Ergodic Theory and Geometry of Nilpotent Groups

BY

MICHAEL A. CANTRELL
B.S., Yale University, 2009

THESIS
Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Chicago, 2016

Chicago, Illinois

Defense Committee:
Alex Furman, Chair and Advisor
Kevin Whyte
Christian Rosendal
Steve Hurder
Kate Juschenko, Northwestern University
Contribution of Authors

Chapter 1 is an introduction and statement of the results that follow. Chapter 2 is essentially an unaltered version of a paper co-authored by Alex Furman and myself that has been submitted for publication but not accepted, as of the writing of this thesis. Chapter 3 is essentially an unaltered version of a paper of which I am the sole author. It has been submitted, but not accepted, as of the writing of this thesis. Chapter 4 contains my synthesis of the research presented in this thesis and a discussion of potential further research directions.
Acknowledgements

I would like to thank my thesis committee members, Alex Furman, Kevin Whyte, Christian Rosendal, Steve Hurder and Kate Juschenko. I would especially like to thank my advisor Alex Furman for teaching me how to do mathematical research.

Finally, I thank my partner Ellie Hong for her constant support and encouragement.

MAC
Summary

We investigate the ergodic theory and geometry of finitely generated virtually nilpotent groups. Let $\Gamma$ be a finitely generated virtually nilpotent group. In the first part of the thesis, we consider three closely related problems: (i) convergence to a deterministic asymptotic cone for an equivariant ergodic family of inner metrics on $\Gamma$, generalizing a theorem of Pansu; (ii) the asymptotic shape theorem for First Passage Percolation for general (not necessarily independent) ergodic processes on edges of a Cayley graph of $\Gamma$; (iii) the sub-additive ergodic theorem over a general ergodic $\Gamma$-action. The limiting objects are given in terms of a Carnot-Carathéodory metric on the graded nilpotent group associated to the Mal’cev completion of $\Gamma$.

In the second part of the thesis we prove an analog for integrable measurable cocycles of Pansu’s differentiation theorem for Lipschitz maps between Carnot-Carathéodory spaces. This yields an alternative, ergodic theoretic proof of Pansu’s quasi-isometric rigidity theorem for nilpotent groups, answers a question of Tim Austin regarding integrable measure equivalence between nilpotent groups, and gives an independent proof and strengthening of Austin’s result that integrable measure equivalent nilpotent groups have bi-Lipschitz asymptotic cones. The main tools for this part are a nilpotent-valued cocycle ergodic theorem and a Poincaré recurrence lemma for nilpotent groups, which may be of independent interest.
Contents

Contribution of Authors ii
Acknowledgements iii
Summary iv
List of Abbreviations vi
Chapter 1. Introduction 1
  1. Introduction and Statement of Main Results 1
Chapter 2. Asymptotic Shapes for Ergodic Families of Metrics on Nilpotent Groups 13
  1. The Carnot group as the asymptotic cone 14
  2. Preparation for the main proofs 23
  3. Proof of Theorems A, B 34
Chapter 3. Differentiability Of Integrable Measurable Cocycles Between Nilpotent Groups 43
  1. Introduction 43
  2. Background and Notation 44
  3. Asymptotic Behavior Along Iterates 54
  4. Asymptotic Behavior Along Arbitrary Elements 64
  5. Construction of Φ and Proof of Main Theorem 72
Chapter 4. Conclusion 83
Cited Literature 85
Vita 87
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.g.</td>
<td>finitely generated</td>
</tr>
<tr>
<td>pmp</td>
<td>probability measure preserving</td>
</tr>
<tr>
<td>ME</td>
<td>measure equivalence</td>
</tr>
<tr>
<td>IME</td>
<td>integrable measure equivalence</td>
</tr>
<tr>
<td>whp</td>
<td>with high probability</td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction

1. Introduction and Statement of Main Results

Nilpotent groups arise naturally in mathematics from abstract algebra and Lie theory to subRiemmanian geometry and control theory. They are simple to define and are only one level of complexity beyond the abelian groups. Nevertheless, they bridge mathematical disciplines and basic questions about them remain open. In this thesis, we investigate aspects of the ergodic theory and geometry of countable infinite nilpotent groups.

Classical ergodic theory studies the long term evolution of a single measure preserving transformation. In recent years, it has become clear that studying a group of measure preserving transformations bears information on many mathematical fields, from the theory of smooth manifolds and Lie theory to algebraic geometry and number theory.

Birkhoff’s pointwise ergodic theorem is, arguably, the starting point of modern ergodic theory.

**Theorem 1.1 (Birkhoff [4]).** Let \((X, \mu)\) be a probability measure space and \(T : X \to X\) be an ergodic measure preserving transformation. Then for every \(f \in L^1(X, \mu)\)

\[
\frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) \to \int_X f d\mu
\]

for \(\mu\)-a.e. \(x\).

In 2001, Lindenstrauss extended Birkhoff’s theorem to the class of amenable groups.

**Theorem 1.2 (Lindenstrauss [19]).** Let \(G\) be an amenable group acting ergodically on a probability measure space \((X, \mu)\) and let \(F_n\) be a tempered Folner sequence. Then for every \(f \in L^1(X, \mu)\)

\[
\frac{1}{|F_n|} \sum_{g \in F_n} f(gx) \to \int_X f d\mu
\]

A fundamental theorem of Kingman generalizes Birkhoff’s theorem to subadditive processes of a single transformation.
1. Introduction and Statement of Main Results

Theorem 1.3 (Kingman [17]). Let $(X, \mu)$ be a probability measure space, $T : X \to X$ be an ergodic measure preserving transformation and $c : \mathbb{N} \times X \to \mathbb{R}$ be an integrable, measurable subadditive cocycle. Then for $\mu$-a.e. $x$

$$\frac{1}{n} c(n, x) \to \inf \frac{1}{n} \int_X c(n, x) d\mu(x).$$

Here, the subadditivity of the cocycle means that

$$c(n + m, x) \leq c(n, T^m x) + c(m, x) \quad n, m \in \mathbb{N} \text{ a.e. } x \in X.$$

This begs the following natural question.

Question 1.4. What is the subadditive generalization of Lindenstrauss’s ergodic theorem for amenable groups?

That is, given a general amenable group $G$, an ergodic probability measure preserving action of the group $G \curvearrowright (X, \mu)$, and a subadditive cocycle, i.e. a measurable map $c : G \times X \to \mathbb{R}$ satisfying

$$c(g_1 g_2, x) \leq c(g_1, g_2 \cdot x) + c(g_2, x) \quad \forall g_1, g_2 \in G, \mu - a.e. x$$

how can one describe the asymptotics of the cocycle?

The first result of this thesis is an answer to this question for the class of countable nilpotent groups. Before we can state the result, we must recall some facts about nilpotent groups.

Upon passing to a finite index subgroup and dividing by a finite normal subgroup, we assume hereafter that our group $\Gamma$ is a torsion-free nilpotent group with torsion-free abelianization $\Gamma^{ab} \cong \mathbb{Z}^d$; this adjustment does not affect the questions that follow- see Chapter 2 §1.4 and Chapter 3 §2.5 below. By the classical work of Mal’cev [20], a finitely generated, torsion-free, nilpotent group $\Gamma$ can be embedded as a discrete subgroup of a connected, simply connected, nilpotent real Lie group $G$ so that $G/\Gamma$ is compact. Moreover, such an embedding $\Gamma < G$ is unique up to automorphisms of $G$. This $G$ is often called the Mal’cev completion of $\Gamma$. Associated with $G$ one has a graded nilpotent connected, simply connected, real Lie group $G_{\infty}$, that is constructed from the quotient spaces $g^i/g^{i+1}$ of the descending central series $g = g^1 > g^2 > \cdots > g^{r+1} = \{0\}$ of the Lie algebra of $G$ (see Chapter 2 §1). In particular, one can identify the abelianizations $G^{ab} := G/[G, G]$ and $G^{ab}_{\infty} = G_{\infty}/[G_{\infty}, G_{\infty}]$ via $g/g^2 \cong g_{\infty}/g^2_{\infty}$. The graded Lie group $G_{\infty}$ admits a one parameter
family \( \{ \delta_t \mid t > 0 \} \) of automorphisms that induce the linear homotheties \( \times t \) on the real vector space \( G^{ab}_\infty \cong \mathfrak{g}^{ab} \cong \mathfrak{g}^{ab}_\infty \). Such a group \( G_\infty \) (with the family of homotheties) is sometimes called a Carnot group.

**Example 1.5.** The integral Heisenberg group \( H_\mathbb{Z} \) embeds in the 3-dimensional real Heisenberg group

\[
H_\mathbb{R} = \left\{ M_{x,y,z} = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \mid x, y, z \in \mathbb{R} \right\}
\]

by restricting \( x, y, z \) to be integers. In this case \( G = H_\mathbb{R} \) is itself graded: \( G = G_\infty \). The abelianization \( G^{ab} \) is two dimensional, and \( G \to G^{ab} \) is given by \( M_{x,y,z} \mapsto (x, y) \). The similarities are given by

\[
\delta_t(M_{x,y,z}) = M_{tx,ty,tz^2}.
\]

In [24] Pansu showed that the associated Carnot group \( G_\infty \) is naturally associated to \( \Gamma \) as the underlying space which realizes all of the asymptotic cones of \( \Gamma \). In general, the asymptotic cone of a metric space is the metric space ‘seen from far away.’ Formally, it is a Gromov-Hausdorff limit of the rescaled metric space. Pansu showed that the asymptotic cone of a countable nilpotent group equipped with a word norm is particularly nice, in that every asymptotic cone is the associated Carnot group equipped with a Carnot-Carathéodory metric.

**Theorem 1.6 (Pansu [24]).** Let \( \Gamma \) be a finitely generated nilpotent group and \( G_\infty \) its associated Carnot group. Let \( d \) be an inner right-invariant metric\(^1\) on \( \Gamma \), e.g. a word metric: \( d(\gamma_1, \gamma_2) = |\gamma_1 \gamma_2^{-1}|_S \), where \( |\gamma|_S \) is the length of a shortest word representing \( \gamma \) using elements of a fixed generating set \( S \) for \( \Gamma \). Then there is a right-invariant proper metric \( d_\infty \) on \( G_\infty \), that is homogeneous in the sense that

\[
d_\infty(\delta_t(g), \delta_t(g')) = t \cdot d_\infty(g, g') \quad (g, g' \in G_\infty, \ t > 0)
\]

and such that there is Gromov-Hausdorff convergence

\[
(\Gamma, \frac{1}{t} \cdot d, e) \to (G_\infty, d_\infty, e).
\]

I.e. \((G_\infty, d_\infty, e)\) is the asymptotic cone of \((\Gamma, d, e)\).

---

\(^1\)One often considers left-invariant metrics; our choice of right-invariance is dictated by our notation for sub-additive cocycles.
To state our first result we need to fix some further notations. Let $\Gamma$ be a finitely generated, torsion-free, nilpotent group, denote by $G$ its Mal’cev completion, and by $G_\infty$ the associated Carnot group with homotheties $\{\delta_t \mid t > 0\}$. Fix a right-invariant inner metric $d$ on $\Gamma$, e.g. a word metric as above, and let $d_\infty$ on $G_\infty$ be the associated Carnot-Carathéodory metric as in Pansu’s theorem.

Given a function $f : \Gamma \to \mathbb{R}$ one can consider an asymptotic cone of its graph in $\Gamma \times \mathbb{R}$, i.e. possible Gromov-Hausdorff limits of

$$\text{Graph}(f) = \{(\gamma, f(\gamma)) \mid \gamma \in \Gamma\} \subset \Gamma \times \mathbb{R}$$

with $(e, 0)$ being the marked point. The functions $f$ that will appear below will be special in several ways:

(f1) the rescaled graphs $\text{Graph}(f)$ actually have a unique Gromov-Hausdorff limit,

(f2) this limit is given by a graph $\text{Graph}(\Phi)$ of a function $\Phi : G_\infty \to \mathbb{R}$,

(f3) the function $\Phi : G_\infty \to \mathbb{R}$ appears in a Carnot-Carathéodory construction; in particular,

it is homogeneous: $\Phi(\delta_t(g)) = t \cdot \Phi(g)$ for $g \in G_\infty$ and $t > 0$.

The convergence in (f2) implies that

$$t_i^{-1} \cdot f(\gamma_i) \to \Phi(g) \quad \text{whenever} \quad \text{scl}_i(\gamma_i) \to g \in G_\infty,$$

where the latter relates to the Gromov-Hausdorff limit in Chapter 2 §1 with $t_i \to \infty$. Let us say that two functions $f, f' : \Gamma \to \mathbb{R}$ are asymptotically equivalent if

$$f(\gamma) - f'(\gamma) = o(|\gamma|_s).$$

Then $f$ satisfies (f1)-(f3) with $\Phi$ iff $f'$ does. One might say that $\Phi$ is the unique homogeneous representative of the asymptotic equivalence class of $f$ (here the uniqueness statement follows from the fact that different homogeneous functions cannot be asymptotically equivalent). Now we can state our first result, which is an answer to Question 1.4 for the class of countable nilpotent groups.

**Theorem A.**

Let $\Gamma$ be a finitely generated virtually nilpotent group, $\Gamma \curvearrowright (X, m)$ an ergodic probability measure-preserving action, and $c : \Gamma \times X \to \mathbb{R}_+$ a measurable subadditive cocycle. Assume that

(1) For some $0 < k \leq K < +\infty$ one has $k \cdot |\gamma|_s \leq c(\gamma, x) \leq K \cdot |\gamma|_s$ for a.e. $x \in X$. 


(2) For a.e. $x \in X$ for every $\epsilon > 0$ there is a finite set $F \subset \Gamma$ so that for every $x' \in \Gamma x$ any $\gamma \in \Gamma$ can be written as $\gamma = \delta_n \cdots \delta_2 \delta_1$ with $\delta_i \in F$ and

$$c(\delta_1, x') + c(\delta_2, \delta_1, x') + \cdots + c(\delta_n, \delta_{n-1} \cdots \delta_1, x') \leq (1 + \epsilon) \cdot c(\gamma, x').$$

Then for a full measure set of $x \in X$ the functions $c(-, x) : \Gamma \to \mathbb{R}$ are asymptotically equivalent to each other and are represented by a unique homogeneous function $\Phi : G_\infty \to \mathbb{R}$, that is obtained in the following construction.

**Construction 1.7.** Let $c : \Gamma \times X \to \mathbb{R}_+$ be a subadditive cocycle over an ergodic action $\Gamma \curvearrowright (X, m)$ of a finitely generated virtually nilpotent group $\Gamma$.

- Up to finite index and finite kernel (once $\Gamma \curvearrowright (X, m)$ and $c : \Gamma \times X \to \mathbb{R}_+$ are adjusted accordingly) we are reduced to the case that $\Gamma$ is a finitely generated nilpotent group that is torsion free and has torsion-free abelianization $\Gamma^{ab}$.
- Define a subadditive function $\overline{c} : \Gamma \to \mathbb{R}_+$ by integration:

$$\overline{c}(\gamma) := \int_X c(\gamma, x) \, dm(x).$$

- Define a subadditive function $f : \Gamma^{ab} \to \mathbb{R}_+$ by minimizing $F$ over fibers:

$$f(\gamma^{ab}) := \inf \{ \overline{c}(\gamma_1) \mid \gamma^{ab} = \gamma_1 \}.$$

- Define $\phi : g^{ab}_\infty \to \mathbb{R}_+$ by viewing $\Gamma^{ab}$ as a lattice in the vector space $\Gamma^{ab} \otimes \mathbb{R}$ and observing that there is a unique homogeneous subadditive function (a possibly asymmetric norm)

$$\phi : \Gamma^{ab} \otimes \mathbb{R} \to \mathbb{R}_+$$

representing $f : \Gamma^{ab} \to \mathbb{R}_+$.
- Define $\Phi : G_\infty \to \mathbb{R}_+$ to be the homogeneous function associated to $\phi$ viewed as an asymmetric norm on $\Gamma^{ab} \otimes \mathbb{R} \cong G^{ab} \cong g^{ab}_\infty \cong g^{ab}_\infty$ and applying the Carnot-Carathéodory construction.

For more details see Chapter 2 §1.2–1.4.

Our next result, which is essentially a restatement of Theorem A, can be seen as a generalization of Pansu’s Theorem 1.6 to the setting of certain random metrics.
Recall that a metric $d$ on a metric space $M$ is called **inner** if given $\epsilon > 0$ there is $R < \infty$ so that for any $p, q \in M$ one can find $n \in \mathbb{N}$ and $p_0, \ldots, p_n$ so that: $p_0 = p$, $p_n = q$, $d(p_{i-1}, p_i) < R$ for $1 \leq i \leq n$, and

$$d(p_0, p_1) + d(p_1, p_2) + \cdots + d(p_{n-1}, p_n) \leq (1 + \epsilon) \cdot d(p, q).$$

**Theorem B.**

Let $\Gamma$ be a finitely generated virtually nilpotent group, $\Gamma \curvearrowright (X, m)$ an ergodic p.m.p. action, and let $\{d_x \mid x \in X\}$ be a measurable family of inner metrics on $\Gamma$ that is right-equivariant:

$$(1.2) \quad d_x(\gamma_1, \gamma_2) = d_{\gamma.x}(\gamma_1 \gamma^{-1}, \gamma_2 \gamma^{-1}) \quad (\gamma, \gamma_1, \gamma_2 \in \Gamma),$$

and satisfies a uniform bi-Lipschitz estimate $0 < a \leq d_x/d \leq b < \infty$ where $d$ is some right-invariant word metric on $\Gamma$.

Then there exists a right-invariant homogeneous metric $d_\phi$ on $G_\infty$ so that for a.e. $x \in X$ there is Gromov-Hausdorff convergence

$$(\Gamma, \frac{1}{t} \cdot d_x, e) \underset{GH}{\to} (G_\infty, d_\phi, e).$$

Here $d_\phi(g_1, g_2) = \Phi(g_2 g_1^{-1})$ with $\Phi$ from Construction I.7 corresponding to

$$(1.3) \quad c(\gamma, x) := d_x(e, \gamma).$$

One can also start from a sub-additive cocycle $c : \Gamma \times X \to \mathbb{R}_+$ and define

$$(1.4) \quad d_x(\gamma_1, \gamma_2) := c(\gamma_2 \gamma_1^{-1}, \gamma_1, x) \quad (x \in X, \gamma_1, \gamma_2 \in \Gamma).$$

The resulting measurable family of functions is equivariant (as in [1.2]), and each is a (possibly asymmetric) metric on $\Gamma$; condition $A(2)$ on $c$ corresponds to $d_x$ being inner.

A natural example of an equivariant family of metrics as above appears in the following setting, known as the **First Passage Percolation** model. Fix a Cayley graph $(V, E)$ for $\Gamma$ defined by some finite symmetric generating set $S \subset \Gamma$ (so $V = \Gamma$ and $E = \{\gamma, s\gamma \mid \gamma \in \Gamma, s \in S\}$), and fix a $0 < a < b < \infty$. Define $X := [a, b]^E$ – the space of functions $x : E \to [a, b]$; we think of $x(v, v')$ as the time it takes to cross edge $(v, v') \in E$. Since $\Gamma$ acts by automorphisms on $(V, E)$, it also acts continuously on the compact metric space $X$. Let $m$ be some $\Gamma$-invariant ergodic Borel probability measure on $X$, e.g. the Bernoulli measure $m = \mu^E$ where $\mu$ is some probability measure on $[a, b]$. 
Every $x \in X$ defines the time it takes to cross any given edge $e \in E$ and we can define

$$d_x(v, v') = \inf \left\{ \sum_{i=1}^{n} x_{(v_{i-1}, v_i)} \mid v_0 = v, v_n = v', (v_{i-1}, v_i) \in E \right\}$$

to be the minimal travel time from $v$ to $v'$ in the particular realization $x \in X$ of the configuration of passage times of edges. One is now interested in the asymptotic shape as $T \to \infty$ of the set

$$B^T_x(T) := \{ v \in V \mid d_x(e, v) < T \}$$

of vertices that can be reached from the origin $e \in V$ in time $< T$, for a typical configuration $x \in X$.

**Corollary C.**

*With the notations as above, there exists a homogeneous function $\Phi : G_{\infty} \to \mathbb{R}_+$, given in Construction 1.7, so that for m-a.e. $x \in X$ the sets $B^T_x(T)$ are within $o(T)$-approximation from

$$\{ g \in G_{\infty} \mid \Phi(g) < T \}$$

which is a $\delta_T$ image of a fixed set:

$$B^T_x(T) \sim \{ g \in G_{\infty} \mid \Phi(g) < T \} = \delta_T (\{ g \in G_{\infty} \mid \Phi(g) < 1 \}).$$

Thus $\{ g \in G_{\infty} \mid \Phi(g) < 1 \}$ gives the asymptotic shape of a.e. $B^T_x(T)$ rescaled by $T$ for $T \gg 1$.

It follows from Theorem A that for m-a.e. $x \in X$ for any $\epsilon > 0$ for $T > T(x, \epsilon)$

$$\{ g \in G_{\infty} \mid \Phi(g) < 1 - \epsilon \} \subset \text{scl}_T (B^T_x(T)) \subset \{ g \in G_{\infty} \mid \Phi(g) < 1 + \epsilon \}$$

which is equivalent to the statement of the Corollary.

Let us make some remarks about these results.

(1) I. Banjamini and R. Tessera \[3\] recently established the asymptotic shape theorem for the First Passage Percolation model (Corollary \[3\] for the case of an independent distribution on edges, i.e. the measure $m = \mu^E$. In this framework their result is stronger: the assumption is weaker (rather than compact support the distribution $\mu$ is assumed to have a finite exponential moment) and there is statement of a speed for the convergence to the asymptotic shape. However, the proofs, being based on probabilistic techniques, do not seem to apply to the general ergodic case as in Corollary \[3\].

(2) The abelian case $\Gamma = \mathbb{Z}^d$ was considered by Boivin \[6\] in the context of First Passage Percolation as in Corollary \[3\] and then by Björkland \[5\] in the more general context of sub-additive cocycles.
as in Theorem A. Both results are proved under weaker integrability condition, namely $c(\gamma, -) \in L^{d,1}(X, m)$ (Lorentz space). This integrability condition is known to be sharp for sub-additive cocycles over general ergodic $\mathbb{Z}^d$-actions. We note that in no a priori innerness assumption is imposed, but in retrospect it is satisfied.

(3) Assumption (2) in Theorem A (and the corresponding assumption of innerness of metrics in Theorem B) is necessary for the limit object $\Phi$ (and $d_\phi$) to be geodesic. Yet, it will become clear from the proof below that this condition is not needed for the inequality

$$\limsup_{scl(\gamma) \to g} \frac{1}{l} \cdot c(\gamma, x) \leq \Phi(g) \quad (g \in G_\infty)$$

for $m$-a.e. $x \in X$. In fact, the proof of this inequality (see Chapter 2 §3.1) does not require the lower estimate in Theorem A(1); it only uses the inequality $c(\gamma, x) \leq K \cdot |\gamma|_S$, which is equivalent to $c(\gamma, -) \in L^\infty(X, m)$ for $\gamma \in S$ a generating set for $\Gamma$.

(4) It is possible that assumption (1) in Theorem A can be relaxed. Yet, note that already in the Abelian case $\Gamma = \mathbb{Z}^d$ pointwise convergence requires $L^{d,1}(X, m)$-integrability.

(5) Let $\Gamma < G$ and $G_\infty$ be as above. Theorems A and B show that asymptotic shapes are classified by $\Phi$ (and $d_\phi$) for some unique, possibly asymmetric, norm $\phi : g^a_{\infty} \to \mathbb{R}_+$. The converse also holds: for every asymmetric norm $\phi$ the associated Carnot-Carathéodory $\Phi$ and $d_\phi$ arise as an asymptotic shape for some cocycle over $\Gamma$, in fact from a subadditive function $F : \Gamma \to \mathbb{R}_+$. However, the question of which asymptotic shapes (equivalently norms) can appear in First Passage Percolation with independent distribution on edges remains widely open.

We would like to emphasize the following remark.

**Remark 1.8.** An important example of subadditive cocycles over group actions are

$$c(\gamma, x) = \log \|A(\gamma, x)\|$$

where $A : \Gamma \times X \to SL_d(\mathbb{R})$ is a matrix valued cocycle, i.e. satisfies $A(\gamma_1 \gamma_2, x) = A(\gamma_1, \gamma_2, x)A(\gamma_2, x)$. If $\Gamma$ is not Abelian, then the results of this chapter do not apply to such cocycles – they systematically fail the innerness assumption. Yet, for any amenable group $\Gamma$ (in particular, nilpotent) one can describe the asymptotic behavior of such cocycles: they are asymptotically equivalent to a homogeneous subadditive function, namely the pull-back of a norm $\phi$ on the abelianization $\Gamma_1^{ab} \otimes \mathbb{R}$.
for some finite index subgroup $\Gamma_1 < \Gamma$. More precisely, the norm has the form

$$\phi(\gamma) = \max_{1 \leq j \leq d} |\chi_j(\gamma^{ab})|$$

for some characters $\chi_1, \ldots, \chi_d : \Gamma_1^{ab} \otimes \mathbb{R} \to \mathbb{R}$. In particular, if $\Gamma$ is a non-abelian nilpotent group, such homogeneous functions do not grow along the commutator subgroup unlike Carnot-Carathéodry metrics. This can be shown by applying a form of Zimmer’s Cocycle Reduction lemma (using the fact that $\Gamma$ is amenable) that allows one to bring the cocycle to an upper triangular form and read off the growth from the diagonal.

The second part of this thesis uses the insight developed in the ergodic theory of nilpotent groups above to address a geometric question about nilpotent groups. Studying finitely generated groups up to the equivalence relation of quasi-isometry has been an impressively active research area in recent years with surprising connections across mathematics \[12, 14, 21, 22\]. A particularly interesting line of inquiry is determining when certain algebraic or analytic properties of groups imply some sort of rigidity of their quasi-isometry classes - i.e. that every group in their quasi-isometry class shares the algebraic or analytic property.

A motivating example of quasi-isometry is that any two cocompact lattices in the same Lie group are quasi-isometric. Recall again that every finitely generated torsion-free nilpotent group $\Gamma$ embeds as a cocompact lattice in a unique connected, simply connected nilpotent Lie group called the Mal’cev completion of $\Gamma$. Thus, one source of quasi-isometry between finitely generated nilpotent groups is that they have isomorphic Mal’cev completions. A major open question is to determine if this is the only source of quasi-isometry between nilpotent groups.

**Question 1.9.** If finitely generated nilpotent groups $\Gamma$ and $\Lambda$ are quasi-isometric, then are their Mal’cev completions isomorphic?

In \[25\] Pansu proved the following seminal partial answer to this question.

**Theorem 1.10 (Pansu \[25\]).** Finitely generated quasi-isometric nilpotent groups have isomorphic associated Carnot groups.

His proof relies on the following, independently interesting, Rademacher-type differentiation theorem for Carnot spaces.
Theorem 1.11 (Pansu [25]). A bi-Lipschitz map between Carnot groups is differentiable almost everywhere. Moreover, the derivative induces a group isomorphism.

Pansu deduces Theorem 1.10 from Theorems 1.11 and 1.6 as follows. By Theorem 1.6, the asymptotic cones of nilpotent groups are Carnot spaces. But a quasi-isometry between groups induces a bi-Lipschitz between their asymptotic cones. Therefore, invoking Theorem 1.11 yields Theorem 1.10.

The main result of the remainder of this thesis is an alternative, ergodic theoretic proof of Theorem 1.10. The idea of the proof will be to consider the equivalence relation of Integrable Measure Equivalence on finitely generated groups which is, for nilpotent groups, finer, and which lends itself naturally to the ergodic theoretic setting. We will then use ergodic theory to prove that the associated Carnot groups is an invariant of Integrable Measure Equivalence.

Measure equivalence is an equivalence relation on groups introduced by Gromov [13] that is a measure-theoretic parallel of quasi-isometry. It has been the object of considerable study: Furman’s survey [11] provides a thorough overview. However, a fundamental result of Ornstein and Weiss [23] implies that measure equivalence collapses all amenable groups into one equivalence class.

A measure equivalence between two groups implicitly defines a pair of measurable cocycles over probability measure preserving actions of those groups. In their study of rigidity of hyperbolic lattices, Bader, Furman and Sauer have sharpened measure equivalence to a finer equivalence relation, called integrable measure equivalence, by considering only those measure equivalences for which these cocycles satisfy an integrability condition.

Recently Austin and Bowen [1] showed that the single ME class of infinite amenable groups splits into many IME classes. Bowen showed that the growth type of a group is preserved by IME, and Austin used Bowen’s result to prove the following.

Theorem 1.12 (Austin [1]). Finitely generated integrable measure equivalent nilpotent groups have bi-Lipschitz asymptotic cones.

Notice that combining Theorems 1.11 and 1.12 one deduces the IME analog of Theorem 1.10.

Theorem 1.13 (Austin [1], Cantrell). Finitely generated integrable measure equivalent nilpotent groups have isomorphic associated Carnot groups.

However this proof is not entirely satisfying as it does not ‘see’ the group isomorphism through the IME. In his proof, Austin considers the measurable cocycle as an equivariant family of random maps.
between the finitely generated groups that induces a sequence of measurable maps \( \kappa_{x,n} \) between the associated Carnot groups indexed by the rescaling \( 1/n \) in the asymptotic cone construction. He then proves that with high probability a subsequence of these maps converge to a bi-Lipschitz map between the Carnot groups. Austin then asks the natural question (Question 5.2 \[1\]): Is there a bi-Lipschitz group isomorphism between the Carnot groups to which this sequence of random maps converge with high probability on bounded sets? We answer this question in the affirmative.

**Theorem D.**

Suppose \( \Gamma \) and \( \Lambda \) are IME f.g. nilpotent groups with associated Carnot groups \( G_\infty \) and \( H_\infty \). Let \( \kappa_{x,n} \) be the maps as in Question 5.2 \[1\]. Then there is a bi-Lipschitz group isomorphism \( \Phi : G_\infty \rightarrow H_\infty \) to which \( \kappa_{x,n} \) converge on bounded sets with high probability as \( n \rightarrow \infty \)

\[ \kappa_{x,n} \rightarrow \Phi. \]

**Remarks 1.14.**

1. Theorem 1.14 implies Pansu's Theorem 1.10. Indeed, Shalom \[27\] keenly observed that amongst finitely generated amenable groups, quasi-isometry implies uniform measure equivalence, which in particular implies IME. While we do not rely logically on Theorem 1.11, we do use the idea of the Pansu derivative.

2. Our proof is independent of Austin's.

3. One might say that the isomorphism \( \Phi \) is the Pansu derivative of the given measurable cocycle. Indeed, in the deterministic case \( \Phi \) is the usual Pansu derivative.

4. Theorem 1.14 is for any Carnot-Carathéodory metrics on \( G_\infty \) and \( H_\infty \). All Carnot-Carathéodory metrics on a given Carnot group are bi-Lipschitz, so in what follows we may not specify the metric. Moreover \( \Phi \) being a group isomorphism implies it is bi-Lipschitz.

Theorem 1.14 is an immediate consequence of Theorem 1.15, which has the spirit of a nilpotent-valued cocycle ergodic theorem.

**Theorem E.**

Let \( \Gamma, \Lambda \) be f.g. IME nilpotent groups with associated cocycles \( \alpha : \Gamma \times X \rightarrow \Lambda \) and \( \beta : \Lambda \times Y \rightarrow \Gamma \), and let \( G_\infty \) and \( H_\infty \) be the associated Carnot groups of \( \Gamma \) and \( \Lambda \). Then there exists a bi-Lipschitz group isomorphism \( \Phi : G_\infty \rightarrow H_\infty \) so that for all \( g \in G_\infty \)
\[(n, \gamma_n) \longrightarrow g \quad \text{implies} \quad (n, \alpha(\gamma_n, x)) \longrightarrow \Phi(g)\]

where the convergence is in the sense of the asymptotic cone, and the second convergence is in measure. The same is true after exchanging the roles of \(\Gamma, \Lambda, \alpha, \beta, \text{ and } \Phi, \Phi^{-1}\).

See Chapter 2 §1.1 for the definition of convergence in the asymptotic cone.

Remarks 1.15.

(1) Convergence in measure is the best one can hope for given the \(L^1\) integrability assumption.

As remarked above, to have pointwise convergence even in case \(\Gamma = \Lambda = \mathbb{Z}^d\) one must assume \(L^{d,1}\) (Lorentz-space) integrability. The correct integrability assumption for pointwise convergence of ergodic theorems for nilpotent groups is commonly believed to be related to the growth type of the group.

(2) All of the theorems stated above are true for f.g. polynomial growth groups, which by [12] are those groups with finite index nilpotent subgroups. Theorem E is insensitive to finite index and finite kernels, so we reduce to the torsion-free nilpotent case. See [2,5].

We remark that while it is almost immediate that the limiting map \(\Phi\) is a homomorphism, the nilpotent Poincaré recurrence Lemma 5.9 is needed to show that \(\Phi\) has (the obvious candidate as) an inverse.

In Chapter II we prove Theorems A, B and Corollary C. In Chapter III we prove Theorems D and E.
CHAPTER 2

Asymptotic Shapes for Ergodic Families of Metrics on Nilpotent Groups

Let $\Gamma$ be a finitely generated virtually nilpotent group. The topic of this chapter may be viewed from three slightly different perspectives:

Theorem A: As a Subadditive Ergodic Theorem over a general ergodic probability measure preserving action $\Gamma \acts (X, m)$. Given a measurable function $c : \Gamma \times X \to \mathbb{R}$, satisfying
$$c(\gamma_1 \gamma_2, x) \leq c(\gamma_1, \gamma_2.x) + c(\gamma_2, x) \quad (\gamma_1, \gamma_2 \in \Gamma),$$
and some additional conditions, we show that for a.e. $x \in X$ there is a unique limit to $c(\gamma, x)$ suitably normalized; the limit is described on the Carnot group $G_\infty$ using a Carnot-Carathéodory construction.

Theorem B: As a generalization of the result of Pansu [24] showing that the asymptotic cone of an invariant inner metric $d$ on $\Gamma$ is the Carnot group $G_\infty$ (the graded nilpotent Lie group associated with the Mal’cev completion $G$ of $\Gamma$) equipped with a certain Carnot-Carathéodory metric $d_\infty$. Here we show that if one replaces a single invariant metric $d$ by an equivariant ergodic family $\{d_x : x \in X\}$ of inner metrics on $\Gamma$, then a.e. $(\Gamma, d_x, e)$ has the same asymptotic cone which is the Carnot group $G_\infty$ equipped with a fixed Carnot-Carathéodory metric associated to certain averages of the family $\{d_x : x \in X\}$.

Corollary C: As a result about the asymptotic shape in the First Passage Percolation model over $\Gamma$ driven by a general ergodic process $\Gamma \acts (X, m)$. (The case of independent times was recently studied by Benjamini and Tessera [3].)

Outline of the Chapter.

In Section [1] we recall some background on graded nilpotent Lie groups, the Carnot-Carathéodory
construction, Pansu’s fundamental result on the asymptotic cone of nilpotent groups, and the construction of $\phi$, $\Phi$ and $d_\phi$ associated with the sub-additive cocycle $c : \Gamma \times X \to \mathbb{R}_+$. Section 2 contains two basic preliminary results needed for the proofs of the main theorems. One result concerns approximation of admissible curves in the asymptotic cone $G_\infty$ by expressions of the form $T_{n_k} \cdots T_{n_2} T_{n_1}$ that we call \textit{polygonal paths} in $\Gamma$. The second result (Theorem 2.3) is of independent interest; it is an ergodic theorem for sub-additive cocycles along above mentioned polygonal paths. With these preparations at hand we prove Theorems A and B in Section 3.

1. The Carnot group as the asymptotic cone

In this section we recall Pansu’s construction of the asymptotic cone $(G_\infty, d_\infty)$ of a finitely generated nilpotent group and give our construction of $(G_\infty, d_\phi)$, the almost sure asymptotic cone of the random (pseudo) metric space $(\Gamma, d_x)$.

1.1. The graded Lie algebra/group.

Let $\Gamma$ be a finitely generated, torsion-free, nilpotent group and $G$ be its Mal’cev completion. In this subsection we construct the associated Carnot group. Since the Lie groups here are connected and simply connected, one can work with the Lie algebras. Let $\mathfrak{g}$ be the Lie algebra of $G$, and set

$$\mathfrak{g}^1 := \mathfrak{g}, \quad \mathfrak{g}^{i+1} := [\mathfrak{g}, \mathfrak{g}^i].$$

Being nilpotent, $G$ satisfies $\mathfrak{g}^{r+1} = \{0\}$ for some $r \in \mathbb{N}$. Since $[\mathfrak{g}^i, \mathfrak{g}^j] \subset \mathfrak{g}^{i+j}$ (and in particular $[\mathfrak{g}^{i+1}, \mathfrak{g}^j], [\mathfrak{g}^i, \mathfrak{g}^{j+1}] \subset \mathfrak{g}^{i+j+1}$) the Lie bracket on $\mathfrak{g}$ defines a bilinear map

$$\left( \mathfrak{g}^i / \mathfrak{g}^{i+1} \right) \otimes \left( \mathfrak{g}^j / \mathfrak{g}^{j+1} \right) \rightarrow \left( \mathfrak{g}^{i+j} / \mathfrak{g}^{i+j+1} \right),$$

which can then be used to define the Lie bracket $[-, -]_\infty$ on

$$\mathfrak{g}_\infty := \bigoplus_{i=1}^r \mathfrak{v}_i, \quad \text{where} \quad \mathfrak{v}_i := \mathfrak{g}^i / \mathfrak{g}^{i+1} \tag{1.1}$$

by extending the above maps linearly. The resulting pair $(\mathfrak{g}_\infty, [-, -]_\infty)$ is called the \textbf{graded Lie algebra} associated to $\mathfrak{g}$. Note that the linear maps

$$\delta_t : \mathfrak{g}_\infty \rightarrow \mathfrak{g}_\infty, \quad \delta_t(v_1, \ldots, v_r) = (t \cdot v_1, t^2 \cdot v_2, \ldots, t^r \cdot v_r),$$

satisfy $\delta_t([v, w]_\infty) = [\delta_t(v), \delta_t(w)]_\infty$ and $\delta_{ts} = \delta_t \circ \delta_s$ for $v, w \in \mathfrak{g}_\infty$, $t, s > 0$. Hence $\{\delta_t \mid t > 0\}$ is a one-parameter family of automorphisms of the Lie algebra $\mathfrak{g}_\infty$, and therefore define a one-parameter
family of automorphisms of the Lie group $G_{\infty} := \exp_{\infty}(g_{\infty})$, that we will still denote by $\{\delta_t | t > 0\}$. (Here we denote the exponential map $g_{\infty} \to G_{\infty}$ by $\exp_{\infty}$ to distinguish it from $\exp : g \to G$).

The graded Lie algebra naturally appears in the following limiting procedure. Choose a splitting of $g$ as a direct sum of vector subspaces

$$g = V_1 \oplus \cdots \oplus V_r,$$

so that $g^i = V_i \oplus \cdots \oplus V_r,$

and choose a vector space identification $L : g \to g_{\infty}$ so that $L(V_i) = v_i$ the $i$th summand of $g_{\infty}$. For $t > 0$ define the vector space automorphism $\sigma_t$ of $g$ by setting $\sigma_t(v) = t^i \cdot v$ for $v \in V_i (i = 1, \ldots, r)$. Then the Lie brackets $[-,-]_t$ on $g$, given by

$$[v, w]_t := \sigma_1^t ([\sigma_t(v), \sigma_t(w)],$$

defines a Lie algebra structure on $g$ that is isomorphic to the original $[-,-] = [-,-]_1$ via $\sigma_t$.

However, one has

$$[L(v), L(w)]_\infty = \lim_{t \to \infty} [v, w]_t$$

due to the fact that for $v \in V_i$, $w \in V_j$ the “leading term” of $[v, w]$ lies in $V_{i+j}$, while the higher terms that belong to $V_{i+j+1} \oplus \cdots \oplus V_r$ become insignificant under the rescaling (see [24]).

Using the $\log : G \to g$ and $\exp_{\infty} : g_{\infty} \to G_{\infty}$ maps we obtain a family of maps

$$scl_t(\cdot) : \Gamma \subseteq G \xrightarrow{\log} g \xrightarrow{\sigma_t^{-1}} g \xrightarrow{L} g_{\infty} \xrightarrow{\exp_{\infty}} G_{\infty} \quad (t > 0)$$

that explains the asymptotic cone description of Pansu [24] as follows. Let $d$ be an inner right-invariant metric $d$ on $\Gamma$ and

$$(\Gamma, \frac{1}{t} \cdot d, e) \xrightarrow{GH} (G_{\infty}, d_{\infty}, e)$$

the Gromov-Hausdorff convergence. Then a sequence $\gamma_i \in \Gamma$, rescaled by $t_i^{-1}$ with $t_i \to \infty$ as $i \to \infty$, converges to $g \in G_{\infty}$ iff $scl_t(\gamma_i) \to g$ in $G_{\infty}$. We shall often write

$$g = \lim_{i \to \infty} \frac{1}{t_i} \cdot \gamma_i$$

instead of $scl_t(\gamma_i) \to g$.

The metric part of the statement shows that for $t_i \to \infty$ and $\gamma_i, \gamma_i' \in \Gamma$

$$g = \lim_{i \to \infty} \frac{1}{t_i} \cdot \gamma_i, \quad g' = \lim_{i \to \infty} \frac{1}{t_i} \cdot \gamma_i' \quad \implies \quad d_{\infty}(g, g') = \lim_{i \to \infty} \frac{1}{t_i} \cdot d(\gamma_i, \gamma_i').$$
1. THE CARNOT GROUP AS THE ASYMPOTIC CONE

The limiting distance $d_\infty$ on $G_\infty$ is **homogeneous** in the sense that

$$d_\infty(\delta_s(g), \delta_s(g')) = s \cdot d_\infty(g, g') \quad (g, g' \in G_\infty, \ s > 0).$$

This distance is right-invariant (this follows from Lemma 1.2). The distance $d_\infty$ appears in the sub-Finsler Carnot-Carathéodory construction discussed below. Meanwhile let us point out two Lemmas.

**Lemma 1.1.** For $\gamma \in \Gamma$ one has

$$\lim_{n \to \infty} \frac{1}{n} \cdot \gamma^n = \exp_\infty(L \circ \pi \circ \log(\gamma)) = \exp_\infty(\pi_\infty \circ L \circ \log(\gamma)),$$

where $\pi : g \to V_1$ and $\pi_\infty : g_\infty \to v_1$ are the linear projection corresponding to (1.2), (1.1).

**Proof.** Denote by $\pi_k : g \to V_k$ ($k = 1, \ldots, r$) the linear projections according to (1.2), so $\pi = \pi_1$. Then

$$\frac{1}{n} \cdot \gamma^n = \exp_\infty\left(\sum_{k=1}^{r} \frac{1}{n^k} \cdot L \circ \pi_k \circ \log(\gamma^n)\right) = \exp_\infty\left(L \circ \pi_1 \circ \log(\gamma) + \sum_{k=2}^{r} \frac{1}{n^k-1} \cdot L \circ \pi_k \circ \log(\gamma)\right)$$

and, since $n^{-k+1} \cdot L \circ \pi_k \circ \log(\gamma) \to 0$ for $2 \leq k \leq r$, the statement is clear. □

**Lemma 1.2.**

Given sequences $t_i \to \infty$, $\gamma_i, \gamma'_i \in \Gamma$ with $\frac{1}{t_i} \cdot \gamma_i \to g$ and $\frac{1}{t_i} \cdot \gamma'_i \to g'$ then $\frac{1}{t_i} \cdot \gamma_i \gamma'_i \to gg'$.

**Proof.** This follows from the Baker-Campbell-Hausdorff formula (cf. §3.3 and the proof of Lemma 5.5 in [8]). □

### 1.2. Carnot-Carathéodory constructions.

We follow [24] (17)-(20)]. Denote by $g_\infty^{ab}$ the abelianization of the graded Lie algebra $g_\infty$. It is isomorphic to the abelianization $g^{ab}$ of $g$, and can also be identified with the direct summand $v_1$ of $g_\infty$:

$$g_\infty^{ab} \cong g^{ab} \cong v_1 < \bigoplus_{i=1}^{r} v_i = g_\infty.$$

Vectors in $v_1 < g_\infty$ are called **horizontal**. A tangent vector $v \in T_g G_\infty$ at $g \in G_\infty$ is horizontal if its right-translate under $g^{-1}$ is in $v_1 < g_\infty = T_e G_\infty$. Hence the horizontal vectors form a sub-bundle of the tangent bundle $TG_\infty$; this is a totally non-integrable sub-bundle because $g_\infty$ is generated as a Lie algebra by $v_1$. Let us say that a continuous piecewise smooth curve $\xi : [a, b] \to G_\infty$ whose
tangent vectors $\xi'(t)$ are horizontal for Lebesgue a.e. $t \in [a,b]$ are admissible. Any two points $g_1, g_2 \in G_\infty$ can be connected by an admissible curve – this follows from total non-integrability of the sub-bundle of horizontal vectors by Chow’s theorem.

Let $\phi : g^{ab}_\infty \to \mathbb{R}_+$ be an asymmetric norm (or rather a not necessarily symmetric norm), that is assume $\phi$ satisfies for all $v, w \in g^{ab}_\infty$, $t > 0$, and some $0 < a \leq b < \infty$:

$$\begin{align*}
\phi(v + w) &\leq \phi(v) + \phi(w), \\
\phi(t \cdot v) &= t \cdot \phi(v), \\
a \cdot \|v\| &\leq \phi(v) \leq b \cdot \|v\|
\end{align*}$$

(1.5)

for some reference Euclidean norm $\| - \|$. Such an asymmetric norm $\phi$ can be used to measure horizontal vectors in $TG_\infty$ by right-translating them back to $v_1 < g_\infty = T_e G_\infty$. Given a curve $\xi : [\alpha, \beta] \to G_\infty$ as above its $\phi$-length is defined to be

$$(1.6) \quad \text{length}_\phi(\xi) := \int_\alpha^\beta \phi(\xi'(t)\xi(t)^{-1}) \, dt.$$  

We define the $\phi$-distance by

$$d_\phi(g_1, g_2) := \inf \{ \text{length}_\phi(\xi) \mid \xi \text{ is an admissible curve from } g_1 \text{ to } g_2 \}.$$  

Starting from a fixed Euclidean norm $\| - \|$ on $g^{ab}_\infty$, one obtains the sub-Riemannian metric $d_{\| - \|}$ on $G_\infty$, also known as a Carnot-Carathéodory metric; it is right-invariant, homogeneous with respect to the homotheties $\{ \delta_t \mid t > 0 \}$, and defines the usual topology on $G_\infty$.

For a general asymmetric norm $\phi : g^{ab}_\infty \to \mathbb{R}_+$ as in (1.5) we obtain

$$d_\phi : G_\infty \times G_\infty \to \mathbb{R}_+,$$

that is a right-invariant, homogeneous, asymmetric metric, bi-Lipschitz to a Carnot-Carathéodory metric:

$$\begin{align*}
d_\phi(g_1 g, g_2 g) &= d_\phi(g_1, g_2), \\
d_\phi(\delta_t(g_1), \delta_t(g_2)) &= t \cdot d_\phi(g_1, g_2), \\
d_\phi(g_1, g_2) &\leq d_\phi(g_1, h) + d_\phi(h, g_2), \\
a \cdot d_{\| - \|}(g_1, g_2) &\leq d_\phi(g_1, g_2) \leq b \cdot d_{\| - \|}(g_1, g_2).
\end{align*}$$

(1.7)
Being right-invariant $d_\phi$ is completely determined by the function

$$\Phi : G_\infty \to \mathbb{R}, \quad \Phi(g) := d_\phi(e, g), \quad d_\phi(g_1, g_2) = \Phi(g_2g_1^{-1}).$$

This function $\Phi$ is sub-additive, homogeneous, and bi-Lipschitz to a Carnot-Carathéodory norm

$$\Phi(\delta_t(g)) = t \cdot \Phi(g),$$

(1.8)

$$\Phi(g_1g_2) \leq \Phi(g_1) + \Phi(g_2),$$

$$a \cdot d_{\parallel-\parallel}(e, g) \leq \Phi(g) \leq b \cdot d_{\parallel-\parallel}(e, g).$$

If $\phi$ is actually a norm, i.e. $\phi(-v) = v$, then $\Phi$ and $d_\phi$ are also symmetric: $\Phi(g^{-1}) = \Phi(g)$ and $d_\phi(g_1, g_2) = d_\phi(g_2, g_1)$. In this case $d_\phi$ is a sub-Finsler Carnot-Carathéodory metric on $G_\infty$ defined by the norm $\phi$. Hereafter we shall use the term **Carnot-Carathéodory metric** (or just a **CC-metric**) when referring to a possibly asymmetric $d_\phi$ associated to $\phi$ as in (1.5). Pansu’s metric $d_\infty$ on $G_\infty$, referred to in the previous section, is the Carnot-Carathéodory metric associated to a certain norm on $g_\infty$, that itself is determined by the given inner right-invariant metric $d$ on $\Gamma$ [24].

The proof does not really use the symmetry assumption, so it can be applied almost verbatim to asymmetric norms. The infimum in the definition of $d_\phi(g_1, g_2)$ is achieved by a (unique) curve, that will be called a $d_\phi$-**geodesic**. But we shall use this fact only in reference to $d_\infty$ (or the classical $d_{\parallel-\parallel}$).

The notion of $\phi$-length can be extended to curves $\xi : [0, 1] \to G_\infty$ that are $d_\infty$-**rectifiable**, i.e. ones for which

$$\sup \left\{ \sum_{j=1}^{n} d_\infty(\xi(s_{j-1}), \xi(s_j)) \mid n \in \mathbb{N}, \ 0 = s_0 < s_1 < \cdots < s_n = 1 \right\} < +\infty.$$

Pansu shows [24] that such a curve is absolutely continuous, a.e. differentiable on $[0, 1]$, and that its derivative is a.e. horizontal, so the integral (1.6) makes sense. The $\phi$-length of such curves can also be defined by

$$\text{length}_\phi(\xi) = \sup \left\{ \sum_{j=1}^{n} d_\phi(\xi(s_{j-1}), \xi(s_j)) \mid n \in \mathbb{N}, \ 0 = s_0 < s_1 < \cdots < s_n = 1 \right\}.$$
1.3. From a sub-additive function $F : \Gamma \to \mathbb{R}_+$ to a CC-metric on $G_\infty$.

Consider a sub-additive function $F : \Gamma \to \mathbb{R}_+$ that is bi-Lipschitz to a word metric, i.e. satisfies
\begin{equation}
F(\gamma_1 \gamma_2) \leq F(\gamma_1) + F(\gamma_2),
\end{equation}
\begin{equation}
a \cdot d(e, \gamma) \leq F(\gamma) \leq b \cdot d(e, \gamma),
\end{equation}
for some constants $0 < a \leq b < \infty$. Note that the upper linear bound $F(\gamma) \leq bd(e, \gamma)$ follows automatically from subadditivity and the fact that $\Gamma$ is finitely generated; so the content of the second assumption is the lower linear bound for $F : \Gamma \to \mathbb{R}_+$.

Such a function induces a subadditive function
\[ f : \Gamma^{ab} \to \mathbb{R}_+ \]
using the following general construction.

**Lemma 1.3.** Let $1 \to \Delta \to \Gamma \to \Lambda \to 1$ be a short exact sequence of groups, and $F : \Gamma \to \mathbb{R}_+$ a subadditive function. Then the function
\[ f : \Lambda \to \mathbb{R}_+ \quad \text{defined by} \quad f(\gamma \Delta) := \inf\{F(\gamma \delta) \mid \delta \in \Delta\} \]
is subadditive.

**Proof.** Given $\lambda_1, \lambda_2 \in \Lambda$ and $\epsilon > 0$ choose $\gamma_1, \gamma_2 \in \Gamma$ so that $\lambda_i = \gamma_i \Delta$ and $F(\gamma_i) \leq f(\lambda_i) + \epsilon$ for $i = 1, 2$. Then $\lambda_1 \lambda_2 = \gamma_1 \gamma_2 \Delta$, so
\[ f(\lambda_1 \lambda_2) \leq F(\gamma_1 \gamma_2) \leq F(\gamma_1) + F(\gamma_2) \leq f(\lambda_1) + f(\lambda_2) + 2\epsilon. \]
Since $\epsilon > 0$ is arbitrary we get $f(\lambda_1 \lambda_2) \leq f(\lambda_1) + f(\lambda_2)$. \qed

Now recall that $\Gamma$ is a uniform lattice in its Mal’cev completion $G$. In fact, viewing $G$ as the $\mathbb{R}$-points $G = G_\mathbb{R}$ of a $\mathbb{Q}$-algebraic group $G$, we may think of $\Gamma$ as (commensurable to) $G_\mathbb{Z}$. Taking the abelianization is a $\mathbb{Q}$-algebraic operation, hence $\Gamma^{ab}$ is (commensurable to) $G^{ab}_\mathbb{Z}$, a lattice in $G^{ab}_\mathbb{R} = G^{ab}$. Hence $\Gamma^{ab}$, that is abstractly isomorphic to $\mathbb{Z}^d$, is a lattice in $G^{ab}$, that is continuously isomorphic to $\mathbb{R}^d$ with $d = \dim \mathfrak{v}_1$. One often writes
\[ G^{ab} = \Gamma^{ab} \otimes \mathbb{R} \]
to emphasize that $\Gamma^{ab}$ is a lattice in the real vector space $G^{ab}$. 
Lemma 1.4. Let $\Lambda$ be a lattice in a finite dimensional real vector space $V$, and $f : \Lambda \to \mathbb{R}_+$ be a subadditive function. Then there exists a unique homogeneous subadditive function $\phi : V \to \mathbb{R}_+$ so that $f$ is asymptotically equivalent to $\phi|_{\Lambda}$; in particular

$$\phi(\lambda) = \lim_{n \to \infty} \frac{1}{n} f(n\lambda) = \inf_{n \geq 1} \frac{1}{n} f(n\lambda).$$

Moreover, if $c_1 \leq \frac{f(\lambda)}{\|\lambda\|} \leq c_2$ on $\Lambda$, then $c_1 \leq \frac{\phi(v)}{\|v\|} \leq c_2$ on $V \setminus \{0\}$.

This is an easy and well known fact; but see Burago’s [9] for much finer results in case of a coarsely geodesic metric.

Remark 1.5. It follows that any subadditive function $f : \mathbb{Z}^d \to \mathbb{R}_+$ is automatically inner in the following sense: given $\epsilon > 0$ there is $R < \infty$ so that any $\lambda \in \mathbb{Z}^d$ can be written as $\lambda = \lambda_1 + \cdots + \lambda_n$ with

$$f(\lambda_i) \leq R \quad (i = 1, \ldots, n), \quad f(\lambda_1) + \cdots + f(\lambda_n) \leq (1 + \epsilon) \cdot f(\lambda).$$

Indeed, this is clear for the asymmetric norm $\phi : \mathbb{R}^d \to \mathbb{R}_+$ associated with $f$ in Lemma 1.4 and translates to $f$ by the virtue of the approximation.

Lemma 1.6. Let $F : \Gamma \to \mathbb{R}_+$ be a subadditive function, $f : \Gamma^{ab} \to \mathbb{R}_+$ and $\phi : \Gamma^{ab} \otimes \mathbb{R} \to \mathbb{R}_+$ be defined by Lemmas 1.3 and 1.4. Then for any $\gamma \in \Gamma$ one has

$$\lim_{n \to \infty} \frac{1}{n} F(\gamma^n) = \inf_{n \geq 1} \frac{1}{n} F(\gamma^n) = \lim_{n \to \infty} \frac{1}{n} f((\gamma^{ab})^n) = \inf_{n \geq 1} \frac{1}{n} f((\gamma^{ab})^n) = \phi(\gamma^{ab})$$

for any $\gamma \in \Gamma$ with $\gamma^{ab} = [\Gamma, \Gamma]$ denoting the image in $\Gamma^{ab}$.

Proof. The sequence $a_n = F(\gamma^n)$ is sub-additive, i.e. $a_{n+m} \leq a_n + a_m$ for all $n, m \in \mathbb{N}$; hence $a_n/n$ converges to $\inf a_n/n$. Next we note that Lemmas 1.1 and relation 1.4 imply that whenever $\gamma_1, \gamma_2 \in \Gamma$ satisfy $\gamma_1^{ab} = \gamma_2^{ab}$ one has

$$\lim_{n \to \infty} \frac{1}{n} \cdot d(\gamma_1^n, \gamma_2^n) = 0.$$ 

Since any sub-additive function is automatically Lipschitz with respect to the word metric, it follows that $\lim F(\gamma^n)/n = \lim F(\gamma_2^n)/n$. Thus this limit of $F(\gamma^n)/n$ depends only on $\gamma^{ab}$, and is easily seen to be $\lim f((\gamma^{ab})^n)/n$, i.e. $\phi(\gamma^{ab})$. □
We can now apply the Carnot-Carathéodory construction to define a (possibly asymmetric) metric $d_\phi$ on $G_\infty$ by

\[ d_\phi(g, g') := \inf \{ \text{length}_\phi(\xi) \mid \xi \text{ is an admissible curve from } g \text{ to } g' \}. \]

### 1.4. From a cocycle $c : \Gamma \times X \to \mathbb{R}_+$ to the CC-metric.

Let $\Gamma$ be a finitely generated, virtually nilpotent group, $\Gamma \curvearrowright (X, m)$ an ergodic p.m.p. action, and $c : \Gamma \times X \to \mathbb{R}_+$ a subadditive cocycle with $c(\gamma, -) \in L^\infty(X, m)$ for every $\gamma \in \Gamma$. We start with a couple of remarks about passing to finite index subgroups and dividing by finite kernels.

Let $\Gamma' < \Gamma$ be a subgroup of finite index. The action of $\Gamma'$ on $(X, m)$ has at most $[\Gamma : \Gamma']$-many ergodic components permuted by the $\Gamma$-action. Let $c' : \Gamma' \times X' \to \mathbb{R}_+$ be the restriction of $c$ to one of the $\Gamma'$-ergodic components $X' \subset X$. If one shows that there is some function $\Phi : G_\infty \to \mathbb{R}_+$ so that for a.e. $x' \in X'$ the function $c'(-, x') : \Gamma' \to \mathbb{R}_+$ is asymptotically equivalent to $\Phi$, then the same would apply to $c(-, x) : \Gamma \to \mathbb{R}_+$ for a.e. $x \in X$. Indeed choosing representatives $\gamma_1, \ldots, \gamma_n$ for $\Gamma' \backslash \Gamma$ one can write

\[ c(\gamma, x) = c(\gamma' \gamma_i, x) \leq c(\gamma', \gamma_i, x) + \|c(\gamma_i, x)\|_\infty \]

for some $\gamma' \in \Gamma'$; similarly

\[ c(\gamma', x) = c(\gamma \gamma_i^{-1}, x) \leq c(\gamma, \gamma_i^{-1}, x) + \|c(\gamma_i^{-1}, x)\|_\infty. \]

Hence $c(\gamma, x)$ is at uniformly bounded distance from $c(\gamma', \gamma_i^{\pm 1} x)$, and therefore has the same asymptotic behavior.

Let $N$ be a finite normal subgroup of $\Gamma$. Then $\Gamma_1 := \Gamma/N$ acts ergodically by p.m.p. transformations on $(X_1, m_1) := (X, m)/N$. A subadditive cocycle $c : \Gamma \times X \to \mathbb{R}_+$ defines $c_1 : \Gamma_1 \times X_1 \to \mathbb{R}_+$ by

\[ c_1(\gamma_1, x_1) := \max\{c(\gamma, x) \mid \text{pr}(\gamma) = \gamma_1, \text{pr}(x) = x_1\}. \]

Then $c_1 : \Gamma_1 \times X_1 \to \mathbb{R}_+$ is a sub-additive cocycle, and it is within bounded distance from $c(\gamma, x)$.

Furthermore we note that conditions (1) and (2) of Theorem A pass to $c'$ and $c_1$ as above. (Condition (2) for $c'$ is an easy exercise using subadditivity and innerness; the others are immediate.) Hence in the context of Theorem A (and Theorem B) we may assume without loss of generality that $\Gamma$ itself
is finitely-generated, torsion-free, nilpotent group with torsion-free abelianization $\Gamma^{\text{ab}}$. Hereafter we shall make this assumption.

Let us define the function

$$c : \Gamma \to \mathbb{R}_+$$

by setting

$$c(\gamma) := \int_X c(\gamma, x) \, dm(x).$$

Observe that $c$ is a sub-additive function, because sub-additivity of $c$ and $\Gamma$-invariance of $m$ imply

$$c(\gamma_1 \gamma_2) = \int_X c(\gamma_1 \gamma_2, x) \, dm(x) \leq \int_X c(\gamma_1, \gamma_2 x) \, dm(x) + \int_X c(\gamma_2, x) \, dm(x) = c(\gamma_1) + c(\gamma_2).$$

Moreover one always has an upper linear estimate

$$c(\gamma) \leq K_1 \cdot |\gamma|_S$$

with constants $0 < k \leq K_1 < +\infty$ and any $\gamma \in \Gamma$.

**Remark 1.7.** It does not seem to be obvious why the condition of being *inner* for the sub-additive cocycle $c : \Gamma \times X \to \mathbb{R}_+$ (condition (2) in Theorem A) should imply innerness for the average sub-additive function $\bar{c} : \Gamma \to \mathbb{R}_+$. It will follow from our results that for an $L^\infty$-cocycle $c$ over an *ergodic* $\Gamma$-action the average $\bar{c}$ is indeed inner as it is asymptotically equivalent to a Carnot-Carathéodory function $\Phi$.

We can now summarize the construction

**Proposition 1.8.** Let $\Gamma \curvearrowright (X, m)$ and $c : \Gamma \times X \to \mathbb{R}_+$ be a subadditive cocycle satisfying condition (1) in Theorem A. Then:

- The average function
  $$\bar{c}(\gamma) := \int_X c(\gamma, x) \, dm(x)$$
  is a subadditive function on $\Gamma$, satisfying bi-Lipschitz condition (1.11).

- This subadditive function defines a subadditive, homogeneous $\phi : \Gamma^{\text{ab}} \otimes \mathbb{R} \to \mathbb{R}_+$, such that
  $$\lim_{n \to \infty} \frac{1}{n} \bar{c}(\gamma^n) = \phi(\gamma^{\text{ab}}) \quad (\gamma \in \Gamma).$$
Moreover, for some $0 < a \leq b < \infty$ one has $a \cdot \|v\| \leq \phi(v) \leq b \cdot \|v\|$ for all $v \in \Gamma^{ab} \otimes \mathbb{R} \cong g^{ab}$.

• The Carnot-Carathéodory construction defines an asymmetric distance on $G_\infty$

$$d_\phi : G_\infty \times G_\infty \to \mathbb{R}_+$$

that is right-invariant, homogeneous, and bi-Lipschitz to $d_\infty$ as in (1.7). We denote

$$\Phi(g) := d_\phi(e, g) \quad (g \in G_\infty).$$

2. Preparation for the main proofs

In this section we prepare two tools for the proof of the main results. The first tool is a purely geometric fact that allows one to approximate an admissible curve in the asymptotic cone $G_\infty$ of $\Gamma$ by rescaled sequences of the form $T^n_k \cdots T^n_2 T^n_1$ in $\Gamma$; we call such sequences polygonal paths. The second tool is an ergodic theorem for a sub-additive cocycle along polygonal paths over a general ergodic, p.m.p. action of a nilpotent group.

2.1. Approximating curves in $G_\infty$ by polygonal paths in $\Gamma$.

This subsection concerns purely geometric aspects of the convergence of $\Gamma$ to its asymptotic cone $G_\infty$ (and is unrelated to the action $\Gamma \curvearrowright (X, m)$ and the cocycle $c : \Gamma \times X \to \mathbb{R}_+$).

As before, $\Gamma$ is a finitely generated, torsion-free, nilpotent group with torsion-free abelianization, $d$ is a right-invariant word metric on $\Gamma$, $(G_\infty, d_\infty)$ is the asymptotic cone, and

$$\text{scl}_t (-) : \Gamma \longrightarrow G_\infty \quad (t > 0)$$

are the maps defined in (1.3) that realize the Gromov-Hausdorff convergence

$$(\Gamma, \frac{1}{t} \cdot d, e_\Gamma) \longrightarrow (G_\infty, d_\infty, e).$$

We also fix a (possibly) asymmetric norm

$$\phi : g^{ab}_\infty \longrightarrow \mathbb{R}_+$$

satisfying (1.5) and use it to associate length length$_\phi(\xi)$ to admissible curves $\xi : [0, 1] \to G_\infty$. We denote the balls in $G_\infty$ by

$$B(g, \epsilon) := \{g' \in G_\infty \mid d_\infty(g, g') < \epsilon\}.$$
Proposition 2.1 (Approximation of curves by polygonal paths).

Given a Lipschitz curve $\xi : [0, 1] \to G_\infty$ with $\xi(0) = e$, and $\epsilon > 0$ one can find $k, p, n_0 \in \mathbb{N}$, $T_1, \ldots, T_k \in \Gamma$ so that for $n \geq n_0$ one has:

$$\sum_{j=1}^{k} d_\infty \left( \frac{1}{np} \cdot T^n_j \cdots T^n_2 T^n_1, \xi\left(\frac{j}{k}\right) \right) < \epsilon,$$

and

$$\left| \frac{1}{p} \cdot (\phi(T^n_k) + \cdots + \phi(T^n_1)) - \text{length}_\phi(\xi) \right| < \epsilon.$$

We emphasize the order of the main quantifiers: the elements $T_1, \ldots, T_k$ and $p \in \mathbb{N}$ depend only on the required accuracy $\epsilon > 0$ (and of course the curve $\xi$), and provide $\epsilon$-good approximation at all sufficiently large scales.

We shall need this proposition (in combination with Theorem 2.3) in two cases:

- In § 3.1 we choose $\xi$ to be a $\phi$-geodesic connecting $e$ to some $g$. In this case $\xi$ is a smooth admissible curve and we are interested in the inequality

  $$\frac{1}{p} \cdot (\phi(T^n_k) + \cdots + \phi(T^n_1)) \leq \text{length}_\phi(\xi) + \epsilon = \Phi(g) + \epsilon$$

  while $d_\infty(\frac{1}{np} \cdot T^n_k \cdots T^n_2 T^n_1, g) < \epsilon$.

- In § 3.2 we get a Lipschitz curve $\xi$ connecting $e$ to some $g$. In this case we are interested in the inequality

  $$\frac{1}{p} \cdot (\phi(T^n_k) + \cdots + \phi(T^n_1)) \geq \text{length}_\phi(\xi) - \epsilon \geq \Phi(g) - \epsilon$$

  while requiring

  $$\sum_{j=1}^{k} d_\infty \left( \frac{1}{np} \cdot T^n_j \cdots T^n_2 T^n_1, \xi\left(\frac{j}{k}\right) \right) < \epsilon,$$

  which is stronger than just $d_\infty(\frac{1}{np} \cdot T^n_k \cdots T^n_2 T^n_1, g) < \epsilon$.

Proof of Proposition 2.1

First we work in $G_\infty$. Our goal is to find $k \in \mathbb{N}$ and horizontal vectors

$$v_1, \ldots, v_k \in v_1 \subset \mathfrak{g}_\infty$$
so that, denoting \( h_j := \exp_\infty(v_j) \) one has

\[
(2.1) \quad \sum_{j=1}^k d_\infty(h_j \cdots h_1, \xi(j_{\frac{j}{k}})) < \frac{1}{2} \epsilon, \quad \sum_{j=1}^k |\phi(v_j) - \text{length}_\phi(\xi)| < \frac{1}{2} \epsilon.
\]

For a fixed \( k \in \mathbb{N} \), that we will take to be sufficiently large, we define \( v_1, \ldots, v_k \) inductively as follows:

set \( v_1 := \pi_\infty \circ \log_\infty(\xi(\frac{1}{k})) \), \( h_1 := \exp_\infty(\xi(\frac{1}{k})) \).

Assuming \( v_1, \ldots, v_{j-1} \) were chosen, set

\[
v_j := \pi_\infty \circ \log_\infty(\xi(\frac{1}{k})(h_{j-1} \cdots h_1 \xi(0))^{-1}), \quad h_j := \exp_\infty(v_j).
\]

Here \( \pi_\infty : g_\infty \to v_1 \) is the linear projection corresponding to the decomposition \( g_\infty = \bigoplus_{i=1}^r v_i \).

Let us now show that by choosing \( k \) large enough we can guarantee (2.1). To this end we need the fact (\ref{lemme18}) that in the unit ball in \( G_\infty \) the “horizontal component” gives an approximation with at most quadratic error. More precisely, there is a constant \( C_1 \) so that for all \( g \in B(e, 1) \):

\[
d_\infty(g, \exp_\infty \circ \pi_\infty \circ \log_\infty(g)) \leq C_1 \cdot d_\infty(e, g)^2.
\]

Hence for large \( k \) one has for \( j = 1, \ldots, k \):

\[
d_\infty(h_j \cdots h_1, \xi(j_{\frac{j}{k}})) \leq C_1 \cdot d_\infty(h_{j-1} \cdots h_1, \xi(j_{\frac{j}{k}}))^2\]

\[
\leq C_1 \cdot \left( d_\infty(h_{j-1} \cdots h_1, \xi(j_{\frac{j-1}{k}})) + d_\infty(\xi(j_{\frac{j-1}{k}}), \xi(j_{\frac{j}{k}})) \right)^2 \leq C_2 \cdot \frac{1}{k^2}
\]

for some \( C_2 \) depending on \( C_1 \) and the Lipschitz constant of \( \xi \). Hence for all \( k \) large enough

\[
\sum_{j=1}^k d_\infty(h_j \cdots h_1, \xi(j_{\frac{j}{k}})) < C_2 \cdot k \cdot \frac{1}{k^2} = \frac{C_2}{k} < \frac{1}{2} \epsilon.
\]

The second fact that we want to use is that a Lipschitz curve \( \xi : [0, 1] \to G_\infty \) is rectifiable. Therefore

\[
\text{length}_\phi(\xi) = \lim_{k \to \infty} \sum_{j=1}^k d_\phi(\xi(j_{\frac{j-1}{k}}), \xi(j_{\frac{j}{k}})).
\]

One also has a constant \( C \) so that

\[
|d_\phi(g, g') - \phi \circ \pi_\infty \circ \log_\infty(g' g^{-1})| \leq C \cdot d_\infty(g, g')^2
\]
whenever \( g' \in B(g, 1) \). Thus for all sufficiently large \( k \) and for each \( j = 1, \ldots, k \), we have

\[
\begin{align*}
d_{r}(\xi(\frac{j-1}{k}), \xi(\frac{j}{k})) - \phi(v_j) &\leq d_{r}(h_{j-1} \cdot h_1, \xi(\frac{j}{k})) - \phi(v_j) \\
&+ d_{r}(h_{j-1} \cdot h_1, \xi(\frac{j-1}{k})) \leq C_3 \cdot \frac{1}{k^2}
\end{align*}
\]

and the second inequality in (2.1) follows.

We have now found \( k \in \mathbb{N} \) and horizontal vectors \( v_1, \ldots, v_k \in v_1 \) satisfying (2.1), and need to find \( T_1, \ldots, T_k \in \Gamma, p \in \mathbb{N}, \) and \( n_0 > 0 \) as in the Proposition. We need the following

**Lemma 2.2.**

*Given a horizontal vector \( v \in v_1 < g_\infty \) and \( \epsilon' > 0 \) there exist \( \tau \in \Gamma, p \in \mathbb{N} \) and \( n_0 \) so that

\[
\frac{1}{p}(\tau^{ab}) - \phi(v) < \epsilon', \quad d_{\infty}(\frac{1}{np} \bullet \tau^n, \exp_\infty(v)) < \epsilon' \quad (n > n_0)
\]

where \( \exp_\infty : g_\infty \to G_\infty \) is the exponential map on \( G_\infty \).*

**Proof.** Since \( \frac{1}{p} \bullet \Gamma \) becomes denser and denser in \( G_\infty \) as \( p \to \infty \), one can find \( p \in \mathbb{N} \) and \( \gamma \in \Gamma \) so that \( \frac{1}{p} \bullet \gamma \) is close to \( \exp_\infty(v) \). Recall that

\[
\begin{align*}
\frac{1}{p} \bullet \gamma &= \exp_\infty \left( \frac{1}{p} \cdot L \circ \pi_1 \circ \log(\gamma) + \frac{1}{p^2} \cdot L \circ \pi_2 \circ \log(\gamma) + \cdots + \frac{1}{p^n} \cdot L \circ \pi_r \circ \log(\gamma) \right)
\end{align*}
\]

where \( \pi_j : g \to V_j = L^{-1}(v_j) \) are the linear projections. Since \( v \in v_1 = L(\pi_1(g)) \), it follows that

\[
\begin{align*}
d_{\infty} \left( \exp_\infty \left( \frac{1}{p} \cdot L \circ \pi_1 \circ \log(\gamma) \right), \exp_\infty(v) \right) < \epsilon', \quad \left| \phi \left( \frac{1}{p} \cdot L \circ \pi_1 \circ \log(\gamma) \right) - \phi(v) \right| < \epsilon'.
\end{align*}
\]

Now considering powers \( \frac{1}{n} \bullet \gamma^n \) we are done by applying Lemma 1.1. This proves Lemma 2.2. \( \square \)

Choose \( \epsilon' \in (0, \epsilon/2k) \) small enough to ensure that whenever \( h'_1, \ldots, h'_k \in G_\infty \) are \( \epsilon' \)-close to \( h_1, \ldots, h_k \), respectively, one has

\[
\sum_{j=1}^{k} d_{\infty}(h'_j \cdot h_1, h_j \cdot h_2) < \frac{1}{2} \epsilon.
\]

Let us now apply Lemma 2.2 with \( \epsilon' > 0 \) as above to obtain elements \( \tau_1, \ldots, \tau_k \) and \( p_1, \ldots, p_k \in \mathbb{N} \) so that the pairs \( (\tau_j, p_j) \) satisfy

\[
\frac{1}{p_j} \phi(\tau_j^{ab}) - \phi(v_j) \leq \epsilon' < \frac{\epsilon}{2k}, \quad d_{\infty}(\frac{1}{np_j} \bullet \tau_j^n, h_j) < \epsilon' \quad (j = 1, \ldots, k).
\]
Replacing a pair \((\tau_j, p_j)\) by \((\tau_j^q, q \cdot p_j)\) with any \(q \in \mathbb{N}\), the above inequalities clearly remain valid. So taking \(p := p_1 \cdots p_k\) and replacing \((\tau_j, p_j)\) by \((T_j := \tau_j^{p_j}, p)\) we get elements \(T_1, \ldots, T_k \in \Gamma\) so that for \(n \gg 1\)

\[
d_\infty\left(\frac{1}{np} \cdot T_j^n \cdot h_j, h_j \right) < \epsilon' \quad \text{and} \quad \left| \frac{1}{p} \phi(T_j^{ab}) - \phi(v_j) \right| < \epsilon' < \frac{\epsilon}{2k} \quad (j = 1, \ldots, k).
\]

In view of Lemma 1.2 we know that for every \(j = 1, \ldots, k\)

\[
d_\infty\left(\frac{1}{np} \cdot (T_j^n, T_1^n), \left(\frac{1}{np} \cdot T_j^n \right) \cdots \left(\frac{1}{np} \cdot T_1^n \right)\right) \to 0.
\]

Thus for all \(n\) large enough, we have

\[
\sum_{j=1}^k d_\infty\left(\frac{1}{np} \cdot T_j^n \cdot h_j \cdot h_1 \right) < \frac{1}{2} \epsilon,
\]

while

\[
\left| \frac{1}{p} \cdot \sum_{j=1}^k \phi(T_j^{ab}) - \sum_{j=1}^k \phi(v_j) \right| < \frac{1}{2} \epsilon.
\]

Combined with (2.1) this establishes the required inequalities. This completes the proof of Proposition 2.1.

2.2. An Ergodic Theorem along polygonal paths.

The goal of this subsection is to prove the following result that might have an independent interest.

**Theorem 2.3** (Ergodic theorem along polygonal paths).

Let \(\Gamma\) be a finitely generated torsion-free nilpotent group with torsion-free abelianization, \(\Gamma \actson (X, m)\) an ergodic p.m.p. action, \(c : \Gamma \times X \to \mathbb{R}_+\) a measurable, non-negative, subadditive cocycle with \(c(\gamma, -) \in L^\infty(X, m)\) for every \(\gamma \in \Gamma\), and \(\overline{c}\) and \(\phi\) as above. Then for any \(T_1, \ldots, T_k \in \Gamma\) one has \(m\)-a.e. and \(L^1(X, m)\)-convergence

\[
\lim_{n \to \infty} \frac{1}{n} \cdot c(T_j^n, T_{j-1}^n \cdots T_1^n x) = \phi(T_j^{ab})
\]

for each \(j = 1, \ldots, k\), and consequently

\[
\lim_{n \to \infty} \frac{1}{n} \left( c(T_k^n, T_{k-1}^n \cdots T_1^n x) + \cdots + c(T_2^n, T_1^n x) + c(T_1^n, x) \right) = \phi(T_k^{ab}) + \cdots + \phi(T_1^{ab}).
\]

The \(L^1\) convergence holds under a weaker assumption: \(c(\gamma, -) \in L^1(X, m)\).
The case $k = 1$ was shown by Austin \cite{Austin} under the weaker assumption that $c(\gamma, -) \in L^1(X, m)$ for every $\gamma \in \Gamma$. For reader's convenience we include a proof.

**Theorem 2.4 (Austin \cite{Austin}).**

Let $c: \Gamma \times X \to \mathbb{R}^+$ be a subadditive cocycle with $c(\gamma, -) \in L^1(X, m)$ for every $\gamma \in \Gamma$. Then for any $T \in \Gamma$ one has

$$
\lim_{n \to \infty} \frac{1}{n} c(T^n, x) = \phi(T^{ab})
$$

for $m$-a.e. $x \in X$ and in $L^1(X, m)$.

**Proof.** Kingman’s subadditive ergodic theorem, applied to the sub-additive cocycle $h_n(x) := c(T^n, x)$ over $(X, m, T)$, gives an $m$-a.e. and $L^1$ convergence

$$
\lim_{n \to \infty} \frac{1}{n} c(T^n, x) = h(x),
$$

where $h(x)$ is a measurable $T$-invariant function, satisfying

$$
\int_X h(x) \, dm(x) = \lim_{n \to \infty} \frac{1}{n} \int_X h_n(x) \, dm(x) = \lim_{n \to \infty} \frac{1}{n} \tau(T^n) = \phi(T^{ab}).
$$

(We used Lemma \ref{lemma:subadditive} in the last equality). Fix $\gamma \in \Gamma$ and denote $\gamma_n := T^n \gamma T^{-n}$. Since $\gamma_n T^n = T^n \gamma$ we have

$$
(2.2) \quad c(T^n, x) - c(\gamma_n^{-1}, \gamma_n T^n, x) \leq c(\gamma_n T^n, x) = c(T^n \gamma, x) \leq c(T^n, \gamma, x) + c(\gamma, x).
$$

Denote $f_n(x) = n^{-1}(c(\gamma_n, x) + c(\gamma_n^{-1}, \gamma_n T^n, x))$ and observe that since one has $|\gamma_n^{-1}|_S = |\gamma_n|_S = o(n)$ (cf. Breuillard \cite{Breuillard} Lemma 5.6)

$$
\|f_n\|_1 \leq \frac{K_1}{n} \cdot 2|\gamma_n|_S \to 0, \quad \text{where} \quad K_1 = \max_{s \in S} \|c(s, -)\|_1.
$$

Thus there is a sequence $n_i \to \infty$ so that $f_{n_i}(x) \to 0$ for $m$-a.e. $x \in X$. Dividing (2.2) by $n$, and taking the limit along the subsequence $n_i$, one obtains

$$
h(x) \leq h(\gamma, x).
$$

Since this is true a.e. for every $\gamma \in \Gamma$, $h$ is $\Gamma$-invariant. By ergodicity it is constant. This constant is $\phi(T^{ab})$ by integration. \hfill \square

In the general case of $k \geq 2$ the term $n^{-1} \cdot c(T^n_1, x)$ converges to $\phi(T^{ab}_1)$ by the above, but dealing with the next terms, such as $n^{-1} \cdot c(T^n_2, T^n_1, x)$, one faces a “moving target” problem. We shall overcome
this difficulty by finding regions $Z_2, \ldots, Z_k \subset X$ where $n^{-1}c(T^n, z) = \phi(T^n) + o(n)$ for $z \in Z_k$, and perturbing the polygonal path $T^n_k \cdots T^n_2 T^n_1$ slightly to make sure to land in the appropriate regions at appropriate times. We need several lemmas.

**Lemma 2.5 (Parallelogram inequality).**

*Given* $\alpha, \beta, \tau, \tau' \in \Gamma$ *one has*

$$|c(\tau, \alpha.x) - c(\tau', \beta.x)| \leq K \cdot (d(\alpha, \beta) + d(\tau, \tau'))$$

*for* $K = \max_{s \in S} \|c(s, -)\|_{\infty}$.

**Proof.** Let us write $\beta = \delta \alpha$ and $\tau' = \omega \tau \delta^{-1}$, so

$$|\delta| = |\delta^{-1}| = d(\alpha, \beta), \quad |\omega| = |\omega^{-1}| = d(\tau, \tau').$$

Since $\tau' = \omega \tau \delta^{-1}$ we have

$$c(\tau', \beta.x) = c(\omega \tau \delta^{-1}, \beta.x) \leq c(\omega, \tau \alpha.x) + c(\tau, \alpha.x) + c(\delta^{-1}, \beta.x)$$

$$\leq c(\tau, \alpha.x) + K \cdot (|\omega| + |\delta^{-1}|).$$

Conversely

$$c(\tau, \alpha.x) = c(\omega^{-1} \tau' \delta, \alpha.x) \leq c(\omega^{-1}, \tau' \delta.x) + c(\tau', \delta.x) + c(\delta, \alpha.x)$$

$$\leq c(\tau', \beta.x) + K \cdot (|\omega^{-1}| + |\delta|).$$

□

**Lemma 2.6.** Let $T \in \Gamma$, $\delta > 0$, and a measurable subset $E \subset X$ be given. Then the set

$$E^*: = \left\{ x \in X \mid \lim_{n \to \infty} \inf \frac{\#\{n' < \delta \cdot n \mid T^{n-n'.x} \in E \}}{\delta \cdot n} > 0 \right\}$$

has $m(E^*) \geq m(E)$. Moreover, given $\epsilon > 0$ there is $N$ so that the set

$$E^*_N: = \left\{ x \in X \mid \forall n \geq N, \quad \#\{n' < \delta \cdot n \mid T^{n-n'.x} \in E \} > 0 \right\}$$

has $m(E^*_N) > m(E) - \epsilon$. 

---

2. PREPARATION FOR THE MAIN PROOFS 29
2. Preparation for the Main Proofs

Proof. Given a function $f \in L^1(X, m)$ and integers $1 \leq k < n$ consider the averaged function

$$A_n^k f(x) := \frac{1}{n-k} \cdot \sum_{j=k}^{n-1} f(T^j x).$$

Birkhoff’s pointwise ergodic theorem asserts $m$-a.e. convergence

$$\lim_{n \to \infty} A_n^0 f = \mathbb{E}(f \mid B^T)$$

to the conditional expectation of $f$ with respect to the sub-$\xi$-algebra of $T$-invariant sets $B^T$. (The conditional expectation is defined only up to null sets, but so is the above convergence). We observe that since

$$A_n^0 f(x) = \frac{k}{n} \cdot A_n^k f(x) + \frac{n-k}{n} \cdot A_n^0 f(x),$$

taking $k = \lceil (1 - \delta)n \rceil$ with $0 < \delta < 1$ fixed and letting $n \to \infty$, it follows that for $m$-a.e. $x \in X$

$$\frac{1}{n} \sum_{\lfloor (1-\eta)n \rfloor}^n f \circ T^j \to \mathbb{E}(f \mid B^T).$$

Applying this to the characteristic function $f = 1_E$ of $E \subset X$, we deduce that for $m$-a.e. $x \in X$

$$\lim_{n \to \infty} \# \{|(1-\delta)n| \leq j \leq n \mid T^j x \in E\} = h_E(x),$$

where $h_E := \mathbb{E}(E \mid B^T)$ is the the conditional expectation of $1_E$. Since $0 \leq h_E(x) \leq 1$ a.e. while

$$\int h_E = m(E),$$

it follows that the set $\{x \in X \mid h_E(x) > 0\}$ has measure $\geq m(E)$. Yet the set

$$\{x \mid h_E(x) > 0\}$$

is, up to null sets, precisely $E^*$. Hence $m(E^*) \geq m(E)$.

For the second statement, note that $\{E^*_N\}$ is an increasing sequence of measurable sets whose union (=limit) is $E^*$.

□

Lemma 2.7 (Small perturbations of polygonal paths).

Given $T_1, \ldots, T_k \in \Gamma$ and $\epsilon > 0$, there is $\delta > 0$ and $N$ so that for all $n \geq N$ we have:

$$\frac{1}{n} \cdot d(T_k^{n-k} T_2^{n-n_2} T_1^{n-n_1}, T_k^{n-k} T_{k-1}^{n-k+1} \cdots T_2^{n-n_2} T_1^{n}) < \epsilon$$

for any $0 \leq n_1, \ldots, n_k \leq \delta \cdot n$.

This Lemma can also be shown by rescaling and passing to the Gromov-Hausdorff limit in $G_\infty$ and relying on Lemma 1.2. Here we give a more direct argument.
PROOF. It suffices to show that for fixed $k \in \mathbb{N}$, $T, T_1, \ldots, T_k \in \Gamma$, $\epsilon' > 0$, there is $\delta > 0$ so that for $n \gg 1$ one has:

\begin{equation}
(2.3) \quad d(T^n_k \cdots T^n_2 T_1^n T^{-m}; T^n_k \cdots T^n_2 T_1^n) < \epsilon' \cdot n \quad (\forall m < \delta \cdot n).
\end{equation}

Indeed, applying such an argument to $T_{j+1}, \ldots, T_l$ and $T = T_j$ with $\epsilon' = \epsilon/k$ we get

\[
\frac{1}{n}(T^n_l \cdots T^n_{j+1} T^n_{j-1} \cdots T^n_1 T^n_i \cdots T^n_{j-1} T^{-m} T^n_i \cdots T^n_1) < \epsilon',
\]

and summing these inequalities over $j = 1, \ldots, \ell - 1$, we get the estimate $\epsilon \cdot \epsilon' \leq \epsilon$ as required.

To establish (2.3) use the general group-theoretic identity $ba = a[a^{-1}, b]$ to push terms from the right to the left creating some commutator factors. More precisely

\[
T^n_k \cdots T^n_2 T_1^n T^{-m} = T^n_k \cdots T^n_2 \cdot (T^{-m} \cdot [T^n_m, T^n_1]) \cdot T^n_1
\]

\[
= T^n_k \cdots T^n_3 (T^{-m} \cdot [T^n_m, T^n_2] \cdot [T^n_m, T^n_1] \cdot [T^n_m, T^n_1]^{-1} \cdot T^n_2) \cdot T^n_2 T^n_1
\]

\[
= (T^{-m}[T^n_m, T^n_k] \cdots [T^n_m, T^n_1] \cdots) \cdot T^n_k \cdots T^n_2 T^n_1
\]

where the expression in the parentheses is a product of $O(k)$ factors each being a higher commutator of the form

\[
\cdots[[T^n_m, T^n_{j_1}]^{-1} \cdot T^n_m \cdots, T^n_{j_s}].
\]

We need to show that the word length of the expression in parentheses is $< \epsilon n$, and it suffices to show that each of the $O(k)$-commutator expressions has length $< \epsilon' n$, where $\epsilon'$ depends on $\epsilon$ and $k$.

Iterated commutators of order $s$ above the nilpotency degree $r$ give identity. For $s \leq r$ one has (cf. [8] Lemma 3.8)

\begin{equation}
(2.4) \quad [\cdots [T^n_m, T^n_{j_1}]^{-1}, \cdots, T^n_{j_s}] = [\cdots [T, T_{j_1}], \cdots, T_{j_s}]^{\pm m \cdot n'}
\end{equation}

For each one of the finitely many elements $\gamma = [\cdots [T, T_{j_1}], \cdots, T_{j_s}]$ as above, we have

\[
|\gamma^n|_s \leq C_\gamma \cdot p^{\frac{m}{m+1}} \quad (p \geq 1)
\]

because such $\gamma$ lies in the $(s+1)$-term of the lower central series $\Gamma^{s+1} = [\Gamma, \Gamma^s] = [\Gamma, [\Gamma, \ldots]]$, and the growth rate on this subgroup is asymptotically scaled by $t^{s+1}$ (recall that in the asymptotic cone $G_\infty$ the homothety $\delta_t$ acts by multiplication by $t^j$ on the $\mathfrak{g}^j/\mathfrak{g}^{j+1}$-subspace of $\mathfrak{g}_\infty$). Therefore the length of the elements in (2.4) is bounded by

\[
C(m \cdot n^n)^{\frac{1}{m+1}} < C(\delta \cdot n^{s+1})^{\frac{1}{s+1}} = C\delta^{\frac{1}{s+1}} \cdot n
\]
which can be made $< 2^{-k}\epsilon$ by choosing $\delta > 0$ small enough. \hfill \Box

Finally, we are ready for the proof of the Ergodic Theorem along Polygonal Paths.

**Proof of Theorem 2.3**

Fix $\epsilon > 0$ and let $\delta > 0$ and $N$ be as in Lemma 2.7.

Choose a small $\eta > 0$ and let $M \in \mathbb{N}$ be large enough so that for each $j = 1, \ldots, k$ the set

$$Y_j := \left\{ y \in X \mid \forall n \geq M : \left| \frac{1}{n} \cdot c(T^n_j, y) - \phi(T^{ab}_j) \right| < \epsilon \right\}$$

has $m(Y_j) > 1 - \eta$.

Let $Z_k := Y_k$, and apply Lemma 2.6 with $E = Z_k$ to find $M_k \in \mathbb{N}$ so that the set

$$(Z_k)^*_{M_k} := \left\{ z \in X \mid \forall n > M_k, \frac{\# \{ n' < \delta \cdot n \mid T^{n-n'}_k \cdot z \in Z_k \}}{\delta \cdot n} > 0 \right\}$$

satisfies $m((Z_k)^*_{M_k}) > m(Z_k) - \eta > 1 - 2\eta$. Define $Z_{k-1} := Y_{k-1} \cap (Z_k)^*_{M_k}$, and observe that

$$m(Z_{k-1}) > 1 - 3\eta.$$  

One then continues inductively to define $Z_{j-1} := (Z_j)^*_{M_j} \cap Y_j$ (for $j = k - 1, \ldots, 3, 2$), where $M_j \in \mathbb{N}$ is chosen large enough to ensure that the set

$$(Z_j)^*_{M_j} := \left\{ z \in X \mid \forall n > M_j, \frac{\# \{ n' < \delta \cdot n \mid T^{n-n'}_j \cdot z \in Z_j \}}{\delta \cdot n} > 0 \right\}$$

has

$$m((Z_j)^*_{M_j}) > m(Z_j) - \eta.$$  

The sets $Z_1, Z_2, \ldots, Z_k$ that are defined in this manner satisfy

$$m(Z_1) > m(Z_2) - 2\eta > m(Z_3) - 4\eta > \cdots > m(Z_k) - 2(k-1)\eta > 1 - (2k-1)\eta.$$  

Let $N := \max(M, M_1, \ldots, M_k)$. Then for every $n > N$ and every $z \in Z_j$, there is $n_j < \delta \cdot n$ so that $T^{n-n_j} z \in Z_{j+1}$ and

$$\left| \frac{1}{n} \cdot c(T^n_j, z) - \phi(T^{ab}_j) \right| < \epsilon.$$  

Thus for $z$ in a set $Z_1$ of size $> 1 - (2k-1)\eta$ and every $n \geq N$, there exist $n_1, \ldots, n_k$ all bounded by $\delta \cdot n$, so that

$$\left| \frac{1}{n} \cdot c(T^n_j, T^{n-n_j-1}_1 \cdots T^{n-n_2}_2 T^{n-n_1}_1 \cdot z) - \phi(T^{ab}_j) \right| < \epsilon.$$
Applying Lemma 2.7 we have for each $j = 1, \ldots, k$:
\[
d(T^n_j, T^n_{j-1} \cdots T^n_1) < n\epsilon,
\]
\[
d(T^{n-n_j-1}_j \cdots T^{n-n_2}_2 T^{n-n_1}_1, T^n_j T^n_{j-1} \cdots T^n_1) < n\epsilon.
\]
Therefore for every $x \in \mathcal{Z}_1$ and $n > N$ one has:
\[
\left| c(T^n_j, T^n_{j-1} \cdots T^n_1.x) - c(\tau, \beta.x) \right| < (2K + 1)\epsilon \quad (j = 1, \ldots, k).
\]
Applying this argument with a sequence of $\eta \to 0$, m.a.e. $x \in X$ would belong to at least one of the sets $\mathcal{Z}_1$, and therefore would satisfy (2.6) for all $n > N(x, \epsilon)$. As $\epsilon > 0$ was arbitrary, this proves that for m.a.e. $x \in X$
\[
\lim_{n \to \infty} \frac{1}{n} \cdot c(T^n_j, T^n_{j-1} \cdots T^n_1.x) = \phi(\tau) \quad (j = 1, \ldots, k)
\]
which in turn gives the convergence of the sum over $j = 1, \ldots, k$ to $\phi(\tau) + \cdots + \phi(\tau)$. The $L^1$-convergence here follows by Lebesgue’s Dominated convergence, because under the $L^\infty$-assumption the terms are uniformly bounded.

However, the latter conclusion of $L^1$-convergence does not require the assumption $c(\gamma, -) \in L^\infty(X, m)$, and holds under the weaker assumption $c(\gamma, -) \in L^1(X, m)$ for $\gamma \in \Gamma$. In the pointwise convergence argument, for every $x$ from a set $\mathcal{Z}_1$ of large measure, for all $n$ large enough we compared the values of the cocycle along a polygonal path with that for a perturbed path (2.5) and used Lemma 2.5 to show that the values are close. In the $L^1$-context it is more natural to compare a polygonal path with the average of all perturbations:
\[
c(T^n_j, T^n_{j-1} \cdots T^n_1.x) - \frac{1}{(\delta n)^{j-1}} \cdot \sum_{n_j = 0}^{\delta n} \cdots \sum_{n_1 = 0}^{\delta n} c(T^n_j, T^{n-n_j-1}_j \cdots T^{n-n_1}_1.x)
\]
and replace Lemma 2.5 by its $L^1$-version:
\[
\int_X |c(\tau, \alpha.x) - c(\tau', \beta.x)| dm(x) \leq K_1 \cdot (d(\alpha, \beta) + d(\tau \alpha, \tau' \beta))
\]
where $K_1 := \max\{\|c(s, -)\|_1 | s \in S\}$. We leave out the rather obvious details for this argument, as it is not needed here.
3. Proof of Theorems A, B

Throughout this section $\Gamma, \Gamma \act (X, m)$, and $c : \Gamma \times X \to \mathbb{R}_+$ are as in Theorem A and

$$\phi : \Gamma^{ab} \otimes \mathbb{R} \cong \mathfrak{g}_\infty \to \mathbb{R}_+, \quad \Phi : G_\infty \to \mathbb{R}_+, \quad d_\phi : G_\infty \times G_\infty \to \mathbb{R}_+$$

are as in Proposition 1.8. We denote by $d_\infty$ the corresponding right-invariant, homogeneous metric on $G_\infty$ that appears in Pansu’s Carnot-Carathéodory construction. We denote

$$B(g, \epsilon) := \{ g' \in G_\infty | d_\infty(g, g') < \epsilon \}$$

the corresponding balls in $G_\infty$. Consider the functions

$$c^*(g, x) := \lim_{\epsilon \searrow 0} \limsup_{t \to \infty} \sup_{\text{scl}(\gamma) \in B(g, \epsilon)} \frac{1}{t} c(\gamma, x),$$

$$c_*(g, x) := \lim_{\epsilon \searrow 0} \liminf_{t \to \infty} \inf_{\text{scl}(\gamma) \in B(g, \epsilon)} \frac{1}{t} c(\gamma, x).$$

While this is not necessary for our argument, it is impossible to ignore the fact that $c^*(g, -)$ and $c_*(g, -)$ are a.e. constant.

**Lemma 3.1.** For each $g \in G_\infty$ the functions $c^*(g, -), c_*(g, -)$ are $m$-a.e. constants, denoted $c^*(g), c_*(g)$, respectively.

**Proof.** For any fixed $g \in G_\infty$ the functions $c_*(g, -), c^*(g, -) : X \to \mathbb{R}_+$ are measurable. Fix $\gamma_0 \in \Gamma$. Then for any $\epsilon > 0$ for all $t > t(g, \gamma_0, \epsilon)$ one has

$$\frac{1}{t} \bullet \gamma \in B(g, \epsilon) \implies \frac{1}{t} \bullet \gamma \gamma_0, \frac{1}{t} \bullet \gamma_0^{-1} \in B(g, 2\epsilon).$$

Since for every $x \in X$

$$\frac{1}{t} c(\gamma_0, x) \leq \frac{1}{t} c(\gamma, x) + \frac{1}{t} c(\gamma_0, x)$$

it follows that $c^*(g, x) \leq c^*(g, \gamma_0, x)$ and $c_*(g, x) \leq c_*(g, \gamma_0, x)$. Applying the same argument to $\gamma_0^{-1}$ and $\gamma_0, x$ we observe that $c_*(g, -)$ and $c^*(g, -)$ are measurable $\Gamma$-invariant functions. Hence they are a.e. constants, because $\Gamma \act (X, m)$ is ergodic. \qed

In the following subsections we shall proceed in the following steps:

1. Show that $c^*(g) \leq \Phi(g)$ for all $g \in G_\infty$.
2. Show that $\Phi(g) \leq c_*(g)$ for all $g \in G_\infty$. 
The obvious inequality \( c_* \leq c^* \) combined with the above implies that \( t_i^{-1} \cdot c(\gamma_i, x) \rightarrow \Phi(g) \) whenever \( \text{scl}_{t_i}(\gamma_i) \rightarrow g \) in \( G_\infty \). We shall show that for a.e. \( x \in X \) the above convergence is uniform over \( g \in B(e, 1) \) and will deduce Theorem A by rescaling.

We will prove Theorem B by combining the ideas of the previous steps.

Let \( X_0 \subset X \) be the set of \( x \in X \) for which \( c^*(g, x) = c^*(g), c_*(g, x) = c_*(g) \), and Theorem 2.3 holds for all \( k \in \mathbb{N} \) and every choice of \( T_1, \ldots, T_k \in \Gamma \). We imposed countably many condition where each holds \( m \)-a.e., therefore \( m(X \setminus X_0) = 0 \).

### 3.1. The upper bound: \( c^*(g) \leq \Phi(g) \).

Fix \( x \in X_0 \), and assume, towards contradiction, that there exists \( \eta > 0 \) and sequences \( t_i \rightarrow \infty \) and \( \gamma_i \in \Gamma \) so that

\[
(3.1) \quad \lim_{i \to \infty} \frac{1}{t_i} \cdot \gamma_i = g, \quad \text{while} \quad \frac{1}{t_i} c(\gamma_i, x) > d_\phi(e, g) + \eta.
\]

Fix a small \( \epsilon > 0 \), namely \( \epsilon = \frac{\eta}{K + 3} \), where \( K \) is as in Theorem A(1). Choose a \( \phi \)-geodesic \( \xi : [0, 1] \to G_\infty \), i.e. a smooth admissible curve such that

\[
\xi(0) = e, \quad \xi(1) = g, \quad \text{length}_\phi(\xi) = \Phi(g)
\]

(we could choose any smooth curve from \( e \) to \( g \) with \( \text{length}_\phi(\xi) < \Phi(g) + \epsilon \) with a sufficiently small \( \epsilon > 0 \)). Applying Proposition 2.1 we find \( k \in \mathbb{N} \), elements \( T_1, \ldots, T_k \in \Gamma \), and a multiple \( p \in \mathbb{N} \) that give \( \epsilon \)-approximation to the curve \( \xi \). Set

\[
n_i := \lfloor \frac{t_i}{p} \rfloor, \quad S_i := T_{n_i} \cdots T_2 T_1.
\]

Note that

\[
\limsup_{i \to \infty} \frac{1}{t_i} \cdot d(S_i, \gamma_i) = \limsup_{i \to \infty} d_\infty(\frac{1}{t_i} \bullet S_i, \frac{1}{t_i} \bullet \gamma_i) = \limsup_{i \to \infty} d_\infty(\frac{1}{t_i} \bullet S_i, g) < \epsilon.
\]

Since \( c(-, x) : \Gamma \to \mathbb{R}_+ \) is \( K \)-Lipschitz, we have for all sufficiently large \( i \gg 1 \).

\[
\frac{1}{n_i p} \cdot \sum_{j=1}^{k} c(T_j^{n_j} \cdots T_2^{n_2} T_1^{n_1}, x) \geq \frac{1}{n_i p} \cdot c(S_i, x) \geq \frac{1}{t_i} \cdot c(S_i, x) - \frac{1}{n_i p} \cdot c(S_i, x) - \epsilon
\]

\[
> \frac{1}{t_i} \cdot c(\gamma_i, x) - K \cdot \epsilon - \epsilon > \Phi(g) + (\eta - (K + 1)\epsilon).
\]
The above inequalities use sub-additivity, the fact that \( n_i p / t_i \to 1 \), the Lipschitz property of \( c(\cdot, x) \), and the assumption (3.1) that we try to refute. Applying Theorem 2.3 we have

\[
\lim_{i \to \infty} \frac{1}{n_i p} \sum_{j=1}^{k} c(T_{n_i}^{n_i}, T_{n_i-1}^{n_i} \cdots T_1^{n_i}, x) = \frac{1}{p} \left( \phi(T_k^a) + \cdots + \phi(T_1^a) \right).
\]

However, by part (ii) of Proposition 2.1, one also has

\[
\frac{1}{p} \left( \phi(T_k^a) + \cdots + \phi(T_1^a) \right) \leq \text{length}_{\phi}(\xi) + \epsilon < \Phi(g) + \epsilon.
\]

This leads to a contradiction, due to our choice of \( \epsilon = \eta / (K + 3) \). Thus (3.1) is impossible.

### 3.2. The lower bound: \( c_*(g) \geq \Phi(g) \)

Let us now prove the inequality \( c_*(g) \geq \Phi(g) \). Fix \( g \in G_{\infty}, x \in X_0 \), and assume, towards contradiction, that there exists \( \eta > 0 \) and sequences \( t_i \to \infty \) and \( \gamma_i \in \Gamma \) so that

\[
\lim_{i \to \infty} \frac{1}{t_i} \cdot \gamma_i = g \quad \text{while} \quad \frac{1}{t_i} c(\gamma_i, x) < \Phi(g) - \eta.
\]

We take a small \( \epsilon > 0 \) and an associated finite set \( F \subset \Gamma \) as in condition (2) of Theorem A. Apply the following argument to each \( \gamma_i \) from the sequence satisfying (3.2).

Each \( \gamma_i \) can be written as a product

\[
\gamma_i = \delta_{i,s_i} \cdots \delta_{i,2} \delta_{i,1},
\]

where \( \delta_{i,j} \in F \) for all \( 1 \leq j \leq s_i \) and

\[
\sum_{j=1}^{s_i} c(\delta_{i,j}, \delta_{i,j-1} \cdots \delta_{i,1}, x) < (1 + \epsilon) \cdot c(\gamma_i, x).
\]

Consider the sequence of points

\[
g_{i,j} := \frac{1}{t_i} \cdot \delta_{i,j} \cdots \delta_{i,1} \quad (j = 1, \ldots, s_i).
\]

Define a piecewise \( d_\infty \)-geodesic curve

\[
\xi_i : [0, 1] \longrightarrow G_{\infty}
\]

connecting \( e \) to \( \frac{1}{t_i} \cdot \gamma_i = g_{i,s_i} \) via the points \( g_{i,j} \), which are to be visited at times

\[
\xi_i \left( \frac{c_{i,1} + \cdots + c_{i,j}}{c_{i,1} + \cdots + c_{i,s_i}} \right) = g_{i,j}, \quad \text{where} \quad c_{i,j} := c(\delta_{i,j}, \delta_{i,r-1} \cdots \delta_{i,1}, x).
\]
Between these times $\xi_i(-)$ follows an appropriately rescaled $g_\infty$-geodesic. So $\xi_i$ traces in $G_\infty$ the points associated to partial products representing a discrete path from $e$ to $\gamma_i$, with time parameter chosen according to the $c_{i,j}$-steps.

The bi-Lipschitz condition for $c(-, x)$ in terms of $d$ (condition (1) in Theorem A), implies that $\xi_i : [0, 1] \to G_\infty$ is a uniformly Lipschitz sequence of maps with $\xi_i(0) = e$. Hence by Arzela-Ascoli, upon passing to a subsequence, we may assume that $\xi_i$ converge (uniformly) to a Lipschitz curve

$$\xi : [0, 1] \to G_\infty, \quad \xi(0) = e, \quad \xi(1) = g.$$ 

Since we are working towards a contradiction to (3.2) which holds for sub-sequences, we may assume that $\xi_i \to \xi$ without complicating our notations any further.

With the Lipschitz curve $\xi$ at hand and small $\epsilon > 0$, Proposition 2.1 provides $k \in \mathbb{N}, T_1, \ldots, T_k \in \Gamma$ and $p \in \mathbb{N}$ that give $\epsilon$-good approximation for the curve $\xi$: In particular, for large $i \gg 1$ and $n_i := \lfloor t_i / p \rfloor$ one has

$$\sum_{j=1}^{k} d_\infty \left( \frac{1}{n_i p} \cdot (T_j^{n_i} \cdots T_2^{n_i} T_1^{n_i}), \xi \left( \frac{j}{k} \right) \right) < \epsilon.$$ 

For each $i \in \mathbb{N}$, choose $0 = r_i(0) < r_i(1) < \cdots < r_i(k) = s_i$ so that for $j = 1, \ldots, k$:

$$\frac{c_{i,r_i(j-1)+1} + \cdots + c_{i,r_i(j)}}{c_{i,1} + \cdots + c_{i,s_i}} \to \frac{1}{k},$$

and write $\gamma_i = \pi_{i,k} \cdots \pi_{i,2} \pi_{i,1}$ with $\pi_{i,j} := \delta_{i,r_i(j)} \cdots \delta_{i,r_i(j-1)+1}$. Then

$$\pi_{i,j} \cdots \pi_{i,2} \pi_{i,1} = \delta_{i,r_i(j)} \cdots \pi_{i,2} \pi_{i,1} \quad (j = 1, \ldots, k).$$

We have

$$\xi \left( \frac{j}{k} \right) = \lim_{i \to \infty} \xi_i \left( \frac{j}{k} \right) = \lim_{i \to \infty} \frac{1}{n_i p} \cdot \pi_{i,j} \cdots \pi_{i,2} \pi_{i,1}.$$ 

Thus for all large enough $i$:

$$\sum_{j=1}^{k} d_\infty \left( \frac{1}{n_i p} \cdot T_j^{n_i} \cdots T_2^{n_i} T_1^{n_i}, \frac{1}{n_i p} \cdot \pi_{i,j} \cdots \pi_{i,2} \pi_{i,1} \right) < \epsilon$$

and therefore for all large enough $i$:

$$\sum_{j=1}^{k} \frac{1}{n_i p} \cdot d \left( T_j^{n_i} \cdots T_2^{n_i} T_1^{n_i}, \pi_{i,j} \cdots \pi_{i,2} \pi_{i,1} \right) < \epsilon.$$ 

We now apply Lemma 2.5 with

$$\alpha = T_{j-1}^{n_i} \cdots T_1^{n_i}, \quad \tau = T_j^{n_i}, \quad \beta = \pi_{i,j-1} \cdots \pi_{i,1}, \quad \tau' = \pi_{i,j}.$$
to deduce that for all large enough $i$:

$$
\sum_{j=1}^{k} \frac{1}{n_i^p} \cdot \left| c(T_j^{n_i}, T_{j-1}^{n_i} \cdots T_1^{n_i}, x) - c(\pi_{i,j}, \pi_{i,j-1} \cdots \pi_{i,1}, x) \right| < 2K\epsilon.
$$

For $i \gg 1$ we have

$$
\frac{1}{t_i} \cdot c(\gamma_i, x) > \frac{n_i^p}{(1+\epsilon)t_i} \cdot \frac{1}{n_i^p} \sum_{r=1}^{\delta_i} c(\delta_i, \delta_{i,r-1} \cdots \delta_{i,1}, x)
$$

$$
\geq (1-\epsilon) \cdot \frac{1}{n_i^p} \sum_{j=1}^{k} \sum_{r=r(j)-1+1}^{\delta_j} c(\delta_i, \delta_{i,r-1} \cdots \delta_{i,1}, x)
$$

$$
\geq (1-\epsilon) \cdot \frac{1}{n_i^p} \sum_{j=1}^{k} c(\pi_{i,j}, \pi_{i,j-1} \cdots \pi_{i,1}, x)
$$

$$
> (1-\epsilon) \cdot \left( \frac{1}{n_i^p} \sum_{j=1}^{k} c(T_j^{n_i}, T_{j-1}^{n_i} \cdots T_1^{n_i}, x) - 2K\epsilon \right)
$$

using sub-additivity of $c$ in the third inequality. Theorem 2.3 gives

$$
\lim_{i \to \infty} \frac{1}{n_i^p} \cdot \sum_{j=1}^{k} c(T_j^{n_i}, T_{j-1}^{n_i} \cdots T_1^{n_i}, x) = \frac{1}{p} \left( \phi(T_k^{ab}) + \cdots + \phi(T_1^{ab}) \right) > \text{length}_{\phi}(\xi) - \epsilon.
$$

Since $\xi$ is only one of many possible admissible curves connecting $\xi(0) = e$ to $\xi(1) = g$ (and most likely is sub-optimal in terms of the $\phi$-length), one has

$$
\text{length}_{\phi}(\xi) \geq d_{\phi}(e, g) = \Phi(g).
$$

Therefore we deduce

$$
\liminf_{i \to \infty} \frac{1}{t_i} \cdot c(\gamma_i, x) \geq (1-\epsilon) \cdot (\Phi(g) - (2K + 1)\epsilon).
$$

A choice of small enough $\epsilon > 0$ contradicts (3.2). This proves the claimed inequality

$$
\Phi(g) \leq c_*(g).
$$

### 3.3. Proof of Theorem A

The results of the two previous subsections giving $c^*(g) \leq \Phi(g) \leq c_*(g)$, combined with the trivial inequality $c_*(g) \leq c^*(g)$, show

$$
c_*(g) = c^*(g) = \Phi(g).
$$
3.3. Proof of Theorems A, B

Equivalently

\[ \lim_{\epsilon \to 0} \limsup_{t \to \infty} \sup \left\{ \frac{1}{t} \cdot c(\gamma, x) - \Phi(g) \mid : \frac{1}{t} \cdot \gamma \in B(g, \epsilon) \right\} = 0. \]

We need to prove that for \( m \)-a.e. \( x \in X \) (or rather every \( x \in X_0 \)) one has

\[ \forall \epsilon > 0, \ \exists R < \infty : \ |\gamma|_S \geq R \implies |c(\gamma, x) - \Phi(scl_1(\gamma))| < \epsilon \cdot |\gamma|_S. \]

Indeed, if the claim were not true, we could find \( \epsilon_0 > 0 \) and a sequence \( \gamma_n \in \Gamma \) with \( |\gamma_n|_S \to \infty \), so that

\[ |c(\gamma_n, x) - \Phi(scl_1(\gamma_n))| \geq \epsilon_0 \cdot |\gamma_n|_S. \]

The sequence

\[ g_n := \frac{1}{|\gamma_n|_S} \cdot \gamma_n \]

has

\[ \limsup_{n \to \infty} d_\infty(g_n, \epsilon) \leq 1. \]

Hence \( \{g_n \mid n \in \mathbb{N}\} \) is bounded. Since balls in \( G_\infty \) are precompact, there is a subsequence \( \gamma_{n_i} \) converging to some \( g \in G_\infty \) (in fact \( g \in B(e, 1) \)). Denote \( t_i := |\gamma_{n_i}|_S \). We note that

\[ d_\infty(\frac{1}{t_i} \cdot \gamma_{n_i}, \delta_{\frac{1}{t_i}}(scl_1(\gamma_{n_i}))) \to 0 \]

and therefore

\[ \lim_{i \to \infty} \frac{1}{t_i} \cdot \Phi(scl_1(\gamma_{n_i})) = \lim_{i \to \infty} \Phi(\delta_{\frac{1}{t_i}}(scl_1(\gamma_{n_i}))) = \Phi(\lim_{i \to \infty} \frac{1}{t_i} \cdot \gamma_{n_i}) = \Phi(g). \]

Finally \( \text{(3.3)} \) implies

\[ \frac{1}{t_i} |c(\gamma_{n_i}, x) - t_i \cdot \Phi(g)| \to 0 \]

contrary to the assumption. This proves Theorem [A]

3.4. Proof of Theorem [B]

The main claim is that given any \( g, g' \in G_\infty \) and sequences \( t_i \to \infty, \gamma_{i}, \gamma_{i}' \in \Gamma \), so that

\[ \lim_{i \to \infty} \frac{1}{t_i} \cdot \gamma_{i} \to g, \quad \lim_{i \to \infty} \frac{1}{t_i} \cdot \gamma_{i}' = g' \]

one necessarily has for every \( x \in X_0 \):

\[ \lim_{i \to \infty} \frac{1}{t_i} \cdot c(\gamma_{i}' \gamma_{i}^{-1}, \gamma_{i}; x) = d_\phi(g, g'). \]
To show this we employ a variant on the upper bound argument §3.1 and on the lower bound argument §3.2. In both of these arguments we use a fixed admissible curve $\xi_0$ connecting $e$ to $g$ in $G_\infty$, concatenated with an appropriate curve $\xi$ connecting $g$ to $g'$ in $G_\infty$.

Denote by $\xi_1 : [0, 2] \to G_\infty$ the curve that connects $e$ to $g'$ via $g$:

$$\xi_1(0) = e, \quad \xi_1(1) = g, \quad \xi_1(2) = g'; \quad \xi_1(s + 1) = \xi(s).$$

Fix a small $\epsilon > 0$, and apply Proposition 2.1 to $\xi_1$ to find $T_1, \ldots, T_{2k} \in \Gamma, \quad p \in \mathbb{N},$

so that

\begin{equation}
\left| \frac{1}{p} \cdot \sum_{j=k+1}^{2k} \phi(T_{j}^{ab}) - \text{length}_\phi(\xi) \right| < \epsilon,
\end{equation}

while for all $n \geq n_0$

\begin{equation}
\sum_{j=1}^{k} d_\infty \left( \frac{1}{np} \cdot T^n_{j+k} \cdots T^n_2 T_1, \xi(\frac{j}{k}) \right) < \epsilon.
\end{equation}

Note that the last condition is a trivial consequence of the estimate on

$$\sum_{j=1}^{k} d_\infty \left( \frac{1}{np} \cdot T^n_{j+k} \cdots T^n_2 T_1, \xi(\frac{j}{k}) \right) + \sum_{j=1}^{k} d_\infty \left( \frac{1}{np} \cdot T^n_{j+k} \cdots T^n_2 T_1, \xi(\frac{j}{k} + 1) \right),$$

while (3.5) can be obtained from approximating the $\phi$-lengths of $\xi_1$ and $\xi_0$ by

$$\frac{1}{p} \cdot \sum_{j=1}^{2k} \phi(T_{j}^{ab}), \quad \text{and} \quad \frac{1}{p} \cdot \sum_{j=1}^{k} \phi(T_{j}^{ab})$$

and the obvious relation

$$\text{length}_\phi(\xi) = \text{length}_\phi(\xi_1) - \text{length}_\phi(\xi_0).$$

Next, choosing $\xi$ to be a $\phi$-geodesic connecting $g$ to $g'$, and taking $n_i = \lfloor t_i / p \rfloor$, we get

$$\lim_{i \to \infty} d_\infty \left( \frac{1}{n_i p} \cdot \gamma_i, g \right) = \lim_{i \to \infty} d_\infty \left( \frac{1}{n_i p} \cdot \gamma'_i, g' \right) = 0$$

and

$$\limsup_{i \to \infty} \frac{1}{n_i p} \cdot d(\gamma_i, T^{n_i}_k \cdots T_1) \leq \epsilon, \quad \limsup_{i \to \infty} \frac{1}{n_i p} \cdot d(\gamma'_i, T^{n_i}_k \cdots T_1) \leq \epsilon.$$
Following the same argument as in §3.1 (using Lemma 2.5 and Theorem 2.3), we have
\[
\limsup_{i \to \infty} \frac{1}{t_i} \cdot c(\gamma_i^{-1}, \gamma_i, x) \leq \limsup_{i \to \infty} \frac{1}{n_{i, p}} \cdot c(T_{2k}^{n_i} \cdot \cdots \cdot T_{k+1}^{n_i}, T_{k+1}^{n_i}, T_1^{n_i}, x) + 2K \epsilon
\]
\[
\leq \lim_{i \to \infty} \frac{1}{n_{i, p}} \cdot \sum_{j=1}^{k} c(T_{k+j}^{n_i} \cdot T_{k+j-1}^{n_i} \cdot \cdots \cdot T_{1}^{n_i}, x) + 2K \epsilon = \frac{1}{p} \cdot \sum_{j=1}^{k} \phi(T_{k+j}^{n_i}) + 2K \epsilon
\]
\[
< \text{length}_\phi(\xi) + (2K + 1) \epsilon = d_\phi(g, g') + (2K + 1) \epsilon.
\]
Since \( \epsilon > 0 \) was arbitrary, this shows the upper bound:
\[
\limsup_{i \to \infty} \frac{1}{t_i} \cdot c(\gamma_i^{-1}, \gamma_i, x) \leq d_\phi(g, g').
\]
The lower bound,
\[
\liminf_{i \to \infty} \frac{1}{t_i} \cdot c(\gamma_i^{-1}, \gamma_i, x) \geq d_\phi(g, g')
\]
is trivial if \( g = g' \). Hence we assume \( g \neq g' \) which implies that \( |\gamma_i^{-1}|_S \to \infty \). We now use the innerness assumption (corresponding to condition (2) in Theorem A). Fix an arbitrary small \( \epsilon > 0 \) and rewrite \( \gamma_i^{-1} \) as a product of
\[
\gamma_i^{-1} = \delta_i s_i \cdots \delta_{i, 1}, \quad \text{while} \quad \sum_{r=1}^{s_i} c(\delta_{i, r}, \delta_{i, r-1} \cdots \delta_{i, 1}, \gamma_i, x) < (1 + \epsilon) \cdot c(\gamma_i^{-1}, \gamma_i, x).
\]
where \( \delta_{i, j} \) belong to a fixed finite set \( F \subset \Gamma \) (depending on \( \epsilon \) and \( x \in X_0 \)). One then proceeds as in §3.2 to construct a uniformly Lipschitz sequence of piecewise geodesic curves connecting \( g \) to \( \approx g' \), and to use Arzela-Ascoli to pass to a convergent subsequence that produces a Lipschitz curve
\[
\xi : [0, 1] \to G_\infty, \quad \xi(0) = g, \quad \xi(1) = g'.
\]
We are going to concatenate \( \xi_0 \) with \( \xi \) to get \( \xi_1 : [0, 2] \to G_\infty \) as before. The long products \( \gamma_i^{-1} = \delta_i s_i \cdots \delta_{i, 1} \) can be sub-partitioned so that
\[
\gamma_i^{-1} = \pi_{i, k} \cdots \pi_{i, 1},
\]
while for \( j = 1, \ldots, k \) one has
\[
\xi(\frac{j}{k}) = \xi_0(\frac{k+j}{k}) = \lim_{i \to \infty} \frac{1}{t_i} \cdot (\pi_{i, j} \cdots \pi_{i, 1}, \gamma_i).
\]
We now invoke the \( T_1, \ldots, T_{2k} \) and \( p \in \mathbb{N} \) satisfying (3.5) and (3.6). One has
\[
\limsup_{i \to \infty} \sum_{j=1}^{k} \frac{1}{n_{i, p}} \cdot d(T_{k+j}^{n_i} \cdot \cdots \cdot T_1^{n_i}, \pi_{i, j} \cdots \pi_{i, 1}, \gamma_i) \leq \epsilon.
\]
The sub-additivity gives
\[
\sum_{j=1}^{k} c(\pi_{i,j}, \pi_{i,j-1} \cdots \pi_{i,2} \pi_{i,1} \gamma_i, x) \leq \sum_{r=1}^{s_i} c(\delta_{i,r}, \delta_{i,r-1} \cdots \delta_{i,1} \gamma_i, x).
\]

Combining these facts, one shows that for a subsequence of the given \( t_i, \gamma_i, \gamma'_i \) one has:
\[
\liminf_{i \to \infty} \frac{1}{t_i} \cdot c(\gamma'_i \gamma_i^{-1}, \gamma_i, x) \geq \liminf_{i \to \infty} \frac{1}{1 + \epsilon} \cdot \frac{1}{n_{i} p} \sum_{j=1}^{k} c(\pi_{i,j}, \pi_{i,j-1} \cdots \pi_{i,2} \pi_{i,1}, \gamma_i, x)
\]
\[
\geq (1 - \epsilon) \cdot \left( \liminf_{i \to \infty} \frac{1}{n_{i} p} \cdot \sum_{j=1}^{k} c(T_{k+j}^{n_i}, T_{k+j-1}^{n_i} \cdots T_{1}^{n_i}.x) - 2K \epsilon \right)
\]
\[
= (1 - \epsilon) \cdot \left( \frac{1}{p} \cdot (\phi(T_{2k}^{ab}) + \cdots + \phi(T_{k+1}^{ab})) - 2K \epsilon \right)
\]
\[
\geq (1 - \epsilon) \cdot (\text{length}_\phi(\xi) - (2K + 1)\epsilon) \geq (1 - \epsilon) \cdot (d_\phi(g, g') - (2K + 1)\epsilon).
\]

Since \( \epsilon > 0 \) is arbitrary, and any subsequence of \( t_i, \gamma_i, \gamma'_i \) contains a sub-sub-sequence satisfying the above, it follows
\[
\liminf_{i \to \infty} \frac{1}{t_i} \cdot c(\gamma'_i \gamma_i^{-1}, \gamma_i, x) \geq d_\phi(g, g').
\]

In view of the lim sup inequality, the lower bound is also proven. As in the proof of Theorem A one can easily deduce that for \( m \)-a.e. \( x \in X \) for every \( \epsilon > 0 \) there is \( R < \infty \) so that for \( |\gamma|_S, |\gamma'|_S > R \) one has
\[
|c(\gamma'_i \gamma_i^{-1}, \gamma, x) - d_\phi(\text{scl}_1(\gamma), \text{scl}_1(\gamma'))| < \epsilon \cdot \max(|\gamma|_S, |\gamma'|_S).
\]

This completes the proof of Theorem B.
CHAPTER 3

Differentiability Of Integrable Measurable Cocycles Between Nilpotent Groups

1. Introduction

In this chapter, we build on ideas developed in the last chapter to

Theorem D: Prove an integrable measurable cocycle analog of Pansu’s Rademacher-type differentiation theorem. This answers a question of Tim Austin \[1\] and gives an independent proof of and strengthens the main results in \[1\].

QI/IME Rigidity: Deduce, as a corollary, an alternative, ergodic theoretic proof of Pansu’s quasi-isometric rigidity stating that the associated Carnot group is a quasi-isometry invariant for nilpotent groups. This follows immediately from the \textit{a priori} stronger result that the associated Carnot group is an integrable measure equivalence invariant.

Theorem E: Prove a nilpotent-valued cocycle ergodic theorem. Together with the nilpotent Poincaré Recurrence Lemma \[5.9\], this implies Theorem D.

In the previous chapter we analyzed \textit{subadditive} cocycles from a nilpotent group in to the positive reals. That is, given an ergodic pmp action \(\Gamma \actson (X, m)\) we analyzed

\[c : \Gamma \times X \to \mathbb{R}_+ \text{ satisfying } c(\gamma_1 \gamma_2, x) \leq c(\gamma_1, \gamma_2 x) + c(\gamma_2, x) \quad (\gamma_1, \gamma_2 \in \Gamma, \text{ a.e.} x.)\]

In this chapter, we shift our focus to measurable cocycles between nilpotent groups. That is, we will be concerned with measurable maps

\[\alpha : \Gamma \times X \to \Lambda \text{ satisfying } \alpha(\gamma_1 \gamma_2, x) = \alpha(\gamma_1, \gamma_2 \cdot x) \alpha(\gamma_2, x) \quad (\gamma_1, \gamma_2 \in \Gamma \text{ a.e.} x.)\]
We will think of \( \alpha(-,x) : \Gamma \to \Lambda \) as a random map between \( \Gamma \) and \( \Lambda \), and we will show that, in an appropriate sense, with high probability the \( \alpha(-,x) \) limit asymptotically to an isomorphism between the associated Carnot groups of \( \Gamma \) and \( \Lambda \). We remark that, in this chapter we revert to the more common convention of working with left-invariant metrics, since our arguments allow it.

The idea of the proof of Theorem \( \mathbb{E} \) is that, following Pansu [24], the large scale geometry of f.g. nilpotent groups depends only on the behavior of the projection to abelianization. Therefore, to understand the large scale geometric behavior of a random map \( \alpha(\cdot,x) : \Gamma \to \Lambda \), we project it to the abelianization and integrate. Since a section of the abelianization generates the whole group, we can write all elements in terms of that section. We then use the cocycle identity to decompose arbitrary elements into a product of those coming from (a section of) the abelianization, which allows us to promote the cocycle ergodic theorem for cocycles with values in \( \mathbb{R}^d \), which is easy, to the desired cocycle ergodic theorem with values in \( \Lambda \).

The rest of this chapter is organized as follows. Section 2 sets notation, gathers background information regarding nilpotent groups, asymptotic cones, and measure equivalence, and reduces to the torsion-free nilpotent case. In section 3 we study asymptotics along iterates of a single element. In section 4 we combine the results of Section 3 with some nilpotent geometry from Section 2 to understand asymptotics along arbitrary elements. Finally in section 5 we define the isomorphism \( \Phi : G_\infty \cong H_\infty \), prove Theorem \( \mathbb{E} \) and deduce Theorem \( \mathbb{D} \).

### 2. Background and Notation

#### 2.1. Integrable Measure Equivalence.

Two infinite discrete countable groups \( \Gamma, \Lambda \) are measure equivalent if there exists an infinite measure space \((\Omega, m)\) with a measurable, measure preserving action of \( \Gamma \times \Lambda \) so that the actions \( A : \Gamma \curvearrowright (\Omega, m) \) and \( B : \Lambda \curvearrowright (\Omega, m) \) admit finite measure fundamental domains \( Y, X \subset \Omega \). The space \((\Omega, m)\) together with the \( \Gamma \times \Lambda \) action is called a measurable coupling of \( \Gamma \) and \( \Lambda \). By restricting attention to an ergodic component, one may always assume that \( m \) is ergodic for the \( \Gamma \times \Lambda \) action.

The fundamental domains \( Y \) and \( X \) for the \( G \) and \( H \) actions give rise to functions

\[
\alpha : \Gamma \times X \to \Lambda \quad \text{and} \quad \beta : \Lambda \times Y \to \Gamma
\]

defined uniquely by requiring
\[ B(\lambda)y \in A(\beta(\lambda, y)^{-1})Y \quad \text{and} \quad A(\gamma)x \in B(\alpha(\gamma, x)^{-1})X \quad \forall x \in X \forall y \in Y. \]

There are associated finite measure preserving actions \( \Gamma \actson (X, m|_X) \) and \( \Lambda \actson (Y, m|_Y) \) (whose actions we denote by \( \cdot \)) defined by requiring that

\[ A(\gamma)x = B(\alpha(\gamma, x)^{-1})(\gamma \cdot x) \quad \text{and} \quad B(\lambda)y = A(\beta(\lambda, y)^{-1})(\lambda \cdot y). \]

If \( m \) is ergodic for \( \Gamma \times \Lambda \) then the actions \( \Gamma \actson (X, m|_X) \) and \( \Lambda \actson (Y, m|_Y) \) are ergodic. We may assume after renormalizing that both \( m|_X \) and \( m|_Y \) are probability measures. Finally, \( \alpha \) and \( \beta \) are measurable cocycles over the pmp actions in the sense that

\[ \alpha(\gamma_1, \gamma_2, x) = \alpha(\gamma_1, \gamma_2 \cdot x) \alpha(\gamma_2, x) \quad \text{and} \quad \beta(\lambda_1, \lambda_2, y) = \beta(\lambda_1, \lambda_2 \cdot y) \beta(\lambda_2, y) \]

for all \( \gamma_1, \gamma_2 \in \Gamma, \lambda_1, \lambda_2 \in \Lambda \) and for \( m \) a.e. \( x \in X, y \in Y \). Most of our reasoning will be about these cocycles.

Replacing the fundamental domain \( Y \) with one of its \( H \) translates only translates the cocycle \( \beta \). Since countably many translates of \( Y \) cover \( \Omega \), we may therefore assume that \( m(X \cap Y) > 0 \). Moreover, (see [1] for more details) if

\[ x \in X \cap Y \cap \gamma^{-1}(X \cap Y) \]

then

\[ \beta(\alpha(\gamma, x), x) = \gamma. \]

Given finitely generated groups \( \Gamma, \Lambda \), a cocycle \( \alpha : \Gamma \times X \to \Lambda \) over a pmp action \( \Gamma \actson (X, \mu) \) is integrable if, for some (any) choice of finite generating set for \( \Lambda \)

\[ |||\alpha(\gamma, \cdot)||_1|_\Lambda| = \int_X |\alpha(\gamma, x)|_\Lambda d\mu(x) < \infty \quad \forall \gamma \in \Gamma \]

where \( |||\cdot||_\Lambda \) is the word norm associated to the generating set. The subadditivity of \( |||\cdot||_\Lambda \) implies
\[ \|\alpha(\gamma, \cdot)\|_\Lambda \leq |\gamma|_\Gamma \cdot \max_{s \in S} \|\alpha(s, \cdot)\|_\Lambda \leq |\gamma|_\Gamma \cdot \max_{s \in S} \|\alpha(s, \cdot)\|_\Lambda \]

where \(|\cdot|_\Gamma\) is any word norm associated to a finite generating set for \(\Gamma\).

Finally, finitely generated groups \(\Gamma\) and \(\Lambda\) are integrably measure equivalent if they admit a measurable coupling so that the associated cocycles (2.1) are integrable. This is an equivalence relation independent of choice of generating sets. For more details, see [11].

Recall that measurable events \(E_n \subset (X, m)\) occur with high probability if \(m(E_n) \to 1\) as \(n \to \infty\). We say that a sequence of measurable functions \(f_n : X \to [0, \infty)\) is \(o(n)\) in probability (or whp) if for all \(\epsilon > 0\) one has \(m(f_n(x)/n < \epsilon) \to 1\) as \(n \to \infty\). Thus for example \(d_\Lambda(\alpha(\gamma^n, x), \lambda) = o(n)\) in probability means that for all \(\epsilon, \delta > 0\) there is \(N\) so that for all \(n \geq N\) one has \(m(d_\Lambda(\alpha(\gamma^n, x), \lambda) < n\delta) < \epsilon\).

Similarly for \(O(n)\).

### 2.2. Logarithmic Coordinates.

We keep the notation from Chapter 2 §1.1. In particular, \(\Gamma\) is a finitely generated torsion-free nilpotent group with torsion-free abelianization, \(G\) is its Mal’cev completion and \(G_\infty\) is the associated Carnot group.

We will use the so called logarithmic coordinates throughout this chapter, which are described as follows. Choose a real basis \(\{X_1, \ldots, X_m\}\) for \(\mathfrak{g}\) that respects the decomposition from (1.2) in Chapter 2. When we write \(g = (x_1, \ldots, x_m) \in G\) we mean that \(g = \exp(x_1 X_1 + \cdots + x_m X_m)\). These are the logarithmic coordinates of \(G\). Thus if \(g = (x_1, \ldots, x_m)\) and \(h = (x'_1, \ldots, x'_m)\) then the product \(gh = (y_1, \ldots, y_m)\) where

\[
\exp(x_1 X_1 + \cdots + x_m X_m) \exp(x'_1 X_1 + \cdots + x'_m X_m) = \exp(y_1 X_1 + \cdots + y_m X_m).
\]

In light of the vector space isomorphism \(L : \mathfrak{g} \to \mathfrak{g}_\infty\) the basis for \(\mathfrak{g}\) yields a basis for \(\mathfrak{g}_\infty\) that respects the decomposition (1.1). Throughout this chapter we will think of \(\mathfrak{g}\) and \(\mathfrak{g}_\infty\) as occupying the same real vector space, only with different Lie brackets \([-,-]\) and \([-,-]_\infty\). We also use the logarithmic coordinates for \(\mathfrak{g}_\infty\), the only difference in definition being the Lie bracket.

Let \(d = \dim V_1\). Then there exist constants \(\eta_i \in \mathbb{N}\) for \(i = 1, \ldots, m\) so that \(\Gamma\) embeds in \(G\) in logarithmic coordinates as

\[
\Gamma = \{(a_1, \ldots, a_d, \eta_{d+1} a_{d+1}, \ldots, \eta_m a_m) : a_i \in \mathbb{Z}\} < G.
\]
Thus we have identified $\Gamma < G \equiv \mathbb{R}^m \equiv G_\infty$. Therefore we think of $\Gamma < G$ and $G_\infty$ as occupying the same copy of $\mathbb{R}^m$. We denote the group product in $\Gamma < G$ by $g \cdot h$ or simply by $gh$, and the group product in $G_\infty$ by $g \star h$. We will always denote a word norm on a discrete nilpotent group $\Gamma$ or $\Lambda$ by $|\cdot|_\Gamma$ or $|\cdot|_\Lambda$, a word norm on a nilpotent Lie group $G$ or $H$ by $|\cdot|_G$ or $|\cdot|_H$ and a Carnot-Carathéodory norm on a graded nilpotent Lie group $G_\infty$ or $H_\infty$ by $|\cdot|_\infty$, and their associated metrics $d_\Gamma$, $d_\Lambda$, $d_H$, $d_H$, and $d_\infty$. Thus we can without notational ambiguity omit the linear identification $L : G \equiv G_\infty$.

For example if $\gamma, \sigma \in \Gamma$ then $|\gamma|_\infty$ means unambiguously $|L\gamma|_\infty$ and $\gamma \star \sigma$ means $L\gamma \star L\sigma$.

Since $V_1 \cong g/g^2$, the sets

$$\{(x_1, \ldots, x_d, 0, \ldots, 0) \in G\} \cong \mathbb{R}^d \quad \text{and} \quad \{(a_1, \ldots, a_d, 0, \ldots, 0) \in \Gamma\} \cong \mathbb{Z}^d$$

are complete sets of coset representatives for $G/G^2$, $G_\infty/G^2_\infty$ and $\Gamma/\Gamma^2$. We will use these choices of coset representatives in the arguments that follow. We define the projections on to the abelian and commutator coordinates for $\Gamma$, $G$, and $G_\infty$ by

$$\pi_{ab}(a_1, \ldots, a_m) = (a_1, \ldots, a_d, 0, \ldots, 0)$$

$$\pi_{com}(a_1, \ldots, a_m) = (0, \ldots, 0, a_{d+1}, \ldots, a_m).$$

2.3. Some Nilpotent Geometry. We now collect some basic nilpotent geometry facts. We make no claim to originality in this subsection.

We will use the following Lemma of Guivarc’h repeatedly throughout this chapter to simplify our arguments. Since asymptotic statements are not sensitive to quasi-isometry, the Guivarc’h Lemma allows us to prove asymptotic statements for only one of $(H, d_H)$ or $(H_\infty, d_\infty)$.

**Lemma 2.1** (Guivarc’h [15]; see also [8] Theorem 3.7). Let $K$ be a compact neighborhood of the identity in a simply connected nilpotent Lie group $G$ and $d_G(g, h) = \inf\{n \geq 1 : g^{-1}h \in K^n\}$. Then for any homogeneous quasi-norm $|\cdot|$ on $G$ there is a constant $C > 0$ so that

$$\frac{1}{C}|g| \leq d_G(e, g) \leq C|g| + C.$$
In particular the linear identification $L : G \to G_\infty$ is a quasi-isometric embedding for every compactly generated word metric $d_G$ on $G$ and every Carnot-Carathéodory metric $d_\infty$ on $G_\infty$; i.e. there exist $c, c' > 0$ (depending on $d_\infty$ and $d_G$) so that for all $g_1, g_2 \in G$

$$\frac{1}{c} d_G(g_1, g_2) - c' \leq d_\infty(Lg_1, Lg_2) \leq cd_G(g_1, g_2) + c'.$$

We now use the Guivarc’h Lemma to give succinct proofs of several nilpotent geometric facts, which could also be proved by induction on nilpotency class. All of the statements are true independent of choice of symmetric generating set, but we work with a fixed generating set $S$ with associated norm $|\cdot|_\Gamma$ and metric $d_\Gamma$ to be concise. All constants depend on $\Gamma$ and $S$. The following lemma is a natural statement regarding the asymptotic word growth of each coordinate in a nilpotent group.

**Lemma 2.2.** For each $1 \leq i \leq m$ there exist $0 < c_1 < c_2 < \infty$ so that for all $n \in \mathbb{Z}$

$$c_1 n^{1/d_i} \leq |(0, \ldots, n, \ldots, 0)|_\Gamma \leq c_2 n^{1/d_i},$$

where the non zero term is in the $i$-th coordinate.

Moreover, if $[X_i, \ldots, [X_{i-1}, X_i] \ldots] = cX_t$ where $i_r \in \{1, \ldots, m\}$ and $c \neq 0$, then

$$\sum_{r=1}^{i} d_{i_r} \leq d_t.$$

**Proof.** The following is a quasi-norm on $G$

$$|(x_1, \ldots, x_m)|_m := \max_i |x_i|^{1/d_i}.$$

$(G, |\cdot|_G)$ restricted to $\Gamma$ is quasi-isometric to $(\Gamma, |\cdot|_\Gamma)$, while by the Guivarc’h Lemma, $(G, |\cdot|_G)$ is quasi-isometric to $(G, |\cdot|_m)$. But $|(0, \ldots, n, \ldots, 0)|_m = n^{1/d_i}$. Since $\Gamma$ is discrete we may absorb the additive factors. The moreover statement is obvious from the definitions. \[\square\]

**Lemma 2.3.** For each $1 \leq i \leq m$ set

$$f_i(n) = |(0, \ldots, 0, n, 0, \ldots, 0)|_\Gamma,$$

$$g_i(n) = \min_{a_j} |(a_1, \ldots, a_{i-1}, n, a_{i+1}, \ldots, a_m)|_\Gamma.$$
where the non-zero coordinate is in the \( i \)-th coordinate. Then there exists \( 1 \leq C < \infty \) so that for all \( n \in \mathbb{N} \)

\[
f_i(n) \leq c g_i(n).
\]

**Proof.**

\[
f_i(n) \leq c n^{1/d_i} \leq c \min_{a_j} |(a_1, \ldots, a_{i-1}, n, a_{i+1}, \ldots, a_m)|_m
\]

\[
\leq c c_1 \min_{a_j} |(a_1, \ldots, a_{i-1}, n, a_{i+1}, \ldots, a_m)|_\Gamma + c_2
\]

\[
\leq (cc_1 + c_2) \min_{a_j} |(a_1, \ldots, a_{i-1}, n, a_{i+1}, \ldots, a_m)|_\Gamma \quad (n \neq 0)
\]

where we have used Lemma 2.2, Proposition 2.4 and the Lemma of Guivarc’h.

\[\square\]

The next lemma says that projecting to the commutator coordinates only reduces word norm by a universal multiplicative constant.

**Lemma 2.4.** There is a constant \( C > 0 \) so that for all \( \gamma \in \Gamma \)

\[
|\gamma|_\Gamma \geq C |\pi_{com} \gamma|_\Gamma.
\]

**Proof.** Consider the quasi-norm \(|\cdot|_m\) defined above. For all \( \gamma \in \Gamma \) we have trivially

\[
|\gamma|_m \geq |\pi_{com} \gamma|_m.
\]

The Guivarc’h Lemma and the discreteness of \( \Gamma \) finish the proof.

\[\square\]

**Lemma 2.5.** There exists \( l > 0 \) so that for all \( \gamma \in \Gamma - \Gamma^2 \) and for all \( n |\gamma^n|_\Gamma \geq ln. \)

**Proof.** If \( \gamma \notin \Gamma^2 \) then \( |\gamma^n|_m \geq n \). The Guivarc’h Lemma and the discreteness of \( \Gamma \) finish the proof.

\[\square\]
**Lemma 2.6.**

\[
|\gamma_n|_\Gamma = |(a_{n,1}, \ldots, a_{n,m})|_\Gamma = o(n) \iff |a_{n,t}| = o(n^{d(t)}) \quad \forall 1 \leq t \leq m
\]

\[
|\gamma_n|_\Gamma = |(a_{n,1}, \ldots, a_{n,m})|_\Gamma = O(n) \iff |a_{n,t}| = O(n^{d(t)}) \quad \forall 1 \leq t \leq m
\]

\[
|g_n|_G = |(a_{n,1}, \ldots, a_{n,m})|_G = o(n) \iff |a_{n,t}| = o(n^{d(t)}) \quad \forall 1 \leq t \leq m
\]

\[
|g_n|_G = |(a_{n,1}, \ldots, a_{n,m})|_G = O(n) \iff |a_{n,t}| = O(n^{d(t)}) \quad \forall 1 \leq t \leq m
\]

where \(a_{n,j} \in \mathbb{Z} (a_{n,j} \in \mathbb{R})\) is the \(j\)-th coordinate of \(\gamma_n \in \Gamma (g_n \in G)\).

**Proof.** Using the quasi-isometry from the Guivarc’h Lemma it is clear that

\[
|g_n|_G = o(n) \iff |g_n|_m = o(n) \iff |a_{n,i}| = o(n^{d_i}).
\]

The quasi-isometry between \((\Gamma, d_G)\) and \((\Gamma, d_\Gamma)\) gives the result for \(\Gamma\). The \(O(n)\) case is similar. \(\square\)

**Lemma 2.7.** If \(g_n \in G\) is a sequence such that

1. \(|\pi \text{com} g_n|_G = o(n)|
2. \(|\pi ab g_n|_G = O(n)|

then

\[
|\pi ab g_n|_G = o(n).
\]

**Proof.** Let \(g_n = (a_{n,1}, \ldots, a_{n,m})\) so that \(\pi ab g_n = (a_{n,1}, \ldots, a_{n,d}, 0, \ldots, 0)\) and

\((\pi ab g_n)^{-1} = (-a_{n,1}, \ldots, -a_{n,d}, 0, \ldots, 0)\). Using the Baker-Campbell-Hausdorff formula, the nilpotency of \(G\) and linearity of the bracket

\[
(\pi ab g_n)^{-1} g_n = a_{n,d+1}X_{d+1} + \cdots + a_{n,m}X_m + h.o.t.
\]

where \(h.o.t.\) are precisely the terms involving at least one bracket in the product

\[
(-a_{n,1}X_1 - \cdots - a_{n,d}X_d) \cdot (a_{n,d+1}X_{d+1} + \cdots + a_{n,m}X_m).
\]
Since the abelian coordinates of \((\pi_{ab}g_n)^{-1}g_n\) are all zero, by Lemma 2.6 it suffices to show that the \(t\)-th coordinate of \((\pi_{ab}g_n)^{-1}g_n\) is \(o(n^d)\) for every \(d < t \leq m\). By assumption \(|\pi_{com}g_n|_G = o(n)|\), so it suffices to show that the contributions from (2.2) to each \(t\) coordinate are \(o(n^d t)\), for \(d < t \leq m\).

Using Baker-Campbell-Hausdorff again, for fixed \(d < t \leq m\) the contribution is a sum of finitely many terms of the form

\[ c[a_{n,i_1}X_{i_1}, \ldots, [a_{n,i_{i-1}}X_{i_{i-1}}, a_{n,i_i}X_{i_i}], \ldots] \]

where \(c\) is a constant from the Baker-Campbell-Hausdorff formula, \(i_1, \ldots, i_l \in\{1, \ldots, m\}\) and for at least one \(r, i_r \in\{d+1, \ldots, m\}\). Since the number of such terms depends only on \(G\), it suffices to show that

\[ a_{n,i_1} \cdots a_{n,i_r} = o(n^d), \]

which follows immediately from the fact that at least one \(i_r \in\{d+1, \ldots, m\}\) and

- \(|a_{n,i_r}| = O(n)|\) for \(1 \leq i_r \leq d\)
- \(|a_{n,i_r}| = o(n^{d_r})|\) for \(d < i_r \leq m\)
- \(\sum_{r=1}^{d} d_i \leq d_t\).

\[ \square \]

**Lemma 2.8.** Let \(g_n, h_n \in G\). If \(|g_n|_G = o(n)|\) and \(|h_n|_G = O(n)|\) then

\[ |g_n^{-1}h_n^{-1}g_nh_n|_G = o(n)|. \]

Moreover the same is true of any Carnot-Carathéodory norm on \(G_{\infty}\) instead of \(G\).

**Proof.** Let \(g_n = (a_{n,1}, \ldots, a_{n,m}), h_n = (b_{n,1}, \ldots, b_{n,m})\) and suppose \(|g_n|_G = o(n)|\) and \(|h_n|_G = O(n)|\). By Lemma 2.6

\[ |a_{n,t}| = o(n^d) \quad 1 \leq t \leq m \]

\[ |b_{n,t}| = O(n^d) \quad 1 \leq t \leq m. \]

Say \(g_n = \exp v_n = \exp(a_{n,1}X_1 + \cdots + a_{n,m}X_m)\) and \(h_n = \exp w_n = \exp(b_{n,1}X_1 + \cdots + b_{n,m}X_m)\), so
where the dots stand for terms involving three or more brackets. Let us examine the coefficient $c_r$ of $X_r$ in (2.3); it is a sum of finitely many terms of the form

$$ca_{n,i_1} \cdots a_{n,i_s} b_{n,j_1} \cdots b_{n,j_t}$$

where $c$ is a (possibly zero) constant depending only on $G$, and $s \neq 0$, i.e. there is at least one $a_{n,i}$ term. Employing Lemma 2.6 again it suffices to show that each of these possible coefficients is $o(n^{d_r})$. Indeed there is a constant $c$ (coming from the $O(n^{d(3s)})$) so that for all $\epsilon > 0$ and all sufficiently large $n$

$$|a_{n,i_1} \cdots a_{n,i_s} b_{n,j_1} \cdots b_{n,j_t}| \leq cn^{d'_t} \leq c\epsilon n^{d_r}.$$

2.4. Notation. All of the above was true of a general finitely generated torsion-free nilpotent group $\Gamma$, though of course the groups $G, G_\infty$, as well as the corresponding dimension of the abelianization $d = \dim(G/[G,G])$, the nilpotency step $s$ and the vector space dimension $m$ all depend on $\Gamma$.

Let us fix two finitely generated torsion-free nilpotent groups $\Gamma$ and $\Lambda$ that are integrably measure equivalent with integrable cocycles as in (2.1) for which the action $\Gamma \curvearrowright (X,m)$ is pmp ergodic. We denote their Mal’cev Lie groups $G$ and $H$ and their graded lie groups $G_\infty$ and $H_\infty$, respectively. Let us now fix finite generating sets $S$ and $T$ for $\Gamma$ and $\Lambda$ respectively. We will denote their respective word norms $|\cdot|_\Gamma$ and $|\cdot|_\Lambda$ and the left-invariant metrics $d_\Gamma$ and $d_\Lambda$. Let us also fix a compact generating set $K \subset H$ and denote the corresponding word norm and left-invariant metric $|\cdot|_H$ and $d_H$. Finally, there are the unique left-invariant Carnot-Carathéodory metrics on $H_\infty$ and $G_\infty$ associated to $d_\Gamma$ and $d_\Lambda$ by [24]. Let us denote both by $d_\infty$, as no confusion can arise.

Keep in mind that, since we are not assuming Pansu’s Theorem 1.11 a priori we do not know whether $G_\infty$ and $H_\infty$ are isomorphic groups or that the dimensions of their abelianization are the same. So let us say that in logarithmic coordinates
\[ \Lambda < H \equiv H_\infty \equiv \mathbb{R}^m \quad \text{dim}(\mathfrak{h}/\mathfrak{h}^2) = d \]
\[ \Gamma < G \equiv G_\infty \equiv \mathbb{R}^{m'} \quad \text{dim}(\mathfrak{g}/\mathfrak{g}^2) = d'. \]

We will only work in the Lie algebras \( \mathfrak{h} \) and \( \mathfrak{h}_\infty \) of \( H \) and \( H_\infty \). Let us identify as in Chapter 2 (1.2) \( V = V_1 \oplus \cdots \oplus V_s = \mathfrak{h} = \mathfrak{h}_\infty \)
with Lie brackets \([-, -]_H\) and \([-, -]_\infty\). The projections we will use are for \( \Lambda, H \) and \( H_\infty \):
\[
\pi_{ab}(a_1, \ldots, a_m) = (a_1, \ldots, a_d, 0, \ldots, 0)
\]
\[
\pi_{com}(a_1, \ldots, a_m) = (0, \ldots, 0, a_{d+1}, \ldots, a_m).
\]

We will think of the image \( \pi_{ab}(H) \equiv \mathbb{R}^d \) in order to integrate, but for notational ease we suppress the identification. Now we may define two maps essential to what follows:
\[
\alpha_{ab} : \Gamma \times X \to H
\quad \alpha_{ab}(\gamma, x) = \pi_{ab} \circ \alpha(\gamma, x)
\]
\[
\overline{\alpha}_{ab} : \Gamma \to H
\quad \overline{\alpha}_{ab}(\gamma) = \int_X \alpha_{ab}(\gamma, x) dm(x).
\]

2.5. Reduction to torsion-free nilpotent groups. Here we reduce Theorem E to the case of torsion-free nilpotent groups. Finitely generated polynomial growth groups have finite index nilpotent subgroups, which themselves have finite normal torsion subgroups. Let \( \Gamma' < \Gamma \) be a finite index subgroup. The action \( \Gamma' \curvearrowright (X, m) \) has at most \( [\Gamma : \Gamma'] \)-many ergodic components permuted by the \( \Gamma \) action. Let \( \tau_1, \ldots, \tau_l \in \Gamma \) be a complete set of representatives for \( \Gamma' \setminus \Gamma \). Consider an ergodic component \( X' \) and the integrable cocycle \( \alpha' : \Gamma' \times X' \to \Lambda \) obtained by restriction. Suppose \( \frac{1}{n} \cdot \gamma_n \to g \in G_\infty \). For each \( n \) write \( \gamma_n = \gamma'_n \tau_{n_i} \) where \( \tau_{n_i} \in \{\tau_1, \ldots, \tau_l\} \) and \( \gamma'_n \in \Gamma' \). Then \( \frac{1}{n} \cdot \gamma'_n \to g \) so by Theorem E \( \frac{1}{n} \cdot \alpha(\gamma'_n, x) \to \Phi(g) \) for some \( \Phi \) that a priori depends on the ergodic component \( X' \). Now the cocycle equality \( \alpha(\gamma_n, x) = \alpha(\gamma'_n \tau_{n_i}, x) = \alpha(\gamma'_n, \tau_{n_i} x) \alpha(\tau_{n_i}, x) \) implies
\[
d_\Lambda(\alpha(\gamma_n, x), \alpha(\gamma'_n, \tau_{n_i} x)) = |\alpha(\tau_{n_i}, x)|_\Lambda
\]
which is bounded by a constant independent of \( n \) with high probability by Markov’s inequality. Therefore

\[
d_{\infty}(\frac{1}{n} \cdot \alpha(\gamma_n', x), \frac{1}{n} \cdot \alpha(\gamma_n, x)) = o(n) \quad \text{whp.}
\]

Now let \( N \) be a finite normal subgroup of \( \Gamma \). Then \( \Gamma/N \) acts ergodically by pmp transformations on \((X, m)/N\). Since \( N \) is finite, we can find a measurable section \( s : X/N \rightarrow X \) of \( \pi : X \rightarrow X/N \). For every \( x \in X \), there is \( n_x \in N \) so that \( n_x \cdot s\pi(x) = x \). Define

\[
f : X \rightarrow \Lambda \quad f(x) = \alpha(n_x, s\pi(x))
\]

and the cocycle cohomologous to \( \alpha \) via \( f \)

\[
\alpha^{f}(\gamma, x) = f(\gamma x)^{-1} \alpha(\gamma, x) f(x).
\]

Notice that \( f \) takes finitely many values, so \( \alpha^{f} \) is integrable. A direction computation shows that \( \alpha^{f} \) restricted to \( N \) is the trivial map, so \( \alpha^{f} \) descends to a cocycle

\[
\alpha^{f} : \Gamma/N \times X/N \rightarrow \Lambda.
\]

Finally, if \( \gamma_n \in \Gamma \) is such that \( \frac{1}{n} \cdot \gamma_n \rightarrow g \), then also \( \frac{1}{n} \cdot \gamma_n \rightarrow g \) where \( \gamma = \gamma N \in \Gamma/N \). Thus

\[
\frac{1}{n} \cdot \alpha^{f}(\gamma_n, \pi x) \rightarrow \Phi(g).
\]

Again since \( f \) takes finitely many values, another application of the Markov inequality shows that

\[
d_{\Lambda}(\alpha^{f}(\gamma_n, x), \alpha(\gamma_n, x)) = o(n) \quad \text{whp}
\]

which finishes the proof.

### 3. Asymptotic Behavior Along Iterates

In this section we analyze the asymptotic behavior of \( \alpha(\gamma^n, x) \) as \( n \rightarrow \infty \) for a given \( \gamma \in \Gamma \). In the following section, we combine Proposition 3.3 and the results of this section to understand the asymptotic behavior of an arbitrary \( \alpha(\gamma, x) \). The idea in this section is to use the cocycle identity to see that \( \alpha(\gamma^n, x) \) typically behaves like a homomorphism in to a nilpotent group. Crucially, iterates
of individual elements in nilpotent groups experience an asymptotic decay in the higher order terms (commutator coordinates). In this section we use ergodicity to extend this phenomenon to a cocycle. Moreover, the position in the abelian coordinates stabilizes asymptotically, so that we have a perfect picture of the asymptotics of iterates: the higher order terms vanish, and the abelian coordinates tend to their average value.

The main results of this section are the following propositions.

**Proposition 3.1.** For every $\gamma \in \Gamma$

$$d_H(\alpha(\gamma^n, x), \delta_n \overline{\alpha_{ab}}(\gamma)) = o(n) \quad \text{in probability.}$$

Using the Guivarc'h Lemma, Proposition 3.1 immediately implies the following.

**Proposition 3.2.** For every $\gamma \in \Gamma$

$$d_\infty(\alpha(\gamma^n, x), \delta_n \overline{\alpha_{ab}}(\gamma)) = o(n) \quad \text{in probability}.$$

We will prove Proposition 3.1 by analyzing the abelian and commutator coordinates separately.

### 3.1. Abelianization Direction

In this subsection we prove the following lemma describing the asymptotic behavior of $\alpha$ along iterates in the abelianization.

**Lemma 3.3.** For a.e. $x \in X$ and every $\gamma \in \Gamma$

$$\frac{1}{n} \alpha_{ab}(\gamma^n, x) \to \overline{\alpha_{ab}}(\gamma)$$

where the convergence is of vectors in $\mathbb{R}^d$.

The proof of the lemma is an easy application of Chapter 2 Theorem 2.4 which deals with the more general sub-additive case. See also [1].

**Proposition 3.4.** Suppose $c : \Gamma \times X \to \mathbb{R}$ is a measurable cocycle over $\Gamma \acts (X, m)$ which is pmp ergodic. Then for a.e. $x \in X$ and every $\gamma \in \Gamma$

$$\frac{1}{n} c(\gamma^n, x) \to \int_X c(\gamma, x) dm(x).$$
3. ASYMMETRIC BEHAVIOR ALONG ITERATES

Proof of Lemma 3.3. \( \alpha_{ab} \) is itself a cocycle taking values in \( \mathbb{R}^d \) which we can decompose as \( d \) independent cocycles with values in \( \mathbb{R} \). Indeed there are cocycles \( \alpha_i : \Gamma \times X \to \mathbb{R} \) for \( 1 \leq i \leq d \) so that

\[
\alpha_{ab}(\gamma, x) = (\alpha_1(\gamma, x), \ldots, \alpha_d(\gamma, x), 0, \ldots, 0).
\]

We can similarly decompose the averages

\[
\overline{\alpha_{ab}}(\gamma) = (\int_X \alpha_1(\gamma, x) dm(x), \ldots, \int_X \alpha_d(\gamma, x), 0, \ldots, 0).
\]

Applying Proposition 3.4 to each of the \( \alpha_i \) finishes the proof. \( \square \)

3.2. Commutator Direction. The purpose of this subsection is to prove the following lemma describing the asymptotic behavior of \( \alpha \) along iterates in the commutator direction.

Lemma 3.5. For every \( \gamma \in \Gamma \)

\[
|\pi_{com} \circ \alpha(\gamma^n, x)|_\Lambda = o(n) \quad \text{in probability.}
\]

The proof of the lemma requires some preparation. The idea is to use the cocycle equation to write \( \alpha(\gamma^{nk}, x) = \alpha(\gamma^n, x_1)\alpha(\gamma^n, x_2) \cdots \alpha(\gamma^n, x_k) \) where \( x_{i+1} = \gamma^{ni}x \). Using Lemma 3.3 whp the abelianization of each of the \( \alpha(\gamma^n, x_i) \approx nv \) for some \( v \), so that the commutator of \( \alpha(\gamma^{nk}, x) \) is roughly the sum of the commutators of the \( \alpha(\gamma^n, x_i) \). This allows us to promote a linear bound on the commutator to an \( o(n) \) bound since the commutator direction ‘should’ grow at least quadratically.

To begin, we use a weakened form of Proposition 3.2 from \( [1] \) to obtain the \( O(n) \) bound. Recall that given a pmp action \( \Gamma \curvearrowright (X, m) \) a map \( c : \Gamma \times X \to \mathbb{R}_+ \) is a subadditive cocycle if

\[
c(\gamma_1\gamma_2, x) \leq c(\gamma_1, \gamma_2 \cdot x) + c(\gamma_2, x) \quad \forall \gamma_1, \gamma_2 \in \Gamma \quad m - a.e. \ x \in X.
\]

Proposition 3.6 (\([1]\)). Given a subadditive cocycle \( c : \Gamma \times X \to \mathbb{R}_+ \), there is \( M \geq 1 \) such that for any \( \epsilon > 0 \) there is \( C = C(\epsilon) \) such that

\[
|c|_\Gamma \geq C \quad \implies \quad m(|c(\gamma, x)| \geq M|\gamma|_\Gamma) < \epsilon.
\]
We would like to use Proposition 3.6 to draw conclusions about the size of the commutator of $\alpha(\gamma, x)$. To do this, we use Lemma 2.4 which says that projection to the commutator increases word norm by at most a universal multiplicative constant, and Lemma 2.5 which says that the norm of iterates of an element with nontrivial abelianization grows linearly up to a multiplicative constant. Combining this with Proposition 3.6 we easily deduce the following $O(n)$ bound on the commutator growth.

Since the word length of iterates of $\gamma \in \Gamma^2$ does not grow linearly, we must deal with this easy case separately.

**Lemma 3.7.** For every $\gamma \in \Gamma - \Gamma^2$ there is $M' \geq 1$ so that for any $\epsilon > 0$ there is $N$ so that for all $n \geq N$

$$m(|\pi_{com} \circ \alpha(\gamma^n, x)|_\Lambda > M'n) < \epsilon.$$  

**Proof.** We apply Proposition 3.6 to the subadditive cocycle $c : \Gamma \times X \to [0, \infty)$ defined by $c(\gamma, x) = |\alpha(\gamma, x)|_\Lambda$. We obtain $M$ and set $M' = M|\gamma|_\Gamma/k$ where $k$ is from Lemma 2.4. Fix $\epsilon > 0$. Then there is $C$ so that

$$|\gamma|_\Gamma \geq C \implies m(|\alpha(\gamma, x)|_\Lambda \geq M|\gamma|_\Gamma) < \epsilon.$$  

Set $N = C/l$ where $l$ is from Lemma 2.5. Then since $|\gamma^n|_\Gamma \leq n|\gamma|_\Gamma$,

$$n \geq N \implies |\gamma^n|_\Gamma \geq C \implies m(|\alpha(\gamma^n, x)|_\Lambda \geq Mn|\gamma|_\Gamma) < \epsilon.$$  

Finally, by Lemma 2.4

$$n \geq N \implies m(|\pi_{com} \circ \alpha(\gamma^n, x)|_\Lambda \geq M'n) < \epsilon.$$  

The proof of Lemma 3.5 is easy in case $\gamma \in \Gamma^2$.

**Lemma 3.8.** If $\gamma \in \Gamma^2$ then

$$|\pi_{com} \circ \alpha(\gamma^n, x)|_\Lambda = o(n) \text{ in probability}.$$  

**Proof.** By Markov’s inequality there is $\kappa = \max_{s \in S} \|\alpha(s, \cdot)|_\Lambda\|_1$ so that for every $M \in \mathbb{N}$
m(|α(γ^n,x)|_Λ > Mκ|γ^n|_Γ) < 1/M.

For γ ∈ Γ^2 there is a constant c > 0 so that for all n ∈ N we have |γ^n|_Γ ≤ c√n (Lemma 2.2 and Lemma 2.6). Thus for such γ we have |α(γ^n,x)|_Λ = o(n) whp. Lemma 2.4 completes the proof. □

We need one more lemma before we can prove Lemma 3.5. Let us illustrate the idea behind the lemma through the example of the Heisenberg group. Recall that in logarithmic coordinates, the multiplication in the Heisenberg group is

$$(x, y, z)(x', y', z') = (x + x', y + y', z + z' + 1/2(xy' - x'y)).$$

The non-linear growth in the z-coordinate is given by the area enclosed by the triangle formed by (x, y), (x + x', y + y') and (0, 0). So, if a pair of elements have very similar abelianizations, the z-coordinate of their product is approximately the sum z + z'. Now suppose we have k elements with uniformly controlled z-coordinates and very similar abelianizations. Then the z-coordinate of their product grows approximately linearly. Thus the z-coordinate is o(k) since the z-coordinate 'should' grow quadratically. The following lemma generalizes this idea to general finitely generated torsion-free nilpotent groups.

We define the projection on to the first t commutator coordinates

$$\pi_t : \Lambda \rightarrow \Lambda \quad \pi_t(a_1, \ldots, a_m) = (0, \ldots, 0, a_{d+1}, \ldots, a_t, 0, \ldots, 0).$$

Let d_1 be the l^1 metric on R^d and |·|_1 be the l^1 norm, so that |(x_1, \ldots, x_d)|_1 = |x_1| + ⋯ + |x_d|.

**Lemma 3.9.** Fix 0 < M < ∞ and v ∈ R^d. For each d ≤ t < m for all δ > 0 there exists K ∈ N and δ' > 0 so that for all k ≥ K and η ≥ 1, whenever there exist λ_1, \ldots, λ_k ∈ Λ such that

\begin{align*}
(3.1) & \quad d_1(\pi_{ab}λ_i, v) < ηδ'|v|_1 \\
(3.2) & \quad |\pi_tλ_i|_Λ < ηδ' \\
(3.3) & \quad |\pi_{com}λ_i|_Λ < ηM
\end{align*}

then
\[|\pi_{t+1}\lambda_1 \cdots \lambda_k|_\Lambda < \eta k\delta.\]

**Proof of Lemma 3.5** Fix \(\gamma \in \Gamma - \Gamma^2\). We obtain \(M\) as in Lemma 3.7 and set \(v = \overline{\alpha ab}(\gamma)\). We prove by induction that for every \(d \leq t \leq m\)

\[|\pi_t\alpha(\gamma^n, x)|_\Lambda = o(n) \quad \text{in probability.}\]

For \(t = d\) there is nothing to show. Suppose the result is known for \(t\). Fix \(\epsilon > 0\) and \(\delta > 0\). We apply Lemma 3.9 with the given \(\delta, M\) and \(v\) to obtain \(k = K\) and \(\delta'\). Let \(N\) be as in Lemma 3.7 applied to \(\epsilon/k\), so that for all \(\eta \geq N\) we have with probability at least \(1 - \epsilon/k\)

\[|\pi_{com}\alpha(\gamma^n, x)|_\Lambda < \eta M.\]

By taking \(N\) larger if necessary, applying the inductive hypothesis to \(\delta'/3\) and \(\epsilon/k\) we obtain \(N\) so that for all \(\eta \geq N\) we have with probability at least \(1 - \epsilon/k\) we have

\[|\pi_t\alpha(\gamma^n, x)|_\Lambda < \eta\delta'/3.\]

By taking \(N\) larger again if necessary, by Lemma 3.3 for all \(\eta \geq N\) with probability at least \(1 - \epsilon/k\)

\[d_1(\pi_{ab}\alpha(\gamma^n, x), \eta v) < \eta\delta'|v|/3.\]

Since the \(\Gamma\) action on \((X, \mu)\) is measure preserving, the previous three statements remain true if we replace any instance of \(x\) with \(gx\) for any \(g \in \Gamma\).

Finally, let \(N\) be larger if necessary so that \(k \leq \delta'N\). Now let \(p \geq kN\). Write \(p = \eta k + r\) where \(0 \leq r < k\) and \(\eta \geq N\). Using the cocycle equation

\[\alpha(\gamma^{k\eta + r}, x) = \alpha(\gamma^n, x)\alpha(\gamma^n, \gamma^n x) \cdots \alpha(\gamma^n, \gamma^{(k-2)n} x)\alpha(\gamma^{n+r}, \gamma^{(k-1)n} x).\]

Since \(\eta, \eta + r \geq N\), with probability at least \(1 - 3\epsilon\) we have simultaneously for all \(0 \leq i \leq k - 2\)
\[ |\pi_{\text{com}} \alpha(\gamma^n, \gamma^m x)|_\Lambda < \eta M \]
\[ |\pi_t \alpha(\gamma^n, \gamma^m x)|_\Lambda < \eta \delta'/3 \]
\[ d_1(\pi_{ab} \alpha(\gamma^n, \gamma^m x), \eta v) < \eta \delta|v|_1/3 \]

and

\[ |\pi_{\text{com}} \alpha(\gamma^{n+\tau}, \gamma^{(k-1)n} x)|_\Lambda < (\eta + r) M \]
\[ |\pi_t \alpha(\gamma^{n+\tau}, \gamma^{(k-1)n} x)|_\Lambda < (\eta + r) \delta'/3 \]
\[ d_1(\pi_{ab} \alpha(\gamma^{n+\tau}, \gamma^{(k-1)n} x), (\eta + r)v) < (\eta + r) \delta'|v|_1/3. \]

Since \( r \leq \delta' \eta \) the final three inequalities imply

\[ |\pi_{\text{com}} \alpha(\gamma^{n+\tau}, \gamma^{(k-1)n} x)|_\Lambda < 2\eta M \]
\[ |\pi_t \alpha(\gamma^{n+\tau}, \gamma^{(k-1)n} x)|_\Lambda < \eta \delta' \]
\[ d_1(\pi_{ab} \alpha(\gamma^{n+\tau}, \gamma^{(k-1)n} x), \eta v) < \delta'|v|_1 \]

where for the final inequality we have used the triangle inequality with intermediate term \((\eta + r)v\).

Therefore with probability at least \( 1 - 3\epsilon \) we apply Lemma 3.9 and obtain

\[ |\pi_{t+1} \alpha(\gamma^p, x)|_\Lambda < k\eta \delta < p\delta. \]

**Proof of Lemma 3.9** Fix \( 0 < M < \infty, v \in \mathbb{R}^d, d \leq t < m, \delta > 0 \) and \( 1 > \delta' > 0. \) We will show in the proof how to choose \( \delta' \) as a function of \( \delta, |v|_1, t. \) Choose \( K \) large so that \( M/\sqrt{K} \leq \delta^2, \) and fix \( k \geq K \) and \( \eta \geq 1. \) Suppose we have \( \lambda_1, \ldots, \lambda_k \) satisfying conditions (3.1) (3.2) (3.3). Let us denote \( \lambda_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,m}) \) for each \( 1 \leq i \leq k, \) keeping in mind that only \( a_{i,1}, \ldots, a_{i,t+1} \) are relevant. Throughout this proof \( c \) will denote an ever-changing constant this is independent of \( \delta, \delta' \) and \( \eta. \)
We are concerned with the absolute value of the \( t+1 \) coordinate of the product \( \lambda_1 \cdots \lambda_k \). By Lemma 2.2 it suffices to show that the absolute value of this coordinate is at most \( c(\eta k \delta)^{d(t+1)} \). The estimate we seek will follow from the Baker-Campbell-Hausdorff equation and the following constraints on the \( a_{i,j} \) implied by conditions (3.1), (3.2) and (3.3):

\[
|a_{i,j} - a_{i',j}| \leq c \eta \delta' |v_1| \\
\forall 1 \leq i, i' \leq k \\
\forall 1 \leq j \leq d 
\]

(3.4) \hspace{2cm} (3.5) \hspace{2cm} (3.6) \hspace{2cm} (3.7)

\[
|a_{i,j}| \leq c(\eta \delta')^{d(j)} \\
\forall 1 \leq i \leq k \\
\forall d < j \leq t 
\]

\[
|a_{i,t+1}| \leq c(\eta M)^{d(t+1)} \\
\forall 1 \leq i \leq k 
\]

\[
|a_{i,j}| \leq c \eta^{d(j)} \\
\forall 1 \leq i \leq k \\
\forall 1 \leq j \leq t. 
\]

Indeed, setting \( v = (v_1, \ldots, v_d) \), from (3.1) we have \( \sum_{j=1}^d |a_{i,j} - v_j| \leq \eta \delta' |v_1| \) which implies in particular \( |a_{i,j} - v_j| \leq \eta \delta' |v_1| \) for all \( i \), giving (3.4). Combining (3.2), Lemma 2.3 and Lemma 2.2 we immediately arrive at (3.5). Similarly combining (3.3), Lemma 2.3 and Lemma 2.2 we arrive at (3.6). It only remains to prove (3.7) in the case \( 1 \leq j \leq d \), which follows from \( |a_{i,j} - v_j| \leq \eta \delta' |v_1| \) above and \( |v_j| \leq |v_1| \).

By the Baker-Campbell-Hausdorff equation we can express the product \( \lambda_1 \cdots \lambda_k \) as a sum of terms of the form

\[
c[\lambda_{i_1}, \ldots, [\lambda_{i_{l-1}}, \lambda_{i_l}], \ldots] 
\]

where \( i_j \in \{1, \ldots, k\} \) for each \( 1 \leq j \leq l \leq m \). We emphasize that it is possible that the indices are repeated, i.e. that \( i_j = i_{j'} \) while \( j \neq j' \). We are only interested in the brackets that contribute to the coefficient of \( X_{t+1} \). We replace each \( \lambda_i \) with \( \sum_{j=1}^m a_{i,j} X_j \) in each of the summands (3.8) above. Using linearity of the Lie bracket, the result is a sum of terms of the form

\[
c[a_{i_1,j(i_1)} X_{j(i_1)}, \ldots, [a_{i_{l-1},j(i_{l-1})} X_{j(i_{l-1})}, a_{i_l,j(i_l)} X_{j(i_l)}], \ldots] 
\]

where for each \( i_r \) we have chosen \( j(i_r) \in \{1, \ldots, t+1\} \). By Lemma 2.2 we have that
so that in particular $l \leq t+1$. We will show that each such term is small by analyzing the possibilities for the choices $j(i_r)$ above. We consider three cases.

For the first case we consider all terms with $j(i_r) = t+1$ for some $r$. Note that in this case, in view of (3.10) in fact (3.9) becomes

$$ca_{i_1, t+1}X_{t+1}.$$  

In view of (3.6), summing these over all $1 \leq i_1 \leq k$, the total contribution to the $t+1$ term from this case is, in absolute value, at most

$$ck(\eta M)^{d_{t+1}} \leq cn^{d_{t+1}}k^{1+d_{t+1}/2}$$

by our choice of $k$. This suffices since we may assume $d_{t+1} \geq 2$.

For the second case, we consider all terms in which at least one of the $j(i_r) \in \{d, \ldots, t+1\}$. By linearity we pull out all of the constants $a_{i, j}$ and consider the size of their product. By our assumption and (3.5) one of the terms is at most $c(\eta\delta')^{d_{j(i_r)}}$ and by (3.7) the rest of the terms are at most $c\eta^{d_{j(i_r)}}$. Therefore their product is at most

$$c\delta' \eta \sum_{r=1}^l d_{j(i_r)} \leq c\delta' \eta^{d_{t+1}}.$$  

Since there are finitely many such terms independent of $\delta'$, by taking $\delta'$ small as a function of $\delta, c, t$ and the number of such terms, the total contribution to the $t+1$ coordinate of the product $\lambda_1 \cdots \lambda_k$ from terms of the second type is as desired.

For the third and final case we group each term into pairs and use antisymmetry, as follows. We may assume all terms $i(j_r) \in \{1, \ldots, d\}$. In particular the inner most term $[a_{i_{t-1}, j(i_{t-1})}X_{j(i_{t-1})}, a_{i_t, j(i_t)}X_{j(i_t)}]$ has $j(i_{t-1}) = s, j(i_t) = t$ for some $s, t \in [1, \ldots, d]$. We pair the terms for which $j(i_{t-1}) = s, j(i_t) = t$ with that for which $j(i_{t-1}) = t, j(i_t) = s$, and all other $j(i_r)$ equal. By anti-symmetry of the bracket, the sum of these two terms is
By properties (3.4) and (3.7) and the triangle inequality we have

\[ a_{i_1,i_1}X_{j(i_1)}, \ldots, [a_{i_{l-1},j(i_{l-1})}X_{j(i_{l-1})}, a_{i_{l-1},j(i_{l-1})}X_{j(i_{l-1})}] \ldots \]

+ \[ [a_{i_{l-1},j(i_{l-1})}X_{j(i_{l-1})}, \ldots, [a_{i_{l-1},j(i_{l-1})}X_{j(i_{l-1})}, a_{i_{l-1},j(i_{l-1})}X_{j(i_{l-1})}] \ldots \]

= \[ [a_{i_1,j(i_1)}X_{j(i_1)}, \ldots, (a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})} - a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})}) ]X_{j(i_{l-1})}, X_{j(i_{l-1})}] \ldots \]

Pulling the constants out and considering the absolute value of the coefficient, we are concerned with the absolute value of

\[(3.11) a_{i_1,j(i_1)} \cdots a_{i_{l-1},j(i_{l-1})}(a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})} - a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})}).\]

By properties (3.4) and (3.7) and the triangle inequality we have

\[
\left|(a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})} - a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})})\right| \leq \left|a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})} - a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})}\right| \\
+ \left|a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})} - a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})}\right| \\
\leq \left|a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})} - a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})}\right| \\
+ \left|a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})} - a_{i_{l-1},j(i_{l-1})}a_{i_{l-1},j(i_{l-1})}\right| \\
\leq c\eta\lambda'\delta' |v|_1 + c\eta\lambda' \delta' |v|_1 = c\eta^2 \delta' |v|_1. 
\]

Now by (3.7) each of the other terms in the product (3.11) has absolute value at most \(c\eta\). Putting this together with the preceding and noting that \(l \leq d_{\delta+1}\), the absolute value of (3.11) is at most \(c\eta^{d+1} \delta' |v|_1\). Since there are a finite number of such terms independent of \(\delta'\), by taking \(\delta'\) small as a function of \(\delta, c, |v|_1\), the total contribution to the absolute value of the \(t+1\) coordinate of \(\lambda_1 \cdots \lambda_k\) from terms of the third case is as desired. This finishes the proof.

\[\square\]

3.3. Proof of Proposition 3.1. Finally we can combine Lemmas 2.7, 3.3 and 3.5 to prove Proposition 3.1.

Proof of Proposition 3.1. Fix \(\gamma \in \Gamma\). Chow’s Theorem, the homogeneity of \(d_\infty\) and Lemma 3.3 imply
which implies in particular that

\begin{equation}
|\alpha_{ab}(\gamma^n, x)|_H = O(n) \quad \text{in probability.}
\end{equation}

Now we use the triangle inequality

\[
d_H(\alpha(\gamma^n, x), n\pi_{ab}(\gamma)) \leq d_H(\alpha(\gamma^n, x), \alpha_{ab}(\gamma^n, x)) + d_H(\alpha_{ab}(\gamma^n, x), n\pi_{ab}(\gamma)).
\]

The second summand is \(o(n)\) by (3.12). For the first summand, we apply Lemma 2.7 with \(h_n(x) = \alpha(\gamma^n, x); \) by (3.13), \(|\pi_{ab}h_n(x)|_H = O(n)\) in probability, while Lemma 3.5 implies \(|\pi_{com}h_n(x)|_H = o(n)\) in probability.

\[
\square
\]

4. Asymptotic Behavior Along Arbitrary Elements

The goal of this section is to prove the following. Note that by Proposition 3.3 in [1] such a \(K\) always exists.

**Theorem 4.1.** Fix an arbitrary \(K \in \mathbb{N}\) and suppose that for each \(\gamma \in \Gamma\) we have a fixed word in the generators evaluating to \(\gamma\)

\[
\gamma = s_1^{a_1} \cdots s_k^{a_k}
\]

where \(s_i \in S, a_i \in \mathbb{N},\) and \(k \leq K.\) Then whp

\[
d_{\infty}(\alpha(\gamma, x), \delta_{a_1} \alpha_{ab}(s_1) \star \cdots \star \delta_{a_k} \alpha_{ab}(s_k)) = o(\sum a_i).
\]
There is an obvious way to compare the two points above. Using the cocycle equation along the given word representing \( \gamma \in \Gamma \) we write

\[
\alpha(\gamma, x) = \alpha(s_{a_1}^{a_1}, x_1) \cdots \alpha(s_{a_k}^{a_k}, x_k)
\]

where \( x_i := s_{a_i}^{a_i+1} \cdots s_{a_k}^{a_k} x \). Proposition 3.2 relates \( \alpha(s_{a_i}^{a_i}, x_i) \) and \( \delta_{a_{a_1}}(s_1) \cdots \delta_{a_{a_k}}(s_k) \). In this section we use the uniform boundedness \( k \leq K \) to extend Proposition 3.2 to Theorem 4.1.

**Theorem 4.1** follows immediately from the triangle inequality and the two following Lemmas.

**Lemma 4.2.**

\[
d_\infty(\alpha(s_{a_1}^{a_1}, x_1) \cdots \alpha(s_{a_k}^{a_k}, x_k), \delta_{a_{a_1}}(s_1) \cdots \delta_{a_{a_k}}(s_k)) = o(\sum a_i) \quad \text{in probability}
\]

where \( a_i \in \mathbb{N}, s_i \in S \) and \( k \leq K \).

**Lemma 4.3.**

\[
d_\infty(\alpha(s_{a_1}^{a_1}, x_1) \cdots \alpha(s_{a_k}^{a_k}, x_k), \alpha(s_{a_k}^{a_k}, x_1) \cdots \alpha(s_{a_1}^{a_1}, x_k)) = o(\sum a_i) \quad \text{in probability}
\]

where \( a_i \in \mathbb{N}, s_i \in S \) and \( k \leq K \).

The main subtlety throughout this section is that as \( \sum a_i \to \infty \) there is no guarantee that each of the \( a_i \to \infty \).

**4.1. Proof of Lemma 4.2** Lemma 4.2 is an ergodic-theoretic formulation of Lemma 4.4, which is pure nilpotent geometry. Here is the idea. Suppose \( g_1, \ldots, g_k, h_1, \ldots, h_k \in H_\infty \) are pairwise very close asymptotically; i.e. \( d_\infty(g_i, h_i) \leq \delta \min(|g_i|_\infty, |h_i|_\infty) \). Then the distance between \( g_1 \cdots g_k \) and \( h_1 \cdots h_k \) will be asymptotically small as well; i.e.

\[
d_\infty(g_1 \cdots g_k, h_1 \cdots h_k) \leq \tau \min(|g_1 \cdots g_k|_\infty, |h_1 \cdots h_k|_\infty). \]

The number of elements \( k \) affects the relationship between \( \delta \) and \( \tau \). Furthermore if we were to allow some of the pairs \( g_j \) and \( h_j \) to not be close but to instead be small in size relative to the products \( g_1 \cdots g_k, h_1 \cdots h_k \), then the conclusion would remain true. This fact relies on the slow commutator growth in nilpotent groups, Lemma 2.8.

**Lemma 4.4.** Let \( K \in \mathbb{N} \). For all \( \tau > 0 \) there exist \( \delta > 0 \) so that for all \( A, B \in \mathbb{N} \) there exists \( N \in \mathbb{N} \) so that whenever \( g_1, \ldots, g_k, h_1, \ldots, h_k \in H_\infty, a_1, \ldots, a_k \in \mathbb{N} \) are such that \( k \leq K \) and such that for all \( 1 \leq i \leq k \)
4. ASYMPTOTIC BEHAVIOR ALONG ARBITRARY ELEMENTS

(1) \( a_i < A \implies |g_i|_\infty \leq B a_i \)
(2) For all \( i \) \( |h_i|_\infty \leq B a_i \)
(3) \( a_i \geq A \implies d_\infty (g_i, h_i) < \delta a_i \)
(4) \( \sum_{i=1}^{k} a_i \geq N \)

Then

\[ d_\infty (g_1 \cdots g_k, h_1 \cdots h_k) \leq \tau \sum_{i=1}^{k} a_i. \]

Moreover, the same is true with \( H \) in place of \( H_\infty \).

Let us now explain how to deduce Lemma 4.2 from Lemma 4.4. We want to show that whp as \( \sum a_i \to \infty \) the elements \( g_i := \alpha(s_i^{a_i}, x_i) \) and \( h_i := \delta a_i \overline{\alpha_{ab}}(s_i) \) satisfy (1) – (3) from Lemma 4.4.

Notice that (2) is immediate.

We pick a large threshold \( A \in \mathbb{N} \) so that when \( a_i \geq A \) then Proposition 3.2 applies to all \( s \in S \) whp. Thus we have (3). In order to control the small \( a_i \) we use the following Markov inequality to obtain (1).

**Lemma 4.5.** There exists \( c > 1 \) so that for all \( \gamma \in \Gamma \) and for all \( M \geq 1 \)

\[ m(|\alpha(\gamma, x)|_\infty \geq c M |\gamma|_\Gamma) \leq 1/M \]

\[ m(|\alpha(\gamma, x)|_H \geq c M |\gamma|_\Gamma) \leq 1/M. \]

Proof of Lemma 4.2. Fix \( \epsilon, \tau > 0 \) and obtain \( \delta > 0 \) from Lemma 4.4. Fix \( M \) large so that \( 1/M \leq \epsilon \). By Lemma 4.5

\[ m(|\alpha(s_i^{a_i}, x_i)|_\infty \leq c M a_i) \geq 1 - \epsilon. \]

By Proposition 3.2 there exists \( A < \infty \) so that for all \( s_i \in S \) whenever \( a_i \geq A \)

\[ m(d_\infty (\alpha(s_i^{a_i}, x_i), \delta a_i \overline{\alpha_{ab}}(s_i) < a_i \delta) \geq 1 - \epsilon. \]
Set $B = cM(1 + \max_{s \in S} |\overline{\alpha_{ab}}(s)|_\infty)$ and recall that there is a fixed $K$ so that $k \leq K$. Combining (4.1) and (4.2), for any given $\gamma \in \Gamma$, with probability at least $1 - K\epsilon$, the following occurs simultaneously

\begin{equation}
|a(s_i^{a_i}, x_i)|_\infty \leq B a_i \quad \forall a_i < A
\end{equation}

\begin{equation}
|\delta_a, \overline{\alpha_{ab}}(s_i)|_\infty \leq B a_i \quad \forall a_i
\end{equation}

\begin{equation}
d_\infty(\alpha(s_i^{a_i}, x_i), \delta_a, \overline{\alpha_{ab}}(s_i)) < a_i \delta \quad \forall a_i \geq A.
\end{equation}

Clearly we may assume $\sum a_i \geq N$ by taking $\sum a_i$ large. Therefore $g_i = \alpha(s_i^{a_i}, x_i)$ and $h_i = \delta_a, \overline{\alpha_{ab}}(s_i)$ satisfy the conditions of Lemma 4.4 with probability at least $1 - K\epsilon$.  

**Proof of Lemma 4.4** By the triangle inequality and the left invariance of the metric

\begin{equation}
d_\infty(g_1 \cdots g_k, h_1 \cdots h_k) \leq d_\infty(g_1 \cdots g_k, g_1 \cdots g_{k-1} \ast h_k) + \\
+ d_\infty(g_1 \cdots g_{k-1} h_k, g_1 \cdots g_{k-2} h_{k-1} h_k) + \cdots + \\
+ d_\infty(g_1 \cdots g_i h_{i+1} \cdots h_k, g_1 \cdots g_{i-1} h_{i+1} \cdots h_k) + \cdots + \\
+ d_\infty(g_1 h_2 \cdots h_k, h_1 \cdots h_k)
\end{equation}

\begin{align*}
&\leq \sum_{i=1}^{k} |(h_{i+1} \cdots h_k)^{-1} (h_i^{-1} \ast g_i) (h_{i+1} \cdots h_k)|_\infty \\
&\leq \sum_{i=1}^{k} |h_i^{-1} \ast g_i|_\infty \ast |(h_{i+1} \cdots h_k)^{-1} (h_i^{-1} \ast g_i)(h_{i+1} \cdots h_k)(h_i^{-1} \ast g_i)^{-1}|_\infty \\
&= \sum_{i=1}^{k} |\gamma_i|_\infty \ast |\eta_i^{-1} \ast \gamma_i \ast \eta_i \ast \gamma_i^{-1}|_\infty
\end{align*}

where we have set $\gamma_i = h_i^{-1} \ast g_i$ and $\eta_i = h_{i+1} \cdots h_k$. Consider first the terms $|h_i^{-1} \ast g_i|_\infty$. If $a_i \geq A$ then $|h_i^{-1} \ast g_i|_\infty < \delta \sum a_i$ by (3). So take $\delta < \tau$. If $a_i < A$ then $|g_i|_\infty \leq AB$ by (1) while $|h_i|_\infty \leq AB$ by (2). Therefore $|h_i^{-1} \ast g_i|_\infty \leq 2AB \leq \tau \sum a_i$ for large enough $\sum a_i$.

To control the term $|\eta_i^{-1} \ast \gamma_i \ast \eta_i \ast \gamma_i^{-1}|_\infty$ we employ Lemma 2.8. The conclusion of the above paragraph is that $|\gamma_i|_\infty = o(\sum a_i)$. On the other hand $|\eta_i|_\infty = |h_{i+1} \cdots h_k|_\infty \leq \sum_{i+1}^{k} |h_j|_\infty \leq B \sum a_i = O(\sum a_i)$. This finishes the proof. The proof for $H$ in place of $H_\infty$ is identical with the obvious changes.  

□
4. ASYMPTOTIC BEHAVIOR ALONG ARBITRARY ELEMENTS

**Proof of Lemma 4.5.** By Markov’s inequality and the integrability assumption there exists $c' > 0$ so that for all $\gamma \in \Gamma$

$$m(|\alpha(\gamma,x)|_\Lambda \geq Mc'|\gamma|_T) \leq 1/M.$$ 

Since $d_\Lambda$ and $d_H$ are quasi-isometric on $\Lambda$ and by Guivarc’h’s Lemma $d_H$ and $d_\infty$ are quasi-isometric on $H$ and $H_\infty$, there exist $c_1, c_2 > 0$ so that for all $\lambda \in \Lambda$

$$|\lambda|_\Lambda \geq \frac{1}{c_1} |\lambda|_\infty - c_2$$
$$|\lambda|_\Lambda \geq \frac{1}{c_1} |\lambda|_H - c_2.$$

Therefore we may take $c = Mc' + c_1c_2$. □

**4.2. Proof of Lemma 4.3.** Lemma 4.3 is an ergodic-theoretic formulation of Lemma 4.6 which is pure nilpotent geometry.

**Lemma 4.6.** Fix $K \in \mathbb{N}$. For all $\tau > 0$ there exists $\delta > 0$ so that for all $A, B \in \mathbb{N}$ there exist $N \in \mathbb{N}$ whenever there exist $g_1, \ldots, g_k, h_1, \ldots, h_k \in H$ and $a_1, \ldots, a_k \in \mathbb{N}$ with $k \leq K$ such that for all $1 \leq i \leq k$

1. $a_i < A \implies |g_i|_H, |g_i|_\infty < Ba_i$
2. $\forall i \quad |h_i|_H, |h_i|_\infty \leq B$
3. $a_i \geq A \implies d_H(g_i, \delta_{a_i} h_i), d_\infty(g_i, \delta_{a_i} h_i) \leq \delta a_i$
4. $\sum a_i \geq N$

then

$$d_\infty(g_1 \cdots g_k, g_1 \cdots g_k) \leq \tau \sum a_i.$$

**Proof of Lemma 4.3.** Fix $\epsilon, \tau > 0$. Obtain $\delta$ from Lemma 4.6. Fix $M$ be large so that $1/M \leq \epsilon$. By Lemma 4.5 for all $a_i$

$$m(|\alpha(s_i^{a_i}, x)\rangle_{\infty} \leq cMa_i) \geq 1 - \epsilon$$

$$m(|\alpha(s_i^{a_i}, x)\rangle_H \leq cMa_i) \geq 1 - \epsilon.$$
By Propositions 3.1 and 3.2 there exists $A < \infty$ so that for all $s_i \in S$ whenever $a_i \geq A$

\[(4.8) \quad m(d_H(\alpha(s_i^{a_i}, x_i), \delta_{a_i} \overline{\alpha_{ab}}(s_i)) < a_i \delta) \geq 1 - \epsilon \]

\[(4.9) \quad m(d_\infty(\alpha(s_i^{a_i}, x_i), \delta_{a_i} \overline{\alpha_{ab}}(s_i)) < a_i \delta) \geq 1 - \epsilon. \]

Recall that there is a fixed $K$ so that $k \leq K$. Set $B = cM(1 + \max_{s \in S} |\overline{\alpha_{ab}}(s)|_\infty + \max_{s \in S} |\overline{\alpha_{ab}}(s)|_H)$.

Combining (4.6), (4.7), (4.8) and (4.9), with probability at least $1 - 4K\epsilon$, the following occurs simultaneously

$$d_H(\alpha(s_i^{a_i}, x_i), \delta_{a_i} \overline{\alpha_{ab}}(s_i)), d_\infty(\alpha(s_i^{a_i}, x_i), \delta_{a_i} \overline{\alpha_{ab}}(s_i)) < a_i \delta$$

with probability at least $1 - 4K\epsilon$. Therefore, for $\sum a_i$ large, with probability at least $1 - 4K\epsilon$

$$d_\infty(g_1 \cdots g_k, g_1 \ast \cdots \ast g_k) \leq \tau \sum a_i.$$

The proof of Lemma 4.6 requires the following natural lemma. It is a generalization to several possibly different scaling sequences of the fact, originally proved implicitly in \[24\], that while the group law in $\Lambda$ does not appear to be related to the law of its associated graded Lie group $H_\infty$, in fact the embedding $L : \Lambda \rightarrow H_\infty$ is asymptotically a homomorphism (cf. \[8\] Lemma 6.12).

**Lemma 4.7.** For all $k \in \mathbb{N}$ and all $h_1, \ldots, h_k \in H$

$$d_\infty(\delta_{a_1} h_1 \cdots \delta_{a_k} h_k, \delta_{a_1} h_1 \ast \cdots \ast \delta_{a_k} h_k) = o(\max a_i)$$
Proof of Lemma 4.6. Fix $A, B, K \in \mathbb{N}$ and $\tau > 0$. For any $g_i, h_i \in H$, by the triangle inequality

$$d_\infty(g_1 \cdots g_k, g_1 \star \cdots \star g_k) \leq d_\infty(g_1 \cdots g_k, \delta a_1 h_1 \cdots \delta a_k h_k)$$

$$+ d_\infty(\delta a_1 h_1 \cdots \delta a_k h_k, \delta a_1 h_1 \star \cdots \star \delta a_k h_k)$$

$$+ d_\infty(\delta a_1 h_1 \star \cdots \star \delta a_k h_k, g_1 \star \cdots \star g_k).$$

The second term is $o(\sum a_i)$ by Lemma 4.7. For the first and third terms, we choose $\delta > 0$ and $N \in \mathbb{N}$ according to Lemma 4.4. For the first term we must also use the Lemma of Guivarc'h. □

It remains to prove Lemma 4.7. We begin with a single term in the Lie algebra.

Lemma 4.8. For all $k \in \mathbb{N}$ and all $X_1, \ldots, X_k \in V$

$$\lim_{\max(a_i) \to \infty} \delta^{-1}_{\max(a_i)}([\delta a_1 X_1, [\delta a_2 X_2, \ldots, [\delta a_{k-1} X_{k-1}, \delta a_k X_k]_H \cdots ]]_H$$

$$- [\delta a_1 X_1, [\delta a_2 X_2, \ldots, [\delta a_{k-1} X_{k-1}, \delta a_k X_k]_\infty \cdots ]_\infty) = 0.$$

Proof. By linearity it suffices to prove the lemma under the assumption that $X_1 \in V_1, \ldots, X_k \in V_k$ where $V = V_1 \oplus \cdots \oplus V_s$ is the decomposition of $V$ in (1.2). Set $m = \max(n_1, \ldots n_k)$ and $r = i_1 + \cdots + i_k$. Then on the one hand

$$[\delta a_1 X_1, [\delta a_2 X_2, \ldots, [\delta a_{k-1} X_{k-1}, \delta a_k X_k]_H \cdots ]]_H$$

$$= [a_1^{i_1} X_1, [a_2^{i_2} X_2, \ldots, [a_{k-1}^{i_{k-1}} X_{k-1}, a_k^{i_k} X_k]_H \cdots ]]_H$$

$$= a_1^{i_1} \cdots a_k^{i_k} [X_1, X_2, \ldots, [X_{k-1}, X_k]_H \cdots ]_H$$

$$= a_1^{i_1} \cdots a_k^{i_k} (Z_r + Z_{r+1} \cdots + Z_s)$$

where $Z_t \in V_t$.

On the other hand
\[ [\delta a_1 X_1, [\delta a_2 X_2, \ldots, [\delta a_{k-1} X_{k-1}, \delta a_k X_k]_\infty]_\infty \]
\[ = [a_1^i X_1, [a_2^i X_2, \ldots, [a_{k-1}^i X_{k-1}, a_k^i X_k]_\infty]_\infty \]
\[ = a_1^i \cdots a_k^i [X_1, [X_2, \ldots, [X_{k-1}, X_k]_\infty]_\infty \]
\[ = a_1^i \cdots a_k^i Z_r. \]

Therefore taking the difference and applying \( \delta_m^{-1} \) we have

\[ \delta_m^{-1}([\delta a_1 X_1, [\delta a_2 X_2, \ldots, [\delta a_{k-1} X_{k-1}, \delta a_k X_k]_H \cdots]_H \]
\[ - [\delta a_1 X_1, [\delta a_2 X_2, \ldots, [\delta a_{k-1} X_{k-1}, \delta a_k X_k]_\infty \cdots]_\infty) \]
\[ = a_1^i \cdots a_k^i m^{-r-1} Z_r + \cdots + a_1^l \cdots a_k^l m^{-s} Z_s \to 0 \text{ as } m \to \infty. \]

Multilinearity of the bracket and the above imply the following, which, by Chow’s Theorem, is sufficient to prove Lemma 4.7.

**Lemma 4.9.** For all \( h_1, \ldots, h_k \in H \)

\[ \lim_{\max a_i \to \infty} \delta_m^{-1} a_i \delta a_1 h_1 \cdots \delta a_k h_k = 0 \]

where the convergence is of vectors in \( \mathbb{R}^m \).

**Proof of Lemma 4.7.** Let \( h_1 = \exp_H X_1, \ldots, h_k = \exp_H X_k \). Then by the definition of \( \delta_n \) on \( H \) and the Baker-Campbell-Hausdorff formula

\[ \delta a_1 h_1 \cdots \delta a_k h_k = \log_H (\delta a_1 \exp_H X_1 \cdots \delta a_k \exp_H X_k) \]
\[ = \log_H (\exp_H \delta a_1 X_1 \cdots \exp_H \delta a_k X_k) \]
\[ = \delta a_1 X_1 + \delta a_2 X_2 + \cdots + \delta a_k X_k + \frac{1}{2} \sum_{1 \leq i < j \leq k} [\delta a_i X_i, \delta a_j X_j]_H \]
\[ + \sum_{2 \leq t \leq k} c_t \sum_{1 \leq i_1 < i_2 < \cdots < i_t \leq k} [\delta a_{i_1} X_{i_1}, \cdots, [\delta a_{i_{t-1}} X_{i_{t-1}}, \delta a_{i_t} X_{i_t}]_H \cdots]_H \]
where \(c_l\) are constants. The same calculation holds for the product in \(H_\infty\), except that the \(H\)-bracket is replaced with the \(H_\infty\)-bracket.

\[
\delta_{a_1} h_1 \cdots \delta_{a_k} h_k - \delta_{a_1} h_1 \cdots \delta_{a_k} h_k = \\
= \frac{1}{2} \sum_{1 \leq i < j \leq k} [\delta_{a_i} X_i, \delta_{a_j} X_j]_H - [\delta_{a_i} X_i, \delta_{a_j} X_j]_\infty + \\
+ \sum_{2 < l \leq k} c_l \sum_{1 \leq i_1 < i_2 < \cdots < i_l \leq k} [\delta_{a_{i_1}} X_{i_1}, \cdots [\delta_{a_{i_{l-1}}} X_{i_{l-1}}, \delta_{a_{i_l}} X_{i_l}]]_H \cdots ]_\infty
\]

Subtracting we obtain

\[
\delta_{a_1} h_1 \cdots \delta_{a_k} h_k = \\
= \frac{1}{2} \sum_{1 \leq i < j \leq k} [\delta_{a_i} X_i, \delta_{a_j} X_j]_H - [\delta_{a_i} X_i, \delta_{a_j} X_j]_\infty + \\
+ \sum_{2 < l \leq k} c_l \sum_{1 \leq i_1 < i_2 < \cdots < i_l \leq k} [\delta_{a_{i_1}} X_{i_1}, \cdots [\delta_{a_{i_{l-1}}} X_{i_{l-1}}, \delta_{a_{i_l}} X_{i_l}]]_H \cdots ]_\infty - \\
[\delta_{a_{i_1}} X_{i_1}, \cdots [\delta_{a_{i_{l-1}}} X_{i_{l-1}}, \delta_{a_{i_l}} X_{i_l}]]_\infty \cdots ]_\infty
\]

Let \(m = \max(a_1, \ldots, a_k)\). Then applying Lemma \[4.9\] and multilinearity we have

\[
\lim_{m \to \infty} \delta_{a_1}^{-1} h_1 \cdots \delta_{a_k} h_k - \delta_{a_1} h_1 \cdots \delta_{a_k} h_k = 0.
\]

\[\square\]

5. Construction of \(\Phi\) and Proof of Main Theorem

In this section we construct \(\Phi\), prove Theorem \[E\] and deduce Theorem \[D\]. We begin by proving an analog of \[1\] Proposition 3.3 which says that in a finitely generated nilpotent group all elements can be written in a uniformly bounded number of ‘straight lines.’ Our analog, Proposition \[5.3\] is for graded nilpotent lie groups, and in place of powers of generators, we use homotheties.

**Definition 5.1.** Let \((G_\infty, \delta_t)\) be a graded nilpotent lie group with its one-parameter family of automorphisms. A finite symmetric subset \(S \subset G_\infty\) generates \(G_\infty\) with respect to \(\delta_t\), \(t \geq 0\), if for every \(g \in G_\infty\) there exist \(k \in \mathbb{N}, s_1, \ldots, s_k \in S\) and \(a_1, \ldots, a_k \in \mathbb{R}_+\) so that
5. CONSTRUCTION OF $\Phi$ AND PROOF OF MAIN THEOREM

\[ g = \delta_{a_1}s_1 \cdots \delta_{a_k}s_k. \]

**Example 5.2.** In the Malcev coordinates on $G_\infty$, the set of \( d' = \dim(G_\infty/G^2_\infty) \) elements

\[ \{(1,0,\ldots,0), (0,1,0,\ldots), (0,\ldots,1,0,\ldots)\} \]

together with their inverses form a finite symmetric generating set for $G_\infty$ with respect to the homotheties $\delta_t$.

More generally any finite symmetric set with real span containing $V_1 = g/g^2$ generates $G_\infty$ with respect to $\delta_t$. Indeed, the group generated by $\exp_\infty(V_1)$ is a connected subgroup of $G_\infty$, so by the Lie correspondence, its Lie algebra is a sub algebra of $g_\infty$ containing $V_1$. Since $V_1$ generates $g_\infty$ as a Lie algebra, the group generated by $\exp_\infty(V_1)$ is all of $G_\infty$.

Let $G_1^\infty := G_\infty$ and set $G^{i+1}_\infty := [G_\infty, G_i^\infty]$, so that $G_c^\infty = 1$ if $c$ is the nilpotency class of $G_\infty$.

**Proposition 5.3 (cf. [1] Proposition 3.3).** Let $G_\infty$ be a graded nilpotent Lie group and let $S \subset G_\infty$ be a finite subset generating $G_\infty$ with respect to $\delta_t$. Then there exists $K = K(G_\infty, S)$ so that for every $g \in G_\infty$ there is an expression

\[ g = \delta_{a_1}s_1 \cdots \delta_{a_k}s_k \]

where $s_i \in S$, $a_i \geq 0$, and $k \leq K$.

**Proof.** The proof is by induction on the nilpotency step of $G_\infty$. If $G_\infty$ is abelian the statement is trivial. So suppose $G_\infty$ has nilpotency step $m \geq 2$. Let $S$ be a finite symmetric generating set for $G_\infty$ and let $\bar{S}$ be the image of $S$ under the map

\[ G_\infty \rightarrow G_\infty/G^m_\infty. \]

Notice that $G_\infty/G^m_\infty$ is a stratified nilpotent Lie group and that the one parameter family of homotheties $\delta_t$ preserve $G^m_\infty$ (in fact they are automorphisms of $G^m_\infty$). Therefore the automorphisms $\delta_t$ descend to automorphisms $\bar{\delta}_t$ of $G_\infty/G^m_\infty$. Denoting the image of an element $g$ under the quotient map above by $\bar{g}$, we have $\bar{\delta}_t \bar{s} = \delta_t \bar{s}$. Therefore $\bar{S}$ generates $G_\infty/G^m_\infty$ with respect to the automorphisms $\bar{\delta}_t$, $t \geq 0$.

Invoking the inductive hypothesis, there exists $K = K(G_\infty/G^m_\infty, \bar{S})$ so that for every $\bar{g} \in G_\infty/G^m_\infty$
\[ \bar{g} = \bar{\delta}_{a_1} \bar{s}_1 \star \cdots \star \bar{\delta}_{a_k} \bar{s}_k \]

where \( k \leq K \). Now let \( g \in G_\infty \) and \( \bar{g} = gG_\infty^m \). Lifting this back to \( G_\infty \) we have

\[ g = \delta_{a_1} s_1 \star \cdots \star \delta_{a_k} s_k \star h \]

for some \( h \in G_\infty^m \), where now \( s_i \in S \). It remains to write \( h \) in a uniform number of straight lines.

\( G_\infty^m \) is a finite dimensional real abelian Lie group. Without loss of generality say \( g_\infty^m \) is spanned by \( t_1, \ldots, t_n \). So

\[ h = \exp_\infty(c_1 t_1) \star \exp_\infty(c_2 t_2) \star \cdots \star \exp_\infty(c_n t_n) \]

for some \( c_i \in \mathbb{R} \).

Since \( S \) generates \( G_\infty > G_\infty^m \), we may write

\[ \exp_\infty t_j = \delta_{a_{i_1}^{l(j)}} s_{i_1} \star \cdots \star \delta_{a_{i_l(j)}^{l(j)}} s_{i_l(j)} \]

for some \( s_{i_l} \in S \), \( a_{i_l} \in \mathbb{R}_+ \), and \( l(j) \in \mathbb{N} \).

Thus if \( c_j > 0 \) we have

\[ \exp_\infty(c_j t_j) = \delta_{c_j^{-1}/m} \exp_\infty(t_j) = \delta_{c_j^{-1}/m a_{i_1}^{l(j)}} s_{i_1} \star \cdots \star \delta_{c_j^{-1}/m a_{i_l(j)}^{l(j)}} s_{i_l(j)} \]

while if \( c_j < 0 \) we have

\[ \exp_\infty(c_j t_j) = \delta_{c_j a_{i_1}^{l(j)}} \exp_\infty(t_j)^{-1} = \delta_{c_j a_{i_1}^{l(j)}} s_{i_1}^{-1} \star \cdots \star \delta_{c_j a_{i_l(j)}^{l(j)}} s_{i_l}^{-1} \]

Thus every element of \( G_\infty^m \) can be written in at most \( \sum_{1 \leq j \leq n} l(j) \) straight lines. This finishes the proof.

\[ \square \]

We can now give a definition of \( \Phi \) that will a priori depend on a choice of representation of each \( g \in G_\infty \) as in Proposition 5.3. Later on we will prove that there was in fact no choice involved.
Let $S \subset G_\infty$ be the set of $2d'$ elements from Example 5.2. By Proposition 5.3 there exists $K < \infty$ so that for every $g \in G_\infty$ there exists an expression

$$g = \delta_{a_1} s_1 \ast \cdots \ast \delta_{a_k} s_k$$

where $k \leq K$, $s_i \in S$ and $0 < a_i < \infty$. We now define $\Phi : G_\infty \to G_\infty$.

**Definition 5.4** (First Definition of $\Phi$).

$$\Phi(g) = \delta_{a_1} \alpha_{ab}(s_1) \ast \delta_{a_2} \alpha_{ab}(s_2) \ast \cdots \ast \delta_{a_k} \alpha_{ab}(s_k)$$

where

$$g = \delta_{a_1} s_1 \ast \cdots \ast \delta_{a_k} s_k$$

is a fixed choice of representation as guaranteed by Proposition 5.3.

**Proposition 5.5.** For each $g \in G_\infty$ there is a sequence $\gamma_n \in \Gamma$ so that

- $\frac{1}{n} \cdot \gamma_n \to g$
- $\frac{1}{n} \cdot \alpha(\gamma_n, x) \to \Phi(g)$ with high probability as $n \to \infty$.

**Proof.** Let $K$ be as guaranteed by Proposition 5.3. Fix $g \in G_\infty$ and the fixed choice of representation of $g$

$$g = \delta_{a_1} s_1 \ast \cdots \ast \delta_{a_k} s_k$$

where $s_i \in S$, $a_i > 0$ and $k \leq K$. For each $n \in \mathbb{N}$ and each $1 \leq i \leq k$ set $m_{n,i} = \lfloor na_i \rfloor$, the greatest integer less than or equal to $na_i$. Then for each $1 \leq i \leq k$ as $n \to \infty$

$$\frac{m_{n,i}}{n} \to a_i.$$ (5.1)

Now define for $n \in \mathbb{N}$
\[ \gamma_n = s_1^{m_{n,1}} s_2^{m_{n,2}} \cdots s_k^{m_{n,k}}. \]

First notice that for each \(1 \leq i \leq k\) we have

\[ \frac{1}{n} \cdot s_i^{m_{n,i}} \rightarrow \delta_{u_i} s_i. \]

Therefore by Chapter 2 Lemma 1.2

\[ \frac{1}{n} \cdot \gamma_n \rightarrow g, \]

giving the first item. For the second item we invoke Theorem 4.1 which says that whp

\[ d_\infty(\alpha(\gamma_n, x), \delta_{m_{n,1}/n} a_{ab}(s_1) \ast \cdots \ast \delta_{m_{n,k}/n} a_{ab}(s_k)) = o(\sum m_{n,i}). \]

By (5.1) and the uniform bound \(k \leq K\) the right hand side is \(o(n)\). Thus whp as \(n \rightarrow \infty\)

\[ d_\infty(\delta_{1/n} \alpha(\gamma_n, x), \delta_{m_{n,1}/n} a_{ab}(s_1) \ast \cdots \ast \delta_{m_{n,k}/n} a_{ab}(s_k)) \rightarrow 0 \]

But as \(n \rightarrow \infty\)

\[ \delta_{m_{n,1}/n} a_{ab}(s_1) \ast \cdots \ast \delta_{m_{n,k}/n} a_{ab}(s_k) \rightarrow \Phi(g) \]

which finishes the proof. \(\square\)

The next Proposition says that \(\frac{1}{n} \cdot \alpha(\sigma_n, x) \rightarrow \Phi(g)\) uniformly as \(\frac{1}{n} \cdot \sigma_n \rightarrow g\).

**Proposition 5.6.** Fix \(g \in G_{\infty}\). For all \(\epsilon_1, \epsilon_2 > 0\) there exist \(\delta > 0\) and \(N \in \mathbb{N}\) so that whenever \(\sigma \in \Gamma\) and \(n \geq N\) are such that \(d_{G_{\infty}}(\frac{1}{n} \cdot \sigma, g) < \delta\), then with probability at least \(1 - \epsilon_1\) we have

\[ d_{H_{\infty}}(\frac{1}{n} \cdot \alpha(\sigma, x), \Phi(g)) < \epsilon_2. \]

In particular, for any sequence \(\frac{1}{n} \cdot \sigma_n \rightarrow g\) we have \(\frac{1}{n} \cdot \alpha(\sigma_n, x) \rightarrow \Phi(g)\) in probability.
Proof. Fix \( g \in G_\infty \) and \( \epsilon_1, \epsilon_2 > 0 \). Choose \( \delta > 0 \) small so that \( \kappa(1 + \epsilon_2)2\delta/\epsilon_1 < \epsilon_2 \) where \( \kappa = \max_{s \in S} \|\alpha(s, \cdot)\|_1 \). Let \( \gamma_n \) be the sequence from Proposition 5.5. Choose \( N \) large so that for all \( n \geq N \), \( d_{G_\infty}(\frac{1}{n} \bullet \gamma_n, g) < \delta \) and so that \( d_{H_\infty}(\frac{1}{n} \bullet \alpha(\gamma_n, x), \Phi(g)) < \epsilon_2 \) with probability at least \( 1 - \epsilon_1 \). Choose \( N \) larger if necessary so that the maps \( \text{scl}_n^\Gamma \) and \( \text{scl}_n^\Lambda \) are \( (1 + \epsilon_2) \)-bi-Lipschitz for all \( n \geq N \).

Now suppose \( d_{G_\infty}(\frac{1}{n} \bullet \sigma, g) < \delta \) where \( n \geq N \). Then \( d_{G_\infty}(\frac{1}{n} \bullet \sigma, \frac{1}{n} \bullet \gamma_n) < 2\delta \), which implies \( d_{\Gamma}(\sigma, \gamma_n) < (1 + \epsilon_2)n2\delta \). Set \( \tau = \sigma^{-1}\gamma_n \), so \( |\tau| < (1 + \epsilon_2)n2\delta \). By Markov’s inequality

\[
m(|\alpha(\tau, x)|_\Lambda \geq \kappa |\tau|/\epsilon_1) \leq \epsilon_1
\]

Thus by our choice of \( \delta \), with probability at least \( 1 - \epsilon_1 \), we have

\[
|\alpha(\tau, x)| \leq n\epsilon_2.
\]

Using the cocycle equation \( \alpha(\gamma_n, x) = \alpha(\sigma, \tau x)\alpha(\tau, x) \) and that \( \text{scl}_n^\Lambda \) is \( (1 + \epsilon_2) \)-bi-Lipschitz we have

\[
d_{H_\infty}(\frac{1}{n} \bullet \alpha(\gamma_n, x), \frac{1}{n} \bullet \alpha(\sigma, \tau x)) < (1 + \epsilon_2)\epsilon_2
\]

with probability at least \( 1 - \epsilon_1 \). Since \( d_{H_\infty}(\frac{1}{n} \bullet \alpha(\gamma_n, x), \Phi(g)) < \epsilon_2 \) with probability at least \( 1 - \epsilon_1 \), we are done. \( \square \)

The next Corollary says that the definition of \( \Phi \) is independent of the choice of representation of \( g \) as a product of straight lines.

Corollary 5.7. Fix \( K' \in \mathbb{N} \) and for each \( g \in G_\infty \) fix \( k'(g) \leq K' \), \( s_1', \ldots, s_{k'}' \in S \) and \( a_1', \ldots, a_{k'}' \in \mathbb{R}_+ \) so that

\[
g = \delta_{a_1'} s_1' \ast \cdots \ast \delta_{a_{k'}'} s_{k'}'.
\]

Now define

\[
\Phi'(g) = \delta_{a_1} \overline{\nu}_{ab}(s_1') \ast \cdots \ast \delta_{a_{k'}} \overline{\nu}_{ab}(s_{k'}').
\]

Then \( \Phi(g) = \Phi'(g) \).
Proof. Repeat the proof of Proposition 5.5 with $\Phi'$ in place of $\Phi$ noting that Theorem 4.1 applies for $K'$ as well as $K$. Doing so we obtain $\gamma'_n \in \Gamma$ so that $\frac{1}{n} \alpha(\gamma'_n, x) \to \Phi'(g)$ in probability. By Proposition 5.6 $\delta_{1/n} \alpha(\gamma'_n, x) \to \Phi(g)$ in probability. Therefore $\Phi'(g) = \Phi(g)$. □

5.1. $\Phi$ is a bi-Lipschitz group automorphism. We can now show that $\Phi$ is a group isomorphism. Since any two Carnot-Carathéodory metrics on the same Carnot group are bi-Lipschitz to one another, we deduce that $\Phi$ is bi-Lipschitz. Let $\Psi$ denote the result of the above construction applied to the cocycle $\beta$ instead of $\alpha$. By symmetry, all of the results above apply equally to $\Psi$. We will see that $\Psi$ and $\Phi$ are inverses.

Proposition 5.8. $\Phi$ is a homomorphism.

Proof. Fix $g, h \in G_\infty$ and $\frac{1}{n} \gamma_n \to g$ and $\frac{1}{n} \sigma_n \to h$. Then by Chapter 2 Lemma 1.2 and Proposition 5.6

- $\frac{1}{n} \alpha(\gamma_n, \sigma_n x) \to \Phi(g)$ in probability
- $\frac{1}{n} \alpha(\sigma_n, x) \to \Phi(h)$ in probability
- $\frac{1}{n} \gamma_n \sigma_n \to gh$
- $\frac{1}{n} \alpha(\gamma_n \sigma_n, x) \to \Phi(gh)$ in probability.

Invoking Lemma 1.2 in $\Lambda$, with high probability

$$
\frac{1}{n} \alpha(\gamma_n, \sigma_n x) \alpha(\sigma_n, x) \to \Phi(g) \ast \Phi(h).
$$

Combining this with the fourth item, the proof is complete. □

To show that $\Phi$ and $\Psi$ are inverse maps, we need the following nilpotent group variant of Poincaré recurrence.

Lemma 5.9 (Poincaré recurrence for nilpotent groups). Fix $g \in G_\infty$ and let $A \subset X$ with $m(A) > 0$. Then
5. CONSTRUCTION OF $\Phi$ AND PROOF OF MAIN THEOREM

\[ m\{x \in A : \exists n_k \in \Gamma \exists n_k \in \mathbb{N} \text{ such that } \frac{1}{n_k} \cdot \gamma_{n_k} \to g \text{ and } \gamma_{n_k} \cdot x \in A \} = m(A). \]

**Lemma 5.10.** Fix \( g \in G_\infty \), let \( A \subset X \) with \( m(A) > 0 \) and let \( \delta > 0 \). Then

\[ m\{x \in A : \exists \gamma \in \Gamma \exists n \in \mathbb{N} \text{ such that } d_\infty(\frac{1}{n} \cdot \gamma, g) < \delta \text{ and } \gamma \cdot x \in A \} = m(A). \]

**Proof that Lemma 5.10 implies Lemma 5.9.** Set

\[ A_\delta = \{ x \in A : \exists n_k \in \Gamma \exists n_k \in \mathbb{N} \text{ such that } d_\infty(\frac{1}{n_k} \cdot \gamma, g) < \delta \text{ and } \gamma \cdot x \in A \} \]

which has measure \( m(A) \) by 5.10. Then

\[ A' = \bigcap_{l=1}^\infty A_{1/l} \]

again has measure \( m(A) \), and has the desired property.

**Proof of Lemma 5.10.** Suppose for contradiction that there is \( g \in G_\infty \), \( A \subset X \) with \( m(A) > 0 \), \( \delta > 0 \) and \( E \subset A \) with \( m(E) > 0 \) such that for a.e. \( x \in E \), \( d_\infty(\frac{1}{n} \cdot \gamma, g) < \delta \) implies \( \gamma \cdot x \notin A \).

We claim that there exist infinitely many \((n_k, \gamma_{n_k}) \in \mathbb{N} \times \Gamma\) such that

\[ d_\infty(\frac{1}{n_k} \cdot \gamma_{n_k}, g) < \delta \text{ and such that if } k_i < k_j \text{ then} \]

\[ d_\infty(\frac{1}{n_{k_j}} \cdot \gamma_{n_k}^{-1} \gamma_{n_{k_j}}, g) < \delta. \]

Indeed, pick any \((n_1, \gamma_{n_1}) \) so that \( d_\infty(\frac{1}{m} \cdot \gamma_{n_1}, g) < \delta \). Now consider any sequence \( \frac{1}{m} \cdot \gamma_m \to g \).

Since \( \frac{1}{m} \cdot \gamma_m^{-1} \to id \) as \( m \to \infty \), Chapter 2 Lemma 1.2 implies that

\[ \frac{1}{m} \cdot \gamma_{n_1}^{-1} \gamma_m \to g. \]

Thus we may pick \( n_2 := m \) large to satisfy the claim. Continuing in this way, the claim is proved.

Now we see that the sets \( \gamma_{n_k}E \) are pairwise disjoint: indeed, if not, then

\[ m(\gamma_{n_k}^{-1} \gamma_{n_{k_j}} E \cap E) > 0. \]
which implies that there is a positive measure set of \( x \in E \) so that \( \gamma_{n_k}^{-1} \gamma_{n_k} \cdot x \in E \subset A \) while \( d_{\infty}(\frac{1}{m_k} \cdot \gamma_{n_k}^{-1} \gamma_{n_k}, g) < \delta \), contradicting the definition of \( E \).

Thus the sets \( \gamma_{n_k}E \) are pairwise disjoint. But as \( m(E) > 0 \), this is also impossible. \( \square \)

Notice that, while one can formulate the Lemmas 5.9 and 5.10 for any group together with one of its asymptotic cones, the key ingredient that fails for groups that are not nilpotent is Chapter 2 Lemma 1.2. This is easily seen in the free group.

**Proposition 5.11.** \( \Phi \) and \( \Psi \) are inverse maps. Consequently, they are group isomorphisms.

Recall (\( \S 2.1 \)) that the fundamental domains \( X \) and \( Y \) satisfy \( m(X \cap Y) > 0 \) and that \( x \in X \cap Y \cap \gamma^{-1}(X \cap Y) \) implies that \( \beta(\alpha(\gamma, x), x) = \gamma \).

**Proof.** Fix \( g \in G_\infty \) and \( \epsilon > 0 \). We will show that \( d_{\infty}(\Psi(\Phi(g)), g) < 2\epsilon \). Using the symmetry of \( \alpha \) and \( \beta \), we apply Proposition 5.6 to the cocycle \( \beta \), the map \( \Psi \) and the element \( \Phi(g) \) to obtain \( N \in \mathbb{N} \) and \( \delta > 0 \) so that for any \( \gamma_n \in \Gamma \) with \( n \geq N \) and any \( x \in X \), for a positive measure set of \( y \in X \cap Y \)

\[
(5.2) \quad d_{H_\infty}(\frac{1}{n} \cdot \alpha(\gamma_n, x), \Phi(g)) < \delta \implies d_{G_\infty}(\frac{1}{n} \cdot \beta(\alpha(\gamma_n, x), y), \Psi(\Phi(g))) < \epsilon.
\]

Now applying Proposition 5.6 to \( \alpha, \Phi \) and \( g \) we obtain \( \delta' > 0 \) and \( N' \in \mathbb{N} \) so that whenever \( n \geq N \)

\[
d_{G_\infty}(\frac{1}{n} \cdot \gamma_n, g) < \delta'
\]

implies that for a positive measure subset of \( X \cap Y \) both (5.2) occurs and

\[
d_{H_\infty}(\frac{1}{n} \cdot \alpha(\gamma_n, x), \Phi(g)) < \delta.
\]

Choose \( \delta' < \epsilon \) if necessary, and set \( N = \max(N, N') \). Then with positive probability in \( X \cap Y \), for \( n \geq N \)

\[
(5.3) \quad d_{G_\infty}(\frac{1}{n} \cdot \gamma_n, g) < \delta' \implies d_{G_\infty}(\frac{1}{n} \cdot \beta(\alpha(\gamma_n, x), x), \Psi(\Phi(g))) < \epsilon.
\]
Now we invoke Lemma 5.9 (Poincaré Recurrence) applied to \( X \cap Y, g \) and \( \delta' \) to assert that with positive probability in \( X \cap Y \) there exists \( n \geq N \) and \( \gamma_n \in \Gamma \) with \( d_{G,\infty}(\frac{1}{n} \cdot \gamma_n, g) < \delta' \), such that \( \gamma_n x \in X \cap Y \) and such that (5.3) occurs. Therefore with positive probability

\[
d_{G,\infty}(\frac{1}{n} \cdot \gamma_n, \Psi(\Phi(g))) < \epsilon \quad \text{and} \quad d_{G,\infty}(\frac{1}{n} \cdot \gamma_n, g) < \epsilon.
\]

\( \square \)

5.2. Theorem E implies Theorem D

**Proof.** We recall the definition of the maps \( \kappa_{x,n} \). For each \( n \in \mathbb{N} \) the maps \( \text{sc}l_{n}^{G,\infty}(\cdot) : \Gamma \to G_{\infty} \) map \( \Gamma \) more and more densely into \( G_{\infty} \) and similarly for \( \text{sc}l_{n}^{H,\infty}(\cdot) : \Lambda \to H_{\infty} \) (see Chapter 2 §1.1). For every \( g \in G_{\infty} \) and every \( n \in \mathbb{N} \) let \( j_{n}(g) \in \Gamma \) be an element of \( \Gamma \) minimizing the distance between \( \text{sc}l_{n}^{G,\infty}(\gamma_{n}) \) and \( g \). Then for \( g \in G_{\infty} \) we define

\[
\kappa_{x,n}(g) = \text{sc}l_{n}^{H,\infty}(\alpha(j_{n}(g), x)).
\]

Now fix \( R > 0, \delta > 0 \) and \( \epsilon > 0 \). Let \( B_{R,\infty}^{G}(e) \) denote the ball of radius \( R > 0 \) in \( (G_{\infty}, d_{\infty}) \) about the identity. By Theorem E for every \( g \in G_{\infty} \) there is \( \tau = \tau(g) > 0 \) so that whenever \( \text{sc}l_{n}^{G,\infty}(\gamma_{n}) \in B_{R,\infty}^{G}(e) \), with probability at least \( 1 - \delta \) we have

\[
d_{H,\infty}(\Phi(g), \text{sc}l_{n}^{H,\infty}(\alpha(\gamma_{n}, x))) < \epsilon.
\]

By the compactness of \( B_{R,\infty}^{G}(e) \) we obtain a finite set \( F \subset B_{R,\infty}^{G}(e) \) with the property that for every \( g \in B_{R,\infty}^{G}(e) \) there is \( g_{0} \in F \) so that \( d_{G,\infty}(g, g_{0}) < \epsilon \) and so that

\[
g \in B_{\tau(g_{0})/2}(g_{0}).
\]

Now set \( \tau = \min_{F} \tau(g) \) and choose \( N \) large so that for all \( n \geq N \), for all \( g \in B_{R,\infty}^{G}(e) \) we have

\[
d_{G,\infty}(\text{sc}l_{n}^{G,\infty}(j_{n}(g)), g) < \tau/2.
\]

Then for all \( n \geq N \) and every \( g \in B_{R,\infty}^{G}(e) \) there is \( g_{0} \in F \) so that with probability at least \( 1 - \delta \)
\[ d_{H,\infty}(\Phi(g_0), \text{sc} \overline{H}_n(\alpha(j_n(g), x))) < \epsilon \]

and

\[ d_{G,\infty}(\Phi(g), \Phi(g_0)) < L\epsilon \]

where \( L \) is the Lipschitz constant for \( \Phi \). This finishes the proof. \( \square \)
CHAPTER 4

Conclusion

In this thesis we have proven several theorems about the ergodic theory and geometry of nilpotent groups. These results are satisfactory in many ways; nevertheless, there are several generalizations and related questions that are left unanswered.

In Chapter 2, we gave an answer to Question 1.4 for the class of countable nilpotent groups. This answer was satisfactory in that it generalized the very natural framework of First Passage Percolation, considered also by Benjamini and Tessera \[3\]. Lending further credit to our answer is the fact that our sub-additive cocycle ergodic theorem was directly interpretable as a generalization to random metrics of Pansu’s theorem on the convergence of the asymptotic cone of a nilpotent group, which interesting in its own right.

It remains, however, to prove the same theorems under weaker integrability assumptions on the cocycles (see Chapter 1 §1 Remark (4)). Indeed, it would be nice to know the precise integrability assumption required.

The inner assumption on the metrics may not be necessary, either. We know that in order for the limit object to be geodesic that innerness is required. However, there may be convergence to a non-geodesic limit in greater generality. One example of a non-inner metric on a nilpotent group is a convex combination of an inner metric on the group with the pullback of the projection to the abelianization of an inner metric. Nevertheless, in this case one can readily describe the asymptotics of such a metric: it is a convex sum of a Carnot-Carathéodory metric on the associated Carnot group and of a Carnot-Carathéodory metric on the abelianization. This leads one to ask if all left-invariant, not necessarily inner, metrics on nilpotent groups can be asymptotically described in terms of a sum of Carnot-Carathéodory metrics on the sum of the terms $\theta^i/\theta^{i+1}$.

There are many natural questions regarding the asymptotic shapes arising in the setting of Chapter 2. For example, what is the asymptotic shape associated to choosing lengths 1 and 2 with equal probability on the standard generators on the integer Heisenberg group? The same questions appear to be very difficult even in the setting of $\mathbb{Z}^2$. While it is possible that somehow some of the questions become easier in the more general setting of nilpotent groups, it seems unlikely.
One blatant shortcoming of our answer to Question 1.4 is the wide array of amenable groups that are not nilpotent. One hopes for an answer to the question that applies to all amenable groups. There is little hope of generalizing the theorems in this thesis directly, by simply replacing ‘nilpotent’ with ‘amenable.’ Indeed, a finitely generated group is virtually nilpotent if and only if all of its asymptotic cones are locally compact. The asymptotic cone not being locally compact is a major obstruction to studying the asymptotics of a sub-additive cocycle in terms of the asymptotic cone. A different notion of asymptotics is probably needed, and one expects to average over a Følner sequence.

In Chapter 3, we gave an alternative, ergodic theoretic proof of Pansu’s quasi-isometric rigidity theorem. The author hopes to use a similar line of reasoning to address the more general open Question 1.9. On the one hand, Chapter 3 is proof positive of a general mathematical phenomenon: given a problem, it is sometimes easier to randomize it, answer the question for ‘typical’ behavior, and then deduce the answer to the original deterministic question. On the other hand, ergodic theory studies asymptotic behavior, and it appears challenging to deduce information about the original nilpotent Lie group from asymptotic information.

Nevertheless, let us remark that we can equivalently define the limiting map $\Phi$ without reference to the homotheties $\delta_t$ (see Definition 5.4). Indeed, starting from an IME between nilpotent groups one can equivalently define $\Phi$ by

$$
\Phi(g) = \alpha_{ab}(s_1^{a_1}) \cdots \alpha_{ab}(s_k^{a_k})
$$

where $g = s_1^{a_1} \cdots s_k^{a_k}$. In this way, we extend the map to nilpotent Lie groups that are not necessarily Carnot, and have a candidate for an isomorphism. At the writing of this thesis, the author does not know whether or not this map can be an isomorphism for a non-Carnot nilpotent Lie group.

Finally, let us mention that the nascent notions of $L^p$ measure equivalence, and its relationship to quasi-isometry is largely unexplored. For example, we do not know if IME and quasi-isometry are the same equivalence class for amenable groups, or specifically for nilpotent groups. Furthermore, it would be interesting to have examples of groups that are $L^p$ measure equivalent but not $L^q$ measure equivalent for some $1 \leq p < q < \infty$, and to understand how the different values of $p$ relate to analytic and geometric properties of the groups in question.

We conclude by remarking that the main statements and arguments of this entire thesis seem likely to go through with mostly obvious changes if one considers nilpotent Lie groups instead of countable nilpotent groups.
Cited Literature


Vita

NAME: Michael Aaron Cantrell

EDUCATION: B.S., Mathematics, Yale University, 2009
Ph.D., Mathematics, University of Illinois, Chicago, 2016

PUBLICATIONS: M. Cantrell,
Differentiability of Integrable Measurable Cocycles Between Nilpotent Groups.

M. Cantrell and A. Furman,
Asymptotic Shapes For Ergodic Families Of Metrics On Nilpotent Groups.

M. Cantrell and J. Palagallo,
Self-Intersection Points of Generalized Koch Curves.

HONORS: Teaching Assistant of the Year Award, University of Illinois, Chicago, 2012
Research Assistant with Alex Furman, University of Illinois, Chicago 2012-2015
Deforest Prize for proficiency in mathematics, Yale College 2009
Stanley Prize for excellence in pure and applied mathematics, Yale College 2008
Yale College Dean’s Research Scholarship, Yale College 2008
Barge Prize for solution of original problems in mathematics, Yale College 2007

SERVICE: Chicago Youth Math Symposium Judge 2012-present
Graduate Student Mentor Volunteer 2012-2014
Research Talk for Prospective UIC PhD Students Invited 2014
Graduate Student Seminar on Expander Graphs Co-organizer 2014
Graduate Student Colloquium Co-organizer 2013-2014
Graduate Geometry, Topology, Dynamics Seminar Organizer 2013-2014

TEACHING: Teaching Assistant, University of Illinois, Chicago
Math 330 - Abstract Algebra Fall 2016
Math 181 - Calculus II Fall 2016
Math 181 - Calculus II Fall 2015
Math 090 - Intermediate Algebra Spring 2015
Math 121 - Precalculus Fall 2014
Math 180 - Calculus I Spring 2012
Math 121 - Precalculus Fall 2011

INVITED TALKS: Random Spatial Processes and Dynamics
Texas A&M, April 2016
Geometry and Topology Seminar
University of Wisconsin, Madison, March 2016
Geometry Seminar
Purdue University, February 2016
Topology Seminar
University of Michigan, December 2015
Geometric and Probabilistic Methods in Group Theory and Dynamical Systems,
Texas A&M, November 2015
Geometry Seminar
Indiana University, November 2015
Dynamics Seminar
Northwestern University, November 2015
Dynamics Seminar
University of Chicago, November 2015
Geometry, Groups and Dynamics/GEAR Seminar
University of Illinois Urbana Champaign, November 2015
Ergodic and Symbolic Actions of Amenable Groups
AMS Sectional Meeting, Loyola University, October 2015
Probability on Graphs and Groups
École Polytechnique Fédérale de Lausanne, August 2015
Ergodic and Geometry Group Theory Seminar
École Polytechnique Fédérale de Lausanne, August 2015

Group Theory Seminar
Université catholique de Louvain, June 2015

Geometric Group Theory Working Seminar
Université Paris-Sud, May 2015

Operator Algebras and Approximation Properties of Groups
NCGOA Vanderbilt University, April 2015

Oberwolfach Arbeitsgemeinschaft on Superrigidity
Oberwolfach, April 2014

Joint Math Meetings, 2008

PROFESSIONAL MEMBERSHIP: American Mathematical Society