convergence theorem that for any $f \in C_b^2(\mathbb{R}^d)$,

$$\lim_{t\to 0} \frac{1}{t}\mathbb{E}\,[f(Z_t) - f(0)] = \frac{1}{2} \int_{\mathbb{R}^d\setminus\{0\}} \Big(f(z) + f(z^{-1}) - 2f(0) \Big) \, \bar{\nu}(z). \tag{7.16}$$

For $\varepsilon \in (0, 1)$, define

$$U_\varepsilon = \{\sigma \in N : \|\sigma\| < \varepsilon\} = \Big\{ \sigma \in N : \sum_{i=1}^d x^i(\sigma)^2 < \varepsilon^2 \Big\},$$

where $\|\sigma\|$ is the symmetric norm of $\sigma \in N$ as defined by (7.2) and $(x^1(\sigma), \ldots x^d(\sigma))$ is the exponential coordinates of $\sigma \in N$. By the Lipschitz equivalents (3.2)-(3.3) between Φ and the exponential coordinates, there is a constant $\lambda_0 \geq 1$ so that

$$\lambda_0^{-1} \|\Phi^{-1}(\sigma)\|_2 \leq \|\sigma\| \leq \lambda_0 \|\Phi^{-1}(\sigma)\|_2 \quad \text{for } \sigma \in N \text{ with } \|\sigma\| \leq 1.$$

Consequently,

$$U_\varepsilon \subset \big\{ \sigma \in N : \|\Phi^{-1}(\sigma)\|_2 < \lambda_0 \varepsilon \big\} \quad \text{for every } \varepsilon \in (0, 1).$$

For $\varepsilon \in (0, 1)$, let $f_\varepsilon \in C_c^2(\mathbb{R}^d)$ so that $f_\varepsilon(z) = \|z\|_2^2$ for $|z\|_2 < \lambda_0\varepsilon$, $f_\varepsilon(z) = 0$ for $\|z\|_2 \geq 2\lambda_0\varepsilon$, $0 \leq f_\varepsilon(z) \leq 2\lambda_0^2\varepsilon^2$ and $|Df_\varepsilon| + |D^2 f_\varepsilon| \leq C$ for some constant $C > 0$ independent of ε. Then we have by (7.16) and the Taylor expansion, that

$$\begin{aligned}
&\limsup_{t\to 0} \frac{1}{t}\mathbb{E}\, \big[\mathbb{1}_{\{\Phi(Z_t)\in U_\varepsilon\}} \|\Phi(Z_t)\|^2 \big] \\
&\leq \limsup_{t\to 0} \frac{\lambda_0^2}{t}\mathbb{E}\, \big[\mathbb{1}_{\{\|Z_t\|_2 \leq \lambda_0\varepsilon\}} \|Z_t\|_2^2 \big] \\
&\leq \limsup_{t\to 0} \frac{\lambda_0^2}{t} \left(\mathbb{E}\, f_\varepsilon(Z_t) - f_\varepsilon(Z_0) \right) \\
&= \lambda_0^2 \int_{\mathbb{R}^d} f_\varepsilon(z) \bar{\nu}(dz) \\
&\leq \lambda_0^2 \|D^2 f_\varepsilon\|_\infty \int_{\{\|z\|_2 \leq 2\lambda_0\varepsilon\}} \|z\|_2^2 \, \bar{\nu}(dz),
\end{aligned} \tag{7.17}$$

which tends to 0 as $\varepsilon \to 0$.

For $\varepsilon \in (0,1)$ so that ∂U_ε has zero ν-measure, it follows from (7.13) that

$$\lim_{t\to 0} \frac{1}{t} \int_{U_\varepsilon^c} \varphi_j(z)\nu_t(dz) = \int_{U_\varepsilon^c} \varphi_j(z)\nu(dz) = 0. \tag{7.18}$$

This together with (7.14) shows that

$$\lim_{t\to 0} \frac{1}{t} \int_{U_\varepsilon} \varphi_j(\sigma)\nu_t(d\sigma) = 0. \tag{7.19}$$

Properties (7.13), (7.17) and (7.18)-(7.19) show that the conditions of Theorems 7.2 are all satisfied for the sequence of random walks on N whose one-step increment distributions are $\nu_n := \widetilde{\nu}_{1/n}$ for $n \in \mathcal{N}$ with $(a_{ij}) = 0$ and $b = 0$. Thus for each $T > 0$, the random walk

$$Z_t^{(n)} := \Phi(Z_{1/n})\,\Phi(Z_{2/n} - Z_{1/n}) \cdots \Phi(Z_{[nt]/n} - Z_{([nt]-1)/n})$$

converges weakly in the Skorokhod space $\mathbb{D}([0,T];G)$ as $n \to \infty$ to a symmetric Lévy process $Y = \{Y_t; t \in [0,T]\}$ on N with Lévy measure ν in the following sense: Denote by $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ the infinitesimal generator of Y. Then $C^2 \subset \mathcal{D}(\mathcal{L})$ and for any $f \in C^2$,

$$\mathcal{L} f(\sigma) = \int_{N\setminus\{e\}} \Big(f(\sigma z) - f(\sigma) - \sum_{j=1}^d \varphi_j(z) X_i f(\sigma) \Big) \nu(dz).$$

We next show that Y has the same distribution as the Lévy process X on N defined through the Dirichlet form $(\mathcal{E}, \mathcal{F})$ of (7.1) on $L^2(N; dx)$. Denote by $\mathcal{L}^0$ the L^2-generator of the symmetric Lévy process X. It is easy to check by definition (cf. [18, 27]) that $C^2 \subset \mathcal{D}(\mathcal{L}^0)$ and

$$\begin{aligned} \mathcal{L}^0 f(\sigma) &= \text{p.v.} \int_{N\setminus\{e\}} \left(f(\sigma z) - f(\sigma) \right) \nu(dz) \\ &= \int_{N\setminus\{e\}} \left(f(\sigma z) - f(\sigma) - \sum_{i=1}^d \varphi_j(z) X_i f(\sigma) \right) \nu(dz) \\ &= \mathcal{L} f(\sigma). \end{aligned}$$

By the uniqueness of infinitesimal generator characterization of Lévy processes on N (see, e.g., [42, Theorem 1] due to Hunt), we conclude that the Lévy processes X and Y have the same law. This completes the proof of the theorem. □

Remark 7.4 When Φ is the exponential coordinate system of the first type for N, Theorem 7.3 follows from Theorem 4.2 and the proof of Theorem 4.1 of [41]. The main point of Theorem 7.3 is that it is valid for any global coordinate system Φ of N, not just the exponential coordinate system of first type. This is important in applications as many times it is more natural or convenient to work in other global coordinate systems such as the matrix coordinate system for Heisenberg groups. In theory, one could translate the global coordinate system into exponential coordinate system, apply Kunita's result in exponential coordinate system, and then translate the results back to the original global coordinate system. But this is not always easy to carry out and it needs to be performed on a case by case basis. Interested reader may try the following two exercises.

Exercise 7.5 Let N be the continuous Heisenberg group $\mathbb{H}_3(\mathbb{R})$ and ν be a Lévy measure on $\mathbb{H}_3(\mathbb{R})$ whose expression under the matrix coordinate (x, y, z) is given by

$$\bar{\nu}(dx, dy, dz) = \frac{\kappa_1}{|x|^{1+\alpha_1}} dx \otimes \delta_0(dy) \otimes \delta_0(dz) + \frac{\kappa_2}{|y|^{1+\alpha_2}} \delta_0(dx) \otimes dy \otimes \delta_0(dz)$$

for some positive constants $\alpha_i \in (0, 2)$ and $\kappa_i > 0$, $i = 1, 2$. What is the expression of ν in the exponential coordinates (x^1, x^2, x^3) of the first type for $\mathbb{H}_3(\mathbb{R})$? The group isomorphism between the matrix coordinate system and the exponential coordinate system on $\mathbb{H}_3(\mathbb{R})$ is given in (3.10). □

Exercise 7.6 Repeat Exercise 7.5 with the Lévy measure ν on N being replaced in the matrix coordinate system by

$$\bar{\nu}(dx, dy, dz) = \frac{\kappa_1}{(|x|^2 + |y|^2)^{1+\beta_1}} dx \otimes dy \otimes \delta_0(dz) + \frac{\kappa_2}{(|y|^2 + |z|^2)^{1+\beta_2}} \delta_0(dx) \otimes dy \otimes dz$$

for some positive constants $\beta_i \in (0, 1)$ and $\kappa_i > 0$, $i = 1, 2$. □

7.3 Examples

To illustrate the main results of this work, in this section, we first revisit Example 1.5 of random walks on the Heisenberg group $\mathbb{H}_3(\mathbb{Z})$. Here, we will not use the limit results for operator stable processes from the literature; that is, we will not use Propositions 1.1 and 1.3. We will use instead Theorems 5.10 and 7.3 developed in this monograph. We will then present some more examples.

Exampe 1.5 (revisited) We use the matrix coordinate system Φ on the discrete Heisenberg group $\mathbb{H}_3(\mathbb{Z})$, through which it is identified with $\mathbb{Z}^3$. Denote by e_1, e_2 and e_3 the elements in $\mathbb{H}_3(\mathbb{Z})$ that has matrix coordinates $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$, respectively. Recall that μ_α is the probability measure on $\mathbb{H}_3(\mathbb{Z}) = \mathbb{Z}^3$ given by

$$\mu_\alpha(g) = \sum_{i=1}^{3} \sum_{n \in \mathbb{Z}} \frac{\kappa_i}{(1 + |n|)^{1+\alpha_i}} \mathbb{1}_{\{e_i^n\}}(g), \quad g \in \mathbb{H}_3(\mathbb{Z}),$$

where $0 < \alpha_j < 2$ and κ_j, $1 \leq j \leq 3$, are positive constants. The measure μ_α is in $\mathcal{SM}$ on $\mathbb{H}_3(\mathbb{Z})$ and the matrix coordinate system Φ is an exponential coordinate system of the second kind described in Section 9.5. The dilation structures $\{\delta_t; t > 0\}$

considered below in this example are straight approximate group dilations of (9.3) adapted to the measure μ_α. So by Chapter 10 below, the conditions **(R1)**-**(R2)**, **(E1)**-**(E2)**, **(T•)** and **(TΓ)** are automatically satisfied for μ_α and these $\{\delta_t; t > 0\}$. For simplicity, we write μ for μ_α. Let $\{\xi_k = (\xi_k^{(1)}, \xi_k^{(2)}, \xi_k^{(3)}); k \geq 1\}$ be a sequence of i.i.d random variables taking values in $\mathbb{H}_3(\mathbb{Z})$ of distribution μ. Then

$$S_n = S_0 \cdot \xi_1 \cdot \ldots \cdot \xi_n, \quad n = 0, 1, 2, \ldots$$

defines a random walk on the Heisenberg group $\mathbb{H}_3(\mathbb{Z})$. Write S_n as (X_n, Y_n, Z_n).

(i) If $1/\alpha_3 < 1/\alpha_1 + 1/\alpha_2$, we consider straight dilation structure $\{\delta_t; t > 0\}$ in matrix coordinates:

$$\delta_t(x, y, z) = \left(t^{1/\alpha_1} x,\ t^{1/\alpha_2} y,\ t^{(1/\alpha_1)+(1/\alpha_2)} z\right).$$

In this case, $\{\delta_t; t > 0\}$ is a straight group dilation structure for the limit group $(G_\bullet, \bullet)$, and $(G_\bullet, \bullet)$ is the continuous Heisenberg group $\mathbb{H}_3(\mathbb{R})$. It is easy to check that $t\delta_{1/t}(\mu)$ converges vaguely on $\mathbb{R}^3 \setminus \{0\}$ to $\bar{\mu}_\bullet(dx, dy, dz)$ as $t \to \infty$, where

$$\bar{\mu}_\bullet(dx, dy, dz) = \frac{\kappa_1}{|x|^{1+\alpha_1}} dx \otimes \delta_0(dy) \otimes \delta_0(dz) + \frac{\kappa_2}{|y|^{1+\alpha_2}} \delta_0(dx) \otimes dy \otimes \delta_0(dz).$$

The measure $\bar{\mu}_\bullet$ defines a Lévy measure measure $\mu_\bullet$ on the continuous Heisenberg group $(G_\bullet, \bullet)$ through the matrix coordinate system; see Remark 5.1(i). In other words, $\bar{\mu}_\bullet$ is the pull-back measure of $\mu_\bullet$ under the matrix coordinate system. When there is no danger of confusions, we simply use the same notation $\mu_\bullet$ for $\bar{\mu}_\bullet$. Thus by Theorem 5.10, for any $T > 0$, the rescaled random walk on $\mathbb{H}_3(\mathbb{Z})$ in matrix coordinates

$$\left\{ \left(n^{-1/\alpha_1} X_{[nt]},\ n^{-1/\alpha_2} Y_{[nt]},\ n^{-1/\alpha_1 - 1/\alpha_2} Z_{[nt]}\right); t \in [0, T]\right\}$$

converges weakly in the Skorohod space $\mathbb{D}([0, T]; \mathbb{R}^3)$ to a Lévy process $X^\bullet$ on $(G_\bullet, \bullet)$ with Lévy measure $\mu_\bullet$ as $n \to \infty$. We next identify the Lévy process $X^\bullet$ in the matrix coordinate system of $(G_\bullet, \bullet)$ by using Theorem 7.3.
By (3.9) and (3.10), in matrix coordinates (x, y, z) for $\sigma \in \mathcal{G}_\bullet$,

$$\sigma_\bullet^{-1} = (-x, -y, -z + xy) \quad \text{and} \quad \|\sigma\| = \sqrt{x^2 + y^2 + (z + \tfrac{1}{2}xy)^2}. \tag{7.20}$$

So $\sigma + \sigma_\bullet^{-1} = (0, 0, xy)$. On the support of $\mu_\bullet$, since $xy = 0$, we have $\|\sigma\| = \|\sigma\|_2$ and $\sigma + \sigma_\bullet^{-1} = (0, 0, 0)$, that is, $\sigma_\bullet^{-1} = -\sigma$. Hence denoting the matrix coordinates for $\sigma \in G_\bullet$ by $(\sigma_1, \sigma_2, \sigma_3) \in \mathbb{R}^d$, it follows from (7.8) that for every $1 \leq i \leq 3$,

$$\bar{b}_j = \frac{1}{2} \int_{\{\sigma \in G_\bullet : \|\sigma\|_2 \leq 1\}} \left(\sigma_i + (\sigma_\bullet^{-1})_i\right) \mu_\bullet(d\sigma) = 0.$$

Let $\bar{X}$ be a symmetric α_1-stable process on $\mathbb{R}$ with Lévy measure $\kappa_1 |z|^{-(1+\alpha_1)}$ and $\bar{Y}$ be a symmetric α_2-stable process on $\mathbb{R}$ with Lévy measure $\kappa_2 |z|^{-(1+\alpha_2)}$

independent of X. Then

$$X^\circ = (\bar{X}, \bar{Y}, 0)$$

is a driftless Lévy process on $\mathbb{R}^3$ with Lévy measure $\bar{\mu}_\bullet$ corresponding to (7.10). By Theorem 7.3, $X^\bullet$ is the weak limit on $\mathbb{D}([0,T]; G_\bullet) = \mathbb{D}([0,T]; \mathbb{R}^3)$ of

$$X_t^{\bullet,n} := \Phi(X_{1/n}^\circ) \bullet \Phi(X_{2/n}^\circ - X_{1/n}^\circ). \bullet \cdots \bullet \Phi(X_{[nt]/n}^\circ - X_{([nt]-1)/n}^\circ).$$

Note that in matrix coordinate system on $G_\bullet$,

$$X_t^{\bullet,n} = \left(\bar{X}_{[nt]/n},\, \bar{Y}_{[nt]/n},\, \sum_{k=1}^{[nt]} \bar{X}_{(k-1)/n} \left(\bar{Y}_{k/n} - \bar{Y}_{(k-1)/n} \right) \right), \quad t \geq 0,$$

which converges weakly in the Skorohod space $\mathbb{D}([0,T]; \mathbb{R}^3)$ equipped with $\mathcal{J}_1$-topology to $\{(\bar{X}_t, \bar{Y}_t, \int_0^t \bar{X}_{s-} d\bar{Y}_s); t \in [0,T]\}$; see, e.g., [44, Theorem 7.19]. This shows that $\{X_t^\bullet; t \in [0,T]\}$ under the matrix coordinate system of $(G_\bullet, \bullet)$ has the same distribution as $\{(\bar{X}_t, \bar{Y}_t, \int_0^t \bar{X}_{s-} d\bar{Y}_s); t \in [0,T]\}$.

(ii) If $1/\alpha_3 = 1/\alpha_1 + 1/\alpha_2$, we consider straight dilation structure $\{\delta_t; t > 0\}$ in matrix coordinates:

$$\delta_t(x, y, z) = \left(t^{1/\alpha_1} x,\, t^{1/\alpha_2} y,\, t^{1/\alpha_3} z \right).$$

As mentioned in (i), $\{\delta_t; t > 0\}$ is a straight group dilation structure the limiting group structure $(G_\bullet, \bullet)$ is the continuous Heisenberg group $\mathbb{H}_3(\mathbb{R})$. It is easy to check in this case that $t\delta_{1/t}(\mu)$ converges vaguely on $\mathbb{R}^3 \setminus \{0\}$ to $\bar{\mu}_\bullet(dx, dy, dz)$ as $t \to \infty$, where

$$\begin{aligned} &\mu_\bullet(dx, dy, dz) \\ &= \frac{\kappa_1}{|x|^{1+\alpha_1}} dx \otimes \delta_0(dy) \otimes \delta_0(dz) + \frac{\kappa_2}{|y|^{1+\alpha_2}} \delta_0(dx) \otimes dy \otimes \delta_0(dz) \\ &\quad + \frac{\kappa_3}{|z|^{1+\alpha_3}} \delta_0(dx) \otimes \delta_0(dy) \otimes dz. \end{aligned}$$

The above measure $\bar{\mu}_\bullet$ is the expression of a symmetric Lévy measure $\mu_\bullet$ under the matrix coordinate system on continuous Heisenberg group $(G_\bullet, \bullet)$. By Theorem 5.10, for any $T > 0$, the rescaled random walk on $\mathbb{H}_3(\mathbb{Z})$ in matrix coordinates

$$\left\{ \left(n^{-1/\alpha_1} X_{[nt]},\, n^{-1/\alpha_2} Y_{[nt]},\, n^{-1/\alpha_3} Z_{[nt]} \right); t \in [0,T] \right\}$$

converges weakly in the Skorohod space $\mathbb{D}([0,T]; \mathbb{R}^3)$ to a symmetric Lévy process $X^\bullet$ on $G_\bullet$ with Lévy measure $\mu_\bullet$ as $n \to \infty$.

To identify the Lévy process $X^\bullet$ on $(G_\bullet, \bullet)$ in matrix coordinates (x, y, z) of $\sigma \in G_\bullet$, note that by (7.20), since $xy = 0$ on the support of $\mu_\bullet$, $\|\sigma\| = \|\sigma\|_2$ and $\sigma_\bullet^{-1} = -\sigma$. Thus we have by (7.8) that

$$\bar{b}_j = \frac{1}{2}\int_{\{\sigma\in G_\bullet:\|\sigma\|_2\le 1\}} \left(\sigma_i + (\sigma_\bullet^{-1})_i\right)\mu_\bullet(d\sigma) = 0 \quad \text{for every } 1 \le i \le 3.$$

Let $\bar{X}, \bar{Y}$ and $\bar{Z}$ be independent one-dimensional symmetric α_1-, α_2- and α_3-stable processes with Lévy measure $\kappa_i|z|^{-(1+\alpha_i)}$, $1 \le i \le 3$. Then

$$X^\circ = (\bar{X}, \bar{Y}, \bar{Z})$$

is a driftless Lévy process on $\mathbb{R}^3$ with Lévy measure $\mu_\bullet$ corresponding to (7.10). By Theorem 7.3, $X^\bullet$ is the weak limit on $\mathbb{D}([0,T]; G_\bullet) = \mathbb{D}([0,T]; \mathbb{R}^3)$ of

$$X_t^{\bullet,n} := \Phi(X^\circ_{1/n}) \bullet \Phi(X^\circ_{2/n} - X^\circ_{1/n}). \bullet \cdots \bullet \Phi(X^\circ_{[nt]/n} - X^\circ_{([nt]-1)/n}).$$

In this case, in matrix coordinates,

$$X_t^{\bullet,n} = \left(\bar{X}_{[nt]/n},\, \bar{Y}_{[nt]/n},\, \bar{Z}_{[nt]/n} + \sum_{k=1}^{[nt]} \bar{X}_{(k-1)/n}\left(\bar{Y}_{k/n} - \bar{Y}_{(k-1)/n}\right)\right),$$

which converges weakly in the Skorohod space $\mathbb{D}([0,T]; \mathbb{R}^3)$ equipped with $\mathcal{J}_1$-topology to $\{(\bar{X}_t, \bar{Y}_t, \bar{Z}_t + \int_0^t \bar{X}_{s-} d\bar{Y}_s); t \in [0,T]\}$. This shows that $\{X_t^\bullet; t \in [0,T]\}$ in the matrix coordinate system of $(G_\bullet, \bullet)$ has the same distribution as $\{(\bar{X}_t, \bar{Y}_t, \bar{Z}_t + \int_0^t \bar{X}_{s-} d\bar{Y}_s); t \in [0,T]\}$.

(iii) If $1/\alpha_3 > 1/\alpha_1 + 1/\alpha_2$, we consider straight dilation structure $\{\delta_t; t > 0\}$ in matrix coordinates:

$$\delta_t(x,y,z) = \left(t^{1/\alpha_1} x,\, t^{1/\alpha_2} y,\, t^{1/\alpha_3} z\right).$$

In this case, we see from Example 3.15 that the limit group structure $(G_\bullet, \bullet)$ is just the additive $\mathbb{R}^3$. It is easy to check that $t\delta_{1/t}(\mu)$ converges vaguely on $\mathbb{R}^3 \setminus \{0\}$ to $\mu_\bullet(dx, dy, dz)$ as $t \to \infty$, where

$$\begin{aligned}
&\bar{\mu}_\bullet(dx, dy, dz) \\
&= \frac{\kappa_1}{|x|^{1+\alpha_1}} dx \otimes \delta_0(dy) \otimes \delta_0(dz) + \frac{\kappa_2}{|y|^{1+\alpha_2}} \delta_0(dx) d \otimes dy \otimes \delta_0(dz) \\
&\quad + \frac{\kappa_3}{|z|^{1+\alpha_3}} \delta_0(dx) \otimes \delta_0(dy) \otimes dz.
\end{aligned}$$

Note that the matrix coordinate system on $\mathbb{R}^3$ is the identity map so the induced Lévy measure $\mu_\bullet$ on the abelian group $(G_\bullet, \bullet)$ is just $\bar{\mu}_\bullet$ itself. By Theorem 5.10, for any $T > 0$, the rescaled random walk on $\mathbb{H}_3(\mathbb{Z})$ in matrix coordinates

$$\left\{\left(n^{-1/\alpha_1} X_{[nt]},\, n^{-1/\alpha_2} Y_{[nt]},\, n^{-1/\alpha_3} Z_{[nt]}\right); t \in [0,T]\right\}$$

converges weakly in the Skorohod space $\mathbb{D}([0,T]; \mathbb{R}^3)$ to a symmetric Lévy process $X^\bullet$ on $G_\bullet$ with Lévy measure $\mu_\bullet$ as $n \to \infty$. Since $(G_\bullet, \bullet)$ is $(\mathbb{R}^3, +)$, we conclude directly that $X^\bullet$ has the same distribution as $X^\circ = (\bar{X}, \bar{Y}, \bar{Z})$, where $\bar{X}$, $\bar{Y}$ and $\bar{Z}$ are independent one-dimensional symmetric α_1-, α_2- and α_3-stable processes with Lévy measure $\kappa_i |z|^{-(1+\alpha_i)}$, $1 \le i \le 3$.

We next present a few more examples.

Example 7.7 Let $\mu = \frac{1}{2}(\mu_1 + \mu_2)$ be the probability measure on $\mathbb{H}_3(\mathbb{Z}) = \mathbb{Z}^3$ with

$$\mu_1(x, y, z) = \frac{c_1}{(1 + |x| + |z|)^{2+\alpha_1}} \quad \text{and} \quad \mu_2(x, y, z) = \frac{c_2}{(1 + |y| + |z|)^{2+\alpha_2}},$$

where $0 < \alpha_1, \alpha_2 < 2$ and $c_j > 0$, $j = 1, 2$, are positive constants. The measure μ is again in $\mathcal{SM}$ on $\mathbb{H}_3(\mathbb{Z})$. Let $\{\xi_k = (\xi_k^{(1)}, \xi_k^{(2)}, \xi_k^{(3)}); k \ge 1\}$ be a sequence of i.i.d random variables taking values in $\mathbb{H}_3(\mathbb{Z})$ of distribution μ. Then

$$S_n = S_0 \cdot \xi_1 \cdot \ldots \cdot \xi_n, \quad n = 0, 1, 2, \ldots$$

defines a random walk on the Heisenberg group $\mathbb{H}_3(\mathbb{Z})$. Write S_n as (X_n, Y_n, Z_n).

We consider straight dilation structure $\{\delta_t; t > 0\}$ in matrix coordinates:

$$\delta_t(x, y, z) = (t^{1/\alpha_1} x,\ t^{1/\alpha_2} y,\ t^{1/\alpha_1 + 1/\alpha_2} z).$$

This dilation structure $\{\delta_t; t > 0\}$ is a straight group dilation of (9.3) adapted to the measure μ so the limiting group $(G_\bullet, \bullet)$ is the continuous Heisenberg group $\mathbb{H}_3(\mathbb{R})$. Since the matrix coordinate system Φ is an exponential coordinate system of the second kind described in Section 9.5, by Chapter 10 below, the conditions **(R1)-(R2)**, **(E1)-(E2)**, **(T•)** and **(TΓ)** are automatically satisfied for μ and $\{\delta_t; t > 0\}$. It is easy to check (cf. Example 4.5) that $t\delta_{1/t}(\mu)$ converges vaguely on $\mathbb{R}^3 \setminus \{0\}$ to $\bar{\mu}_\bullet(dx, dy, dz)$ as $t \to \infty$, where

$$\bar{\mu}_\bullet(dx, dy, dz) = \frac{\kappa_1}{|x|^{1+\alpha_1}} dx \otimes \delta_0(dy) \otimes \delta_0(dz) + \frac{\kappa_2}{|y|^{1+\alpha_2}} \delta_0(dx) \otimes dy \otimes \delta_0(dz).$$

Here $\kappa_i = c_i \int_{\mathbb{R}} (1 + |u|)^{-(2+\alpha_i)} du$ for $i = 1, 2$. The measure $\bar{\mu}_\bullet$ induces a Lévy measure $\mu_\bullet$ on $(G_\bullet, \bullet)$ via the matrix coordinate system Φ. In part (i) of Example 1.5 (revisited), we have already identified the symmetric Lévy process $X^\bullet$ on the continuous Heisenberg group $(G_\bullet, \bullet)$. Thus it follows from Theorem 5.10 that, for any $T > 0$, the rescaled random walk on $\mathbb{H}_3(\mathbb{Z})$ in matrix coordinates

$$\left\{ \left(n^{-1/\alpha_1} X_{[nt]},\ n^{-1/\alpha_2} Y_{[nt]},\ n^{-1/\alpha_1 - 1/\alpha_2} Z_{[nt]} \right); t \in [0, T] \right\}$$

converges weakly in the Skorohod space $\mathbb{D}([0, T]; \mathbb{R}^3)$ to $\{(\bar{X}_t, \bar{Y}_t, \int_0^t \bar{X}_{s-} d\bar{Y}_s); t \in [0, T]\}$ on the continuous Heisenberg group $\mathbb{H}_3(\mathbb{R})$ in matrix coordinates, where $\bar{X}$ and $\bar{Y}$ are independent one-dimensional symmetric α_1- and α_2-stable processes, respectively.

Example 7.8 Let μ be the probability measure on $\mathbb{H}_3(\mathbb{Z}) = \mathbb{Z}^3$ with

$$\mu(x,y,z) = \frac{c}{(1+\sqrt{x^2+y^2+|z-xy|})^{4+\alpha}},$$

where $0 < \alpha < 2$ and $c > 0$ are positive constants.

The measure μ is again in $\mathcal{SM}$ on $\mathbb{H}_3(\mathbb{Z})$. Let $\{\xi_k = (\xi_k^{(1)}, \xi_k^{(2)}, \xi_k^{(3)}); k \geq 1\}$ be a sequence of i.i.d random variables taking values in $\mathbb{H}_3(\mathbb{Z})$ of distribution μ. Then

$$S_n = S_0 \cdot \xi_1 \cdot \ldots \cdot \xi_n, \quad n = 0,1,2,\ldots$$

defines a random walk on the Heisenberg group $\mathbb{H}_3(\mathbb{Z})$. Write S_n as (X_n, Y_n, Z_n).

Consider the dilation

$$\delta_t((x,y,z)) = (t^{1/\alpha}x, t^{1/\alpha}y, t^{2/\alpha}z).$$

This dilation structure $\{\delta_t; t > 0\}$ is a straight group dilation of (9.3) adapted to the measure μ so the limiting group structure $(G_\bullet, \bullet)$ is the continuous Heisenberg group $\mathbb{H}_3(\mathbb{R})$. By Chapter 10 below, the conditions **(R1)**-**(R2)**, **(E1)**-**(E2)**, **(T•)** and **(TΓ)** are automatically satisfied for μ and $\{\delta_t, t > 0\}$. It is easy to check in this case that $t\delta_{1/t}(\mu)$ converges vaguely on $\mathbb{R}^3 \setminus \{0\}$ to $\bar{\mu}_\bullet(dx,dy,dz)$ as $t \to \infty$, where

$$\bar{\mu}_\bullet(dx) = \frac{c}{(\sqrt{x^2+y^2+|z-xy|})^{4+\alpha}} dxdydz.$$

The measure $\bar{\mu}_\bullet$, though itself is not symmetric on $\mathbb{R}^3$, induces a symmetric Lévy measure $\mu_\bullet$ on $(G_\bullet, \bullet)$ via the matrix coordinate system Φ. Thus by Theorem 5.10, for any $T > 0$, the rescaled random walk on $\mathbb{H}_3(\mathbb{Z})$ in matrix coordinates

$$\left\{ \left(n^{-1/\alpha} X_{[nt]},\, n^{-1/\alpha} Y_{[nt]},\, n^{-2/\alpha} Z_{[nt]} \right); t \in [0,T] \right\}$$

converges weakly in the Skorohod space $\mathbb{D}([0,T];\mathbb{R}^3)$ to a purely discontinuous symmetric Lévy process $X^\bullet$ on $(G_\bullet, \bullet)$ with Lévy measure $\mu_\bullet$ as $n \to \infty$. We next identify the Lévy process $X^\bullet$ in the matrix coordinate system of $(G_\bullet, \bullet)$ by using Theorem 7.3.

Recall that for $\sigma = (x,y,z) \in \mathbb{H}_3(\mathbb{R})$,

$$\sigma_\bullet^{-1} = (-x, -y, -z+xy) \quad \text{and} \quad \sigma + \sigma_\bullet^{-1} = (0,0,xy).$$

Since $\|\sigma\|_2$, $\|\sigma_\bullet^{-1}\|_2$ and $\bar{\mu}_\bullet$ are invariant under the transformations $(x,y,z) \mapsto (-x,y,-z)$ and $(x,y,z) \mapsto (x,-y,-z)$, we have by (7.8) that $\bar{b}_j = 0$ for every $1 \leq j \leq 3$. Let $X^\circ = (\bar{X}, \bar{Y}, \bar{Z})$ be the Lévy process on $\mathbb{R}^3$ with Lévy triplet $(0,0,\bar{\mu}_\bullet)$. Note that the Lévy process X° is not symmetric on $\mathbb{R}^3$ as its Lévy measure $\bar{\mu}_\bullet$ is not symmetric on $\mathbb{R}^3$. We conclude from Theorem 7.3 with the same calculation as that in part (ii) of Example 1.5(revisited) that, in matrix coordinate Φ,

$$X_t^\bullet = \left(\bar{X}_t, \bar{Y}_t, \bar{Z}_t + \int_0^t \bar{X}_{s-} d\bar{Y}_s \right) \quad \text{for } t \geq 0.$$

Example 7.9 Consider the group $\mathbb{U}_4(\mathbb{Z})$ of 4 by 4 upper-triangular matrices with diagonal entries equal to 1 given in Example 3.16. That is,

$$\Gamma = \mathbb{U}_4(\mathbb{Z}) = \left\{ \begin{pmatrix} 1 & x_1 & x_4 & x_6 \\ 0 & 1 & x_2 & x_5 \\ 0 & 0 & 1 & x_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} : x_i \in \mathbb{Z} \right\}.$$

In matrix coordinates, $\mathbb{U}_4(\mathbb{R})$ is $\mathbb{R}^6$ with multiplication $(x_i)_1^6(y_i)_1^6 = (z_i)_1^6$ given by

$$z_i = \begin{cases} x_i + y_i & \text{for } i = 1,2,3, \\ x_4 + y_4 + x_1 y_2 & \text{for } i = 4, \\ x_5 + y_5 + x_2 y_3 & \text{for } i = 5, \\ x_6 + y_6 + x_1 y_5 + x_4 y_3 & \text{for } i = 6. \end{cases}$$

This matrix coordinate system Φ is an exponential coordinate system of the second kind described in Section 9.5; see Example 3.3.

We consider two cases.

(i) Let $\mu = \frac{1}{2}(\mu_1 + \mu_2)$ be the probability measure on $\mathbb{U}_4(\mathbb{Z}) = \mathbb{Z}^6$ with

$$\mu_1((x_i)_1^6) = \frac{c_1}{(1 + \sqrt{x_1^2 + x_2^2 + |x_4 - x_1 x_2|})^{4+\alpha_1}} \otimes \mathbb{1}_{(0,0,0)}(x_3, x_5, x_6) \tag{7.21}$$

and

$$\mu_2((x_i)_1^6) = \frac{c_2}{(1 + \sqrt{x_3^2 + x_5^2 + x_6^2})^{3+\alpha_2}} \otimes \mathbb{1}_{(0,0,0)}(x_1, x_2, x_4), \tag{7.22}$$

where $0 < \alpha_1, \alpha_2 < 2$ and c_1, c_2 are appropriate positive normalizing constants.

The measure μ is in $\mathcal{SM}$ on $\mathbb{U}_4(\mathbb{Z})$. Let $\{\xi_k = (\xi_k^{(1)}, \xi_k^{(2)}, \xi_k^{(3)}, \xi_k^{(4)}, \xi_k^{(5)}, \xi_k^{(6)}); k \geq 1\}$ be a sequence of i.i.d random variables taking values in $\mathbb{U}_4(\mathbb{Z})$ of distribution μ. Then

$$S_n = S_0 \cdot \xi_1 \cdot \ldots \cdot \xi_n, \quad n = 0, 1, 2, \ldots$$

defines a random walk on $\mathbb{U}_4(\mathbb{Z})$. Write S_n as

$$(X_n^{(1)}, X_n^{(2)}, X_n^{(3)}, X_n^{(4)}, X_n^{(5)}, X_n^{(6)}).$$

Consider the dilation

$$\delta_t((x_i)_1^6) = (t^{1/\alpha_1} x_1, t^{1/\alpha_1} x_2, t^{1/\alpha_2} x_3, t^{2/\alpha_1} x_4, t^{1/\alpha_1 + 1/\alpha_2} x_5, t^{2/\alpha_1 + 1/\alpha_2} x_6).$$

As noted in Example 3.16, this is a group dilation structure so the limit group $(G_\bullet, \bullet)$ is $\mathbb{U}_4(\mathbb{R})$. It is in fact the straight group dilation of (9.3) adapted to the measure μ Thus by Chapter 10 below, the conditions **(R1)**-**(R2)**, **(E1)**-**(E2)**, **(T•)** and **(TΓ)** are automatically satisfied for μ and $\{\delta_t, t > 0\}$.

The measure $\mu_t = t\delta_t(\mu)$ has vague limit $\bar{\mu}_\bullet$ as $t \to \infty$ given by

$$\bar{\mu}_\bullet(dx) = \frac{c_1}{2\left(x_1^2 + x_2^2 + |x_4 - x_1 x_2|\right)^{(4+\alpha_1)/2}} dx_1 dx_2 dx_4 \otimes \delta_{(0,0,0)}(dx_3, dx_5, dx_6) + \frac{c_2'}{2|x_3|^{1+\alpha_2}} dx_3 \otimes \delta_{(0,0,0,0,0)}(dx_1, dx_2, dx_4, dx_5, dx_6).$$

Note that though the measure $\bar{\mu}_\bullet$ is not symmetric on $\mathbb{R}^6$, it induces a symmetric Lévy measure $\mu_\bullet$ on $(G_\bullet, \bullet)$ through the matrix coordinate system Φ.

By a similar reasoning as in the previous example, one can check that the drift $\bar{b}$ defined by (7.8) is the zero vector in $\mathbb{R}^6$. Let

$$X^\circ = \left(\bar{X}^{(1)}, \bar{X}^{(2)}, \bar{X}^{(3)}, \bar{X}^{(4)}, 0, 0\right),$$

be the Lévy process on $\mathbb{R}^6$ with Lévy triplet $(0, 0, \bar{\mu}_\bullet)$. Note that $(\bar{X}^{(1)}, \bar{X}^{(2)}, \bar{X}^{(4)})$ is a Lévy process on $\mathbb{R}^3$ with Lévy triplet $\left(0, 0, \frac{c_1}{2\left(x_1^2+x_2^2+|x_4-x_1x_2|\right)^{(4+\alpha_1)/2}} dx_1 dx_2 dx_4\right)$ and $\bar{X}^{(3)}$ is a one-dimensional symmetric α_2-stable process with Lévy measure $\frac{c_2'}{2|x_3|^{1+\alpha_2}} dx_3$ independent of $(\bar{X}^{(1)}, \bar{X}^{(2)}, \bar{X}^{(4)})$. Thus we have by Theorem 5.10, for any $T > 0$, the rescaled random walk $\{\delta_{1/n}(S_{[nt]}); t \in [0, T]\}$ on $\mathbb{U}_4(\mathbb{Z})$ converges weakly in the Skorohod space $\mathbb{D}([0, T]; \mathbb{R}^6)$ to a purely discontinuous symmetric Lévy process $X^\bullet$ on $(G_\bullet, \bullet)$ with Lévy measure $\mu_\bullet$ as $n \to \infty$.

We next identify the Lévy process $X^\bullet$ in the matrix coordinate system of $(G_\bullet, \bullet)$ by using Theorem 7.3, through the fact that $X^\bullet$ is the weak limit of

$$X_t^{\bullet,n} := \Phi(X_{1/n}^\circ) \bullet \Phi(X_{2/n}^\circ - X_{1/n}^\circ). \bullet \cdots \bullet \Phi(X_{[nt]/n}^\circ - X_{([nt]-1)/n}^\circ).$$

When $\alpha_1 \le \alpha_2$, $(G_\bullet, \bullet)$ is $\mathbb{U}_4(\mathbb{R})$. By a similar reasoning as in previous examples, we conclude from Theorem 7.3 that in matrix coordinates, the symmetric Lévy process $X_t^\bullet$ on $\mathbb{U}_4(\mathbb{R})$ has the following six coordinates:

$$\bar{X}_t^{(1)},\ \bar{X}_t^{(2)},\ \bar{X}_t^{(3)},\ \bar{X}^{(4)} + \int_0^t \bar{X}_{s-}^{(1)} d\bar{X}_s^{(2)},\ \int_0^t \bar{X}_{s-}^{(2)} d\bar{X}_s^{(3)},$$

and

$$\int_0^t \bar{X}_{s_-}^{(1)} \bar{X}_{s_-}^{(2)} d\bar{X}_s^{(3)} + \int_0^t \left(\bar{X}_{r_-}^{(4)} + \int_{[0,r)} \bar{X}_{s-}^{(1)} d\bar{X}_s^{(2)}\right) d\bar{X}_r^{(3)}.$$

Note that Lévy processes are semimartingales so the above stochastic integrals are all well defined.

(ii) Now let $\mu = \frac{1}{3}(\mu_1 + \mu_2 + \mu_3)$ be the probability measure on $\mathbb{U}_4(\mathbb{Z}) = \mathbb{Z}^6$ with μ_1 and μ_2 given by (7.21) and (7.22) with $\alpha_1 = \alpha_2 \in (0, 2)$, and

$$\mu_3((x_i)_1^6) = \frac{c_3}{\left(1 + \sqrt{x_4^2 + x_5^2 + x_6^2}\right)^{3+\alpha_3}} \otimes \mathbb{1}_{(0,0,0)}(x_1, x_2, x_3)$$

for some $\alpha_3 \in (0, \alpha_1/2)$. This measure μ is in $\mathcal{SM}$ on $\mathbb{U}_4(\mathbb{Z})$. Let

$$\left\{\xi_k = \left(\xi_k^{(1)}, \xi_k^{(2)}, \xi_k^{(3)}, \xi_k^{(4)}, \xi_k^{(5)}, \xi_k^{(6)}\right); k \geq 1\right\}$$

be a sequence of i.i.d random variables taking values in $\mathbb{U}_4(\mathbb{Z})$ of distribution μ. Then

$$S_n = S_0 \cdot \xi_1 \cdot \ldots \cdot \xi_n, \quad n = 0, 1, 2, \ldots$$

defines a random walk on $\mathbb{U}_4(\mathbb{Z})$.

Consider the dilation

$$\delta_t((x_i)_1^6) = \left(t^{1/\alpha_1} x_1,\, t^{1/\alpha_1} x_2,\, t^{1/\alpha_1} x_3,\, t^{1/\alpha_3} x_4,\, t^{1/\alpha_3} x_5,\, t^{1/\alpha_1 + 1/\alpha_3} x_6\right),$$

which is a straight group approximate dilation of (9.3) adapted to the measure μ. As noted in Example 3.16 (the fourth bullet case), this is an approximate group dilation structure for $\mathbb{U}_4(\mathbb{Z})$ and the group law $\bullet$ of the limit group $(G_\bullet, \bullet)$ is the direct product of the 5 dimensional Heisenberg group $\mathbb{H}_5(\mathbb{R})$ and a copy of $\mathbb{R}$, that is,

$$\begin{aligned}&(x_i)_1^6 \bullet (y_i)_1^6 \\ &= (x_1 + y_1, x_2 + y_2, x_3 + y_3, x_4 + y_4, x_5 + y_5, x_6 + y_6 + x_1 y_5 + x_4 y_3).\end{aligned} \qquad \square$$

Clearly, $(G_\bullet, \bullet)$ is different from from $\mathbb{U}_4(\mathbb{R})$. Since the measure μ is in $\mathcal{SM}$ on $\mathbb{U}_4(\mathbb{Z})$, the conditions **(R1)**-**(R2)**, **(E1)**-**(E2)**, **(T$\bullet$)** and **(TΓ)** are automatically satisfied for μ and $\{\delta_t, t > 0\}$ by Chapter 10 below.

The measure $\mu_t = t\delta_t(\mu)$ has vague limit $\bar{\mu}_\bullet$ as $t \to \infty$ given by

$$\begin{aligned}\bar{\mu}_\bullet(dx) =& \frac{\kappa_1}{(x_1^2 + x_2^2)^{(2+\alpha_1)/2}} dx_1 dx_2 \otimes \delta_{(0,0,0,0)}(dx_3, dx_4, dx_5, dx_6) \\ &+ \frac{\kappa_2}{|x_3|^{1+\alpha_1}} dx_3 \otimes \delta_{(0,0,0,0,0)}(dx_1, dx_2, dx_4, dx_5, dx_6) \\ &+ \frac{\kappa_3}{(x_4^2 + x_5^2)^{(2+\alpha_3)/2}} dx_4 dx_5 \otimes \delta_{(0,0,0,0)}(dx_1, dx_2, dx_3, dx_6).\end{aligned}$$

It induces a symmetric Lévy measure $\mu_\bullet$ on $(G_\bullet, \bullet)$ through the matrix coordinate system Φ. It is easy to see that the drift $\bar{b}$ defined by (7.8) is the zero vector in $\mathbb{R}^6$ and the Lévy process X° on $\mathbb{R}^d$ with Lévy triplet $(0, 0, \bar{\mu}_\bullet)$ is

$$X_t^\circ = \left(\bar{X}_t^{(1)}, \bar{X}_t^{(2)}, \bar{X}_t^{(3)}, \bar{X}_t^{(4)}, \bar{X}_t^{(5)}, 0\right),$$

where $(\bar{X}^{(1)}, \bar{X}^{(2)})$ is a two-dimensional isotropic α_1-stable process, $\bar{X}^{(3)}$ is an independent one-dimensional α_1-stable process, and $(\bar{X}^{(4)}, \bar{X}^{(5)})$ is a two-dimensional isotropic α_3-stable process that is independent of $(\bar{X}^{(1)}, X^{(2)}, X^{(3)})$. In a similar way as in previous examples, we can conclude from Theorems 5.10 and 7.3 that for any $T > 0$, the rescaled random walk $\left\{\delta_{1/n}(S_{[nt]}); t \in [0, T]\right\}$ on $\mathbb{U}_4(\mathbb{Z})$ converges weakly in the Skorohod space $\mathbb{D}([0, T]; \mathbb{R}^6)$ to a purely discontinuous symmetric Lévy process $X^\bullet$ on $(G_\bullet, \bullet)$ with Lévy measure $\mu_\bullet$ as $n \to \infty$, which in the matrix coordinate system is given by

$$\left(\bar{X}^{(1)}_t,\ \bar{X}^{(2)}_t,\ \bar{X}^{(3)}_t,\ \bar{X}^{(4)}_t,\ \bar{X}^{(5)}_t,\ \int_0^t \bar{X}^{(1)}_{s-} d\bar{X}^{(5)}_s + \int_0^t \bar{X}^{(4)}_{s-} d\bar{X}^{(3)}_s\right).$$

Example 7.10 In Example 7.9, now consider the probability measure $\mu = \frac{1}{3}(\mu_1 + \mu_2 + \mu_3)$ on $\mathbb{U}_4(\mathbb{Z}) = \mathbb{Z}^6$, where

$$\mu_1((x_i)_1^6) = \frac{c_1}{\left(1 + \sqrt{x_1^2 + x_4^2 + x_6^2|}\right)^{3+\alpha_1}} \otimes \mathbb{1}_{(0,0,0)}(x_2, x_3, x_5),$$

$$\mu_2((x_i)_1^6) = \frac{c_2}{(1 + |x_2|)^{1+\alpha_2}} \otimes \mathbb{1}_{(0,0,0,0,0)}(x_1, x_3, x_4, x_5, x_6),$$

and

$$\mu_3((x_i)_1^6) = \frac{c_3}{\left(1 + \sqrt{x_3^2 + x_5^2 + x_6^2}\right)^{3+\alpha_3}} \otimes \mathbb{1}_{(0,0,0)}(x_1, x_2, x_4),$$

where $0 < \alpha_1, \alpha_2, \alpha_3 < 2$ and c_1, c_2, c_3 are appropriate positive normalizing constants.

The measure μ is in $\mathcal{SM}$ on $\mathbb{U}_4(\mathbb{Z})$. Let

$$\left\{\xi_k = (\xi_k^{(1)}, \xi_k^{(2)}, \xi_k^{(3)}, \xi_k^{(4)}, \xi_k^{(5)}, \xi_k^{(6)});\ k \geq 1\right\}$$

be a sequence of i.i.d random variables taking values in $\mathbb{U}_4(\mathbb{Z})$ of distribution μ. Then

$$S_n = S_0 \cdot \xi_1 \cdot \ldots \cdot \xi_n, \quad n = 0, 1, 2, \ldots$$

defines a random walk on $\mathbb{U}_4(\mathbb{Z})$.

Consider the dilation

$$\delta_t((x_i)_1^6) = \left(t^{1/\alpha_1} x_1, t^{1/\alpha_2} x_2, t^{1/\alpha_3} x_3, t^{1/\alpha_1 + 1/\alpha_2} x_4, t^{1/\alpha_2 + 1/\alpha_3} x_5, t^{1/\alpha_1 + 1/\alpha_2 + 1/\alpha_3} x_6\right).$$

As noted in Example 3.16, this is a group dilation structure so the limiting group $G_\bullet$ is $\mathbb{U}_4(\mathbb{R})$. It is in fact a straight group dilation of (9.3) adapted to the measure μ.

Since measure μ is in $\mathcal{SM}$ on $\mathbb{U}_4(\mathbb{Z})$, the conditions **(R1)-(R2)**, **(E1)-(E2)**, **(T•)** and **(TΓ)** are again automatically satisfied for μ and $\{\delta_t, t > 0\}$ by Chapter 10 below. The measure $\mu_t = t\delta_t(\mu)$ has vague limit $\bar{\mu}_\bullet$ as $t \to \infty$ given by

$$\begin{aligned} \bar{\mu}_\bullet(dx) = & \frac{\kappa_1}{|x_1|^{1+\alpha_1}} dx_1 \otimes \delta_{(0,0,0,0,0)}(dx_2, dx_3, dx_4, x_5, dx_6) \\ & + \frac{\kappa_2}{|x_2|^{1+\alpha_2}} dx_2 \otimes \delta_{(0,0,0,0,0)}(dx_1, dx_3, dx_4, dx_5, dx_6) \\ & + \frac{\kappa_3}{|x_3|^{1+\alpha_3}} dx_3 \otimes \delta_{(0,0,0,0,0)}(dx_1, dx_2, dx_4, dx_5, dx_6). \end{aligned} \tag{7.23}$$

It induces a symmetric Lévy measure $\mu_\bullet$ on $(G_\bullet, \bullet)$ through the matrix coordinate system Φ. It is easy to see that the drift $\bar{b}$ defined by (7.8) is the zero vector in $\mathbb{R}^6$ and the Lévy process X° on $\mathbb{R}^d$ with Lévy triplet $(0, 0, \bar{\mu}_\bullet)$ is

$$X_t^\circ = \left(\bar{X}_t^{(1)}, \bar{X}_t^{(2)}, \bar{X}_t^{(3)}, 0, 0, 0\right),$$

where $\bar{X}^{(i)}$ are one-dimensional symmetric α_i-stable processes with Lévy measure $\kappa_i |z|^{-1-\alpha_i} dz$ for $1 \le i \le 3$, independent to each other. In a similar way as in previous examples, we can conclude from Theorems 5.10 and 7.3 that for any $T > 0$, the rescaled random walk $\left\{\delta_{1/n}(S_{[nt]}); t \in [0, T]\right\}$ on $\mathbb{U}_4(\mathbb{Z})$ converges weakly in the Skorohod space $\mathbb{D}([0, T]; \mathbb{R}^6)$ to a purely discontinuous symmetric Lévy process $X^\bullet$ on $(G_\bullet, \bullet)$ with Lévy measure $\mu_\bullet$ as $n \to \infty$, which in the matrix coordinate system is given by

$$\bar{X}_t^{(1)},\ \bar{X}_t^{(2)},\ \bar{X}_t^{(3)},\ \int_0^t \bar{X}_{s-}^{(1)} d\bar{X}_s^{(2)},\ \int_0^t \bar{X}_{s-}^{(2)} d\bar{X}_s^{(3)},$$

and

$$\int_0^t \bar{X}_{s-}^{(1)} \bar{X}_{s-}^{(2)} d\bar{X}_s^{(3)} + \int_0^t \left(\int_{[0,r)} \bar{X}_{s-}^{(1)} d\bar{X}_s^{(2)}\right) d\bar{X}_r^{(3)}.$$

In the above, if we replace μ_2 by

$$\mu_2'((x_i)_1^6) = \frac{c_2}{\left(1 + \sqrt{x_2^2 + x_3^2 + x_5^2}\right)^{3+\alpha_2}} \otimes \mathbb{1}_{(0,0,0)}(x_1, x_4, x_6),$$

the measure μ is again an $\mathcal{SM}$ measure on $\mathbb{U}_4(\mathbb{Z})$ and $\mu_t = t\delta_t(\mu)$ converges vaguely to the same $\bar{\mu}_\bullet$ of (7.23) as $t \to \infty$. Thus for any $T > 0$, the rescaled random walk $\{\delta_{1/n}(S_{[nt]}); t \in [0, T])$ on $\mathbb{U}_4(\mathbb{Z})$ converges weakly in the Skorohod space $\mathbb{D}([0, T]; \mathbb{R}^6)$ to the same purely discontinuous symmetric Lévy process on $\mathbb{U}_4(\mathbb{R})$.□

Remark 7.11 Since condition **(R2)** is satisfied, the local limit theorem, Theorem 6.2, holds as well for all the examples in this section. □

Chapter 8
Measures in $\mathcal{SM}(\Gamma)$ and their geometries

Abstract For any torsion free finitely generated nilpotent group Γ, this short but essential chapter introduces the set $\mathcal{SM}(\Gamma)$, a set of stable-like probability measures on Γ. For each measure μ in $\mathcal{SM}(\Gamma)$, a particular "geometry" associated to μ is defined. This geometry will later be the key needed to understand how to define norms and appropriate approximate dilations adapted to the measure μ in order to apply the limit theorems of Chapters 5 and 6.

8.1 Probability measures in $\mathcal{SM}$ and $\mathcal{SM}_1$

Consider a subgroup $H \subset \Gamma$. Because Γ is nilpotent, H is automatically finitely generated and we equip H with a finite symmetric generating set S and the associated word-length $|\cdot|_S$. Let $\alpha \in (0,2)$. Let $\mathcal{SM}_H^\alpha(\Gamma)$ be the set of all symmetric probability measures ν on Γ which are supported on H and satisfy

$$\nu(g) \asymp \frac{1}{(1+|g|_S)^\alpha V_{H,S}(|g|_S)} \mathbb{1}_H(g),$$

where $V_{H,S}$ is the volume growth function of the pair (H,S) and the notation $\nu \asymp \mu$ indicates that there are constants $0 < c \leq C < \infty$, which may depend on ν and μ, so that

$$c\mu(g) \leq \nu(g) \leq C\mu(g) \quad \text{for every } g \in \Gamma.$$

Note that since H is a subgroup of the finitely generated nilpotent group, its volume growth is polynomial and there is an integer d_H such that $V_{H,S}(k) \asymp k^{d_H}$, $k = 1,2,\dots$. Note also that the set $\mathcal{SM}_H^\alpha(\Gamma)$ does not depend on the choice of the generating S for H. These symmetric probability measures are the basic building blocks of the set $\mathcal{SM}(\Gamma)$ which we now define.

Definition 8.1 ($\mathcal{SM}(\Gamma)$) The set $\mathcal{SM}(\Gamma)$ is the set of all finite convex combinations μ of probability measures belonging to the union

$$\bigcup_{\alpha\in(0,2)} \bigcup_{H:\,\text{subgroup of }\Gamma} \mathcal{SM}_H^\alpha(\Gamma)$$

such that the support of μ generates Γ. □

Definition 8.2 ($\mathcal{SM}_1(\Gamma)$) The subset $\mathcal{SM}_1(\Gamma)$ of $\mathcal{SM}(\Gamma)$ is the set of all finite convex combinations μ of probability measures in

$$\bigcup_{\alpha\in(0,2)}\bigcup_{h\in\Gamma}\mathcal{SM}^{\alpha}_{\langle h\rangle}(\Gamma)$$

such that the support of μ generates Γ. That is, $\mu\in\mathcal{SM}_1(\Gamma)$ if it is the finite convex combinations of stable-like symmetric probability measures supported on a finite collection of subgroups of Γ, $\langle h_i\rangle$, $1\le i\le k$, with the property that $\{h_i^{\pm1}, 1\le i\le k\}$ generates Γ. □

Definition 8.3 ($\mathcal{SM}^{\alpha}(\Gamma)$, $\alpha\in(0,2)$) For each $\alpha\in(0,2)$, the subset $\mathcal{SM}^{\alpha}(\Gamma)$ of $\mathcal{SM}(\Gamma)$ is the set of all finite convex combinations μ of probability measures in

$$\bigcup_{H:\,\text{subgroup of }\Gamma}\mathcal{SM}^{\alpha}_{H}(\Gamma)$$

such that the support of μ generates Γ. □

So, any probability measure μ in $\mathcal{SM}(\Gamma)$ as the form

$$\mu=\sum_{i=1}^{k}p_i\mu_{H_i,\alpha_i}, \tag{8.1}$$

where $\alpha_i\in(0,2)$, $p_i>0$, $\sum_{i=1}^{k}p_i=1$, each H_i is a subgroup of Γ, and μ_{H_i,α_i} is a probability measure in $\mathcal{SM}^{\alpha_i}_{H_i}(\Gamma)$. In addition, $\Gamma=\langle H_1,\dots,H_k\rangle$. The typical measures in $\mathcal{SM}_1(\Gamma)$ have the more explicit form

$$\mu(g)=\sum_{1}^{k}\sum_{m\in\mathbb{Z}}\frac{p_ic_{\alpha_i}}{(1+|m|)^{1+\alpha_i}}\mathbb{1}_{\{s_i^m\}}(g),$$

where $\alpha_i\in(0,2)$, $p_i>0$, $\sum_1^k p_i=1$, the finite set $\{s_i^{\pm1}:1\le i\le k\}$ is a generating set of Γ, and $c_\alpha^{-1}=\sum_{m\in\mathbb{Z}}\frac{1}{(1+|m|)^{1+\alpha}}$. There are more measures in $\mathcal{SM}_1(\Gamma)$ because the individual component of the convex combination above do not have to be exactly $\sum_{m\in\mathbb{Z}}\frac{c_{\alpha_i}}{(1+|m|)^{1+\alpha_i}}\mathbb{1}_{\{s_i^m\}}(g)$, they only have to be $\asymp$-comparable to such a measure.

Example 8.4 On the Heisenberg group $\mathbb{H}_3(\mathbb{Z})$ viewed as the group of matrix (4.2), consider the measures

$$\mu_4((x_1,x_2,x_3))=\frac{c_\alpha}{\left(1+\sqrt{x_1^2+x_2^2+|x_3-x_1x_2/2|}\right)^{4+\alpha_4}},$$

(this is μ from Example 4.6) and

$$\mu_i((x_1,x_2,x_3))=\frac{c_\alpha\mathbb{1}_{H_i}((x_1,x_2,x_3))}{(1+|x_i|)^{1+\alpha_i}},\ i=1,2,3,$$

where $H_i = \{(x_1, x_2, x_3) : x_j = 0 \text{ if } j \neq i\}$, with $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in (0, 2)$. The measure $\mu = \frac{1}{4}\sum_{i=1}^{4} \mu_i$ is a good example of a measure in $\mathcal{SM}(\mathbb{H}_3(\mathbb{Z}))$. This is because the expression $\sqrt{x_1^2 + x_2^2 + |x_3 - x_1x_2/2|}$ is constant under taking inverse and is comparable to the word-length on $\mathbb{H}_3(\mathbb{Z})$ (e.g., on the natural minimal symmetric generating set). □

Example 8.5 On the Heisenberg group $\mathbb{H}_3(\mathbb{Z})$ viewed as the group of matrix (4.2), for $i = 1, 2$, $\alpha_i \in (0, 2)$, and $H_1 = \{(x_1, 0, x_3) : x_1, x_3 \in \mathbb{Z}\}$, $H_2 = \{(0, x_2, x_3) : x_2, x_3 \in \mathbb{Z}\}$, consider the measures

$$\mu_i((x_1, x_2, x_3)) = \frac{c_{\alpha_i}}{\left(1 + \sqrt{x_i^2 + x_3^2}\right)^{\alpha_i + 2}} \mathbb{1}_{H_i}(x_1, x_2, x_3).$$

The measure $\mu = \frac{1}{2}(\mu_1 + \mu_2)$ is another good example of a measure in $\mathcal{SM}(\mathbb{H}_3(\mathbb{Z}))$. The measure in Example 4.6 is also in $\mathcal{SM}(\mathbb{H}_3(\mathbb{Z}))$. □

8.2 Weight systems on Γ associated to measures in $\mathcal{SM}(\Gamma)$

Let $\mu \in \mathcal{SM}(\Gamma)$ be given by (8.1). From the data defining μ, we extract a long generating tuple

$$\Sigma = (\sigma_1, \ldots \sigma_\ell)$$

by listing one representative of $\{s, s^{-1}\}$ for each $s \in S_i$, $1 \leq i \leq k$, with repetition when the same s, s^{-1} belongs to more than one set S_i. Thus, we can think of each σ_j as carrying a label that tells us from which S_i it comes. Using this label we give each $\sigma_j \in \Sigma$ the positive weight $w(\sigma_j) = 1/\alpha_i$ if σ_j comes from S_i. Now, consider Σ as a finite alphabet and consider the set of all finite length formal commutators over $\Sigma \cup \Sigma^{-1}$ where Σ^{-1} is the set of formal inverse letters. We can proceed inductively. Elements of $\Sigma \cup \Sigma^{-1}$ are length 1 commutators. After formal commutators of length at most n have been defined, the formal commutators of length at most $n + 1$ are all the formal expressions of the form $[\tau, \theta]$ where τ and θ are commutators of length s and t with $s + t \leq n + 1$. Recall that each formal commutator $\sigma^{\pm 1}$ of length 1 has a weight $w(\sigma^{\pm 1}) = 1/\alpha_i$ if σ comes originally from S_i. Extend the weight function w to all formal commutators by setting $w([\tau, \theta]) = w(\tau) + w(\theta)$.

A priori, there are countably many formal commutators but because Γ is nilpotent and we will ultimately consider only the formal commutators that are not trivial when evaluated in Γ, we only have to deal with finitely many formal commutators, whose lengths are at most the nilpotent class of Γ. We now use weighted formal commutators to define a non-increasing equence of subgroups of Γ. Recall that by convention and abuse of notation, each letter σ in Σ is also a group element in Γ. The following definition is essentially from [55] where further details can be found. See Definition 1.4 and Proposition 1.5 in [55].

Definition 8.6 For any $s \geq 0$, let $\Gamma_s^{\Sigma,w}$ be the subgroup of Γ generated by the elements in Γ of all the formal commutators over the alphabet Σ with weight at least s. By construction $\Gamma_t^{\Sigma,w} \subseteq \Gamma_s^{\Sigma,w}$ if $s \leq t$. Also, $[\Gamma_s^{\Sigma,w}, \Gamma_t^{\Sigma,w}] \subseteq \Gamma_{s+t}^{\Sigma,w}$. □

Definition 8.7 There is a greatest t such that $\Gamma_t^{\Sigma,w} = \Gamma$, call it w_1. By induction, having defined w_j, define w_{j+1} to be the largest $t \in (w_j, \infty]$ such that $\Gamma_s^{\Sigma,w} = \Gamma_t^{\Sigma,w}$ for all $w_j < s \leq t$. This defines a finite strictly increasing sequence

$$w_1 < w_2 < \cdots < w_j < w_{j+1} < \cdots < w_{j_*+1} = \infty$$

such that

$$\Gamma_{w_{j+1}}^{\Sigma,w} \subsetneq \Gamma_{w_j}^{\Sigma,w}, \quad \Gamma_s^{\Sigma,w} = \Gamma_{w_j}^{\Sigma,w} \text{ for } s \in (w_{j-1}, w_j], \text{ and } \Gamma_s^{\Sigma,w} = \{e\} \text{ for } s > w_{j_*}.$$

By construction $\left[\Gamma, \Gamma_{w_j}^{\Sigma,w}\right] \subset \Gamma_{w_{j+1}}^{\Sigma,w}$. Call A_{w_j} the abelian group

$$A_{w_j} = \Gamma_{w_j}^{\Sigma,w} / \Gamma_{w_{j+1}}^{\Sigma,w}, \quad 1 \leq j \leq j_*.$$

Definition 8.8 Set $\gamma_0(\Sigma, w) = \sum_1^{j_*} w_j \mathrm{Rank}(A_{w_j})$, where $\mathrm{Rank}(A)$ denotes the torsion free rank of the finitely generated abelian group A. □

That the construction described above and the definition of the positive real $\gamma_0(\Sigma, w)$ is relevant to the study of random walks driven by measures in $\mathcal{SM}(\Gamma)$ is apparent from the following theorem from [55, 20].

Theorem 8.9 ([55, 20]) Let Γ be a finitely generated nilpotent group. For any probability measure μ in $\mathcal{SM}(\Gamma)$ with associated data (Σ, w) as above, there are constants $c(\mu)$ and $C(\mu)$ such that, for all n,

$$c(\mu) n^{-\gamma_0(\Sigma,w)} \leq \mu^{(n)}(e) \leq C(\mu) n^{-\gamma_0(\Sigma,w)}.$$

8.3 Quasi-norms on Γ associated with elements of $\mathcal{SM}(\Gamma)$

The previous section associates to any measure $\mu \in \mathcal{SM}(\Gamma)$ a weight system built on the ℓ-tuple of group elements $\Sigma = (\sigma_1, \dots, \sigma_\ell)$ obtained by listing consecutively with possible repetitions all the elements of the sets S_i, $1 \leq i \leq k$, and the attached weight $w(\sigma) = 1/\alpha_i$ if σ comes from S_i. Recall that in this construction, we view Σ as an abstract alphabet. This data allows us to construct a quasi-norm on the countable group Γ based on the writing of any element g of Γ as a word over the alphabet $\Sigma \cup \Sigma^{-1}$. For any finite word $\omega \in \cup_{m=0}^{\infty} (\Sigma \cup \Sigma^{-1})^m$, set

$$\deg_\sigma(\omega) = \text{number of times the letters } \sigma, \sigma^{-1} \text{ are used in } \omega.$$

The following definition is from [55, 20].

Definition 8.10 Given Γ, $\Sigma = (\sigma_1, \ldots, \sigma_\ell)$ and weight w as above, for each element $g \in \Gamma$, set

$$\|g\|_{\Sigma,w} = \inf\left\{\max_{\sigma\in\Sigma}\{(\deg_\sigma(\omega))^{1/w(\sigma)}\} : \omega \in \cup_{m=0}^{\infty}(\Sigma \cup \Sigma^{-1})^m,\ g = \omega \text{ in } \Gamma\right\}.$$

By convention, $\|e\|_{\Sigma,w} = 0$. □

Remark 8.11 When $w(\sigma) = w_0$ for all $\sigma \in \Sigma$, the quasi-norm $\|\cdot\|_{\Sigma,w}$ satisfies

$$\frac{1}{\ell}|g|_\Sigma \le \|g\|_{\Sigma,w}^{w_0} \le |g|_\Sigma \quad \text{for every } g \in \Gamma,$$

where $|\cdot|_\Sigma$ denotes the usual word length of the finite symmetric generating set $\Sigma \cup \Sigma^{-1} \subset \Gamma$. □

Remark 8.12 It may be worth noting that, in general, it is hard to compute or estimate $\|g\|_{\Sigma,w}$ for a given $g \in \Gamma$. The reference [55] gives many results in this direction and these results will be useful in the sequel. This is related to the use of coordinate systems in so far as the question of estimating $g \in \Gamma$ becomes a precise question only when g is given in terms of some parameter set, i.e., some sort of (possibly partial) coordinate system, see [55, Theorem 2.10]. To help the reader understand this comment, we suggest the following question: given a fixed $g \in \Gamma$, what is the behavior of $\|g^m\|_{\Sigma,w}$ as a function of m? See [55, Proposition 2.17].

Chapter 9
Adapted approximate group dilations

Abstract Given a probability measure μ in $\mathcal{SM}(\Gamma)$ and its associated geometry, this chapter discusses the construction of adapted coordinate systems using the group exponential map (exponential coordinates of the first and second type). Useful properties of these coordinate systems and their relations with the geometry previously associated with the measure μ are discussed. Each of these coordinate systems leads to the definition of an approximate group dilation adapted to the measure μ. In the next chapter, these constructions are used to show that every $\mu \in \mathcal{SM}(\Gamma)$ satisfies all the conditions except condition **(A)** of the limit theorems of Chapters 5 and 6.

9.1 Searching for adapted dilations

The goal of this chapter is to associate to each probability measure μ in $\mathcal{SM}(\Gamma)$ an adapted approximate dilation structure. This includes making the choice of an appropriate polynomial coordinate system for the simply connected nilpotent Lie group $G = (\mathbb{R}^d, \cdot)$ in which Γ embeds as a co-compact discrete subgroup. The given measure μ determines uniquely certain features of the appropriate coordinate systems and associated approximate group dilations but not all. Among the feature that are determined uniquely (in this case, up to an arbitrary multiplicative positive constant) is a vector of non-decreasing weight values b_j, $1 \le j \le d$, so that, in the chosen coordinate system $u = (u_i)_1^d \in \mathbb{R}^d$ for G, the appropriate approximate dilation structure is of the form $\delta_t(u) = (t^{b_i} u_i)_1^d$. Among the exponential coordinates of the first and second kind, the group structure $G_\bullet = (\mathbb{R}^d, \bullet)$ defined by

$$u \bullet u' = \lim_{t \to \infty} \delta_{1/t}(\delta_t(u) \cdot \delta_t(u'))$$

(understood up to isomorphisms) depends only on Γ and μ and not on the particular choice of a coordinate system. An interesting question is if this remains true beyond these exponential coordinate systems.

In the next sections, we describe two key constructions: the construction of adapted exponential coordinates of the first kind and that of adapted coordinate of the second kind. The essential difference between the two constructions is that, in the discussion of exponential coordinates of the first kind, we assume that the simply connected nilpotent Lie group G in which Γ sits as a co-compact subgroup is already given to us together with its Lie algebra and canonical exponential map. All we need to construct is an adapted linear basis of this Lie algebra based on the nature of the measure μ. In the case of exponential coordinates of the second kind, we start from scratch with only the finitely generated torsion free nilpotent group Γ carrying the measure μ and, following Malcev and Hall, we construct a "discrete coordinate system" for Γ which is adapted to μ and, in turn, "generates" for us the simply connected Lie group G and its adapted exponential coordinates of the second type. It is only a posteriori (and with some work) that one can check that certain features of these two constructions are identical.

9.2 Exponential coordinates of the first kind

This section focuses on the situation when the torsion free finitely generated group Γ is given to us as a co-compact discrete subgroup of a simply connected nilpotent Lie group G with Lie algebra $\mathfrak{g} = (\mathbb{R}^d, [\cdot,\cdot])$ and the group G is given in the (canonical) exponential coordinates of the first kind. This identifies the group G with $(\mathbb{R}^d, \cdot)$ where the product $\cdot$ is given by the famous Campbell-Hausdorff formula (3.6).

In the next definition, we are given a probability measure $\mu \in \mathcal{SM}(\Gamma)$ and the associated data Σ, w as in Section 8.2 and we transfer the weight system to the Lie algebra $\mathfrak{g}$. Observe that the tuple (use the same ordering as for Σ)

$$\Sigma_{\mathfrak{g}} = (\varsigma_i \in \mathfrak{g} : \exp(\varsigma_i) = \sigma_i \in \Sigma) = (\varsigma_1, \dots, \varsigma_\ell)$$

must be an algebraically generating set for $\mathfrak{g}$ in that this set together with all iterated brackets of its elements generates $\mathfrak{g}$ linearly. Indeed, because the exponential map is a global diffeomorphism between $\mathfrak{g}$ and G, if $\Sigma_{\mathfrak{g}}$ did not generate $\mathfrak{g}$, Γ would be contained in a proper closed connected Lie subgroup of G. This would contradict the fact that Γ is co-compact in G.

We now trivially transfer the weight function $w : \Sigma \to (0,\infty)$ to a function defined on $\Sigma_{\mathfrak{g}}$ by setting $w^{\mathfrak{g}}(\varsigma) = w(\sigma)$ if $\sigma = \exp(\varsigma)$. This leads to the definition of a weight system $w^{\mathfrak{g}}$ on the formal (Lie) commutators of the ς's in a way that is formally analogous to what we did on Γ.

Definition 9.1 Let $\mathfrak{g}_s^{\Sigma,w}$ be the Lie sub-algebra of $\mathfrak{g}$ generated by the evaluation in $\mathfrak{g}$ of all formal commutators of the $\varsigma \in \Sigma_{\mathfrak{g}}$ whose weight is at least s. By construction,

$$\mathfrak{g}_t^{\Sigma,w} \subseteq \mathfrak{g}_s^{\Sigma,w} \text{ if } s \le t$$

and

$$[\mathfrak{g}_s^{\Sigma,w}, \mathfrak{g}_t^{\Sigma,w}] \subseteq \mathfrak{g}_{s+t}^{\Sigma,w}.$$

Definition 9.2 There is a greatest t such that $\mathfrak{g}_t^{\Sigma,w} = \mathfrak{g}$, call it $w_1^{\mathfrak{g}}$. By induction, having defined $w_j^{\mathfrak{g}}$, there is a greatest $t \in (w_j^{\mathfrak{g}}, \infty]$ such that $\mathfrak{g}_s^{\Sigma,w} = \mathfrak{g}_t^{\Sigma,w}$ for all $s \in (w_j^{\mathfrak{g}}, t]$. Call it $w_{j+1}^{\mathfrak{g}}$. Let $j_\star^{\mathfrak{g}}$ be the largest integer j such that $w_j^{\mathfrak{g}} < \infty$ so that $w_{j_\star^{\mathfrak{g}}+1}^{\mathfrak{g}} = \infty$ and $\mathfrak{g}_{w_{j_\star^{\mathfrak{g}}+1}^{\mathfrak{g}}}^{\Sigma,w} = \{0\}$. This defines a finite strictly decreasing sequence of sub-Lie algebras

$$\mathfrak{g} = \mathfrak{g}^{\Sigma,w}_{w^{\mathfrak{g}}_1} \supset \cdots \supset \mathfrak{g}^{\Sigma,w}_{w^{\mathfrak{g}}_{j^{\mathfrak{g}}_\star}} \supset \{0\}$$

with the property that

$$[\mathfrak{g}, \mathfrak{g}^{\Sigma,w}_{w^{\mathfrak{g}}_j}] \subseteq \mathfrak{g}^{\Sigma,w}_{w^{\mathfrak{g}}_{j+1}}, \quad j = 1, \dots, j^{\mathfrak{g}}_\star.$$

Definition 9.3 (Adapted direct sum decomposition) We say that a direct sum decomposition of $\mathfrak{g}$,

$$\mathfrak{g} = \oplus_1^{j^{\mathfrak{g}}_\star} \mathfrak{n}_j,$$

is adapted to (Σ, w) if, for all $j \in \{1, \dots, j^{\mathfrak{g}}_\star\}$,

$$\mathfrak{g}^{\Sigma,w}_{w^{\mathfrak{g}}_j} = \oplus_{\ell=j}^{j^{\mathfrak{g}}_\star} \mathfrak{n}_\ell.$$

Remark 9.4 To construct an adapted direct sum decomposition, start from the top and set $\mathfrak{n}_{j^{\mathfrak{g}}_\star} = \mathfrak{g}^{\Sigma,w}_{w^{\mathfrak{g}}_{j^{\mathfrak{g}}_\star}}$. By descending induction, having constructed $\mathfrak{n}_j, \dots, \mathfrak{n}_{j^{\mathfrak{g}}_\star}$ so that $\mathfrak{g}^{\Sigma,w}_{w^{\mathfrak{g}}_j} = \oplus_{\ell=j}^{j^{\mathfrak{g}}_\star} \mathfrak{n}_\ell$, pick a linear complement of $\mathfrak{g}^{\Sigma,w}_{w^{\mathfrak{g}}_j}$ inside $\mathfrak{g}^{\Sigma,w}_{w^{\mathfrak{g}}_{j-1}}$ and call it $\mathfrak{n}_{j-1}$. □

Definition 9.5 (Approximate Lie dilation structure (first kind)) Given a direct sum decomposition that is adapted to (Σ, w), consider the group of invertible linear maps

$$\delta_t : \mathfrak{g} \to \mathfrak{g}, \quad t > 0,$$

define by

$$\delta_t(v) = t^{\mathfrak{w}^{\mathfrak{g}}_j} v \quad \text{for all } v \in \mathfrak{n}_j, 1 \le j \le j^{\mathfrak{g}}_\star.$$

Let $\boldsymbol{\varepsilon} = (\varepsilon_i)_1^d$ be a linear basis of $\mathbb{R}^d$ adapted to the direct sum $\mathfrak{g} = \oplus_1^{j^{\mathfrak{g}}_\star} \mathfrak{n}_j$, let $u = (u_i)_1^d$ be the corresponding coordinate system, and let

$$b_i = \mathfrak{w}^{\mathfrak{g}}_j \quad \text{if } \varepsilon_i \in \mathfrak{n}_j$$

so that

$$\delta_t(u) = (t^{b_i} u_i)_1^d.$$

We shall see below in Proposition 9.12 and Corollary 9.13 that the important quantity $\gamma_0(\Sigma, w)$ is given in terms of the sequences $(b_i)_1^d$ and $(\mathfrak{w}^{\mathfrak{g}}_j)_1^{j^{\mathfrak{g}}_\star}$ by

$$\gamma_0(\Sigma, w) = \sum_1^d b_i = \sum_1^{j^{\mathfrak{g}}_\star} \mathfrak{w}^{\mathfrak{g}}_j \dim(\mathfrak{n}_j).$$

Proposition 9.6 The maps $(\delta_t)_{t>0}$ defined above form an approximate Lie dilation structure on $\mathfrak{g}$. In any exponential coordinate system of the first kind adapted to the direct sum decomposition $\mathfrak{g} = \oplus_1^{j_\star^{\mathfrak{g}}} \mathfrak{n}_j$, $(\delta_t)_{t>0}$ is a straight approximate group dilation structure on G.

Proof By linearity, it suffices to prove that for any $v_i \in \mathfrak{n}_{j_i}$, $i = 1, 2$,

$$\delta_t^{-1}([\delta_t(v_1), \delta_t(v_2)])$$

has a limit when t tends to infinity. By construction, $\delta_t(v_i) = t^{w^{\mathfrak{g}}_{j_i}} v_i$ and $[v_1, v_2] \in \mathfrak{g}^{\Sigma,w}_{w^{\mathfrak{g}}_N} = \oplus_{\ell \geq N} \mathfrak{n}_\ell$ where $w^{\mathfrak{g}}_N \geq w^{\mathfrak{g}}_{j_1} + w^{\mathfrak{g}}_{j_2}$, namely,

$$[v_1, v_2] = \sum_{\ell=N}^{j_\star^{\mathfrak{g}}} f_\ell, \quad f_\ell \in \mathfrak{n}_\ell.$$

It follows that

$$\delta_t^{-1}([\delta_t(v_1), \delta_t(v_2)]) = \sum_{\ell=N}^{j_\star^{\mathfrak{g}}} t^{w^{\mathfrak{g}}_{i_1} + w^{\mathfrak{g}}_{i_2} - w^{\mathfrak{g}}_\ell} f_\ell.$$

The limit of this expression when t tends to infinity exists because $w^{\mathfrak{g}}_{i_1} + w^{\mathfrak{g}}_{i_2} \leq w^{\mathfrak{g}}_N \leq w^{\mathfrak{g}}_\ell$ for all $\ell \geq N$. If $w^{\mathfrak{g}}_N > w^{\mathfrak{g}}_{i_1} + w^{\mathfrak{g}}_{i_2}$, the limit is 0. If $w^{\mathfrak{g}}_N = w^{\mathfrak{g}}_{i_1} + w^{\mathfrak{g}}_{i_2}$, the limit is f_N. □

9.3 Building adapted exponential coordinates of the second kinds from Γ and μ

In this section, we start with the given discrete torsion free nilpotent group Γ (described, perhaps, by generators and relations, or as a subgroup of a bigger group) and we explain how to construct the Lie group G using well-known ideas related to exponential coordinate systems of the second kind. This is done in [46, 35] and we refer the reader to the treatment in [21, Theorem 4.9, Section 4.3].

Hall-Malcev coordinates

Theorem 4.9 of [21] asserts that, for any finitely generated torsion free nilpotent group Γ, any descending central series (this means that Γ_i/Γ_{i+1} is central in Γ/Γ_{i+1} for each $1 \leq i \leq n$)

$$\Gamma = \Gamma_1 \triangleright \Gamma_2 \triangleright \cdots \triangleright \Gamma_n \triangleright \Gamma_{n+1} = \{e\}$$

with Γ_i/Γ_{i+1} infinite cyclic, and any sequence of elements $\tau_i \in G$ such that $\Gamma_i = \langle \Gamma_{i+1}, \tau_i \rangle$, each element $\gamma \in \Gamma$ can be written uniquely

$$\gamma = \tau_1^{u_1} \cdot \tau_2^{u_2} \cdot \dots \cdot \tau_n^{u_n}, \quad u_1, u_2, \dots, u_n \in \mathbb{Z}.$$

Moreover, for any $k \in \mathbb{Z}$ and any $\gamma' = \tau_1^{u'_1} \cdot \tau_2^{u'_2} \cdot \dots \cdot \tau_n^{u'_n}$,

$$\gamma^k = \tau_1^{g_1(u,k)} \cdot \tau_2^{g_2(u,k)} \cdot \dots \cdot \tau_n^{g_n(u,k)}$$

and

$$\gamma \cdot \gamma' = \tau_1^{f_1(u,u')} \cdot \tau_2^{f_2(u,u')} \cdot \dots \cdot \tau_n^{f_n(u,u')},$$

where $u = (u_1, \dots, u_n)$, $u' = (u'_1, \dots, u'_n)$, and f_i, g_i, $1 \le i \le n$ are polynomials with rational coefficients in there respective variables.

Furthermore ([21, Theorems 4.11-4.12]), by interpreting these coordinates in $\mathbb{R}^n$ instead of $\mathbb{Z}^n$, one obtains a simply connected nilpotent Lie group of which Γ is a discrete co-compact subgroup.

For our present purpose, the task is to produce a descending central series

$$\Gamma = \Gamma_1 \rhd \Gamma_2 \rhd \dots \rhd \Gamma_n \rhd \Gamma_{n+1} = \{e\}$$

with Γ_i/Γ_{i+1} infinite cyclic, which is adapted to the measure μ. Using the sequence $\Gamma^{\Sigma,w}_{w_j}$, $1 \le j \le j_*$ is a good first guess. If each quotient $A_{w_j} = \Gamma^{\Sigma,w}_{w_j}/\Gamma^{\Sigma,w}_{w_{j+1}}$ is free abelian (i.e., has no torsion), then we can produce a descending central series

$$\Gamma = \Gamma_1 \rhd \Gamma_2 \rhd \dots \rhd \Gamma_n \rhd \Gamma_{n+1} = \{e\}$$

which refines the sequence $\Gamma^{\Sigma,w}_{w_j}$, $1 \le j \le j_*$, and has Γ_i/Γ_{i+1} infinite cyclic. In addition, we can find a sequence of elements $\tau_i \in \Gamma$, each of which is a commutator of the elements in Σ, such that $\Gamma_i = \langle \Gamma_{i+1}, \tau_i \rangle$ and such that

$$w(\tau_i) = w_j \text{ if and only if } \Gamma^{\Sigma,w}_{w_{j+1}} \supset \Gamma_i \supseteq \Gamma^{\Sigma,w}_{w_j}.$$

The problem we face is that it is NOT always the case that the groups A_{w_j} are torsion free (even in the simplest of all cases when $\Gamma = \mathbb{Z}$!).

Modified weight system on Γ

Given a measure $\mu \in \mathcal{SM}(\Gamma)$ as in (8.1), we defined in Section 8.2 a generating set $\Sigma = (\sigma_1, \dots, \sigma_\ell)$ and a weight system w on formal commutators which generates the descending central sequence of subgroups $\Gamma^{\Sigma,w}_{w_j}$, $1 \le j \le j_*$.

Consider the finite set of all formal commutators over the alphabet Σ which are not trivial in Γ and organize that finite set as a long tuple $\Sigma_{\text{com}} = (c_1, \dots, c_L)$. Let $\underline{\Sigma}_{\text{com}} = (\underline{c}_1, \dots, \underline{c}_L)$ be the evaluation of Σ_{com} in Γ.

Let us introduce a (modified) weight function, $\underline{w}$, on $\underline{\Sigma}_{\text{com}}$ by setting, for each $\underline{c}$ appearing in the tuple $\underline{\Sigma}_{\text{com}}$,

$$\underline{w}(\underline{c}) = \max\left\{w_j : \exists\, m \in \mathbb{N},\ \underline{c}^m \in \Gamma^{\Sigma,w}_{w_j}, 1 \le j \le j_*\right\}.$$

For commutators whose evaluation in Γ is trivial, we can set $\underline{w}(c) = \infty$. Following [55, Section 2.2], we set

$$\mathrm{core}(\Sigma, w) = \{\sigma_i : \underline{w}(\sigma_i) = w(\sigma_i), 1 \le i \le \ell\}.$$

The function $\underline{w}$ is no less than w and has the property that, if $c = [c_1, c_2]$ is nontrivial in Γ then

$$\underline{w}(\underline{c}) \ge \underline{w}(\underline{c}_1) + \underline{w}(\underline{c}_2).$$

It follows that the induced weight of a formal commutator $\mathbf{c}$ over the alphabet $\underline{\Sigma}_{\mathrm{com}}$ whose evaluation in Γ is not trivial is actually equal to the $\underline{w}$ weight of the same commutator view as an element of $\underline{\Sigma}_{\mathrm{com}}$. Moreover, $\mathrm{core}(\Sigma_{\mathrm{com}}, \underline{w}) = \Sigma_{\mathrm{com}}$.

Definition 9.7 For any $s \ge 0$, let Γ^{com}_s be the subgroup of Γ generated by the values in Γ of all the formal commutators over the alphabet Σ with $\underline{w}$-weight at least s. By construction $\Gamma^{\mathrm{com}}_t \subseteq \Gamma^{\mathrm{com}}_s$ if $s \le t$. Also, $[\Gamma^{\mathrm{com}}_s, \Gamma^{\mathrm{com}}_t] \subseteq \Gamma^{\mathrm{com}}_{s+t}$. □

Definition 9.8 There is a greatest t such that $\Gamma^{\mathrm{com}}_t = \Gamma$, call it $\underline{w}_1$. By induction, having defined $\underline{w}_j$, define $\underline{w}_{j+1}$ to be the largest $t \in (\underline{w}_j, \infty]$ such that $\Gamma^{\mathrm{com}}_t = \Gamma^{\mathrm{com}}_s$ for all $\underline{w}_j < s \le t$. This defines a finite strictly increasing sequence

$$0 < \underline{w}_1 < \underline{w}_2 < \cdots < \underline{w}_j < \underline{w}_{j+1} < \cdots < \underline{w}_{j^{\mathrm{com}}_*+1} = \infty$$

such that

$$\Gamma^{\mathrm{com}}_{\underline{w}_{j+1}} \subset \Gamma^{\mathrm{com}}_{\underline{w}_j},\quad \Gamma^{\mathrm{com}}_s = \Gamma^{\mathrm{com}}_{\underline{w}_j} \text{ for } s \in (\underline{w}_{j-1}, \underline{w}_j],\quad \Gamma^{\mathrm{com}}_s = \{e\} \text{ for } s > \underline{w}_{j^{\mathrm{com}}_*}.$$

By construction $[\Gamma, \Gamma^{\mathrm{com}}_{\underline{w}_j}] \subset \Gamma^{\mathrm{com}}_{\underline{w}_{j+1}}$. Call $A^{\mathrm{com}}_{\underline{w}_j}$ the abelian group

$$A^{\mathrm{com}}_{\underline{w}_j} = \Gamma^{\mathrm{com}}_{\underline{w}_j} / \Gamma^{\mathrm{com}}_{\underline{w}_{j+1}}, \quad 1 \le j \le j^{\mathrm{com}}_*.$$

The following lemma follows immediately from the construction outlined above.

Lemma 9.9 The groups $A^{\mathrm{com}}_{\underline{w}_j}$, $1 \le j \le j^{\mathrm{com}}_*$ are free abelian and each is generated by a finite subset of the commutators $\underline{c} \in \Sigma_{\mathrm{com}}$. Consequently, there exists a descending central series

$$\Gamma = \Gamma_1 \rhd \Gamma_2 \rhd \cdots \rhd \Gamma_d \rhd \Gamma_{d+1} = \{e\}$$

refining the descending central series

$$\Gamma = \Gamma^{\mathrm{com}}_{\underline{w}_1} \rhd \Gamma^{\mathrm{com}}_{\underline{w}_2} \rhd \cdots \rhd \Gamma^{\mathrm{com}}_{\underline{w}_{j^{\mathrm{com}}_*}} \rhd \Gamma = \Gamma^{\mathrm{com}}_{\underline{w}_{j^{\mathrm{com}}_*+1}} = \{e\}$$

and a sequence $\tau_i = \underline{c}_{\ell_i}$, $1 \le j \le d$, in Σ_{com} such that Γ_i / Γ_{i+1} is an infinite cyclic, $\Gamma_i = \langle \tau_i, \Gamma_{i+1} \rangle$, $1 \le i \le d$, and

$$\Gamma^{\mathrm{com}}_{\underline{w}_j} = \langle \tau_i : \underline{w}(\tau_i) \ge \underline{w}_j \rangle.$$

Because of this lemma, it is clear that [21, Theorems 4.9, 4.11,4.12] apply and provide a set of coordinate of the second kind

$$\Gamma = \left\{\gamma = \tau_1^{u_1} \cdot \tau_2^{u_2} \cdot \dots \cdot \tau_d^{u_d},\ u_1, u_2, \dots, u_d \in \mathbb{Z}\right\}$$

for Γ, as well as an embedding of Γ as a co-compact discrete subgroup a simply connected Lie group G

$$\begin{aligned} G =& \{g = \tau_1^{x_1} \cdot \tau_2^{x_2} \cdot \dots \cdot \tau_d^{x_d},\ x_1, x_2, \dots, x_d \in \mathbb{R}\} \\ =& \{g = \exp(x_1\zeta_1) \cdot \exp(x_2\zeta_2) \cdot \dots \cdot \exp(x_d\zeta_d),\ x_1, x_2, \dots, x_d \in \mathbb{R}\}, \end{aligned}$$

where $\zeta_i = \log \tau_i \in \mathfrak{g}$.

Definition 9.10 (Approximate group dilation structure (second kind)) In the exponential coordinate system of the second kind $(x_i)_1^d$ introduced above, consider the group of straight dilations

$$\delta_t : \mathbb{R}^d \to \mathbb{R}^d, \quad t > 0, \quad x \mapsto \delta_t(x) = (t^{\underline{w}_i} x_i)_1^d.$$

Proposition 9.11 The maps $(\delta_t)_{t>0}$ defined above form an approximate group dilation structure on $G = (\mathbb{R}^d, \cdot)$. □

By the same token, we obtain an associated coordinate system of the first kind

$$(y_1, \dots, y_d) \mapsto \exp\left(\sum_1^d y_i \zeta_i\right),$$

which is compatible with the weight $w^{\mathfrak{g}}$ introduced earlier and such that $b_i = w^{\mathfrak{g}}(\zeta_i) = \underline{w}(\tau_i)$, $1 \le i \le d$. The straight dilation groups we introduced in these two distinct coordinate systems have the same exponents b_i in their respective coordinate systems. Viewed as maps from G to G, they are clearly different in general even so we use the same notation δ_t in both cases; see Example 3.15 for such an example where the matrix coordinate system is an exponential coordinate system of second kind. This is because there are really no good reasons to consider both coordinate systems at the same time, except to understand that this parallel constructions yield compatible results at the end.

9.4 Relations between the filtrations associated with w, $w^{\mathfrak{g}}$ and $\underline{w}$

Although there are great similitudes in the construction of the (discrete group) filtrations $\Gamma^{\Sigma,w}_{w_j}$, $1 \le j \le j_*$, and $\Gamma^{\text{com}}_{\underline{w}_j}$, $1 \le j \le j_*^{\text{com}}$, of the group Γ, and the (Lie algebra) filtration $\mathfrak{g}^{\Sigma,w}_{\mathfrak{w}_j^{\mathfrak{g}}}$, $1 \le j \le j_\star^{\mathfrak{g}}$ of $\mathfrak{g}$, there are also differences.

We start with a comparison of the coordinates of the first and second kind in this context. It is not hard to see that the definitions of the sequences $w_j^{\mathfrak{g}}$, $1 \le j \le j_\star^{\mathfrak{g}}$, and

$\underline{w}_i$, $1 \le i \le j_*^{\mathrm{com}}$, and the above remark concerning the relations between group and Lie algebra commutators, imply that these sequences of weights are actually equal, that is,

$$j_\star^{\mathfrak{g}} = j_*^{\mathrm{com}} \quad \text{and} \quad w_j^{\mathfrak{g}} = \underline{w}_j, \quad 1 \le j \le j_\star^{\mathfrak{g}}. \tag{9.1}$$

More generally, each (discrete group) formal commutator $\boldsymbol{\tau}$ on the alphabet $\Sigma = (\sigma_1, \dots, \sigma_\ell)$ corresponds in an obvious formal way to a formal Lie commutator $\boldsymbol{\theta}$ on the alphabet $\Sigma^{\mathfrak{g}} = (\zeta_1, \dots, \zeta_\ell)$ in such a way that the Campbell-Hausdorff formula provides a formal equality

$$\boldsymbol{\tau} = \exp(\boldsymbol{\zeta}) = \exp(\boldsymbol{\theta} + \mathbf{R}_{\boldsymbol{\tau}}), \tag{9.2}$$

where $\mathbf{R}_{\boldsymbol{\tau}}$ is a formal series of Lie commutators with $w^{\mathfrak{g}}$-weights strictly larger than $\underline{w}(\boldsymbol{\tau}) = w^{\mathfrak{g}}(\boldsymbol{\theta})$. The concrete meaning of this formal identity in the present context is that it is an equality when evaluated over any pair $\Gamma \subset G$ where G is a simply connected nilpotent Lie group with algebra $\mathfrak{g}$, with the formal series $\mathbf{R}_{\boldsymbol{\sigma}}$ reducing to a finite sum. Obviously, the evaluation θ of $\boldsymbol{\theta}$ in $\mathfrak{g}$ belongs to $\mathfrak{g}^{\Sigma,w}_{w^{\mathfrak{g}}(\boldsymbol{\theta})}$. It follows that the evaluation $\zeta = \theta + R_\sigma$ of $\boldsymbol{\zeta}$ in $\mathfrak{g}$ also belongs to $\mathfrak{g}^{\Sigma,w}_{w^{\mathfrak{g}}(\boldsymbol{\theta})}$.

Two choices of exponential coordinate of the first kind

From the discussion above, it becomes clear that there are at least two very natural exponential coordinate systems of the first kind associated with the sequence $(\tau_i)_1^d$ of elements of Γ given by Lemma 9.9.

Choice 1: Lie commutators

Each τ_i is a commutator built on $\Sigma = (\sigma_i)_1^\ell$. Let θ_i be the Lie commutator over $\Sigma^{\mathfrak{g}} = (\varsigma_i)_1^\ell$ that corresponds formally to τ_i. Here $\sigma_i = \exp(\varsigma_i)$ as before and the last sentence means that $\theta = [\varsigma, \varsigma']$ if $\tau = [\sigma, \sigma']$ with $\sigma = \exp(\varsigma), \sigma' = \exp(\varsigma')$. By (9.2) $\underline{w}(\tau_i) = w^{\mathfrak{g}}(\theta_i)$, and the final subsequence of $(\theta_i)_1^n$ corresponding to those i such that $\underline{w}(\tau_i) \ge w_j^{\mathfrak{g}} = \underline{w}_j$ is a linear basis of $\mathfrak{g}^{\Sigma,w}_{w_j^{\mathfrak{g}}}$. In particular,

$$\underline{\mathfrak{n}}_j = \left\{ \zeta \in \mathfrak{g} : \zeta = \sum_{i : \underline{w}(\tau_i) = w_j^{\mathfrak{g}}} z_i \theta_i,\ z_i \in \mathbb{R} \right\}, \quad 1 \le j \le j_\star^{\mathfrak{g}}$$

provides an adapted direct sum decomposition of $\mathfrak{g}$ in the sense of Definition 9.3.

Choice 2: Logarithms of group commutators

Each τ_i can be written uniquely as $\tau_i = \exp(\zeta_i)$, where ζ_i and θ_i are related via (9.2). It follows that the final subsequence of $(\zeta_i)_1^n$ corresponding to those i such that $\underline{w}(\tau_i) \geq w_j^{\mathfrak{g}} = \underline{w}_j$ is (also) a linear basis of $\mathfrak{g}^{\Sigma,w}_{w_j^{\mathfrak{g}}}$. In particular,

$$\underline{\mathfrak{n}}'_j = \left\{ \zeta \in \mathfrak{g} : \zeta = \sum_{i:\underline{w}(\tau_i)=w_j^{\mathfrak{g}}} z_i \zeta_i,\ z_i \in \mathbb{R} \right\}, \quad 1 \leq j \leq j_\star^{\mathfrak{g}}$$

provides an adapted direct sum decomposition of $\mathfrak{g}$ in the sense of Definition 9.3.

From the description of these two related coordinate systems, it follows that, for each $j \in \{1, \ldots, j_*^{\text{com}} = j_\star^{\mathfrak{g}}\}$, the group $\Gamma^{\text{com}}_{\underline{w}_j}$ is a co-compact discrete subgroup of the Lie group $\exp(\mathfrak{g}^{\Sigma,w}_{w_j^{\mathfrak{g}}})$ (recall that $\underline{w}_j = w_j^{\mathfrak{g}}$). Note that, by definition, any two exponential coordinate systems are always related by a linear change of basis in $\mathfrak{g}$. In the present case, these linear changes of coordinates have an obvious triangular form with unit diagonal and they respect the increasing filtration $\mathfrak{g}^{\Sigma,w}_{w_j^{\mathfrak{g}}}$, $1 \leq j \leq j_\star^{\mathfrak{g}}$.

The following proposition records the relations between the objects related to the original weight system w on Γ and those related to the Lie algebra weight $w^{\mathfrak{g}}$. The proof follows classical arguments developed in [46], see also [30, Appendix] and [55, 21]. It is omitted.

Proposition 9.12 The finite sequence of weight-values $w_j^{\mathfrak{g}}$, $1 \leq j \leq j^{\mathfrak{g}}_{j_\star}$, is a subsequence of the increasing finite sequence of weight-values w_j, $1 \leq j \leq j_*$, and $w^{\mathfrak{g}}_{j_\star^{\mathfrak{g}}} = w_{j_*}$. If $w_{i-1}^{\mathfrak{g}} < w_j \leq w_i^{\mathfrak{g}}$ for some $1 \leq i \leq j \leq j_*$, then

$$\Gamma^{\Sigma,w}_{w_j} \subset \exp(\mathfrak{g}^{\Sigma,w}_{\mathfrak{w}_i^{\mathfrak{g}}})$$

and the quotient

$$\exp(\mathfrak{g}^{\Sigma,w}_{\mathfrak{w}_i^{\mathfrak{g}}})/\Gamma^{\Sigma,w}_{w_j}$$

is compact. If $j \in \{1, \ldots, j_*\}$ is such that the value w_j does not appear in $(w_i^{\mathfrak{g}})_1^{j_\star^{\mathfrak{g}}}$, then

$$\Gamma^{\Sigma,w}_{w_{j+1}}/\Gamma^{\Sigma,w}_{w_j} \text{ is a finite abelian group.}$$

The following is an immediate corollary.

Corollary 9.13 $\gamma_0(\Sigma, w) = \sum_1^{j_*} w_j \mathrm{Rank}\big(\Gamma^{\Sigma,w}_{w_j}/\Gamma^{\Sigma,w}_{w_{j+1}}\big) = \sum_1^{j_\star^{\mathfrak{g}}} \mathfrak{w}_j^{\mathfrak{g}} \dim(\mathfrak{n}_j)$. □

An associated exponential coordinate system of the second kind

By the Hall-Malcev construction reviewed in Section 9.3, the sequence $(\tau_i)_1^d$ of elements of Γ given by Lemma 9.9 and the sequence of their logarithm $(\zeta_i)_1^d$ in $\mathfrak{g}$ give us an exponential coordinate system of the second kind in which an element g of the group G is written

$$g = \prod_1^d \exp(y_i \zeta_i), \quad y = (y_i)_1^d \in \mathbb{R}^d.$$

Recall that we also have exponential coordinates $(x)_1^d \in \mathbb{R}^d$ of the first kind such that

$$g = \exp\left(\sum x_i \zeta_i\right).$$

By using the Campbell-Hausdorf formula, we obtain a polynomial map

$$x = M(y) = (M_i(y))_1^d \text{ such that } g = \prod_1^d \exp(y_i \zeta_i) = \exp\left(\sum x_i \zeta_i\right)$$

and this map has a specific triangular structure which can be described as follows. For a multi-index of length q, $I = (i_1, \ldots, i_q) \in \{1, \ldots, d\}^q$, set $\underline{w}_I = \sum_1^q w_{i_j}$. We say that a polynomial p in the coordinate $(y_i)_1^d$ has weight at most w if it can be written as a linear combination of $y^I = y_{i_1} \ldots y_{i_q}$ with $\underline{w}_I \le w$. Then the map M has the form

$$M_i(y) = y_i + m_i(y),$$

where m_i is a polynomial of weight at most $\underline{w}_i$ with no linear terms.

Let us use the notation

$$\delta_t : \mathbb{R}^d \to \mathbb{R}^d,\ u = (u_i)_1^d \mapsto \delta_t(u) = (t^{b_i} u_i)_1^d, \quad b_i = \underline{w}_i,\ t > 0,$$

and note that we can use these dilations in the x coordinate system as a well as in the y coordinate system discussed above. We find that

$$\delta_{1/t} \circ M \circ \delta_t(y) = y_i + t^{-b_i} m_i(\delta_t(y)).$$

Because m_i has weight at most b_i, this expression has a limit when t tends to infinity which is of the form

$$y_i + m_i^\infty(y),$$

where m_i^∞ is a linear combination of terms of weight exactly b_i. This defines a polynomial map

$$M^\infty : \mathbb{R}^d \to \mathbb{R}^d,$$

which is a group isomorphism between the limit groups $G_\bullet^1$ (obtained by using the approximate group dilations δ_t in the exponential coordinates of the first kind) and the group $G^2_\bullet$ (obtained by using the approximate group dilations δ_t in the exponential coordinates of the second kind).

9.5 More choices of coordinate systems

There are many more possible choices of exponential coordinates of first and second kind that suit our needs. The key structure that must be preserved for our purpose is the filtration $\mathfrak{g}_j^{\Sigma,w}$, $1 \le j \le j_\star^{\mathfrak{g}}$, of the Lie algebra $\mathfrak{g}$ which is canonically associate to Σ, w. After that, a number of choices have to be made, the first of which is the choice of the direct sum $\mathfrak{g} = \oplus_{j=1}^{j_\star^{\mathfrak{g}}} \mathfrak{n}_j$ so that

$$\mathfrak{g}_j^{\Sigma,w} = \oplus_{i \ge j} \mathfrak{n}_j.$$

One then need to pick an adapted linear basis $\boldsymbol{\varepsilon} = (\varepsilon_i)_1^d$. Any such choice gives both an exponential coordinate system of the first kind

$$g = \exp\left(\sum_1^d x_i \varepsilon_i\right), \quad x = (x_i)_1^d \in \mathbb{R}^d,$$

and an exponential coordinate system of the second kind

$$g = \prod_{i=1}^d \exp(y_i \varepsilon_i), \quad y = (y_i)_1^d \in \mathbb{R}^d.$$

Each of these choices of coordinates, call it $(u_1, \dots, u_d) \in \mathbb{R}^d$, comes with its own straight approximate group dilations

$$\delta_t : \mathbb{R}^d \to \mathbb{R}^d,\ u = (u_i)_1^d \mapsto \delta_t(u) = (t^{b_i} u_i)_1^d, \quad b_i = \underline{w}_i,\ t > 0. \tag{9.3}$$

Everything that has been said above for the special case $\varepsilon_i = \zeta_i$ applies as well to these other choices (including the properties of the maps M and M^∞). The choice $\boldsymbol{\varepsilon} = \boldsymbol{\zeta}$ is justified mostly by the fact that, in that coordinate system, the discrete group Γ is represented as a set as $\mathbb{Z}^d \subset \mathbb{R}^d$. This is not the case in most other coordinate systems. If one remains in the class of exponential coordinate systems of the first type, moving from one such system to another is captured by a linear change of coordinate in $\mathbb{R}^d = \mathfrak{g}$. If one move from a system of the first kind to one of the second kind or between systems of the second kind, the maps capturing the changes of coordinates are polynomial maps with a special structure reflecting the preservation of the filtration $\mathfrak{g}_j^{\Sigma,w}$, $1 \le j \le j_\star^{\mathfrak{g}}$, of the Lie algebra $\mathfrak{g}$ (the best way to think of a change of coordinates involving at least one system of the second kind is to pass through the associated system of the first kind: this step is described by the map M above).

9.6 Comparison of the quasi-norms on Γ, G and $G_\bullet$

Consider the exponential coordinate systems of first and second type associated with a basis $\boldsymbol{\varepsilon} = (\varepsilon_i)_1^d$ adapted to the filtration $\mathfrak{g}_j^{\Sigma,w}$ of the Lie algebra $\mathfrak{g}$ as considered in the preceding section. It comes with a family of approximate dilations given by (9.3). On $\mathbb{R}^d$, consider the usual Euclidean norm $\|\cdot\|_2$ and the quasi-norm

$$N_w(z) = \max_{1\le i\le d}\{|z_i|^{1/b_i}\},\quad b_i = \underline{w}_i = w_i^{\mathfrak{g}},$$

and note that, for all $z \in \mathbb{R}^d$, $N_w(\delta_t(z)) = tN_w(z)$. The structure of the change of coordinate map M between exponential coordinates of the first ($x = (x_i)_1^d$) and second ($y = (y_i)_1^d$) type shows that there are constants $0 < c \le C < \infty$ such that, if $g = \exp(\sum_1^d x_i\zeta_i) = \prod_1^d \exp(y_i\zeta_i)$ then

$$cN_w(x) \le N_w(y) \le CN_w(x).$$

Lemma 9.14 Referring to the above setup and notation, there is a constant C_* such that for any $R \ge 1$ and any $\zeta_i \in \mathfrak{g}$ with $N_w(\zeta_i) \le R,\ i = 1,2$, we have $\exp(\zeta_1)\exp(\zeta_2) = \exp(\zeta)$ with $N_w(\zeta) \le C_*R$. □

Proof This follows from the Campbell-Hausdorff formula because of the properties of the direct sum decomposition along the subspaces $\mathfrak{n}_j$ and its relation to the weight system w. Note that this is not correct in general for small R. This reflects the fact that the coordinate system and the quasi-norm N_w have been chosen to capture the large scale geometry of the situation. □

The following proposition is one of the important keys to the results presented in this monograph . It relates the geometry of the discrete group Γ equipped with the quasi-norm $\|\cdot\|_{\Sigma,w}$ (Definition 8.10) to the geometry of N_w in the above coordinate systems.

Proposition 9.15 There are constants $c, C \in (0,\infty)$ such that, for any

$$\gamma = \exp\left(\sum_1^d x_i\varepsilon_i\right) \text{ with } x = (x_i)_1^d \in \mathbb{R}^d,$$

$$cN_w(x) \le \|\gamma\|_{\Sigma,w} \le CN_w(x).$$

Similarly, there are constants $c, C \in (0,\infty)$ such that, for any

$$\gamma = \prod_1^d \exp(y_i\varepsilon_i) \text{ with } y = (y_i)_1^d \in \mathbb{R}^d,$$

$$cN_w(y) \le \|\gamma\|_{\Sigma,w} \le CN_w(y).$$

Thanks to earlier considerations, it suffices to prove the first set of inequalities which refers to exponential coordinates of the first kind.

Proof (Proof of $cN_w(\zeta) \le \|\gamma\|_{\Sigma,w}$) To simplify notation, set $N_w = N$. In [55], it is proved that there exists a finite tuple $(i_1,\dots,i_q)$, $i_j \in \{1,\dots,\ell\}$, $1 \le j \le q$, such that any $\gamma \in \Gamma$ with $\|\gamma\|_{\Sigma,w} = R$ can be written as

$$\gamma = \prod_1^q \sigma_{i_j}^{z_j}, \quad |z_j| \le CR^{w(\sigma_{i_j})}. \tag{9.4}$$

Since $\sigma_i = \exp(\varsigma_i)$, $\sigma_i^x = \exp(x\varsigma_i)$ for any $i \in \{1,\dots,\ell\}$. Because ς_i has weight $w^{\mathfrak{g}}(\varsigma_i) = w(\sigma_i)$, by construction, there is a k_i with $w^{\mathfrak{g}}_{w_{k_i}} \ge w(\sigma_i)$ such that $\varsigma_i \in \mathfrak{g}^{w^{\mathfrak{g}}}_{k_i}$. In particular,

$$\varsigma_{i_j} = \sum_{k=k_{i_j}}^{j^{\mathfrak{g}}_\star} \xi_k, \quad \xi_k \in \mathfrak{n}_k,$$

and

$$\exp(z_j\varsigma_{i_j}) = \exp\left(\sum_{k=k_{i_j}}^{j^{\mathfrak{g}}_\star} z_j\xi_k\right)$$

with

$$\|z_j\xi_k\|_2 \le \max_{1\le i\le \ell}\{\|\varsigma_i\|_2\} \times |z_j| \le C'R^{w(\sigma_{i_j})} \le C'R^{w^{\mathfrak{g}}_k},$$

because $R \ge 1$ and $w^{\mathfrak{g}}_k \ge w(\sigma_{i_j})$ for all $k \ge k_{i_j}$. That is, $N(z_j\xi_k) \le C''R$.

Because formula (9.4) gives any γ as a product of at most q elements $\exp(z_j\varsigma_{i_j})$ with $N(z_j\varsigma_{i_j}) \le C'R$, it follows that any $\gamma = \exp(\sum_1^n x_i\varepsilon_i) \in \Gamma$ satisfies

$$N(x) \le C''C_*^{q-1}R = C''C_*^{q-1}\|\gamma\|_{\Sigma,w}.$$

Proof (Proof of $\|\gamma\|_{\Sigma,w} \le CN(\zeta)$) The proof is by induction on the dimension n of $\mathfrak{g}$. If the dimension is 0, there is nothing to prove. Assume that for all cases when the dimension of $\mathfrak{g}$ is less then m, there exists a constant $\widetilde{C}$ such that $\|\tilde\gamma\|_{\widetilde\Sigma,\tilde w} \le \widetilde{C}N(\tilde\zeta)$ for all $\tilde\gamma \in \widetilde\Gamma \subset (\mathbb{R}^m,\cdot) = \widetilde{G}$. Consider $\Gamma, \Sigma, w, G = (\mathbb{R}^{m+1},\cdot)$. Let $g \in \Gamma$ be a non-trivial element of the highest weight w_{j_*} which is a commutator of the elements σ_i forming the tuple Σ (this includes the elements of Σ which are considered commutators of length 1). Let $a \ge 1$ be the length of this commutator and $\sigma_{i_1},\dots,\sigma_{i_a}$, be the list of σ_i used to write g as a commutator of length a with the property that $w_{j_*} = \sum_1^a w(\sigma_{i_a})$. The element g must commute with all elements in Γ and it is of the form $g = \exp(\theta)$ where θ is the Lie commutator over $\Sigma^{\mathfrak{g}}$ associated with the writing of g as a commutator over Σ. Formally, let us use the notation $\mathbf{c}_G(x_1,\dots,x_a)$ to express the formal group commutator in question evaluated at the group elements $x_1,\dots,x_a$ so that $g = \mathbf{c}_G(\sigma_{i_1},\dots,\sigma_{i_a})$. Let $\mathbf{c}_{\mathfrak{g}}$ be the corresponding formal Lie commutator so that $\theta = \mathbf{c}_{\mathfrak{g}}(\zeta_{i_1},\dots,\zeta_{1_a})$. For any a-tuple of reals $t_1,\dots,t_a$, we also have

$$\mathbf{c}_G\left(e^{t_1\zeta_{i_1}},\dots,\dots,e^{t_a\zeta_{i_a}}\right) = \exp(t_1\dots t_a\mathbf{c}_{\mathfrak{g}}(\zeta_{i_1},\dots,\zeta_{i_a})).$$

Let $\Theta = \{\exp(s\theta) : s \in \mathbb{R}\}$ be the central one parameter subgroup of G associated with θ and consider the simply connected nilpotent group $\widetilde{G} = G/\Theta$ and its discrete subgroup $\widetilde{\Gamma}$ which is the image of Γ by the projection map $\pi : G \to \widetilde{G}$. The subgroup $\widetilde{\Gamma}$ is generated by the tuple $\widetilde{\Sigma} = (\pi(\sigma_1), \dots, \pi(\sigma_\ell))$. The dimension of $\tilde{\mathfrak{g}}$ is $m - 1$. We can choose it to be the orthogonal complement of θ in $\mathfrak{g}$ so that $d\pi$ is the orthogonal projection onto $\widetilde{\mathfrak{g}}$.

For any $\gamma = \exp(\zeta) \in \Gamma$ with $N(\zeta) = R$, we have $N(\tilde{\zeta}) \le N(\zeta) = R$. Moreover, applying the induction hypothesis, we can write $\tilde{\gamma} = \exp(\widetilde{\zeta}) = \pi(\gamma) \in \widetilde{\Gamma}$ as a word over the alphabet $\widetilde{\Sigma} \cup \widetilde{\Sigma}^{-1}$ with

$$\|\widetilde{\gamma}\|_{\widetilde{\Sigma},\tilde{w}} \le \widetilde{C} N(\tilde{\zeta}).$$

Using this word representing of $\pi(\gamma)$, replacing each $\tilde{\sigma}_i$ by σ_i to obtain a word over the alphabet $\Sigma \cup \Sigma^{-1}$, and evaluating in G give us an element $\bar{\gamma} \in \Gamma$ and an element $\bar{\zeta} \in \mathfrak{g}$ such that

$$\begin{cases} \bar{\gamma} = \exp(\bar{\zeta}), \\ \pi(\bar{\gamma}) = \tilde{\gamma}, \quad d\pi(\bar{\zeta}) = \tilde{\zeta}, \\ \gamma = \bar{\gamma} \exp(t\theta) \quad \text{for some real } t \le CR^{w_{j_*}}. \end{cases}$$

The estimate on t is from [55, Theorem 2.10] (together with an application of the Campbell-Hausdorff formula in our special system of coordinates). By construction, $\exp(t\theta) \in \Gamma$, and [55, Theorem 2.10] implies that

$$\|\exp(t\theta)\|_{\Sigma,w} \le CR = CN(\zeta).$$

It follows that

$$\|\gamma\|_{\Sigma,w} \le C'(|\bar{\gamma}_{\Sigma,w}| + CN(\zeta)) \le C'(\widetilde{C} + C)N(\zeta).$$

The following proposition captures the fact that N_w is almost a quasi-norm (a quasi-norm at large scale) on $G = (\mathbb{R}^d, \cdot)$ and is a quasi-norm on $G_\bullet = (\mathbb{R}^d, \bullet)$. The first fact follows from the adapted triangular nature of multiplication in the the type of coordinate system considered here. The second fact then follows from the homogeneity of N_w together with the fact that $(\delta_t)_{t>0}$ is a group dilstion structure on $G_\bullet$.

Proposition 9.16 For any exponential coordinate system of the first or second kind adapted to the filtration $\mathfrak{g}_j^{\Sigma,w}$, $1 \le j \le j_\star^{\mathfrak{g}}$, we have, for any $z, z' \in \mathbb{R}^d$,

$$N_w(z \cdot z') \le C(N_w(z) + N_w(z') + 1),$$

where the group law $\cdot$ refers to the multiplication in $G = (\mathbb{R}^d, \cdot)$. Moreover, in the same linear basis for $\mathbb{R}^d$, we have

$$N_w(z \bullet z') \le C(N_w(z) + N_w(z')),$$

where $\bullet$ is the group law on $G_\bullet = (\mathbb{R}^d, \bullet)$ associated with the approximate group dilation $\delta_t(z) = (t^{b_i} z_i)_1^d$. □

Chapter 10
The main results for random walks driven by measures in $\mathcal{SM}(\Gamma)$

Abstract The goal of this monograph is to establish a functional limit theorem and a local limit theorem for long-range random walks driven by appropriate probability measures in $\mathcal{SM}(\Gamma)$. This chapter is devoted to verifying that such probability measures satisfy the properties set forth in Chapters 5 and 6, properties that were proved in those chapters to be sufficient to obtain both a functional limit theorem (Theorem 5.10) and a local limit theorem (Theorem 6.2). This involves using the coordinate systems, approximate group dilations and associated geometries introduced and studied in Chapter 8 and 9. Explicit examples to which the resulting theory applies have already been discussed in Chapter 7.

10.1 The limit theorems for $\mathcal{SM}(\Gamma)$

In this chapter we state our main results concerning measures in $\mathcal{SM}(\Gamma)$. They are direct applications of Theorems 5.10 and 6.2. We state these results in adapted coordinate systems. Namely, given $\mu \in \mathcal{SM}(\Gamma)$ and the simply connected Lie group G containing Γ as a co-compact discrete subgroup, we choose to write $G = (\mathbb{R}^d, \cdot)$ using *one* of the polynomial coordinate systems described in Section 9.5 above. This coordinate system is adapted to the filtration $(\mathfrak{g}_j^{\Sigma,w})_j$ of the Lie algebra $\mathfrak{g}$, itself built from the data describing the measure μ as an element of $\mathcal{SM}(\Gamma)$. In particular, in this coordinate system, we have an approximate group dilation structure given by (9.3) which defines a limit group structure $G_\bullet = (\mathbb{R}^d, \bullet)$. The law $\bullet = \bullet_\mu$ defining this limit structure depends on μ.

Below, we show that for *any* measure $\mu \in \mathcal{SM}(\Gamma)$, there are a suitable approximate group dilation structure $(\delta_t)_{t>0}$ given by (9.3) and a norm $\|\cdot\|$ on Γ so that assumptions (5.3), **(R1)-(R2)-(E1)-(E2)** and **(TΓ)** are all satisfied with the common constant $\beta > 0$. This is in contrast to condition **(A)** which may or may not be satisfied. Recall that condition **(A)** is the requirement that the measure $\mu_t = t\delta_{1/t}(\mu)$, $t \geq 1$, defined by (4.1) converges vaguely on $\mathbb{R}^d \setminus \{0\}$ to a measure $\mu_\bullet$ as t tends to infinity. Because the dilations $(\delta_t)_{t>0}$ have been carefully constructed from μ, the family $(\mu_t)_{t>0}$ is always tight and, if **(A)** is satisfied then **(T$\bullet$)** is satisfied and the support of the limit $\mu_\bullet$ generates G; see the subsection below for the proofs.

Recall that $\{\mathbb{P}_\bullet^x;\ x \in G_\bullet\}$ is the family of probability measures induced by the limit symmetric Lévy process $X^\bullet$ on $\mathbb{D}([0, M_0], \mathbb{R}^d)$.

Fix an arbitrary increasing sequence of reals T_k that tends to infinity, e.g., $T_k = k$, and recall the notation $\hat{X}_t^k,\ t > 0,\ \hat{P}_t^k,\ t > 0$, and $\mathbb{P}_k^{[x]_k},\ x \in G$ associated with the space-time rescaled discrete random walk, see (5.19). In this notation, $[x]_k$ is the closest point of x (any one of, if there are more than one such points) on Γ_{T_k} in the norm $\|\cdot\|$, and

$$\mathbb{P}_k^x(\hat{X}_t^k = y) = \mu^{([tT_k])}((\delta_{T_k}(x))^{-1} \cdot \delta_{T_k}(y)), \quad x, y \in \Gamma_{T_k}.$$

Hence, applying Theorems 5.10 and 6.2, we obtain the following theorem.

Theorem 10.1 Let $\mu \in \mathcal{SM}(\Gamma)$. Referring to the above set-up and notation, assume that condition **(A)** holds true, that is, the measure $\mu_t = t\delta_{1/t}(\mu)$, $t \geq 1$, defined at (4.1) converges vaguely on $\mathbb{R}^d \setminus \{0\}$ to a Radon measure $\mu_\bullet$ on $\mathbb{R}^d \setminus \{0\}$ as t tends to infinity.

(i) For any bounded continuous function f on $\mathbb{R}^d$, $\hat{P}_s^k f$ converges uniformly on compacts to $P_{\bullet,s} f$. Furthermore, for each $M_0 > 0$ and for every $x \in \mathbb{R}^d$, $\hat{\mathbb{P}}_k^{[x]_k}$ converges weakly to $\mathbb{P}_\bullet^x$ on the space $\mathbb{D}([0, M_0], \mathbb{R}^d)$ equipped with $\mathcal{J}_1$-topology.
(ii) For any $U_2 > U_1 > 0$ and $r > 1$,

$$\lim_{k\to\infty} \sup_{x\in\mathbb{R}^d:\|x\|\leq r} \sup_{t\in[U_1,U_2]} \left|\det(\delta_{T_k})\mu_k^{([tT_k])}(\delta_{T_k}([x]_k)) - p_\bullet(t,x)\right| = 0.$$

10.2 The hypotheses (R1)-(R2), and (E1)-(E2) when $\mu \in \mathcal{SM}(\Gamma)$

Using the constructions described in the previous two chapters and the results from [20], we can now show that any probability measure in $\mathcal{SM}(\Gamma)$ satisfies the hypotheses **(R1)-(R2)-(E1)-(E2)** in the context of properly chosen exponential coordinates of the first or second kinds. Let us assume that we are given $\mu \in \mathcal{SM}(\Gamma)$ and the associated data Σ, w as in Section 8.2 and quasi-norm $\|\cdot\|_{\Sigma,w}$ as in Definition 8.10. We assume that Γ is given as a co-compact subgroup of a simply connected Lie group G and that an adapted global exponential coordinate system of the first or second kind has been chosen as explained in Section 9.6 so that $\Gamma \subset G = (\mathbb{R}^d, \cdot)$. Moreover, $(\mathbb{R}^d, \cdot)$ is equipped with a straight approximate dilation structure

$$(\delta_t)_{t>0}: \quad \delta_t(z) = (t^{b_i} z_i)_1^d$$

with the group limit $(\mathbb{R}^d, \bullet)$. Here the basis for $\mathbb{R}^d$ can also be identified as in Section 9.6 with a linear basis of $\mathfrak{g}$ which is compatible with the direct sum decomposition $\mathfrak{g} = \oplus_1^{j_\star^{\mathfrak{g}}} \mathfrak{n}_j$ in Definition 9.3. Each subspace $\mathfrak{n}_j$ is associated with a weight value $\underline{w}_j = w_j^{\mathfrak{g}} > 2$ and for any index i such that the corresponding basis element is in $\mathfrak{n}_j$, $b_i = \mathfrak{w}_j^{\mathfrak{g}}$.

We pick

$$0 < b < \min\{b_i : 1 \le i \le d\} = (\max\{\beta_i : 1 \le i \le d\})^{-1}, \quad \beta_i := 1/b_i$$

and set $|\gamma|_\Gamma = \|\gamma\|_{\Sigma,w}^b$. By construction, this is a norm on Γ, that is, $|\gamma \cdot \gamma'|_\Gamma \le |\gamma|_\Gamma + |\gamma'|_\Gamma$ for all $\gamma, \gamma' \in \Gamma$. We also let $\|\cdot\|$ be a norm on $(\mathbb{R}^d, \bullet)$ (i.e., satisfying the triangle inequality $\|g \bullet g'\| \le \|g\| + \|g'\|$ for all $g, g' \in G_\bullet$) that is equivalent to

$$N_w(z)^b = \max\{|z_i|^{b/b_i} : 1 \le i \le d\} \quad \text{for } z \in \mathbb{R}^d$$

and satisfies the homogeneity condition (5.3) with $\beta = \max_{1 \le i \le d} \beta_i = 1/b$. By [37], such a norm always exists.

By Proposition 9.15, we have a tight comparison between the discrete object $|\cdot|_\Gamma$ and the continuous homogeneous norm $\|\cdot\|$ on $G_\bullet = (\mathbb{R}^d, \bullet)$, namely, there are constant $0 < c, C < \infty$ such that, for any $\gamma \in \Gamma \subset \mathbb{R}^d$,

$$c\|\gamma\| \le |\gamma|_\Gamma \le C\|\gamma\|. \tag{10.1}$$

Conditions (R1)-(R2) and (E1)-(E2)

Recall that condition **(R1)** reads

(R1) There are constants C_1 and κ such that, for any bounded function u defined on Γ and μ-harmonic in $B(r) := \{x \in \mathbb{R}^d : \|x\| < r\}$, we have

$$|u(y) - u(x)| \le C_1 \|u\|_\infty \left(\frac{\|x^{-1} \cdot y\|}{r} \right)^\kappa \quad \text{for } x, y \in B(r/2).$$

For any $\mu \in \mathcal{SM}(\Gamma)$, [20, Corollary 6.10] gives the following Γ-version of this property

(RΓ1) There are constants C_1 and κ such that, for any bounded function u defined on Γ and μ-harmonic in $B_\Gamma(r) := \{x \in \Gamma : |x|_\Gamma < r\}$, we have

$$|u(y) - u(x)| \le C_1 \|u\|_\infty \left(\frac{|x^{-1} \cdot y|_\Gamma}{r} \right)^\kappa \quad \text{for } x, y \in B_\Gamma(r/2).$$

To pass from this Γ version, **(RΓ1)**, to the desired **(R1)** we use the key norm comparison (10.1) and a simple covering argument to adjust the permitted range of x, y from one statement to the other.

Similarly, recall that condition **(R2)** reads

(R2) There are positive constants $C_2 > 0$ and $\beta > 0$ such that, for all $n, m \in \mathbb{N}$ and $x, y \in \Gamma$,

$$|\mu^{(n+m)}(xy) - \mu^{(n)}(x)| \le \frac{C_1}{V(n^{1/\beta})} \left(\frac{m}{n+1} + \sqrt{\frac{\|y\|^\beta}{n+1}} \right), \tag{10.2}$$

where $V(r) := \sharp\{g \in \Gamma : \|g\| \le r\}$.

The fact that **(R2)** holds true for any probability μ in $\mathcal{SM}(\Gamma)$ follows straightforwardly from [20, Theorem 5.5 (3)-(4)] (see also [20, Proposition A.3]), together with (10.1). Regarding related recent results concerning the regularity of stable-like transition kernels in the abelian case, see [15].

Regarding the exit times conditions **(E1)-(E2)**, which are expressed using the norm $\|\cdot\|$ on $\mathbb{R}^d$, for any measure $\mu \in \mathcal{SM}(\Gamma)$, they follow from (10.1) together with [20, Theorem 5.5(5)] (for **(E1)**) and [20, Lemma 6.6] (for **(E2)**), with the exponent $\beta > 0$ being in the same as those in **(R2)**. (In these results of [20], $\beta = 1/w_*$ there, where $0 < w_* := \min\{w(s); s \in \Sigma\}$, which is our b; see Example 2.9 and Proposition 2.11(c) there.)

10.3 Condition (TΓ)

Verifying condition **(TΓ)** for any measure μ in $\mathcal{SM}(\Gamma)$ requires some work. Any $\mu \in \mathcal{SM}(\Gamma)$ is a finite convex combination of probability measures of a certain type and it suffices to prove **(TΓ)** for any such building block, ν. By definition, any such probability measure ν has the following property: there is a subgroup H of Γ with finite, symmetric generating set S, word length $|\cdot|_S$ and volume growth exponent d_H, and an exponent $\alpha \in (0, 2)$, such that

$$\nu(x) \asymp \begin{cases} (1 + |x|_S)^{-\alpha - d_H} & \text{if } x \in H, \\ 0 & \text{otherwise.} \end{cases} \tag{10.3}$$

Note that the discrete subgroup H is contained as a co-compact discrete subgroup in a unique closed connected Lie subgroup $L = L_H$ of G. As in Section 10.2, we assume we have made the choice of an adapted coordinate system for G and of an appropriate approximate dilation structure $(\delta_t)_{t>0}$.

Recall that G is described by a polynomial global coordinate chart $G = (\mathbb{R}^d, \cdot)$ in which the ebesgue measure is a Haar measure for G. Let $m \le d$ be the dimension of L. This closed Lie subgroup can be described parametrically as an embedded sub-manifold of $\mathbb{R}^d$ given by a polynomial map $i_H = i$ from $\mathbb{R}^m$ into $\mathbb{R}^d$:

$$v = (v_1, \dots, v_m) \in \mathbb{R}^m \mapsto i_H(v) = (i_1(v), \dots, i_d(v)) \in \mathbb{R}^d \tag{10.4}$$

with a polynomial inverse on its image. Assume further that this map i is also a group isomorphism on its image, that is,

$$L = (\mathbb{R}^m, \cdot_H) \quad \text{and} \quad i(v) \cdot i(w) = i(v \cdot_H w).$$

In fact, we can use this formula to define $\cdot_H$ on $\mathbb{R}^m$. However, it will be convenient to assume that $L = (\mathbb{R}^m, \cdot)$ is an exponential coordinate system of the first type for L.

The Lie group L is, of course, nilpotent and simply connected, and we assume that the global coordinate system $(\mathbb{R}^m, \cdot)$ is an exponential coordinate system of the first type compatible with the lower central series of L:

$$L_1 = L \supset L_2 = [L, L] \supset \cdots \supset L_j = [L, L_{j-1}] \supset \cdots \supset L_t \supset L_{r_H+1} = \{0\},$$

where r_H is the smallest j such that $L_{j+1} = \{0\}$. Namely, there is a strictly increasing r_H-tuple of integers k_j, $1 \le j \le r_H$, $k_1 = 1, k_{r_H} = m$ such that

$$L_j = \{(0, \ldots, 0, v_{k_j}, \cdots, v_m) : v_{k_j}, \ldots, v_m \in \mathbb{R}\}.$$

In this coordinate system for L, the straight dilation

$$\gamma_t^H(v) = (t^{p_i} v_i)_1^m, \quad p_i = j \;\text{ if }\quad k_j \le i \le k_{j+1} - 1 \tag{10.5}$$

form an approximate group dilation structure with limit $L_* = (\mathbb{R}^m, *)$, a stratified nilpotent Lie group of homogeneous dimension d_H with

$$d_H = \sum_{j=1}^{r_H} j(k_{j+1} - k_j).$$

According to Pansu's theorem, see [51] and [13], the word length $|\cdot|_S$ has the property that there is a norm $|\cdot|_*$ on L_* such that $|\gamma_t^H(v)|_* = t|v|_*$ for all $t > 0$ and $v \in \mathbb{R}^m$, and

$$\lim_{v \in H,\, v \to \infty} |v|_S / |v|_* = 1. \tag{10.6}$$

Moreover,

$$|v|_* \asymp \max_i \{|v_i|^{1/p_i} : v = (v_1, \ldots, v_m)\}.$$

This implies that for any $v \in \mathbb{R}^m$ with $\|v\|_2 \le 1$ and $r \in (0, 1]$ such that $\|\gamma_{1/r}^H v\|_2 = 1$, we have

$$\|v\|_2 \le r. \tag{10.7}$$

Moreover, for any $v \in H$, if we define t_v by $\|\gamma_{1/t_v}^H v\|_2 = 1$, then we have

$$|v|_S \asymp t_v. \tag{10.8}$$

By construction, because the probability measure ν is one of the building blocks of μ, the approximate dilation structure $(\delta_t)_{t>0}$ has the property that

$$\lim_{t \to \infty} \delta_t^{-1} \circ i_H \circ \gamma_{t^{1/\alpha}}^H(v) =: p(v) \tag{10.9}$$

exists for all $v \in \mathbb{R}^m = i_H^{-1}(L)$. This limit is uniform on compact sets and the map p is a continuous map (in fact a smooth map) from $\mathbb{R}^m$ to $\mathbb{R}^d$. In the following Lemma, we consider any approximate dilation structure $(\delta_t)_{t>0}$ such that the limit in (10.9) exists.

Lemma 10.2 (The map p is a group homomorphism) Assume that $(\delta_t)_{t>0}$ is an approximate dilation structure on $G = (\mathbb{R}^d, \dot{,})$, that ν is a probability measure on H

satisfying (10.3), and that the limit p at (10.9) exists for all $v \in \mathbb{R}^m = i_H^{-1}(L)$. If we equip $\mathbb{R}^m$ with the limit group structure $L_* = (\mathbb{R}^m, *)$ associated with the dilations $(\gamma_t^H)_{t>0}$, the map p is a continuous group homomorphism from L_* to $G_\bullet$ (it is typically neither injective nor onto). Let $L_\bullet^* = p(L_*) \subseteq G_\bullet$. For any $x \in L_\bullet^*$ and $u \in L$, $\delta_{t^\alpha}(x) = p(\gamma_t^H u)$. Define $\gamma_t^{L_\bullet^*}$ on $L_\bullet^*$ by $\gamma_t^{L_\bullet^*}(x) = p(\gamma_t^H u)$. This is a group of group diffeomorphisms on $L_\bullet^*$. Namely,

$$\gamma_s^{L_\bullet^*} \circ \gamma_t^{L_\bullet^*} = \gamma_{st}^{L_\bullet^*} \quad \text{and} \quad \gamma_t^{L_\bullet^*}(x \bullet y) = \gamma_t^{L_\bullet^*}(x) \bullet \gamma_t^{L_\bullet^*}(y), \quad s,t > 0,\ x, y \in L_\bullet^*.$$

Moreover, for any $s > 0$, $\delta_{1/s} \circ \gamma_{s^{1/\alpha}}^{L_\bullet^*} = \mathrm{Id}$ on $L_\bullet^*$. □

Proof We can approximate $p(u) \bullet p(v)$ by $\delta_{1/t}\left(\delta_t(p(u)) \cdot \delta_t(p(v))\right)$ with t large enough. Note that

$$\delta_{1/t}\left(\delta_t(p(u)) \cdot \delta_t(p(v))\right) = \delta_{1/t} \circ i_H \circ \gamma_{t^{1/\alpha}}^H \left(\gamma_{1/t^{1/\alpha}}^H(\gamma_{t^{1/\alpha}}^H(u) \cdot \gamma_{t^{1/\alpha}}^H(v))\right).$$

Since, for large s, we can approximate $\gamma_{1/s}^H(\gamma_s^H(u) \cdot \gamma_s^H(v))$ by $u * v$ and the convergence of $\delta_{1/t} \circ i_H \circ \gamma_{t^{1/\alpha}}^H$ to p is uniform on compact sets, it follows that p is a continuous group homomorphism from L_* to $G_\bullet$. The remaining statements are straightforward. □

Lemma 10.3 Let ν be a probability measure on H as in (10.3) and let $(\delta_t)_{t>0}$ be an approximate dilation structure on $G = (\mathbb{R}^d, \dot{})$ satisfying (10.9). Let $\nu_t = t\delta_{1/t}(\nu)$ and let J_t the associated jump kernel from Proposition 4.7 with ν in place of μ there. For any compact subset $K \subset \mathbb{R}^d$,

$$\lim_{\eta \to 0} \limsup_{t \to \infty} \iint_{\{(x,y) \in K \times K : \|x_\bullet^{-1} \bullet y\|_2 \le \eta\}} \|x_\bullet^{-1} \bullet y\|_2^2 J_t(dx, dy) = 0, \tag{10.10}$$

$$\lim_{R \to \infty} \limsup_{t \to \infty} \int_K \int_{B_\bullet(x,R)^c} J_t(dx, dy) = 0. \tag{10.11}$$

□

To prove this lemma, note that, for any $f \ge 0$, $\iint f(x,y) J_t(dxdy)$ is dominated by a constant times

$$t \det(\delta_{1/t}) \sum_{\substack{x,y \in \delta_{1/t}(\Gamma), x \ne y \\ \delta_t(x)^{-1} \cdot \delta_t(y) \in H}} f(x,y) \frac{1}{(1 + |\delta_t(x)^{-1} \cdot \delta_t(y)|_S)^{\alpha + d}}.$$

The two functions f of interest here are

$$f(x,y) = \mathbb{1}_K(x) \mathbb{1}_K(y) \mathbb{1}_{\{\|x_\bullet^{-1} \bullet y\|_2 \le \eta\}}(y) \|x_\bullet^{-1} \bullet y\|_2$$

and $f(x,y) = \mathbb{1}_K(x) \mathbb{1}_{B_\bullet(x,R)^c}(y)$. The results follow from the facts that μ has the form (10.3) and that δ_t is compatible with $\gamma_{t^{1/\alpha}}^H$ in the sense that (10.9) holds true.

Proof (Proof of (10.11)) We need to bound

$$\begin{aligned}
I(R,t) &= t\det(\delta_{1/t}) \sum_{\substack{x,y\in\delta_{1/t}(\Gamma)\\ \delta_t(x)^{-1}\cdot\delta_t(y)\in H}} \frac{\mathbb{1}_K(x)\mathbb{1}_{B_\bullet(x,R)^c}(y)}{(1+|\delta_t(x)^{-1}\cdot\delta_t(y)|_S)^{\alpha+d_H}}\\
&= t\det(\delta_{1/t}) \sum_{\substack{x,y\in\delta_{1/t}(\Gamma)\\ \delta_t(x)^{-1}\cdot\delta_t(y)\in H}} \frac{\mathbb{1}_K(x)\mathbb{1}_{B_\bullet(x,R)^c}(y)\mathbb{1}_{\{|\cdot|_S\le\varepsilon(R^\beta t)^{1/\alpha}\}}(\delta_t(x)^{-1}\cdot\delta_t(y))}{(1+|\delta_t(x)^{-1}\cdot\delta_t(y)|_S)^{\alpha+d_H}}\\
&\quad + t\det(\delta_{1/t}) \sum_{\substack{x,y\in\delta_{1/t}(\Gamma)\\ \delta_t(x)^{-1}\cdot\delta_t(y)\in H}} \frac{\mathbb{1}_K(x)\mathbb{1}_{B_\bullet(x,R)^c}(y)\mathbb{1}_{\{|\cdot|_S>\varepsilon(tR^\beta)^{1/\alpha}\}}(\delta_t(x)^{-1}\cdot\delta_t(y))}{(1+|\delta_t(x)^{-1}\cdot\delta_t(y)|_S)^{\alpha+d_H}}\\
&= I_1(R,t)+I_2(R,t).
\end{aligned}$$

The second sum, $I_2(R,k)$ is the main term and we treat it first by going back to H.

$$\begin{aligned}
I_2(R,k) &\le t\det(\delta_{1/t}) \sum_{\substack{x,y\in\delta_{1/t}(\Gamma)\\ \delta_t(x)^{-1}\cdot\delta_t(y)\in H}} \frac{\mathbb{1}_K(x)\mathbb{1}_{\{|\cdot|_S>\varepsilon(R^\beta t)^{1/\alpha}\}}(\delta_t(x)^{-1}\cdot\delta_t(y))}{(1+|\delta_t(x)^{-1}\cdot\delta_t(y)|_S)^{\alpha+d_H}}\\
&= \left(\det(\delta_{1/t})\sum_{x\in\Gamma}\mathbb{1}_{\delta_t(K)}(x)\right)\left(t\sum_{\substack{z\in H\\ |z|_S>\varepsilon(R^\beta t)^{1/\alpha}}}\frac{1}{(1+|z|_S)^{\alpha+d_H}}\right).
\end{aligned}$$

The first factor is clearly bounded by a constant depending only on K because Γ is a co-compact lattice in G so that

$$\sum_{x\in\Gamma}\mathbb{1}_{\delta_t(K)}(x)\le C(K)\det(\delta_t). \tag{10.12}$$

Using a decomposition by the Dyadic annulus in H, $\{x\in H : 2^k\le|\cdot|_S<2^{k+1}\}$, it is elementary to verify that, for all $R,t>1$, the second factor satisfies

$$t\sum_{\substack{z\in H\\ |z|_S>\varepsilon(R^\beta t)^{1/\alpha}}}\frac{1}{(1+|z|_S)^{\alpha+d_H}}\asymp t(\varepsilon(R^\beta t)^{1/\alpha})^{-\alpha}\asymp\varepsilon^{-\alpha}R^{-\beta}.$$

This proves that $\lim_{R\to\infty}\sup_{t\ge1}\{I_2(R,t)\}=0$.

To finish the proof of (10.11), we show that we can chose $\varepsilon>0$ such that, for any $R\ge1$ and t large enough, $I_1(R,t)=0$. To see that, we will use the description of H as a discrete subgroup of $\mathbb{R}^m$ embedded into $\mathbb{R}^d$ via the $i_H:\mathbb{R}^m\to\mathbb{R}^d$, see (10.4), and the approximate dilation structure $(\gamma^H_t)_{t>0}$. A basic fact about this structure is that, there is a constant C_1 such that, for any $x\in H$ and $t\ge|x|_S$, $\|\gamma^H_{1/t^{1/\alpha}}(x)\|_2\le C_1$, where $\|\cdot\|_2$ is the Euclidean norm on $\mathbb{R}^m$. Now, for any $x\in H$, and any $t>0$, we have

$$N(i_H(x))=tN(\delta_{1/t}\circ i_H(x))=tN\left(\delta_{1/t}\circ i_H\circ\gamma^H_{t^{1/\alpha}}(\gamma^H_{1/t^{1/\alpha}}(x))\right).$$

For large enough t and $\|z\|_2 \leq C_1$, $N\left(\delta_{1/t} \circ i_H \circ \gamma^H_{t^{1/\alpha}}(z)\right) \asymp N(p(z))$, because $\lim_{t\to\infty} \delta_{1/t} \circ i_H \circ \gamma^H_{t^{1/\alpha}} = p$ uniformly on compact sets. So, if $|x|_S$ is large enough and we choose $t^{1/\alpha} = |x|_S$, we obtain $N(i_H(x)) \leq C_2|x|_S^\alpha$.

Now, consider $x, y \in \delta_{1/t}(\Gamma)$ such that $x \in K$, $z = \delta_t(x)^{-1} \cdot \delta_t(y) \in H$ and $|z|_S \leq \varepsilon(R^\beta t)^{1/\alpha}$. Because $x \in K$, we have $N(\delta_t(x)) \leq C_3 t$ and thus

$$N(y) \leq t^{-1}N(\delta_t(y)) \leq C_4 t^{-1}(N(\delta_t(x)) + N(\delta_t(x)^{-1} \cdot \delta_t(y))) \leq C_5 R^\beta,$$

that is, $\|y\| \leq C_5^{1/\beta} R$. Because x and y are confined in a compact set (that depends on R), for t large enough, $x_\bullet^{-1} \bullet y$ is close to $\delta_{1/t}(\delta_t(x)^{-1} \cdot \delta_t(y))$ so that

$$\|x_\bullet^{-1} \bullet y\|^\beta = N(x_\bullet^{-1} \bullet y) \leq C_6 t^{-1} N(\delta_t(x)^{-1} \cdot \delta_t(y)) \leq \varepsilon^\alpha C_7 R^\beta,$$

which implies $\mathbb{1}_{B_\bullet(x,R)^c}(y) = 0$ by taking ε small enough.

This proves that we can find $\varepsilon > 0$ small enough such that, for any R fixed and t large enough, $\mathbb{1}_K(x)\mathbb{1}_{B_\bullet(x,R)^c}(y)\mathbb{1}_{\{|\cdot|_S \leq \varepsilon(R^\beta t)^{1/\alpha}\}}(\delta_t(x)^{-1} \cdot \delta_t(y)) = 0$. This ends the proof of (10.11).

Proof (Proof of (10.10)) We need to bound

$$J(K,\eta,t) = t \det(\delta_{1/t}) \sum_{\substack{x,y \in \delta_{1/t}(\Gamma)\cap K \\ \delta_t(x)^{-1}\cdot\delta_t(y) \in H}} \frac{\mathbb{1}_{\{\|\cdot\|_2 \leq \eta\}}(x_\bullet^{-1} \bullet y)\|x_\bullet^{-1} \bullet y\|_2^2}{(1 + |\delta_t(x)^{-1} \cdot \delta_t(y)|_S)^{\alpha+d_H}}.$$

On $K \times K$ and for t large enough, $x_\bullet^{-1} \bullet y$ is close to $\delta_{1/t}(\delta_t(x)^{-1} \cdot \delta_t(y))$ so that

$$\begin{aligned} J(K,\eta,t) &\leq t \det(\delta_{1/t}) \sum_{\substack{x,y \in \Gamma\cap\delta_t(K) \\ x^{-1}\cdot y \in H}} \frac{\mathbb{1}_{\{\|\cdot\|_2 \leq \eta\}}(\delta_{1/t}(x^{-1} \cdot y))\|\delta_{1/t}(x^{-1} \cdot y)\|_2^2}{(1 + |x^{-1} \cdot y|_S)^{\alpha+d_H}} \\ &\leq \left(\det(\delta_{1/t}) \sum_{x\in\Gamma} \mathbb{1}_{\delta_t(K)}(x)\right)\left(t \sum_{z \in H} \frac{\mathbb{1}_{\{\|\cdot\|_2 \leq \eta\}}(\delta_{1/t}(z))\|\delta_{1/t}(z)\|_2^2}{(1 + |z|_S)^{\alpha+d_H}}\right). \end{aligned}$$

As in (10.12), the first factor is bounded for any fixed compact K. So we are left with inspecting

$$J'(K,\eta,t) = t \sum_{z \in H} \frac{\mathbb{1}_{\{\|\cdot\|_2 \leq \eta\}}(\delta_{1/t}(z))\|\delta_{1/t}(z)\|_2^2}{(1 + |z|_S)^{\alpha+d_H}}.$$

In order to use the dilation structure γ_t^H, we represent H as a discrete set in $\mathbb{R}^m$ which injects into $\mathbb{R}^d$ via the map $i_H : \mathbb{R}^m \to \mathbb{R}^d$. Recall that $\delta_{1/t} \circ i_H(\gamma^H_{t^{1/\alpha}}(z)) \to p(z)$ uniformly on compact sets so that

$$\begin{aligned} \mathbb{1}_{\{\|\cdot\|_2 \leq \eta\}}(\delta_{1/t}(i_H(z)))\|\delta_{1/t}(i_H(z))\|_2^2 &\leq \mathbb{1}_{\{\|\cdot\|_2 \leq \eta\}}(p(\gamma^H_{1/t^{1/\alpha}}(z)))\|p(\gamma^H_{1/t^{1/\alpha}}(z))\|_2^2 \\ &\leq \min\{\eta^2, \|p(\gamma^H_{1/t^{1/\alpha}}(z))\|_2^2\} \end{aligned}$$

$$\le C_1 \min\{\eta^2, \|\gamma^H_{1/t^{1/\alpha}}(z)\|_2^2\}$$
$$\le C_1 C_2 \min\{\eta^2, t^{-2/\alpha}|z|_S^2\},$$

because $\|p(v)\|_2 \le C_1\|v\|_2$ for any $v \in \mathbb{R}^m$ with $\|v\|_2^2 \le 1$ and, by (10.7)-(10.8), $\|\gamma^H_{1/t^{1/\alpha}} z\|_2 \le C_2 t^{-1/\alpha}|z|_S$. It follows that

$$\begin{aligned} J'(K,\eta,t) &\le C_3 t \sum_{z\in H} \frac{\min\{\eta^2, t^{-2/\alpha}|z|_S^2\}}{(1+|z|_S)^{\alpha+d_H}} \\ &\le C_3 t \sum_{z\in H, |z|_S>\eta t^{1/\alpha}} \frac{\eta^2}{(1+|z|_S)^{\alpha+d_H}} + C_3 t^{1-2/\alpha} \sum_{z\in H, |z|_S\le \eta t^{1/\alpha}} \frac{|z|_S^2}{(1+|z|_S)^{\alpha+d}} \\ &\le C_4 \left(t\eta^2(\eta t^{1/\alpha})^{-\alpha} + t^{1-2/\alpha}(\eta t^{1/\alpha})^{2-\alpha}\right) = 2C_4\eta^{2-\alpha}. \end{aligned}$$

This proves that $\lim_{\eta\to 0}\limsup_{t\to\infty} J(K,\eta,t) = 0$ as desired. □

10.4 Condition (T•) holds automatically for measures in $\mathcal{SM}(\Gamma)$

The careful reader will have notice that the title of this section needs additional context because, for a measure μ in $\mathcal{SM}(\Gamma)$ (and a coordinate system as discussed above), we do have an associated approximate group dilation structure $(\delta_t)_{t>0}$ and a limit group structure $G_\bullet = (\mathbb{R}^d, \bullet)$ but, in general, the family of measures $\mu_t = t\delta_{1/t}(\mu)$ does not converge vaguely on $\mathbb{R}^d \setminus \{0\}$ and thus, **(T•)** does not make immediate sense. There is, however, a simple way to correctly interpret the title of this section.

Lemma 10.4 For any $\mu \in \mathcal{SM}(\Gamma)$ and associated approximate group dilation structure $(\delta_t)_{t>0}$ in a coordinate system as above, and any vague sub-limit $\mu_\bullet$ of the family $\{\mu_t\}_{t>0}$ as t tends to infinity, condition **(T•)** holds true. □

To prove this lemma, it suffices to prove the similar statement for each component ν of the measure μ. So, we assume (10.3). Lemma 10.2 shows that any vague sub-limit $\nu_\bullet$ of ν_t is supported on $L^*_\bullet$ and is bounded from above and below by multiples of the measure $\nu_\psi(f) = \int_{\mathbb{R}^m} f(p(u))\psi(u)du$ for $f \in C_c(\mathbb{R}^d \setminus \{0\})$, where $p : \mathbb{R}^m \to \mathbb{R}^d$ is defined by (10.9) and $\psi : \mathbb{R}^m \setminus \{0\} \to [0,\infty)$ is given by $\psi(u) = |u|_*^{-\alpha-d_H}$. Let J_ψ be the associated jump kernel measure. Then, for any compact subset $K \subset V$, we claim that

$$\lim_{\eta\to 0} \iint_{\{(x,y)\in K\times K : \|x_\bullet^{-1}\bullet y\|_2\le\eta\}} \|x_\bullet^{-1}\bullet y\|_2^2 J_\psi(dx,dy) = 0, \qquad (10.13)$$

$$\lim_{R\to\infty} \int_K \int_{B_\bullet(x,R)^c} J_\psi(dx,dy) = 0. \qquad (10.14)$$

Proof (Proof of** (10.13) **and** (10.14)**) By Lemma 10.2, $\delta_t(p(y)) = p(\gamma^H_{t^{1/\alpha}} y)$. It follows that

$$\begin{aligned}
\int_K \int_{B_\bullet(x,R)^c} J_\psi(dx,dy) &= \int_K \int_{B(R)^c} d\nu_\psi(dy)dx \\
&= |K| \int_{\mathbb{R}^m} \mathbb{1}_{B(R)^c}(p(y))\psi(y)dy \\
&= |K| \int_{\mathbb{R}^m} \mathbb{1}_{B(1)^c}(\delta_{1/R^\beta}(p(y)))\psi(y)dy \\
&= |K| \int_{\mathbb{R}^m} \mathbb{1}_{B(1)^c}(p(\gamma^H_{1/R^{\beta/\alpha}}(y)))\psi(y)dy \\
&= |K| R^{d_H\beta/\alpha} \int_{\mathbb{R}^m} \mathbb{1}_{B(1)^c}(p(y))\psi(\gamma^H_{R^{\beta/\alpha}}(y))dy \\
&= |K| R^{-\beta} \int_{\mathbb{R}^m} \mathbb{1}_{B(1)^c}(p(y))\psi(y)dy \quad \overset{R\to\infty}{\longrightarrow} \; 0.
\end{aligned}$$

For the last step, note that p is continuous so that $\mathbb{1}_{B(1)^c}(p(y))$ is equal to 0 in a neighborhood of 0 in $\mathbb{R}^m$ and thus, $\int_{\mathbb{R}^m} \mathbb{1}_{B(1)^c}(p(y))\psi(y)dy < \infty$.

Similarly, consider

$$I(K,\eta) = \iint_{\{(x,y)\in K\times K:\|x_\bullet^{-1}\bullet y\|_2\le\eta\}} \|x_\bullet^{-1}\bullet y\|_2^2 J_\psi(dx,dy).$$

Write

$$\begin{aligned}
I(K,\eta) &= \int_{x\in K}\int_{\{xz\in K:\|z\|_2\le\eta\}} \|z\|_2^2 \nu_\psi(dz)dx \\
&\le |K| \int_{\{\|z\|_2\le\eta\}} \|z\|_2^2 \mu_\bullet(dz) \\
&= |K| \int_{\mathbb{R}^m} \mathbb{1}_{B(C_1\eta^\theta)}(p(u))\|p(u)\|_2^2\psi(u)du \\
&= |K| \int_{\mathbb{R}^m} \mathbb{1}_{B(C_1)}(\delta_{\eta^{-\theta}}(p(u)))\|p(u)\|_2^2\psi(u)du \\
&= |K| \int_{\mathbb{R}^m} \mathbb{1}_{B(C_1)}(p(\gamma^H_{\eta^{-\theta/\alpha}}(u)))\|p(u)\|_2^2\psi(u)du \\
&= |K|\eta^{d_H\theta/\alpha} \int_{\mathbb{R}^m} \mathbb{1}_{B(C_1)}(p(u))\|p(\gamma^H_{\eta^{\theta/\alpha}}(u))\|_2^2\psi(\gamma^H_{\eta^{\theta/\alpha}}(u))du \\
&= |K|\eta^{-\theta} \int_{\mathbb{R}^m} \mathbb{1}_{B(C_1)}(p(u))\|\delta_{\eta^\theta}(p(u))\|_2^2\psi(u)du \\
&\le |K|\eta^{\theta\left(\frac{2}{\beta_+}-1\right)} \int_{\mathbb{R}^m} \mathbb{1}_{B(C_1)}(p(u))\|p(u)\|_2^2\psi(u)du,
\end{aligned}$$

where $\beta_+ = \max\{\beta_i\} < 2$. This integral is finite because $\mathbb{1}_{B(C_1)}(p(u))\|p(u)\|_2^2 \le C_2\|u\|_2^2/(1+\|u\|_2^2)$ and, using "polar coordinates" adapted to the dilation structure

$(\gamma_t^H)_{t>0}$ on $H_* = (\mathbb{R}^m, *)$,

$$\int_{\mathbb{R}^m} \frac{\|u\|_2^2}{1+\|u\|_2^2}\psi(u)du \le 1 + \int_0^1 r^{-1+2-\alpha}dr,$$

because $\|u\|_2 \le r$ if $\|\gamma^H_{1/r}u\|_2 = 1$ and $r \le 1$ (see (10.7)-(10.8)). □

10.5 Sufficient condition for (A) when $\mu \in \mathcal{SM}(\Gamma)$

In this section, we explain why Theorem 10.1 applies to a large class of examples in $\mathcal{SM}(\Gamma)$ that includes the two main examples described in Section 2.3. To give sufficient conditions for a measure in $\mathcal{SM}(\Gamma)$ to satisfy condition **(A)**, we proceed component by component and follow the basic setup of Section 10.3. Namely, we give sufficient conditions on a measure ν satisfying (10.3) for the family $\nu_t = t\delta_{1/t}(\nu)$ to have a vague limit $\nu_\bullet$ on $\mathbb{R}^d \setminus \{0\}$. Recall that ν is supported on a discrete subgroup H contained as a co-compact closed subgroup in a closed Lie subgroup L of G. The groups G and H both have a global coordinate system $G = \mathbb{R}^d$ and $L = \mathbb{R}^m$. See Section 10.3. We consider the following additional conditions:

(SA1) There exists an everywhere defined measurable non-negative function ϕ on $\mathbb{R}^m$ such that $\nu(i_H(v)) = \phi(v)$, where i_H is the polynomial map from $\mathbb{R}^m$ to $\mathbb{R}^d$ defined by (10.4). For $v \in \mathbb{R}^m, v \neq 0$, we set

$$\phi_t^H(v) = t^{1+d_H/\alpha}\phi(\gamma^H_{t^{1/\alpha}}v). \tag{10.15}$$

(SA2) There exists a continuous function $\psi : \mathbb{R}^m \setminus \{0\} \to [0,\infty)$ such that

$$\text{for any } v \in \mathbb{R}^m \setminus \{0\}, \quad |\phi_t^H(v) - \psi(v)| \le \eta(t)\bar{\psi}(v), \tag{10.16}$$

where, $\lim_{t\to\infty}\eta(t) = 0$ and $\bar{\psi}$ is locally bounded on $\mathbb{R}^m \setminus \{0\}$.

Remark 10.5 By the construction, the function ψ must satisfy

$$\psi(v) \asymp |v|_*^{-(\alpha+d_H)} \quad \text{for } v \in \mathbb{R}^m \setminus \{0\},$$

and

$$\psi(\gamma^H_{t^{1/\alpha}}(v)) = t^{-1-d_H/\alpha}\psi(v) \quad \text{for } v \in \mathbb{R}^m \setminus \{0\} \text{ and } t > 0.$$

Remark 10.6 Regarding hypothesis (SA1), two typical examples are:

(i) The function ϕ is a continuous function on $\mathbb{R}^m$ and μ is defined in terms of ϕ. For instance, this covers Example 4.6.

(ii) The function ϕ may not be continuous but satisfies $\phi(xy) = \phi(x)$ for all $x \in H$ and $y \in \Omega_H$, where Ω_H is a relatively compact connected fundamental domain for the action of H on $L_H = \mathbb{R}^m$ (that is, Ω_H is a relatively compact connected subset of L_H so that $L_H = \cup_{h\in H} h\Omega_H$).

In this second case, we can define the function ϕ in terms of μ using the formula $\phi(xy) = \mu(x)$ for $x \in H$ and $y \in \Omega_H$. □

Example 10.7 Consider the case when $\mu(h) = c(1 + |h|_S)^{-\alpha-d_H}\mathbb{1}_H(h)$. Following Remark 10.6(ii) above, we can extend this function defined on H to a function ϕ defined on $L_H = \mathbb{R}^m$ by setting ϕ to be constant on the translates of a precompact fundamental domain. For $x \in L_H$, let $\tilde{x} \in H$ be the representative of x so that $\tilde{x}^{-1}x \in \Omega_H$, $\tilde{x} \in H$. Then,

$$\phi_t^H(v) = c(t^{-1/\alpha} + |\widetilde{i_H(v)}|_S)^{-\alpha-d_H},$$

and, setting $\psi(v) = \frac{c}{|v|_*^{\alpha+d_H}}$, Pansu's theorem (see [13, 51]) gives

$$\lim_{t\to\infty} \phi_t^H(v) = \frac{c}{|v|_*^{\alpha+d_H}} = \psi(v).$$

Furthermore,

$$|\phi_t^H(v) - \psi(v)| \le C\frac{t^{-1/\alpha}}{|v|_*^{\alpha+d_H+1}}.$$

Proposition 10.8 Under assumptions (SA1)-(SA2), the measure

$$\nu_t = t\delta_t^{-1}(\nu): \quad \nu_t(f) := t\sum_{x\in H} f(\delta_t^{-1}(x))\nu(x) \quad \text{for } f \in C_c(\mathbb{R}^d \setminus \{0\}),$$

converges vaguely on $\mathbb{R}^d \setminus \{0\}$ to a symmetric Radon measure $\nu_\bullet$ on $\mathbb{R}^d \setminus \{0\}$ given by

$$\nu_\bullet(f) = \int_{\mathbb{R}^m} f(p(u))\psi(u)du, \quad f \in C_c(\mathbb{R}^d \setminus \{0\}),$$

where $p : \mathbb{R}^m \to \mathbb{R}^d$ is defined by (10.9) and $\psi : \mathbb{R}^m \setminus \{0\} \to [0,\infty)$ by (10.16). □

Proof This follows by a sequence of algebraic manipulation and approximations as follows. We use the notations introduced above and drop the superscript H (if there is one), in particular, $d = d_H$, α, $\gamma_t = \gamma_t^H$, ϕ, $\phi_t = \phi_t^H$, ψ, $i = i_H : \mathbb{R}^m \to \mathbb{R}^d$, and the norm $|\cdot|_*$ on L_*. For any $f \in C_c(\mathbb{R}^d \setminus \{0\})$, we have the scaled down copy of H in $\mathbb{R}^m$

$$\begin{aligned}
\nu_t(f) &= t\sum_{x\in H} f(\delta_t^{-1}(x))\nu(x) = t\sum_{u\in i^{-1}(H)} f(\delta_t^{-1}(i(u)))\phi(u)\\
&= t^{-d/\alpha}\sum_{u\in\gamma_{t^{-1/\alpha}}(i^{-1}(H))} f(\delta_t^{-1}\circ i\circ\gamma_{t^{1/\alpha}}(u))\phi_t(u)\\
&= t^{-d/\alpha}\sum_{u\in\gamma_{t^{-1/\alpha}}(i^{-1}(H))} f(\delta_t^{-1}\circ i\circ\gamma_{t^{1/\alpha}}(u))\psi(u)\\
&\quad + t^{-d/\alpha}\sum_{u\in\gamma_{t^{-1/\alpha}}(i^{-1}(H))} f(\delta_t^{-1}\circ i\circ\gamma_{t^{1/\alpha}}(u))(\phi_t(u) - \psi(u))
\end{aligned}$$

$$
\begin{aligned}
&= t^{-d/\alpha} \sum_{u \in \gamma_{t^{-1/\alpha}}(i^{-1}(H))} f(p(u))\psi(u) \\
&\quad + t^{-d/\alpha} \sum_{u \in \gamma_{t^{-1/\alpha}}(i^{-1}(H))} (f(\delta_t^{-1} \circ i \circ \gamma_{t^{1/\alpha}}(u)) - f(p(u)))\psi(u) \\
&\quad + t^{-d/\alpha} \sum_{u \in \gamma_{t^{-1/\alpha}}(i^{-1}(H))} f(p(u))(\phi_t(u) - \psi(u)) \\
&\quad + t^{-d/\alpha} \sum_{u \in \gamma_{t^{-1/\alpha}}(i^{-1}(H))} (f(\delta_t^{-1} \circ i \circ \gamma_{t^{1/\alpha}}(u)) - f(p(u)))(\phi_t(u) - \psi(u)) \\
&= \Sigma_1(f,t) + \Sigma_2(f,t) + \Sigma_3(f,t) + \Sigma_4(f,t).
\end{aligned}
$$

Now, the multivariate Riemann sum $\Sigma_1(f,t)$ of the continuous function $f \circ p \times \psi$ satisfies $\lim_{t\to\infty} \Sigma_1(f,t) = \int_{\mathbb{R}^m} f(p(u))\psi(u)du$ because, for any large real R,

$$
\begin{aligned}
\Sigma_1(f,t) = t^{-d/\alpha} &\sum_{u \in \gamma_{t^{-1/\alpha}}(i^{-1}(H)); |u|_* \le R} f(p(u))\psi(u) \\
&+ t^{-d_H/\alpha} \sum_{u \in \gamma_{t^{-1/\alpha}}(i^{-1}(H)); |u|_* > R} f(p(u))\psi(u).
\end{aligned}
$$

The first term tends to $\int_{|u|_* \le R} f(p(u))\psi(u)$, whereas the second term is bounded by $CR^{-\alpha}$ because $\psi(u) \asymp |u|_*^{-\alpha-d_H}$. Similarly, $\int_{|u|_*>R} \psi(u)du \le CR^{-\alpha}$ and this proves the stated limit for $\Sigma_1(f,t)$. Using our various hypotheses regarding μ, the limits for $\Sigma_2(f,t), \Sigma_3(f,t)$ and $\Sigma_4(f,t)$ are easily seen to be equal to 0. □

10.6 The illustrative case of measures in $\mathcal{SM}_1(\Gamma)$

The simplest case illustrating the previous section is related to the treatment of measures in $\mathcal{SM}_1(\Gamma)$ when the building blocks have the form

$$
\nu(g) = c_\alpha \sum_{k\in\mathbb{Z}} (1 + |k|)^{-\alpha-1} \mathbb{1}_{\sigma^k}(g)
$$

for some $\sigma \in \Gamma \subset G$, that is, $H = \langle\sigma\rangle \subset \Gamma \subset G$. Here, of course, $d_H = 1$. We use exponential coordinates of the first kind. Recall that the element $\sigma \in H$ is of the form $\sigma = \exp(\zeta) = \zeta$ for some $\zeta = (\zeta_1, \dots, \zeta_d) \in \mathfrak{g} = \mathbb{R}^d$. This is because the exponential map is the identity in our setup. Define the function $\phi : V \to [0,\infty)$ by $\phi(x) = c_\alpha(1 + |s|)^{-\alpha-1}$ if $x = s\zeta$ and $\phi(x) = 0$ otherwise so that $\nu(f) = \sum_{y\in\{\sigma^k : k\in\mathbb{Z}\}} f(y)\phi(y)$. Set $\phi_t(x) = c_\alpha(t^{-1/\alpha} + |s|)^{-\alpha-1}$ if $x = s\zeta$ and 0 otherwise so that $\phi(t^{1/\alpha}s\zeta) = t^{-1-1/\alpha}\phi_t(s\zeta)$. We also set $\psi(s\zeta) = c_\alpha|s|^{-\alpha-1}$ for $s \neq 0$ and $\psi(y) = 0$ if $y \notin \{s\zeta : s \in \mathbb{R}\}$ so that, for each $s \neq 0$,

$$\phi_t(z\zeta) - \psi(s\zeta) = c_\alpha \frac{|s|^{1+\alpha} - (t^{-1/\alpha} + |s|)^{1+\alpha}}{(t^{-1/\alpha} + |s|)^{1+\alpha}|s|^{1+\alpha}} \to 0 \quad \text{as } t \to \infty.$$

Assume, in addition, that we are given an approximate dilation structure δ_t which can be expressed as $\delta_t(x) = (t^{w_1}x_1, \ldots, t^{w_d}x_d)$ in the basis of $\mathbb{R}^d$. We want to understand the limit of $t\delta_t^{-1}(\nu)$ which is given on a continuous function f with compact support in $V \setminus \{0\}$ by

$$\begin{aligned} t\delta_t^{-1}(\nu)(f) &= t \sum_{y=\delta_t^{-1}(k\zeta):k\in\mathbb{Z}} \phi(\delta_t(y))f(y) \\ &= t \sum_{y=\delta_t^{-1}(k\zeta):k\in\mathbb{Z}} \phi(t^{1/\alpha}t^{-1/\alpha}\delta_t(y))f(y) \\ &= t^{-1/\alpha} \sum_{y=t^{-1/\alpha}k\zeta:k\in\mathbb{Z}} \phi_t(y)f(\delta_t^{-1}(t^{1/\alpha}y)). \end{aligned}$$

Now we need to consider different cases depending on how δ_t^{-1} acts on ζ. Indeed, $\delta_t^{-1}(t^{1/\alpha}s\zeta) = (t^{1/\alpha - w_i}\zeta_i)_1^d$. In order to have a vague limit, we need to assume that, for every $i \in \{1, \ldots, d\}$ such that $\zeta_i \neq 0$, $w_i \geq 1/\alpha$. If that is the case, then

$$\lim_{t\to\infty} \delta_t^{-1}(t^{1/\alpha}(s\zeta)) = s(\zeta_i^\infty)_1^d \quad \text{with } \zeta_i^\infty = p(\zeta) = \begin{cases} \zeta_i & \text{if } w_i = 1/\alpha, \\ 0 & \text{otherwise.} \end{cases}$$

Under this assumption (i.e., the approximate dilation structure $(\delta_t)_{t>0}$ is admissible for μ), we write

$$\begin{aligned} t\delta_t^{-1}(\nu)(f) &= t^{-1/\alpha} \sum_{y=t^{-1/\alpha}k\zeta:k\in\mathbb{Z}} \phi_t(y)f(\delta_t^{-1}(t^{1/\alpha}y)) \\ &= c_\alpha t^{-1/\alpha} \sum_{y=t^{-1/\alpha}k\zeta:k\in\mathbb{Z}} \frac{f(t^{-1/\alpha}kp(\zeta))}{(t^{-1/\alpha}|k|)^{1+\alpha}} \\ &\quad + c_\alpha t^{-1/\alpha} \sum_{y=t^{-1/\alpha}k\zeta:k\in\mathbb{Z}} \frac{[f(t^{-1/\alpha}k\delta_t^{-1}(t^{1/\alpha}(\zeta))) - f(t^{-1/\alpha}kp(\zeta))]}{(t^{-1/\alpha}|k|)^{1+\alpha}} \\ &\quad + t^{-1/\alpha} \sum_{y=t^{-1/\alpha}k\zeta:k\in\mathbb{Z}} [f(t^{-1/\alpha}k\delta_t^{-1}(t^{1/\alpha}(\zeta)))(\phi_t(t^{-1/\alpha}k\zeta) - \psi(t^{-1/\alpha}k\zeta))]. \end{aligned}$$

Because f is a compactly supported continuous function in $V \setminus \{0\}$, the second and third sums tend to 0 while the first sum tends to

$$\int_{\mathbb{R}} c_\alpha |s|^{-\alpha-1} f(sp(\zeta))ds = \int_{\mathbb{R}} \psi(s\zeta)f(sp(\zeta))ds.$$

Appendix A
Nilpotent groups

A.1 Definition of nilpotent groups

In Chapter 1, we gave the classical definition of a nilpotent group and we recall it here.

Definition A.1 A nilpotent group is a group G with identity element e which has a central series of finite length, that is, there is a finite sequence of normal subgroups so that

$$\{e\} = K_0 \lhd K_1 \lhd \cdots \lhd K_n = G$$

with K_{i+1}/K_i contained in the center of G/K_i for $0 \le i \le n-1$. See, for example, [21, Definition 2.3]. □

An alternative definition of nilpotent group uses commutators. For two elements x, y of a group G, the commutator of x and y is $[x, y] := x^{-1}y^{-1}xy$. For two subsets A, B of G, $[A, B]$ denotes the group generated by all commutators $[a, b]$ for $a \in A$ and $b \in B$. See [21, Lemma 1.4] for a collection of commutator identities. The lower central series of a group G is defined inductively by setting $G_1 = G$ and $G_{i+1} = [G, G_i]$ for $i \ge 1$. It is a non-increasing sequence of subgroups of G. A group is nilpotent if and only if its lower central series terminates, that is, there is an integer $r \ge 1$ such that $G_i = \{e\}$, for all $i \ge r+1$. The smallest such r is called the nilpotent class of the group G.

Example A.2 In the Heisenberg group of 3 by 3 upper-triangular matrices with diagonal entries equal to 1, any commutator of length 3, $[M_1, [M_2, M_3]]$, is the identity and there are elements that do not commute. Hence the Heisenberg group is nilpotent of class 2. This applies to either the discrete Heisenberg group $\mathbb{H}_3(\mathbb{Z})$ with integers matrix entries or the real Heisenberg group $\mathbb{H}_3(\mathbb{R})$ with real matrix entries.□

A.2 Definition of nilpotent Lie groups and Lie algebras

We refer the reader to [22, Sections 1.1 and 1.2] for a short introduction to nilpotent Lie algebra and connected nilpotent Lie groups. In the case of a Lie algebra $(\mathfrak{g}, [\cdot,\cdot])$, the bracket $[\cdot,\cdot]$ is the key structural operation and the descending lower central series is defined inductively by $\mathfrak{g}_1 = \mathfrak{g}$, and $\mathfrak{g}_{i+1} = [\mathfrak{g}, \mathfrak{g}_i]$ for $i \ge 1$. The Lie algebra is said to be nilpotent if there is an integer $r \ge 0$ so that $\mathfrak{g}_{r+1} = \{0\}$. The smallest such r is the nilpotent class of $\mathfrak{g}$. A connected Lie group is nilpotent according to Definition

A.1 if and only if its Lie algebra is nilpotent. Any simply connected nilpotent Lie group of topological dimension d can be identified via the exponential map with $\mathbb{R}^d$ equipped with a group law given in coordinate by polynomial function. See, e.g., [22, Theorem 1.2.1]. The Campbell-Baker-Hausdorff formula (e.g., [22, Page 11]) expresses the group product in this coordinate system.

A.3 Embeddings into Lie groups

Consider the following two natural questions. When can one embed a finitely generated torsion free nilpotent group Γ as a co-compact subgroup into a nilpotent Lie group G? Which connected simply connected nilpotent Lie group contains a co-compact finitely generated subgroup?

The first question is answered by constructions due to Malcev and P. Hall which provide such embeddings for any finitely generated torsion free nilpotent group. This is the subject of [21, Chapter 4]. This result is used in this monograph, both as a black box, to embed Γ as a co-compact subgroup into a nilpotent Lie group G, and, more concretely, when we construct coordinate systems.

The answer to the second question is negative (there are connected simply connected nilpotent Lie groups that do not admit co-compact discrete subgroups). See, e.g., [22, Theorem 5.1.8 and Example 5.1.13]. This result is not needed for the purpose of this monograph.

A.4 Volume growth

A finitely generated group Γ is naturally equipped with the family of all word distances. A word distance is associated with a finite symmetric generating set S (symmetric means that $g^{-1} \in S$ if $g \in S$). The length $|g|_S$ of an element g is the least number m of elements in S that allow to write g as a product $g = \sigma_1 \dots \sigma_m$ using elements σ_i from S. By convention, $|e|_S = 0$. The associated left-invariant distance is $d_S(g,h) = |g^{-1}h|_S$. Given two finite symmetric generating sets S and T, there are positive constants $a = a(S,T)$ and $A = A(S,T)$ such that

$$a|g|_s \le |g|_T \le A|g|_S \quad \text{for all } g \in \Gamma.$$

The volume growth of Γ with respect to S is

$$V_S(t) = \#\{g \in \Gamma : |g|_S \le t\},$$

the number of points in any closed balls of radius m in $(\Gamma, d_S(\cdot,\cdot))$. If S,T are two generating sets as above then there are positive constants $b = b(S,T)$ and $B = B(S,T)$ such that

$$bV_S(bt) \le V_T(m) \le BV_S(Bt) \quad \text{for all } t > 0.$$

In the case of a finitely generated nilpotent group Γ of nilpotent class r, the behavior of the volume growth function V_S can be understood in terms of the lower central series $\Gamma_1 = \Gamma$, $\Gamma_{i+1} = [\Gamma, \Gamma_i]$ for $i \geq 1$ as follows. The quotient groups Γ_i/Γ_{i+1} are finitely generated abelian groups. As any such group, the quotient Γ_i/Γ_{i+1} is the product of a finite abelian group and $\mathbb{Z}^{\ell_i}$ for some integer $\ell_i = \mathrm{rank}(\Gamma_i/\Gamma_{i+1})$ which is called the torsion-free rank of this abelian group. Set

$$D = D(\Gamma) = \sum_{j=1}^{r} j\,\mathrm{rank}(\Gamma_i/\Gamma_{i+1}). \tag{A.1}$$

Then there are constant $c = C(S)$ and $C = C(S)$ such that,

$$c(1+t)^D \leq V_S(t) \leq C(1+t)^D \quad \text{for all } t \geq 0. \tag{A.2}$$

See, e.g., [25, Theorem VII.C.26] for references and comments on this result.

If the nilpotent Γ above is a discrete co-compact subgroup of a connected Lie group G then, for any fixed left-invariant Riemannian metric on G, the Haar measure $|B(r)|$ of the ball of radius r around the identity element e satisfies

$$c_1 r^D \leq |B(r)| \leq C_1 r^D \quad \text{for all } r \geq 1,$$

where $D = D(\Gamma)$ is as in (A.1). The positive constants c_1, C_1 depend on the choice of the Riemannian metric.

Example A.3 The Heisenberg group $\mathbb{H}_3(\mathbb{Z})$ is a co-compact subgroup of $\mathbb{H}_3(\mathbb{R})$. The elements of $\mathbb{H}_3(\mathbb{Z})$ with at most one non-zero non-diagonal entry in the the top-right corner is the center of $\mathbb{H}_3(\mathbb{Z})$ as well as the commutator subgroup $[\mathbb{H}_3(\mathbb{Z}), \mathbb{H}_3(\mathbb{Z})]$. It follows that the parameter $D = D(\mathbb{H}_3(\mathbb{Z}))$ is equal to $2 + 1 \times 2 = 4$. Any left-invariant Riemannian metric on $\mathbb{H}_3(\mathbb{R})$ has large-scale volume growth of type r^4.

References

1. David Aldous, *Stopping times and tightness*, Ann. Probab. **6** (1978), no. 2, 335–340. MR 474446
2. Georgios K. Alexopoulos, *Centered densities on Lie groups of polynomial volume growth*, Probab. Theory Related Fields **124** (2002), no. 1, 112–150. MR 1929814
3. ______, *Random walks on discrete groups of polynomial volume growth*, Ann. Probab. **30** (2002), no. 2, 723–801. MR 1905856
4. ______, *Sub-Laplacians with drift on Lie groups of polynomial volume growth*, Mem. Amer. Math. Soc. **155** (2002), no. 739, x+101. MR 1878341
5. Martin T. Barlow and Richard F. Bass, *The construction of Brownian motion on the Sierpiński carpet*, Ann. Inst. H. Poincaré Probab. Statist. **25** (1989), no. 3, 225–257. MR 1023950
6. Martin T. Barlow and Ben M. Hambly, *Parabolic Harnack inequality and local limit theorem for percolation clusters*, Electron. J. Probab. **14** (2009), no. 1, 1–27. MR 2471657
7. Richard F. Bass, Moritz Kassmann, and Takashi Kumagai, *Symmetric jump processes: localization, heat kernels and convergence*, Ann. Inst. Henri Poincaré Probab. Stat. **46** (2010), no. 1, 59–71. MR 2641770
8. Richard F. Bass and Takashi Kumagai, *Symmetric Markov chains on $\mathbb{Z}^d$ with unbounded range*, Trans. Amer. Math. Soc. **360** (2008), no. 4, 2041–2075. MR 2366974
9. Richard F. Bass, Takashi Kumagai, and Toshihiro Uemura, *Convergence of symmetric Markov chains on $\mathbb{Z}^d$*, Probab. Theory Related Fields **148** (2010), no. 1-2, 107–140. MR 2653223
10. Jean Bertoin, *Lévy Processes*, Cambridge Tracts in Mathematics, vol. 121, Cambridge University Press, Cambridge, 1996. MR 1406564
11. Luis Boza, Eugenio M. Fedriani, Juan Núñez, and Ángel F. Tenorio, *A historical review of the classifications of Lie algebras*, Rev. Un. Mat. Argentina **54** (2013), no. 2, 75–99. MR 3263653
12. Emmanuel Breuillard, *Random walks on Lie groups*, https://www.imo.universite-paris-saclay.fr/mobile/~breuilla/part0gb.pdf.
13. ______, *Geometry of locally compact groups of polynomial growth and shape of large balls*, Groups Geom. Dyn. **8** (2014), no. 3, 669–732. MR 3267520
14. E. A. Carlen, S. Kusuoka, and D. W. Stroock, *Upper bounds for symmetric Markov transition functions*, Ann. Inst. H. Poincaré Probab. Statist. **23** (1987), no. 2, suppl., 245–287. MR 898496
15. Jamil Chaker and Moritz Kassmann, *Nonlocal operators with singular anisotropic kernels*, Comm. Partial Differential Equations **45** (2020), no. 1, 1–31. MR 4037095
16. Xin Chen, Takashi Kumagai, and Jian Wang, *Random conductance models with stable-like jumps: heat kernel estimates and Harnack inequalities*, J. Funct. Anal. **279** (2020), no. 7, 108656, 51. MR 4108401
17. Zhen-Qing Chen, David A. Croydon, and Takashi Kumagai, *Quenched invariance principles for random walks and elliptic diffusions in random media with boundary*, Ann. Probab. **43** (2015), no. 4, 1594–1642. MR 3353810
18. Zhen-Qing Chen and Masatoshi Fukushima, *Symmetric Markov Processes, Time Change, and Boundary Theory*, London Mathematical Society Monographs Series, vol. 35, Princeton University Press, Princeton, NJ, 2012. MR 2849840
19. Zhen-Qing Chen, Panki Kim, Takashi Kumagai, and Jian Wang, *Heat kernel upper bounds for symmetric Markov semigroups*, J. Funct. Anal. **281** (2021), no. 4, Paper No. 109074, 40. MR 4249776
20. Zhen-Qing Chen, Takashi Kumagai, Laurent Saloff-Coste, Jian Wang, and Tianyi Zheng, *Long range random walks and associated geometries on groups of polynomial growth*, Annales de l'Institut Fourier (to appear).
21. Anthony E. Clement, Stephen Majewicz, and Marcos Zyman, *The Theory of Nilpotent Groups*, Birkhäuser/Springer, Cham, 2017. MR 3729243
22. Lawrence J. Corwin and Frederick P. Greenleaf, *Representations of Nilpotent Lie Groups and Their Applications. Part I*, Cambridge Studies in Advanced Mathematics, vol. 18, Cambridge University Press, Cambridge, 1990, Basic theory and examples. MR 1070979

23. Pierre Crépel and Albert Raugi, *Théorème central limite sur les groupes nilpotents.*, Ann. Inst. H. Poincaré Sect. B **14** (1978), no. 2, 145–164.
24. David A. Croydon and Ben M. Hambly, *Local limit theorems for sequences of simple random walks on graphs*, Potential Anal. **29** (2008), no. 4, 351–389. MR 2453564
25. Pierre de la Harpe, *Topics in geometric group theory*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 2000. MR 1786869
26. William Feller, *An Introduction to Probability Theory and Its Applications. Vol. II*, second ed., John Wiley & Sons, Inc., New York-London-Sydney, 1971. MR 0270403
27. Masatoshi Fukushima, Yoichi Oshima, and Masayoshi Takeda, *Dirichlet Forms and Symmetric Markov Processes*, extended ed., De Gruyter Studies in Mathematics, vol. 19, Walter de Gruyter & Co., Berlin, 2011. MR 2778606
28. Ĭ. Ī. Gīhman and A. V. Skorohod, *The Theory of Stochastic Processes. I*, english ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 210, Springer-Verlag, Berlin-New York, 1980, Translated from the Russian by Samuel Kotz. MR 636254
29. Paweł Głowacki, *Lipschitz continuity of densities of stable semigroups of measures*, Colloq. Math. **66** (1993), no. 1, 29–47. MR 1242643
30. Ben Green and Terence Tao, *The quantitative behaviour of polynomial orbits on nilmanifolds*, Ann. of Math. (2) **175** (2012), no. 2, 465–540. MR 2877065
31. Ulf Grenander, *Probabilities on algebraic structures*, second ed., Almqvist & Wiksell, Stockholm; John Wiley & Sons Inc., New York-London, 1968. MR 0259969
32. Philip S. Griffin, *Matrix normalized sums of independent identically distributed random vectors*, Ann. Probab. **14** (1986), no. 1, 224–246.
33. Alexander Grigor'yan, Xueping Huang, and Jun Masamune, *On stochastic completeness of jump processes*, Math. Z. **271** (2012), no. 3-4, 1211–1239. MR 2945605
34. Mikhael Gromov, *Groups of polynomial growth and expanding maps*, Inst. Hautes Études Sci. Publ. Math. (1981), no. 53, 53–73. MR 623534
35. Philip Hall, *The Collected Works of Philip Hall*, Oxford Science Publications, The Clarendon Press Oxford University Press, New York, 1988, Compiled and with a preface by K. W. Gruenberg and J. E. Roseblade, With an obituary by Roseblade. MR 986732 (90b:01108)
36. Wilfried Hazod and Eberhard Siebert, *Stable Probability Measures on Euclidean Spaces and on Locally Compact Groups: Structural Properties and Limit Theorems*, Springer, 2001.
37. Waldemar Hebisch and Adam Sikora, *A smooth subadditive homogeneous norm on a homogeneous group*, Studia Math. **96** (1990), no. 3, 231–236. MR 1067309
38. Robert Hough, *The local limit theorem on nilpotent Lie groups*, Probab. Theory Related Fields **174** (2019), no. 3-4, 761–786. MR 3980304
39. Zbigniew J. Jurek and J. David Mason, *Operator-limit Distributions in Probability Theory*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1993, A Wiley-Interscience Publication. MR 1243181
40. Hiroshi Kunita, *Convolution semigroups of stable distributions over a nilpotent Lie group*, Proc. Japan Acad. Ser. A Math. Sci. **70** (1994), no. 10, 305–310. MR 1313184
41. ———, *Stable Lévy processes on nilpotent Lie groups*, Stochastic analysis on infinite-dimensional spaces (Baton Rouge, LA, 1994), Pitman Res. Notes Math. Ser., vol. 310, Longman Sci. Tech., Harlow, 1994, pp. 167–182. MR 1415667
42. ———, *Some problems concerning Lévy processes on Lie groups*, Stochastic analysis (Ithaca, NY, 1993), Proc. Sympos. Pure Math., vol. 57, Amer. Math. Soc., Providence, RI, 1995, pp. 323–341. MR 1335479
43. ———, *Stable limit distributions over a nilpotent Lie group*, Proc. Japan Acad. Ser. A Math. Sci. **71** (1995), no. 1, 1–5. MR 1326784
44. Thomas G. Kurtz and Philip E. Protter, *Weak convergence of stochastic integrals and differential equations*, Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), Lecture Notes in Math., vol. 1627, Springer, Berlin, 1996, pp. 1–41. MR 1431298

45. Enrico Le Donne, *A primer on Carnot groups: homogenous groups, Carnot-Carathéodory spaces, and regularity of their isometries*, Anal. Geom. Metr. Spaces **5** (2017), no. 1, 116–137. MR 3742567
46. Anatoliĭ I. Mal'cev, *On a class of homogeneous spaces*, Amer. Math. Soc. Translation **1951** (1951), no. 39, 33. MR 0039734
47. Mark M. Meerschaert and Hans-Peter Scheffler, *Limit Distributions for Sums of Independent Random Vectors*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 2001, Heavy tails in theory and practice. MR 1840531
48. ______, *Limit theorems for continuous time random walks with slowly varying waiting times*, Statist. Probab. Lett. **71** (2005), no. 1, 15–22. MR 2125427
49. Mathav Murugan and Laurent Saloff-Coste, *Transition probability estimates for long range random walks*, New York J. Math. **21** (2015), 723–757. MR 3386544
50. Daniel Neuenschwander, *Probabilities on the Heisenberg group*, Lecture Notes in Mathematics, vol. 1630, Springer-Verlag, Berlin, 1996, Limit theorems and Brownian motion. MR 1439509
51. Pierre Pansu, *Croissance des boules et des géodésiques fermées dans les nilvariétés.*, Ergodic Theory Dynam. Systems **3, no.3** (1983), 415–445.
52. Gyula Pap, *Edgeworth expansions in nilpotent Lie groups*, Probability measures on groups and related structures, XI (Oberwolfach, 1994), World Sci. Publ., River Edge, NJ, 1995, pp. 274–291. MR 1414941
53. Albert Raugi, *Théorème de la limite centrale sur les groupes nilpotents*, Z. Wahrsch. Verw. Gebiete **43** (1978), no. 2, 149–172. MR 499429
54. ______, *Théorème de la limite centrale pour un produit semi-direct d'un groupe de Lie résoluble simplement connexe de type rigide par un groupe compact*, Probability measures on groups (Proc. Fifth Conf., Oberwolfach, 1978), Lecture Notes in Math., vol. 706, Springer, Berlin-New York, 1979, pp. 257–324. MR 536991
55. Laurent Saloff-Coste and Tianyi Zheng, *Random walks on nilpotent groups driven by measures supported on powers of generators*, Groups Geom. Dyn. **9** (2015), no. 4, 1047–1129. MR 3428409
56. Ken-iti Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge Studies in Advanced Mathematics, vol. 68, Cambridge University Press, Cambridge, 2013, Translated from the 1990 Japanese original, Revised edition of the 1999 English translation. MR 3185174
57. Eberhard Siebert, *Contractive automorphisms on locally compact groups*, Math. Z. **191** (1986), no. 1, 73–90. MR 812604
58. Daniel W. Stroock and S. R. S. Varadhan, *Limit theorems for random walks on Lie groups*, Sankhyā Ser. A **35** (1973), no. 3, 277–294. MR 517406
59. Nicholas Th. Varopoulos, Laurent Saloff-Coste, and Thierry Coulhon, *Analysis and Geometry on Groups*, Cambridge Tracts in Mathematics, vol. 100, Cambridge University Press, Cambridge, 1992. MR 1218884 (95f:43008)
60. Donald Wehn, *Limit Distributions on Lie Groups*, ProQuest LLC, Ann Arbor, MI, 1960, Thesis (Ph.D.)–Yale University. MR 2939161
61. ______, *Probabilities on Lie groups*, Proc. Nat. Acad. Sci. U.S.A. **48** (1962), 791–795. MR 153042