NORTHWESTERN UNIVERSITY

Large-scale Geometry of First Passage Percolation on Graphs of Polynomial Growth

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Mathematics

By

Christian Gorski

EVANSTON, ILLINOIS

September 2023

ABSTRACT

Large-scale Geometry of First Passage Percolation on Graphs of Polynomial Growth

Christian Gorski

This dissertation proves several results for first passage percolation on graphs of polynomial growth. The class of limit shapes for first passage percolation with stationary weights on Cayley graphs of virtually nilpotent groups is characterized. Then strict monotonicity theorems for independent first passage percolation on graphs of polynomial growth and quasi-trees are given. Specifically, for such graphs, when we compare the expected passage time metrics with respect to two different weight measures, strict stochastic domination of weight measures implies (an analogue of) strict inequality of the associated "time constants" as long as the dominating measure satisfies an appropriate subcriticality condition. This is proven by showing that in the subcritical regime, long geodesics "use all possible weights linearly often in expectation," which is a result of independent interest. Moreover, a similar strict monotonicity theorem with respect to *variability* of measures holds for such graphs if and only if the graphs satisfy a geometric condition we call *admitting detours*. Lastly, we show that for Cayley graphs of virtually nilpotent groups, in the supercritical regime, there is a nontrivial "percolation cone" where strict

monotonicity with respect to stochastic domination fails; that is, the subcriticality assumption in our strict monotonicity theorems is necessary.

Acknowledgements

I would first like to deeply thank my advisor Antonio Auffinger for his support and guidance throughout graduate school. His kindness and confidence in me have been invaluable to me throughout my pursuit of a PhD.

I would also like to thank my parents Jean and John Gorski for raising me, always supporting me, and believing, in earnest, that I can literally do anything.

Next I would like to thank all the other graduate students in the mathematics department at Northwestern. Their kindness and companionship created an environment where I could thrive inside and outside of mathematics. I especially thank Rin Ray for always being there, and for filling my life with many more interesting activities than I could ever dream of finding on my own.

The mathematical work in this thesis would not have been possible without the help of several people. I thank Nir Avni for providing resources on finitely generated nilpotent groups, Sami Douba for conversations on topics related to this geometric group theory, and Yves Cornulier for providing helpful context by giving a counterexample showing that on nilpotent groups, being bi-Lipschitz to a word metric does not imply innerness. I also thank Alice Kerr, from whom I learned Manning's bottleneck criterion. I thank Tom Hutchcroft for helpful conversations and references regarding recent work on percolation on transitive graphs of polynomial growth. I again thank Antonio Auffinger, for suggesting most of the problems studied in this dissertation, for many helpful conversations, and for critical assistance navigating the literature in first-passage percolation.

The research in this dissertation was partially supported by NSF Grant DMS-1502632 RTG: Analysis on Manifolds.

Table of Contents

List of Figures

CHAPTER 1

Introduction

This thesis exposits some results on the large scale geometry of first passage percolation (FPP) on graphs more general than the "hypercubic lattice" (i.e. the standard Cayley graph of $\mathbb{Z}^d$, or the "d-dimensional grid"), particularly graphs of polynomial growth. First-passage percolation is a natural random metric on a graph which can be thought of as a perturbation of the usual graph metric. Most of the work in FPP has been done on the standard Cayley graph of $\mathbb{Z}^d$, which has polynomial growth of degree d; thus studying more general graphs of polynomial growth is a natural generalization. If one wishes to study *Cayley graphs* of polynomial growth, then by a famous theorem of Gromov [16] these are precisely the Cayley graphs of virtually nilpotent groups, so these also form a natural class of graphs to study.

One way of understanding the large-scale geometry of this random metric is to consider scaling limits. The famous "shape theorem" of Cox and Durett [10] tells us that the random metric given by FPP on $\mathbb{Z}^d$ has a deterministic scaling limit given by a norm on $\mathbb{R}^d$. Moreover, it was shown in work of Benjamini-Tessera [5] and Cantrell-Furman [8] that many classes of random stationary metrics on Cayley graphs of virtually nilpotent groups have deterministic scaling limits given by Carnot-Carathéodory metrics on a nilpotent Lie group. (More information on Carnot-Caratheéodory metrics is given in Section 3.8).

Existence of such scaling limits is guaranteed in rather general settings, but explicit descriptions of the limit metrics remain elusive in almost all non-trivial cases. Many expected qualitative aspects of the limit metric also remain unproven. For instance, it is conjectured that the unit ball of the limit norm associated to any independent FPP on $\mathbb{Z}^d$ with a continuous weight distribution is strictly convex, but this is not known. On the other hand, the limit metric for *stationary* FPP on $\mathbb{Z}^d$ is essentially unrestrained; by a theorem of Häggström and Meester [20], any norm on $\mathbb{R}^d$ is the scaling limit of some FPP metric with stationary weights. Thus, the set of norms attainable as scaling limits of FPP metrics on $\mathbb{Z}^d$ is precisely the set of norms on $\mathbb{R}^d$.

The first main theorem of this dissertation extends this theorem of Häggström and Meester to the setting of arbitrary Cayley graphs of arbitrary virtually nilpotent groups, that is, the theorem characterizes which Carnot-Carathéodory metrics (CC-metrics) on a nilpotent Lie group L_∞ actually arise as the scaling limit of some stationary FPP metric on a particular Cayley graph of a finitely generated virtually nilpotent group Γ. The answer is *all* CC-metrics if Γ is itself nilpotent. If Γ is only *virtually* nilpotent, then the set consists of precisely those CC-metrics which are *conjugation-invariant*. This is proven in Chapter 3. The work in this chapter is adapted from [3] and is joint work with Antonio Auffinger.

Since limit shapes are hard to describe explicitly for a fixed weight measure, another way one might try to understand them is by trying to understanding the *relationship* between different limit shapes as we vary the parameters of our FPP model. The theorems of Chapter 4 concern comparison of limit shapes in the setting of independent FPP. For two weight measures $\tilde{\nu}, \nu$, if we have $\tilde{\nu} \prec \nu$ for some partial order on measures, can we conclude anything about the corresponding limit metrics, e.g. some sort of strict inequality? Theorems to this effect were proven by van den Berg and Kesten [36] in the classical case of the hypercubic lattice. The theorems in Chapter 4 of this dissertation are of this type, but apply to general bounded degree graphs which either have polynomial upper and lower growth bounds of the same degree, or which are quasi-isometric to trees. In this coarse-geometric setting, in a subcritical regime we have "strict monotonicity" with respect to stochastic domination. "Strict monotonicity" with respect to variability in the subcritical regime is shown to be equivalent to a fine-geometric condition which we call "admitting detours." We give sufficient conditions for a Cayley graph of a group to admit detours. In particular, we find that for any Cayley graph of any virtually nilpotent group, if the graph is not isomorphic to the standard Cayley graph of $\mathbb{Z}$, then we have strict monotonicity with respect to variability in the subcritical regime. The proof of the strict monotonicity theorem for stochastic domination proves roughly that "the weight measure is absolutely continuous with respect to the expected empirical measure," that is, "all possible weights are used by the geodesic linearly often in expectation." This chapter is adapted from [14].

Lastly, in Chapter 5, we show that the subcriticality assumption in our strict monotonicity theorems is necessary for Cayley graphs of virtually nilpotent groups, at least in the restricted setting that our weight measures have an exponential moment. That is, if the weight measure has an exponential moment but is supercritical, then there is some direction for which strict monotonicity of the time constant with respect to stochastic domination fails. Failure of strict monotonicity is related to the existence of "percolation cones," roughly speaking, directions in which there exist infinite edge-geodesics with all edges having minimal weight.

CHAPTER 2

Basics of first passage percolation

The purpose of this chapter is primarily to fix definitions and review basic facts related to graphs, Cayley graphs, groups, and first passage percolation.

2.1. Graphs, Cayley graphs, virtually nilpotent groups

By a *graph* we mean a pair $G = (V, E)$ of sets and an "endpoint" or "boundary" map from E to the set of subsets of V of size 2. In particular, we allow more than one edge between each pair of vertices but we do not allow self-loops. (Disallowing self-loops is simply a matter of convenience; virtually all questions considered in this dissertation are easily seen to be equivalent for a graph G with self-loops and the graph G' obtained from G by deleting all self-loops). A graph is called *simple* if there are no parallel edges, that is, each pair of vertices has at most one edge between them. Throughout, the "ambient graph" G is tacitly assumed to be connected, locally finite (i.e. each vertex has finite degree) and infinite (that is, V is countably infinite); we will often however consider subgraphs of G which are finite and/or disconnected.

A path π in G is an alternating sequence of vertices and edges (starting and ending with a vertex) such that the vertices immediately preceding and following an edge comprise the edge's boundary. If π starts at $x \in V$ and ends at $y \in V$, we often write $\pi : x \to y$. We will typically abuse notation and use the same symbol π to refer to the set of edges appearing in the path π (so $\pi \subset E$). $|S|$ denotes the cardinality of the set S, so in particular, if π is a path, $|\pi|$ is the number of edges appearing in the path (again abusing notation and considering π as a subset of E). If π

does not contain any repeated edges, then this agrees with the usual notion of length of a path. In fact, we will mostly be concerned with paths which do not have any repeated vertices; we call such paths *self-avoiding* (or *vertex-self-avoiding*).

One of the most important classes of graphs we will consider are *Cayley graphs* of finitely generated groups. Given a finitely generated group Γ and a finite generating set S, the Cayley graph of Γ with respect to S is the graph which has vertex set $V = \Gamma$ and which has an edge connecting $x, y \in \Gamma$ whenever $x = ys$ for some $s \in S$.[1] For instance, the "hypercubic lattice", or standard Cayley graph of $\mathbb{Z}^d$, is the Cayley graph associated to the standard generating set $\{(1, 0, ...0), (0, 1, 0, ...,), ..., (0, ..., 0, 1)\}$. The triangular lattice is isomorphic to the Cayley graph of $\mathbb{Z}^2$ associated to the generating set $\{(1, 0), (1, 1), (0, 1)\}$.

Cayley graphs are examples of *transitive* graphs. A graph is called (vertex)-*transitive* if the group of automorphisms $\text{Aut}(G)$ of G acts transitively on V, i.e. for any $v, w \in V$, there is an automorphism ϕ of G such that $\phi(v) = \phi(w)$. In other words, the action of $\text{Aut}(G)$ on V only has one orbit. In transitive graphs "all points look the same." We call a graph *almost-transitive* if the action of $\text{Aut}(G)$ on V has only finitely many orbits, that is, "there are only finitely many types of points." Some theorems in this dissertation are restricted to Cayley graphs, but others apply to transitive graphs, almost transitive graphs, or even more general graphs of bounded degree. (Recall that a graph has *bounded degree* if there is some $D < \infty$ such that each vertex has degree at most D, i.e. the vertex is an endpoint of at most D edges).

[1] There are actually two slightly different natural definitions of Cayley graph, depending on whether one wishes to restrict to simple graphs or allow parallel edges, see Section 4.5. All the theorems in this dissertation apply to both reduced and unreduced Cayley graphs, although Chapter 3 was written assuming all Cayley graphs are reduced, for simplicity.

A graph G gives a natural metric on V by

$$d(x, y) := \inf\{|\gamma| : \gamma : x \to y\} = \inf\{|\gamma| : \gamma : x \to y \text{ self-avoiding}\}.$$

We write $B(x, R)$ for the ball $\{y \in V : d(x, y) \leq R\}$ in this metric and write $S(x, R)$ for the sphere $\{y \in V : d(x, y) = R\}$. Note that if G has degree bounded by D, $B(x, R)$ contains at most D^R vertices. An almost-transitive graph is said to have *polynomial growth* if the function $R \mapsto |B(x, R)|$ is bounded above by a polynomial in R. For example, the standard Cayley graph of $\mathbb{Z}^d$ has polynomial growth of degree d. In Chapter 4 we will also consider graphs which are not necessarily almost-transitive but have *strict* polynomial growth, that is, uniform upper and lower bounds on the volume growth by polynomials of the same degree.

A famous theorem of Gromov [16] says that a Cayley graph has polynomial growth if and only if the underlying group is *virtually nilpotent*. A group Γ is *nilpotent* if $\Gamma_k = \{1\}$ for some finite k, where $\Gamma_0 = \Gamma$ and $\Gamma_{i+1} = [\Gamma, \Gamma_i]$ is the lower central series for Γ. (Here the notation $[H, K]$ denotes the subgroup generated by the set of all *commutators* $[h, k] := hkh^{-1}k^{-1}$ with $h \in H$ and $k \in K$). Note that abelian groups have $\Gamma_1 = \{1\}$ and are thus nilpotent. The simplest example of a nonabelian nilpotent group is the Heisenberg group, that is, the set of upper triangular 3×3 matrices with integer coefficients and ones along the diagonal. A group Γ is *virtually nilpotent* if it contains a finite index subgroup Γ' which is nilpotent. (Recall that a subgroup $\Gamma' \leq \Gamma$ has *finite index* if Γ contains only finitely many cosets of Γ'; equivalently, the action of Γ' on Γ by right multiplication has only finitely many orbits).

One feature of nilpotent groups that will be useful for many of our results is that they always have nontrivial *center*. Recall that the *center* $Z(\Gamma)$ of a group Γ is the set of elements that commute with every other element (sometimes called *central elements*), i.e.

$$Z(\Gamma) = \{z \in \Gamma : gz = zg \text{ for all } g \in \Gamma\}.$$

A group Γ is abelian if and only if $\Gamma = Z(\Gamma)$. It will turn out that many geometric constructions on the standard Cayley graph of $\mathbb{Z}^d$ can be adapted to constructions on general Cayley graphs of nilpotent group by using *central* elements of Γ.

2.2. First-passage percolation

First passage percolation (FPP) was introduced by Hammersley and Welsh [22] in 1965 as a model for the spread of a fluid through a porous medium. It is a random perturbation of a given graph distance, where random lengths are assigned to edges of a fixed graph. For a survey on this model, the reader is invited to read [2, 26] and the references therein.

Mathematically, FPP is is simply a natural procedure for producing random metrics on a fixed graph. Let G be a graph with vertex set V and edge set E. Given a random function $w : E \to [0, \infty)$, we say that $w(e)$ is the *weight* of the edge e, and for any path of edges π we define its *total weight* or *passage time* to be

$$T(\pi) := \sum_{e \in \pi} w(e).$$

(Note that if $w(e) = 1$ for all e, we have that $T = d$, the usual graph metric).

The idea is that if $w(e)$ is the amount of time it would take a stream of fluid to cross e, then $T(\pi)$ is the amount of time it would take to pass through the whole path π. We further define the *first passage percolation (FPP) (pseudo)metric* T to be the (pseudo)metric on V given by

$$T(x,y) := \inf\{T(\pi) : \pi : x \to y \text{ is an edge path }\}.$$

So if fluid flows from a source at site x, $T(x,y)$ is the first time it reaches site y. One readily checks that T is symmetric and satisfies the triangle inequality. If $\mathbb{P}(w(e) = 0) > 0$ for some edge then T is a pseudometric, i.e. we have with positive probability that $T(x,y) = 0$ for some $x \neq y$; otherwise T is a genuine metric.

The central focus of the study of first-passage percolation is studying T, particularly its large-scale geometry, and the behavior of T*-geodesics*, that is, paths $\pi : x \to y$ such that $T(\pi) = T(x,y)$.

Without restrictions on w, this problem is, of course, very underdetermined. The most common setting is *independent first passage percolation*, that is, the case that $\{w(e)\}_{e \in E}$ is a family of independent random variables sampled from some common probability measure ν supported on $[0,\infty)$. This is the setting of Chapter 4 and Chapter 5. If (as is the case for Chapter 4) we are considering other possible weight measures, e.g. $\tilde{\nu}$, then the associated random weights are denoted by $\tilde{w} : E \to [0,\infty)$, and the induced FPP metric is denoted by $\tilde{T}$; we proceed similarly for other diacritics.

More generally, one could simply require w to be *stationary*; that is, there may be correlations between the weights $w(e)$ of different edges, but the joint distribution of the $w(e)$ must be invariant under the action of some group Γ which acts on G by graph isomorphisms. This is the setting of Chapter 3. In fact, for that chapter, we restrict to the case that Γ is a finitely generated virtually nilpotent group and G is a Cayley graph for Γ.

CHAPTER 3

Asymptotic shapes for stationary first passage percolation on Cayley graphs of virtually nilpotent groups

3.1. Introduction

3.1.1. Main result

Consider first-passage percolation on $\mathbb{Z}^d$ where the edge weights are i.i.d. random variables. Under suitable moment conditions on the weight distribution, one obtains the famous shape theorem of Cox and Durrett ($d = 2$) [10] and Kesten ($d > 2$) [26]: there exists a norm μ on $\mathbb{R}^d$ such that FPP on $\mathbb{Z}^d$ has almost surely a deterministic scaling limit given by the normed vector space $(\mathbb{R}^d, \mu)$. The limiting norm μ depends on the distribution of the edge weights. It is a famous open question to determine which possible metrics arise as FPP limits on $\mathbb{Z}^d$ with i.i.d. edge weights. In particular, it is expected that the limit unit ball should be strictly convex, ruling out trivial metrics such as ℓ_1 or ℓ_∞.

In 1995, Haggstrom and Meester [20] showed that if the assumption of i.i.d. edge weights on $\mathbb{Z}^d$ is relaxed, some of the expected restrictions on the limit norm disappear. Precisely, they showed that for any norm ρ on $\mathbb{R}^d$ there exist *stationary* edge weights on $\mathbb{Z}^d$ which give a FPP model whose scaling limit is $(\mathbb{R}^d, \rho)$. In this chapter, we explore this direction for FPP in different (non-abelian) graphs.

Benjamini and Tessera [5] explored i.i.d. FPP models on Cayley graphs of finitely generated virtually nilpotent groups. This class of groups is precisely the class of groups with polynomial growth, due to a famous theorem of Gromov [16], and includes the classical example of $\mathbb{Z}^d$. The question of scaling limits of such groups was first answered in the deterministic setting by Pansu [30], who proved that, for a large class of invariant metrics on such groups, the scaling limit is given by a Carnot-Carathéodory metric on a certain nilpotent Lie group. (See Sections 3.1.2 and 3.8 for an explanation of Carnot-Carathéodory metrics).

Benjamini and Tessera prove that, under mild conditions, an i.i.d. FPP on a virtually nilpotent Cayley graph also has a deterministic scaling limit given by a Carnot-Carathéodory metric on a nilpotent Lie group. Later Cantrell and Furman [8] proved an analogous theorem for *stationary* edge weights. Again, in all these cases, the limit shape depends on the distribution of the edge weights, and in the i.i.d. case, restrictions on realizable metrics are conjectured but largely unproven.

A natural question then arises, in the spirit of Haggstrom and Meester [20] : for stationary FPP on virtually nilpotent groups, are all possible limit shapes realizable? What are the required symmetries for the limit metric? More explicitly, given a Cayley graph of some finitely generated virtually nilpotent group and a Carnot-Carathéodory metric on the associated nilpotent Lie group, do there exist stationary edge weights which give a FPP with a scaling limit given by that Carnot-Carathéodory metric? The goal of this chapter is to provide an affirmative answer to this last question in the nilpotent case and to obtain a similar characterization of all limit shapes of stationary FPPs in the virtually nilpotent case. Our main theorem is the following.

Theorem 3.1.1. *Let Γ be a finitely generated virtually nilpotent group with generating set S, and let E be the edge set of the corresponding Cayley graph. Let d_Φ be a Carnot-Carathéodory metric on the associated graded Lie group L_∞. If Φ is conjugation invariant, then there exist*

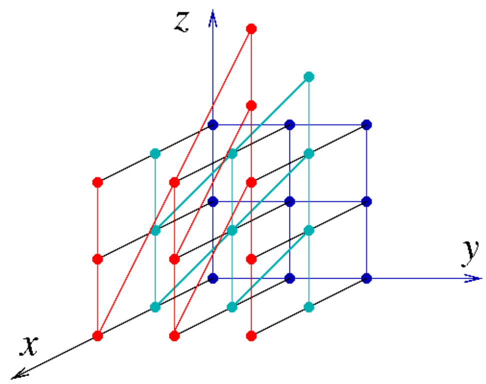

Figure 3.1. A portion of the Cayley graph of $H(\mathbb{Z})$ with respect to the generating set $\{X, Y, Z\}$. Source: Wikipedia; image by Gabor Pete. Colors are for visual contrast only.

stationary weights $w : E \to \mathbb{R}_{\geq 0}$ such that the associated metric space (Γ, T) satisfies

$$\left(\Gamma, \frac{1}{n}T\right) \xrightarrow[n\to\infty]{} (L_\infty, d_\Phi)$$

in the sense of pointed Gromov-Hausdorff convergence.

To make the theorem more concrete, let us consider the example of the Heisenberg group, the simplest nonabelian nilpotent group. The integer Heisenberg group $H(\mathbb{Z})$ has presentation

$$\langle X, Y, Z | [X, Y] = Z, [X, Z] = [Y, Z] = 1 \rangle,$$

and can be realized as the subgroup

$$\left\{ \begin{bmatrix} 1 & a & b \\ & 1 & c \\ & & 1 \end{bmatrix} : a, b, c \in \mathbb{Z} \right\}$$

of $GL_3(\mathbb{R})$. It sits as a cocompact lattice inside the real Heisenberg group $H(\mathbb{R})$, the group of real upper triangular matrices with 1s on the diagonal. Given any norm Φ on the subspace

$$V := \left\{ \begin{bmatrix} & a & \\ & & c \\ & & \end{bmatrix} : a, c \in \mathbb{R} \right\}$$

of the Lie algebra of $H(\mathbb{R})$, there exists a metric called the Carnot-Carathéodory metric d_Φ on $H(\mathbb{R})$ associated to Φ (see Section 3.8). So in the special case of the Heisenberg group, our theorem is as follows:

Theorem 3.1.2. *Let Φ be any norm on V, d_Φ the associated Carnot-Carathéodory metric on $H(\mathbb{R})$. Then, given any Cayley graph of $H(\mathbb{Z})$, there exist stationary edge weights $w : E \to \mathbb{R}_{\geq 0}$ (E the edge set of the Cayley graph) such that the resulting FPP metric T is such that*

$$\left(H(\mathbb{Z}), \frac{1}{n}T\right) \xrightarrow[n\to\infty]{} (H(\mathbb{R}), d_\Phi)$$

in the sense of pointed Gromov-Hausdorff convergence.

3.1.2. Definitions, notations, and background

We now provide the definitions and the setup for Theorem 3.1.1. Let Γ be a finitely generated virtually nilpotent group, and let S be a finite generating set. The Cayley graph associated to (Γ, S) is the graph with vertex set Γ and edge set $E := \{\{g, gs\} : g \in \Gamma, s \in S\}$. For an element $g \in \Gamma$, set

$$|g| := \inf\{n \geq 0 : \exists s_1, ..., s_n \in S \cup S^{-1} \text{ such that } s_1 \cdots s_n = g\},$$

and denote by d the word metric

$$d(x, y) := |x^{-1}y|$$

on Γ. Note that d is a left-invariant metric on Γ. If γ is an edge path in E, we will denote by $|\gamma|$ the number of edges in γ. Thus we have

$$d(x, y) = \inf\{|\gamma| : \gamma \text{ is a path from } x \text{ to } y\}.$$

Let w be a random function $w : E \to [0, \infty)$. We call $w(e)$ the *weight* of the edge e. The collection of weights w is called *stationary* if the distribution is invariant under the left action of Γ, that is, for every finite collection of edges $f_1, ..., f_k \in E$ and every $g \in \Gamma$, the joint distributions of $(w(f_1), ..., w(f_k))$ and $(w(g^{-1}f_1), ..., w(g^{-1}f_k))$ are equal. The weights are called *ergodic* if the underlying probability space is ergodic, that is, if all Γ-invariant events have probability 0 or 1.

For an edge path $\gamma = (f_1, ..., f_k)$, we define

$$T(\gamma) := \sum_{i=1}^{k} w(f_i)$$

and for two $x, y \in \Gamma$ we define the *passage time* from x to y to be

$$T(x, y) := \inf\{T(\gamma) : \gamma \text{ is a path from } x \text{ to } y\}.$$

T is a random pseudo-metric on Γ and the pseudo-metric space (Γ, T) is called *first passage percolation* or *FPP* on Γ. Taking expectations we see that $\mathbb{E}T$ also gives a metric on Γ; if w is stationary, then this metric is left-invariant.

Let N be a finite index normal torsion-free nilpotent subgroup of Γ. Such a subgroup is constructed in Proposition 3.9.1 in Section 3.9. We denote the abelianization $N/[N, N]$ of N by

N^{ab}. This is a finitely generated abelian group, and so its torsion elements form a finite subgroup N^{ab}_{tor}. We define $N^{ab}_{free} := N^{ab}/N^{ab}_{tor}$.

There is a graded nilpotent Lie group L_∞ associated to Γ (via N), and a certain subalgebra of its Lie algebra, which we denote by $\mathfrak{g}^{ab}$, is equipped with a natural isomorphism $N^{ab} \otimes \mathbb{R} \cong \mathfrak{g}^{ab}$. Each norm Ψ on $\mathfrak{g}^{ab}$ determines a metric d_Ψ on L_∞ which is called the Carnot-Carathéodory metric associated to Ψ; conversely, every Carnot-Carathéodory metric on L_∞ comes from a unique norm on $\mathfrak{g}^{ab}$. More explicit descriptions and constructions of these objects can be found in Section 3.8, as well as [8].

Lastly, there is a construction which plays a central role in our proof, which associates a norm on $\mathfrak{g}^{ab}$ to a metric on Γ. Since $|\cdot|$ is a symmetric subadditive function on Γ (i.e. $|ab| \leq |a|+|b|$ for all $a, b \in \Gamma$), and hence a symmetric subadditive function on N, it induces a symmetric subadditive function on $N^{ab}_{free} \cong \mathbb{Z}^d$ via the quotient map $N \to N^{ab}_{free}$, $x \mapsto x^{ab}_{free}$:

$$|y|_{ab} := \inf_{x \in N, x^{ab}_{free} = y} |x|.$$

As a symmetric subadditive function on $N^{ab}_{free} \cong \mathbb{Z}^d$, $|\cdot|_{ab}$ is asymptotically equivalent to a unique seminorm on $\mathbb{R}^d \cong N^{ab}_{free} \otimes \mathbb{R} \cong N^{ab} \otimes \mathbb{R}$. That is, there is a unique seminorm $\|\cdot\|$ on $N^{ab} \otimes \mathbb{R}$ such that

$$\|y\| - |y|_{ab} = o(y)$$

where the in the little-o notation we may take any norm on $N^{ab} \otimes \mathbb{R}$ to measure y. Similarly, assuming our weights are integrable, $\mathbb{E}T(1, \cdot)$ is also subadditive, and hence it induces a subadditive fuction $\tilde{T}$ on N^{ab}_{free} which is asymptotically equivalent to a unique seminorm Φ on $N^{ab} \otimes \mathbb{R}$.

The conjugation action of Γ on N induces an action of Γ on $N^{ab} \otimes \mathbb{R}$, hence induces an action on the set of norms on $N^{ab} \otimes \mathbb{R}$. We call a norm on $N^{ab} \otimes \mathbb{R}$ *conjugation-invariant* if it is invariant under this action. The conjugation action is discussed further in Section 3.4, but in the case that Γ itself is already nilpotent, the action is trivial, and hence in this case all norms on $N^{ab} \otimes \mathbb{R}$ are conjugation invariant. In Section 3.4 (see Proposition 3.4.2), we also show that conjugation-invariance is a necessary restriction, that is, if Φ is a norm associated to an invariant metric (such as $\mathbb{E}T$ when each $T(x, y)$ is integrable), then Φ is necessarily conjugation-invariant.

In the notations above, it is known that $(L_\infty, d_{\|\cdot\|})$ is the scaling limit of (Γ, d) [30] and that (Γ, T) almost surely has scaling limit (L_∞, d_Φ) for many choices of edge weights [5,8]. Theorem 3.1.1 above shows that any Carnot-Carathéodory d_Ψ as in (3.8.1) is the scaling limit of some stationary FPP model on any Cayley graph of Γ, so long as Ψ is conjugation-invariant.

3.1.3. Proof strategy and organization of the chapter

The following theorem of Cantrell and Furman [8] provides a starting point for us:

Theorem 3.1.3. *([8]) Let w be ergodic stationary weights such that T is bi-Lipschitz to d, that is, there exist $0 < k < K < \infty$ such that*

$$kd(x, y) \leq T(x, y) \leq Kd(x, y)$$

for all $x, y \in \Gamma$ almost surely. Let Φ be the norm on $\mathfrak{g}^{ab}$ associated to the metric $\mathbb{E}T$ on Γ, and let d_Φ be the Carnot-Carathéodory metric on L_∞ associated to Φ, as above. Then almost surely

$$\left(\Gamma, \frac{1}{n}T\right) \xrightarrow[n\to\infty]{} (L_\infty, d_\Phi) \tag{3.1.1}$$

is the sense of pointed Gromov-Hausdorff convergence.

Remark 3.1.1. *The fact that the norm Φ we describe above is the same norm constructed in [8] is perhaps not obvious except in the case that $\Gamma = N$ is torsion-free with torsion-free abelianization. A proof that the two constructions do give the same answer is given in Section 3.9.*

Remark 3.1.2. *We take the identity as the base point in the above pointed Gromov-Hausdorff convergence. We omit the base point in our notation throughout.*

Remark 3.1.3. *Cantrell and Furman don't require the random metric T to come from edge weights but require it to be* inner *(see Section 3.9) in addition to being bi-Lipschitz to d. On the other hand, if T comes from edge weights which are uniformly bounded above (implied by the bi-Lipschitz condition on T), then T is inner, so the above statement is implied by the main theorem of [8]. Thus our theorem shows that the collection of scaling limits of FPPs coming from stationary edge weights on a fixed Cayley graph is no smaller than the collection of scaling limits of stationary inner metrics which are bi-Lipschitz to d.*

Remark 3.1.4. *In Section 3.10 we provide a step that was omitted in the proof of Theorem 3.1.3 in [8]. It guarantees that the convergence in* (3.1.1) *is indeed in Gromov-Hausdorff sense. See Remark 3.10.1 for more details.*

In view of Theorem 3.1.3 and the correspondence between Carnot-Carathéodory metrics and norms on $\mathfrak{g}^{ab}$, in order to prove Theorem 3.1.1, it suffices to prove:

Theorem 3.1.4. *Let Γ be a finitely generated virtually nilpotent group with generating set S, and let E be the edge set of the corresponding Cayley graph. Let Ψ be a norm on $N^{ab} \otimes \mathbb{R}$*

which is conjugation-invariant. Then there exist ergodic stationary weights $w : E \to \mathbb{R}$ *such that* T *is bi-Lipschitz to* d*, and such that the subadditive function on* N^{ab}_{free} *induced by* $\mathbb{E}T(1, \cdot)$ *is asymptotically equivalent to* Ψ*.*

Proof of Theorem 3.1.1 given Theorem 3.1.4. Let d_Φ be a Carnot-Carathéodory metric on L_∞ and suppose that the associated norm Φ on $\mathfrak{g}^{ab}$ is conjugation-invariant. Given any Cayley graph of Γ, use Theorem 3.1.4 to choose ergodic stationary weights w such that the resulting T is bi-Lipschitz to d and such that the norm on $\mathfrak{g}^{ab}$ associated to the metric $\mathbb{E}T$ on Γ is equal to Φ. Applying Theorem 3.1.3 to w then gives

$$\left(\Gamma, \frac{1}{n}T\right) \xrightarrow[n\to\infty]{} (L_\infty, d_\Phi)$$

in the sense of pointed Gromov-Hausdorff convergence, as desired. □

Thus, our main theorem is reduced to the problem of constructing stationary weights which induce a given norm Ψ on $\mathfrak{g}^{ab}$. Haggstrom and Meester [20] give a construction for inducing the correct norms in the $\mathbb{Z}^d$ case, and in the simplest case, the core of our work is "lifting" the Haggstrom-Meester construction from the abelianization of the finitely generated nilpotent group to the group itself, and then checking that everything goes through. Therefore, to give an idea of the construction we start by proving Theorem 3.1.4 in this simplest case—namely, the case that $\Gamma = N$ is a torsion-free nilpotent group with torsion-free abelianization, and the generating set S projects to the standard generating set of $\mathbb{Z}^d \cong N^{ab} = \Gamma^{ab}$. As mentioned above, in this case conjugation-invariance does not play a role, and *any* norm Ψ is attainable. This is done in the next two sections.

In Section 3.4, we discuss the restriction of conjugation-invariance and the nontrivial subtleties that arise when treating the general *virtually* nilpotent case. The rest of the chapter is then dedicated to proving Theorem 3.1.4 in full generality. In particular, this involves understanding a *virtually* abelian "almost-abelianization" of Γ, and then again "lifting" a construction from the "almost-abelianization" to Γ. In order to accommodate all possible Cayley graphs as well as the slightly non-abelian nature of the "almost-abelianization", the general construction has a "coarser" flavor than the original construction and requires some non-trivial modifications. This completes the main body of the proof.

The last three sections give supplementary results and information (and formed the appendix of [3]). Section 3.8 provides more background on the associated graded nilpotent Lie group and Carnot-Carathéodory metrics. Section 3.9 shows that the construction at the end of Section 3.1.2 coincides with the construction in Cantrell-Furman's theorem [8]. In Section 3.10, we review the notion of Gromov-Hausdorff convergence and we also provide a missing step in Cantrell-Furman's theorem so that it guarantees Gromov-Hausdorff convergence.

3.2. Construction of the edge weights when Γ is nilpotent and torsion-free with torsion free abelianization

Assume that $\Gamma = N$ is a finitely generated torsion-free nilpotent group with torsion-free abelianization. Moreover, assume that $S = \{s_1, ..., s_d\}$ is such that the image of S under the quotient map $\Gamma \to \Gamma^{ab}$ is a basis, and we choose an isomorphism $\Gamma^{ab} \cong \mathbb{Z}^d$ such that S maps to the standard basis for $\mathbb{Z}^d$. In this and the next section we prove the result of Theorem 3.1.4[1]

[1]Technically we prove a weaker version of Theorem 3.1.4 which still implies the conclusion of Theorem 3.1.1; see Remark 3.2.1 below.

under these extra assumptions, which then implies the result of Theorem 3.1.1 under these extra assumptions, as shown above.

First, let us note that since Γ is nilpotent, we cannot have $d = 0$, and if $d = 1$ then in fact $\Gamma \cong \mathbb{Z}$. (For this latter fact, let $a \in \Gamma$ be such that $\langle a \rangle [\Gamma, \Gamma] = \Gamma$; then also $\langle a \rangle = \Gamma$ by Theorem 16.2.5 in [25]). It is easy to induce any norm on $\mathbb{Z}$ no matter what the finite generating set is using deterministic weights, so from here on we assume $d \geq 2$.

We are given a norm Φ on $\Gamma^{ab} \otimes \mathbb{R} \cong \mathbb{R}^d$. We want to find weights $w : E \to \mathbb{R}_{\geq 0}$ for Γ such that the subadditive function $\tilde{T}$ on $\Gamma^{ab} \cong \mathbb{Z}^d$ induced by $\mathbb{E}T$ via $\Gamma \to \Gamma^{ab}$ is asymptotically equivalent to Φ. Let $B \subset \mathbb{R}^d \cong \Gamma^{ab} \otimes \mathbb{R}$ be the unit ball of Φ. Note that B is a compact, convex, and symmetric (i.e. $x \in B$ implies $-x \in B$) subset of $\mathbb{R}^d$ which contains an open neighborhood of 0. The construction below is a "lift" of the construction of Haggstrom and Meester [20].

We first recall the following geometrc result from [20].

Proposition 3.2.1. *There is a constant C_0 depending only on d such that, for any $u \in \mathbb{R}^d$, if z is a point in $\mathbb{Z}^d$ with minimal Euclidean distance to u, there exists a directed edge path γ from 0 to z in the standard Cayley graph $\mathbb{Z}^d$ with the following properties:*

(1) Any point on γ is at Euclidean distance at most C_0 from some point on the line through 0 and u in $\mathbb{R}^d$

(2) If a subpath of γ starts at $x \in \mathbb{R}^d$ and ends at $y \in \mathbb{R}^d$, then $\langle y - x, u \rangle > 0$.

(3) The number of edges in γ is the least possible, i.e. $\sum_{i=1}^{d} |\pi_i(z)|$, where $\pi_i : \mathbb{R}^d \to \mathbb{R}$ is projection onto the i^{th} coordinate.

We will use the Proposition above as follows. Let $\{b_n\}_{n=1}^{\infty}$ be a countable dense subset of the boundary of $B \subset \mathbb{R}^d$. For each $n \geq 1$, let z_n be a point in $\mathbb{Z}^d$ with minimum possible distance to $2^n \frac{b_n}{||b_n||_2} \in \mathbb{R}^d$, where $|| \cdot ||_2$ is the standard Euclidean norm on $\mathbb{R}^d$. Let γ_n be the path in $\mathbb{Z}^d$ associated to b_n from Proposition 3.2.1. In short, these nice paths γ_n have the property that they (1) stay close to the straight line through b_n, and (2) they travel "monotonically forward" along b_n.

We lift each of these nice paths γ_n to an edge path $\bar{\gamma}_n$ in the Cayley graph of Γ that shares similar properties. The quotient map $\Gamma \to \Gamma^{ab} \cong \mathbb{Z}^d$ induces a covering map of Cayley graphs, so just let $\bar{\gamma}_n$ be the unique lift of γ_n starting at $1 \in \Gamma$. Equivalently, paths in Cayley graphs starting at the identity are naturally in correspondence with words in the generating sets. The path γ_n then corresponds to a word in $e_1, ..., e_d$, which we lift to a word in $s_1, ..., s_d$, which corresponds to a path $\bar{\gamma}_n$ in our Cayley graph for Γ.

For each $n \geq 1$, set $E_n \subset E$ to be the set of edges of the Cayley graph of Γ which share at least one vertex in common with an edge of $\bar{\gamma}_n$. Note that $|E_n| \lesssim 2^n$, where the implied constant depends on $|S|$ but is independent of n.

Now we define a configuration of edge weights $\eta_n : E_n \to \mathbb{R}_+$. First choose $h > 0$ sufficiently small so that $\{x \in \mathbb{R}^d : ||x||_2 \leq h\} \subset B$. Next, choose $K < \infty$ sufficiently large so that $\frac{1}{K - 2h^{-1} \cdot C_0} \leq h$ and $K \geq h^{-1}$. We then define

$$\eta_n(f) = \begin{cases} \frac{|\pi_i(b_n)|}{||b_n||_2^2} & f \in \bar{\gamma}_n, f \text{ labeled by } s_i, \\ K, & \text{otherwise} \end{cases}$$

where π_i is again the projection onto the i^{th} coordinate. If $x \in \Gamma$, then we can also define the translated configuration $T_x\eta_n : xE_n \to \mathbb{R}_+$ by $T_x\eta_n(f) = \eta_n(x^{-1}f)$. The reason for these choices will hopefully become clearer later, but in short we want the weights along the paths $\bar{\gamma}_n$ to yield fast passage times (with correct asymptotic speed) in the direction $\frac{b_n}{||b_n||_2}$. Moreover, $E_n \setminus \bar{\gamma}_n$

forms a “shell” of slow weight K edges around the fast “highway” $\bar{\gamma}_n$; when we have defined our weights, these “shells” will discourage paths from leaving the “highways.”

Let $(Y_x)_{x\in\Gamma}$ and $(Z_x)_{x\in\Gamma}$ be collections of i.i.d. random variables with distributions that satisfy $\mathbb{P}(Y_x = 0) = \frac{1}{2}, \mathbb{P}(Y_x = n) = 3^{-n}$ for $n \geq 1$, and Z_x is uniformly distributed on $[0, 1]$. We also assume that the collections $(Y_x)_{x\in\Gamma}$, $(Z_x)_{x\in\Gamma}$ are independent.

Finally, the weights $w : E \to \mathbb{R}_+$ are defined as follows: if $Y_x = n > 0$, assign the edges in xE_n according to $T_x\eta_n$. If two configurations compete for the same edge, then the configuration with the larger value of n wins; if both configurations have the same value of n, then the one with the larger value of Z_x wins. Any remaining edges with no assigned weight are given weight K.

More formally: for each $f \in E$, let $X_f := \{x \in \Gamma : f \in xE_{Y_x}\}$ be the set of starting points of configurations competing for the edge f. Let $n_f := \max\{Y_x : x \in X_f\}$ be the largest value of n among these competing configurations, and let $x_f \in \Gamma$ be the element of X_f which attains the maximum (that is, $Y_{x_f} = n_f$) and has the largest value of Z_x among such elements, that is, $Z_{x_f} = \max\{Z_x : x \in X_f, Y_x = n_f\}$. Then

$$w(f) = \begin{cases} T_{x_f}\eta_{n_f}(f) & X_f \neq \emptyset \\ K & \text{otherwise.} \end{cases}$$

Note that x_f is a.s. unique since all the Z_x are uniform, and it exists since $|X_f| < \infty$ a.s. by the calculation

$$\mathbb{E}|X_f| = \sum_{x\in\Gamma} \mathbb{P}(f \in xE_{Y_x}) = \sum_{n=1}^{\infty} \sum_{x\in\Gamma} \mathbb{1}_{\{f\in xE_n\}} \mathbb{P}(Y_x = n) \leq \sum_{n=1}^{\infty} |E_n| 3^{-n} \lesssim \left(\sum_{n=1}^{\infty} 2^n \cdot 3^{-n} \right) < \infty.$$

Here we used that Γ acts freely on E and so $\#\{x \in \Gamma : x^{-1} f \in E_n\} \leq |E_n|$. Hence the weights are well-defined. They are also evidently stationary and a.s. bounded above by $K < \infty$. The weights are also ergodic, since we can take our probability space Ω to be $(\mathbb{N} \times [0,1])^\Gamma$, corresponding to the outcomes of Y_x and Z_x, which is clearly ergodic as a direct product of probability spaces over Γ.

Remark 3.2.1. *These weights do* not *give a metric which is bi-Lipschitz to a word metric, since* $\pi_i(b_n)$ *will typically cluster around* 0 *and a uniform lower bound on the edge weights is not available.*

By the remark above, this construction does not suffice to prove Theorem 3.1.4. There are two ways around this. In Section 3.5, we provide a different construction in the general virtually nilpotent case which is bi-Lipschitz to the word metric, and implies Theorem 3.1.4 as stated. Secondly, the weights constructed above *do* satisfy a weaker condition which one might call "bi-Lipschitz away from the diagonal." That is, we have a uniform upper bound K on the edge weights, and there exist some constants $0 < C < \infty$ and $k > 0$ such that for any $x, y \in \Gamma$ with $d(x, y) \geq C$, we have

$$T(x, y) \geq k d(x, y) \tag{3.2.1}$$

almost surely. This fact follows from Lemma 3.7.2 proven in Section 3.7 below. Taking M and k' as in Lemma 3.7.2, and doing a similar analysis as in the next section, one sees that if a path γ with $|\gamma| \geq M$ contains no edges of weight K, then it (or its reverse) is a subpath of a "highway"

$x\bar{\gamma}_n$ ($Y_x = n$) and hence has passage time

$$T(\gamma) = \frac{1}{||b_n||}\left\langle D(\gamma), \frac{b_n}{||b_n||_2}\right\rangle \geq \left(\inf_{b\in B}\frac{1}{||b||_2}\right)k'|\gamma|.$$

On the other hand, if a path γ with $M \leq |\gamma| \leq 2M$ does contain an edge of weight K, then $T(\gamma) \geq K \geq \frac{K}{2M}|\gamma|$. One then concludes (3.2.1) with $k := \min\left(\left(\inf_{b\in B}\frac{1}{||b||_2}\right)k', \frac{K}{2M}\right)$ and $C := M$.

Under this weaker assumption, the proof of Theorem 3.1.3 given in [8] goes through unchanged. Thus, although we prove a weaker version of Theorem 3.1.4 in the next section, namely Theorem 3.1.4 with the conclusion "T is bi-Lipschitz to d" replaced by the conclusion "T is bi-Lipschitz to d away from the diagonal", we can then use the stronger version of Theorem 3.1.3 to still conclude the result of Theorem 3.1.1 in this restricted setting.

3.3. Proof of Theorem 3.1.4 when Γ is nilpotent and torsion-free with torsion free abelianization

Using the weights w defined in the previous section, let T be the metric associated to w as defined in Section 3.1.2. Let $\tilde{T}$ be the subadditive function on Γ^{ab} induced by $\mathbb{E}T$ via the abelianization map $\Gamma \to \Gamma^{ab}$ as above. In order to prove our version of Theorem 3.1.4, all that remains is to show that as $x \in \Gamma^{ab}$ tends to infinity,

$$\tilde{T}(x) - \Phi(x) = o(x),$$

where in the little o notation we may use any norm on $\mathbb{R}^d$ to measure x. We use the following proposition which is used in [20] (where they take $Q = [-1/2, +1/2]^d \subset \mathbb{R}^d$, but the exact form that Q takes does not matter):

Proposition 3.3.1. *To show that $\tilde{T}(x) - \Phi(x) = o(x)$, it suffices to show the following*

(1) For all $y \notin B$, $y \notin \frac{1}{t}\bar{B}(t)$ for all sufficiently large t.

(2) For all y in the interior of B, $y \in \frac{1}{t}\bar{B}(t)$ for all sufficiently large t.

Here we define

$$\bar{B}(t) := \bigcup_{\{x \in \Gamma^{ab} : \tilde{T}(x) \le t\}} x + Q,$$

where $Q \subset \mathfrak{g}^{ab}$ is a compact connected neighborhood of 0 such that the quotient map $Q \to \mathfrak{g}^{ab}/\Gamma^{ab}$ is surjective.

First, we prove (1). To do this, we must establish some facts about the relationship between the T-lengths of paths in E and their "displacements" in Γ^{ab}. In proving these we will repeatedly use the following easily verifiable lemma from [20]:

Lemma 3.3.1. *Let B be a convex subset of $\mathbb{R}^d$ and let $x_1, ..., x_m \in \mathbb{R}^d$, $\alpha_1, ..., \alpha_m \ge 0$ be such that each $\alpha_i^{-1} x_i \in B$. Then $\frac{x_1 + \cdots + x_n}{\alpha_1 + \cdots + \alpha_n} \in B$.*

Let us call an edge $f \in E$ "slow" if $w(f) = K$ and "fast" otherwise. Let us also call an edge path in E "fast" if all its edges are fast and "slow" if all its edges are slow. For an edge path γ in E from $x \in \Gamma$ to $y \in \Gamma$ denote by $D(\gamma)$ its "displacement" $y^{ab} - x^{ab} \in \mathbb{R}^d$. Note that displacement is preserved by left translations:

$$D(z\gamma) = (zy)^{ab} - (zx)^{ab} = (z^{ab} + y^{ab}) - (z^{ab} + x^{ab}) = y^{ab} - x^{ab} = D(\gamma).$$

Let us first consider fast paths γ. Note that by construction of the weights, each fast path is a subpath of $x\bar{\gamma}_n$ for some $x \in \Gamma, n \ge 1$ (because of the "shell" of slow edges surrounding each fast $x\bar{\gamma}_n$). We can then decompose $D(\gamma)$ as

$$D(\gamma) = D_{\parallel}(\gamma) + D_{\perp}(\gamma),$$

where $D_{\|}$ is the orthogonal projection of $D(\gamma)$ onto the line passing through 0 and b_n and $D_{\perp}(\gamma)$ is orthogonal to that line. Note that the construction of the edge weights guarantees precisely that if f is a fast edge in $x\bar{\gamma}_n$ labeled by s_i then

$$\frac{D_{\|}(f)}{T(f)} = \frac{\left\langle \pm e_i, \frac{b_n}{\|b_n\|_2} \right\rangle \frac{b_n}{\|b_n\|_2}}{\frac{|\pi_i(b_n)|}{\|b_n\|_2^2}} = \pm b_n \in B.$$

Then by Lemma 3.3.1 we have

$$\frac{D_{\|}(\gamma)}{T(\gamma)} = \frac{\sum_{f \in \gamma} D_{\|}(f)}{\sum_{f \in \gamma} T(f)} \in B.$$

We also know by Proposition 3.2.1 that

$$\|D_{\perp}(\gamma)\|_2 \leq 2C_0,$$

and hence

$$\frac{D_{\perp}(\gamma)}{h^{-1} \cdot 2C_0} \in \{x \in \mathbb{R}^d : \|x\|_2 \leq h\} \subset B.$$

So again by Lemma 3.3.1,

$$\frac{D(\gamma)}{T(\gamma) + 2h^{-1}C_0} = \frac{D_{\|}(\gamma) + D_{\perp}(\gamma)}{T(\gamma) + h^{-1} \cdot 2C_0} \in B.$$

On the other hand, if f is a slow edge, then by our choice of K

$$\frac{D(f)}{T(f) - 2h^{-1}C_0} \in \{x \in \mathbb{R}^d : \|x\|_2 \leq h\} \subset B,$$

and so for a slow path γ, by Lemma 3.3.1 we have

$$\frac{D(\gamma)}{T(\gamma) - 2|\gamma|h^{-1}C_0} \in B.$$

Now, a general path in E is an alternating concatenation of fast and slow paths. That is, $\gamma = \gamma_f^0\gamma_s^1 \cdots \gamma_s^n\gamma_f^n$, where the γ_f^i are fast, the γ_s^i are slow, and we may take γ_f^0 or γ_f^n to be empty, but all the γ_s^i consist of at least one edge. Then by our previous arguments and Lemma 3.3.1 we have

$$\frac{\sum_{i=0}^n D(\gamma_i^f) + \sum_{i=1}^n D(\gamma_i^s)}{\sum_{i=0}^n (T(\gamma_i^f) + 2h^{-1}C_0) + \sum_{i=1}^n (T(\gamma_i^s) - 2|\gamma_i^s|h^{-1}C_0)} \in B.$$

The numerator in the above expression is $D(\gamma)$, and the denominator is at most $T(\gamma) + 2h^{-1}C_0$, so we have

$$\frac{D(\gamma)}{T(\gamma) + 2h^{-1}C_0} \in B$$

for any path γ in E.

Finally, let $y \notin B$. Since B is closed, there is some $\epsilon > 0$ such that for any $c > 0$, $cB(y, \epsilon) \cap B \neq \emptyset$ implies that $\frac{1}{c} > 1 + \epsilon$. Now for any $t > 0$ let $z \in \Gamma$ be such that $ty \in z^{ab} + Q$, where Q is the fixed compact set in Proposition 3.3.1. If we choose γ to be a T-minimal path from 1 to z in Γ, by our above arguments we have that

$$\frac{z^{ab}}{T(\gamma) + 2h^{-1}C_0} = \frac{t[y - \frac{1}{t}(z^{ab} - ty)]}{T(1, z) + 2h^{-1}C_0} \in B.$$

Therefore, whenever $\frac{\text{diam}(Q)}{t} < \epsilon$, we have $\frac{1}{t}\|z^{ab} - ty\|_2 < \epsilon$ and hence

$$\frac{T(1, z) + 2h^{-1}C_0}{t} > 1 + \epsilon;$$

and so whenever also $\frac{2h^{-1}C_0}{t} < \epsilon/2$, we have

$$\frac{T(1,z)}{t} > 1 + \frac{\epsilon}{2},$$

and then taking expectation gives

$$\frac{\mathbb{E}T(1,z)}{t} > 1 + \frac{\epsilon}{2};$$

since this argument did not depend on our choice of z, we conclude that, for all t sufficiently large, $\tilde{T}(z^{ab}) > t(1+\frac{\epsilon}{2})$ whenever $ty \in z^{ab} + Q$, and hence

$$y \notin \frac{\bar{B}(t)}{t}.$$

Now we prove (2).

It is sufficient to prove that for every $\epsilon > 0$, for all but finitely many n,

$$\frac{\|b_n\|_2 \tilde{T}(z_n)}{2^n} < 1 + \epsilon.$$

Fix $\epsilon > 0$. We give an upper bound on the $\tilde{T}$-distance from 0 to z_n by constructing a path γ from 1 to a lift of z_n in Γ. The lift we choose is the endpoint of the path $\bar{\gamma}_n$, which we denote by $\bar{z}_n$. Note that although the path we construct is random, the endpoints 1 and $\bar{z}_n$ are not.

Denote by Z the center of Γ, and fix a total ordering $<$ on Z such that if $d(1,x_0) < d(1,x_1)$, then $x_0 < x_1$ (recall that here d denotes the word metric on Γ with respect to S). Then choose x to be the least element of Z with respect to this ordering such that $Y_x = n$. Note that x is then a well-defined Z-valued random variable with minimal distance from 1, and that

$$(x = x_0) \Leftrightarrow (Y_{x_0} = n \text{ and } Y_{x_1} \neq n \text{ for all } x_1 < x_0).$$

That is, x is the nearest central starting point of a "highway" in the b_n direction.

Now, to construct our path γ, first, take a path of minimal d-length from 1 to x in Γ. Then, travel along $x\bar{\gamma}_n$ (even if some of the edges are overwritten by slow edges) to $x\bar{z}_n$. Finally, travel back to $x\bar{z}_n x^{-1} = \bar{z}_n$ by traveling backwards along a translate of the path you took from 1 to x. Note that we have used the fact that x is central to conclude that $x\bar{z}_n x^{-1} = \bar{z}_n$ and in particular that the d-distance from $x\bar{z}_n$ to $\bar{z}_n$ is no larger than the d-distance from 1 to x.

If $x\bar{\gamma}_n$ was not overwritten by any slow edges, the passage time of the path would be equal to

$$\sum_{f\in\bar{\gamma}_n} \eta_n(f) = \sum_{f\in\bar{\gamma}_n} \frac{\langle D(f), b_n\rangle}{||b_n||_2^2} = \frac{\langle D(\bar{\gamma}_n), b_n\rangle}{||b_n||_2^2} = \frac{\langle z_n, b_n\rangle}{||b_n||_2^2}.$$

(Here we have used the fact that, by construction, all edges f in γ have positive inner product with b_n.) Since z_n is less than distance $\frac{\sqrt{d}}{2}$ from $2^n \frac{b_n}{||b_n||}$, the above is bounded above by

$$\frac{\left\langle 2^n \frac{b_n}{||b_n||}, b_n\right\rangle}{||b_n||_2^2} + \frac{\frac{\sqrt{d}}{2}\cdot ||b_n||_2}{||b_n||_2^2} = \frac{2^n}{||b_n||}\left(1 + \frac{\sqrt{d}}{2^{n+1}}\right).$$

Taking into account the travel from 1 to x and from $x\bar{z}_n$ to $\bar{z}_n$, as well as the fact that some of the edges of $x\bar{\gamma}_n$ may be overwritten by slow edges, we have

$$\mathbb{E}T(\gamma) \le K[2\mathbb{E}d(1,x) + \mathbb{E}\#\{e \in x\bar{\gamma}_n : e \text{ is slow}\}] + \frac{2^n}{||b_n||}\left(1 + \frac{\sqrt{d}}{2^{n+1}}\right). \tag{3.3.1}$$

To bound the first term, we calculate

$$\mathbb{E}d(1,x) = \sum_{i=0}^{\infty} \mathbb{P}(d(1,x) > i) = \sum_{i=0}^{\infty} \mathbb{P}(Y_\xi \neq n \text{ for all } \xi \in B_d(i) \cap Z).$$

Since we have assumed that $\Gamma \not\cong \mathbb{Z}$, the growth of the center is at least 2-dimensional, that is, we have some $C > 0$ depending only on Γ and S such that

$$|B_d(i) \cap Z| \geq Ci^2$$

for all $i \geq 0$. This is proved in Lemma 3.3.3 below, but for now we take it for granted.

Then, since the Y_ξ are iid, we continue the above computation to get

$$\mathbb{E}d_S(1,x) \leq \sum_{i=0}^{\infty}(1-3^{-n})^{Ci^2} \leq 1 + \int_0^\infty (1-3^{-n})^{Cs^2}ds.$$

Using the substitution $\sigma = \left[\frac{\ln(1-3^{-n})}{\ln(1-3^{-1})}\right]^{1/2} s$, we get

$$\int_0^\infty (1-3^{-n})^{Cs^2}ds = \left[\frac{\ln(1-3^{-n})}{\ln(1-3^{-1})}\right]^{-1/2} \int_0^\infty (1-3^{-1})^{C\sigma^2}d\sigma,$$

which is to say that

$$\mathbb{E}d(1,x) \leq 1 + C'[-\ln(1-3^{-n})]^{-1/2}$$

for some $C' > 0$ independent of n. By convexity, $-\ln(1-s) \geq s$ for all $s < 1$, and so

$$[-\ln(1-3^{-n})]^{-1/2} \leq (3^{-n})^{-1/2} = 3^{n/2},$$

thus

$$\mathbb{E}d(1,x) \lesssim 3^{n/2}, \tag{3.3.2}$$

the implied constant of course independent of n.

Now, we bound

$$\mathbb{E}\#\{e \in x\bar{\gamma}_n : e \text{ is slow }\} = \sum_{e\in\bar{\gamma}_n} \mathbb{P}(xe \text{ is slow});$$

since xe will only be slow if another $T_z E_{Y_z}$ with $Y_z \geq n$ competes for it, the above quantity is bounded above by

$$\begin{aligned}
&\sum_{e\in\bar{\gamma}_n} \mathbb{P}(xe \in zE_{Y_z} \text{ and } Y_z \geq n \text{ for some } x \neq z \in \Gamma) \\
&\leq \sum_{e\in\bar{\gamma}_n} \sum_{x_0\in\Gamma} \sum_{z\in\Gamma\setminus x_0} \sum_{i=n}^{\infty} \mathbb{P}(x = x_0, x_0 e \in zE_i, Y_z = i) \\
&= \sum_{e\in\bar{\gamma}_n} \sum_{x_0\in\Gamma} \sum_{i=n}^{\infty} \sum_{z\in\Gamma : x_0^{-1} ze\in E_i} \mathbb{P}(x = x_0, Y_z = i);
\end{aligned}$$

we claim that for $i \geq n$ and $x_0 \neq z$, $\mathbb{P}(x = x_0, Y_z = i) \leq \frac{3}{2}\mathbb{P}(x = x_0)\mathbb{P}(Y_z = i)$, and hence we continue

$$\begin{aligned}
\mathbb{E}\#\{e \in x\bar{\gamma}_n : e \text{ is slow }\} &\leq \sum_{e\in\bar{\gamma}_n} \sum_{x_0\in\Gamma} \sum_{i=n}^{\infty} \sum_{z\in\Gamma : x_0^{-1} ze\in E_i} \frac{3}{2}\mathbb{P}(x = x_0)\mathbb{P}(Y_z = i) \\
&\leq \frac{3}{2} \sum_{e\in\bar{\gamma}_n} \sum_{x_0\in\Gamma} \sum_{i=n}^{\infty} |E_i| \mathbb{P}(x = x_0)\mathbb{P}(Y_z = i) = \frac{3}{2} \sum_{e\in\bar{\gamma}_n} \sum_{i=n}^{\infty} |E_i| \mathbb{P}(Y_z = i) \\
&\lesssim \sum_{e\in\bar{\gamma}_n} \sum_{i=n}^{\infty} 2^i \cdot 3^{-i} = \sum_{e\in\bar{\gamma}_n} 3\left(\frac{2}{3}\right)^n = 3|\bar{\gamma}_n|\left(\frac{2}{3}\right)^n \\
&\lesssim \left(\frac{4}{3}\right)^n. \qquad (3.3.3)
\end{aligned}$$

To prove the claim, note that for $x_0 \neq z$, $i \geq n$,

$$\mathbb{P}(x = x_0, Y_z = i) = \mathbb{P}(Y_{x_1} \neq n \text{ for all } x_1 < x_0, Y_{x_0} = n, Y_z = i);$$

if $x_0 < z$, then all these events are independent, and hence $\mathbb{P}(x = x_0, Y_z = i) = \mathbb{P}(x = x_0)\mathbb{P}(Y_z = i)$. Otherwise $z < x_0$, and then

$$\mathbb{P}(x = x_0, Y_z = i) = \left(\prod_{x_1 < x_0, x_1 \neq z} \mathbb{P}(Y_{x_1} \neq n)\right)\mathbb{P}(Y_{x_0} = n)\mathbb{P}(Y_z \neq n, Y_z = i).$$

If $i = n$, then this is equal to 0. Otherwise, $i > n$, and

$$\mathbb{P}(Y_z \neq n, Y_z = i) = \mathbb{P}(Y_z = i) = \frac{\mathbb{P}(Y_z = i)}{\mathbb{P}(Y_z \neq n)}\mathbb{P}(Y_z \neq n) \leq \frac{3}{2}\mathbb{P}(Y_z = i)\mathbb{P}(Y_z \neq n),$$

where we used that $\mathbb{P}(Y_z \neq n) = 1 - 3^{-n} \geq \frac{2}{3}$. Hence

$$\begin{aligned}\mathbb{P}(x = x_0, Y_z = i) &\leq \frac{3}{2}\left(\prod_{x_1 < x_0, x_1 \neq z} \mathbb{P}(Y_{x_1} \neq n)\right)\mathbb{P}(Y_{x_0} = n)\mathbb{P}(Y_z \neq n)\mathbb{P}(Y_z = i)\\ &= \frac{3}{2}\mathbb{P}(x = x_0)\mathbb{P}(Y_z = i),\end{aligned}$$

as desired.

Hence, applying (3.3.1), (3.3.2), and (3.3.3),

$$\frac{\|b_n\|_2\tilde{T}(z_n)}{2^n} \leq \frac{\|b_n\|_2\mathbb{E}T(\gamma)}{2^n} \leq K\|b_n\|_2\left[2O\left(\left(\frac{3^{1/2}}{2}\right)^n\right) + O\left(\left(\frac{2}{3}\right)^n\right)\right] + 1 + \frac{\sqrt{k}}{2^{n+1}},$$

which is less than $1 + \epsilon$ for sufficiently large n, as desired.

To tie up the final loose end, we prove that the volume growth of the center of Γ is at least 2-dimensional. This is a simple corollary of the following lemma from the notes of Drutu and Kapovich [11]:

Lemma 3.3.2 (Lemma 14.15 from [11])**.** *Let Γ be a finitely generated nilpotent group of class k and let $C^k\Gamma$ be the last nontrivial term in its lower central series. If S is a generating set* *for* Γ*, and* $g \in C^k\Gamma$*, then there exists a constant* $\lambda = \lambda(S, g)$ *such that for all* $m \geq 0$*,*

$$d_S(1, g^m) \leq \lambda m^{1/k}.$$

Lemma 3.3.3. *Let* Γ *be a nontrivial finitely generated torsion-free nilpotent group which is not isomorphic to* $\mathbb{Z}$*,* S *a finite generating set for* Γ*. Denote the center of* Γ *by* Z*. Then, there exists a constant* $C > 0$ *depending only on* Γ *and* S *such that*

$$\#\{z \in Z : d(1, z) \leq i\} \geq Ci^2$$

for all $i \geq 0$*.*

Proof. We know that Z is a nontrivial finitely generated free abelian group. First, assume that $Z \not\cong \mathbb{Z}$. Then $Z \cong \mathbb{Z}^k$ for some $k \geq 2$. Then the lemma follows, since the quantity in question grows at least as fast as Z does as a finitely generated group. More explicitly, if S' is a finite generating set for $Z \cong \mathbb{Z}^k$, we know that there exists $C' > 0$ depending only on S' such that

$$\#\{z \in Z : d_{S'}(1, z) \leq i\} \geq C' i^k.$$

Take $m = \max_{s \in S'} d(1, s) < \infty$. Then for all $z \in Z$, $d(1, z) \leq m d_{S'}(1, z)$, and hence

$$\#\{z \in Z : d(1, z) \leq i\} \geq \#\left\{z \in Z : d_{S'}(1, z) \leq \frac{i}{m}\right\} \geq \frac{C'}{m^k} i^k.$$

Now, suppose $Z \cong \mathbb{Z}$. Then Γ is not abelian (otherwise we would have $\Gamma = Z \cong \mathbb{Z}$, contradicting our assumption). So Γ is nilpotent of step k for some $k \geq 2$, and $C^k\Gamma$ is a nontrivial subgroup of Z. Take a generator g for $C^k\Gamma$. By Lemma 3.3.2, we get $\lambda = \lambda(g, S) > 0$ such that

$d(1, g^m) \leq \lambda m^{1/k}$ for all $m \geq 0$. Therefore

$$\{z \in Z : d(1,z) \leq i\} \geq \{m \geq 0 : d(1,g^m) \leq i\} \geq \{m \geq 0 : \lambda m^{1/k} \leq i\} \geq \left\lfloor \frac{1}{\lambda^k} i^k \right\rfloor \geq C i^k$$

for some $C > 0$. □

3.4. Restrictions in the virtually nilpotent case

Any finitely generated virtually nilpotent group Γ will contain a finite index subgroup H which is finitely generated, nilpotent, torsion free, and which has torsion-free abelianization (see Section 3.9). We often think of the H and Γ as having the same coarse geometry; indeed:

Proposition 3.4.1. *Let Γ be a group endowed with a metric T, let H be a finite index subgroup, and let (X, D) be a metric space. If $T \lesssim d$ (d the word metric) and $(H, \frac{1}{t}(T|_H)) \xrightarrow{GH} (X, D)$, then also $(\Gamma, \frac{1}{t}T) \xrightarrow{GH} (X, D)$.*

Proof. Since $(H, \frac{1}{t}T|_H)$ is a metric subspace of $(\Gamma, \frac{1}{t}T)$, the Gromov-Hausdorff distance between the two spaces is bounded—up to an absolute constant—by

$$\inf\{\epsilon > 0 : T(g, H) < \epsilon \text{ for all } g \in \Gamma\},$$

which is itself bounded up to a constant by

$$\frac{1}{t}[\Gamma : H] = O(1/t).$$

Thus $(\Gamma, \frac{1}{t}T)$ and $(H, \frac{1}{t}T)$ must tend to the same limit. □

Thus, it might seem trivial to pass from the simplified case we just proved to the general case. However, perhaps surprisingly, the answer to the question we consider is *not* the same for

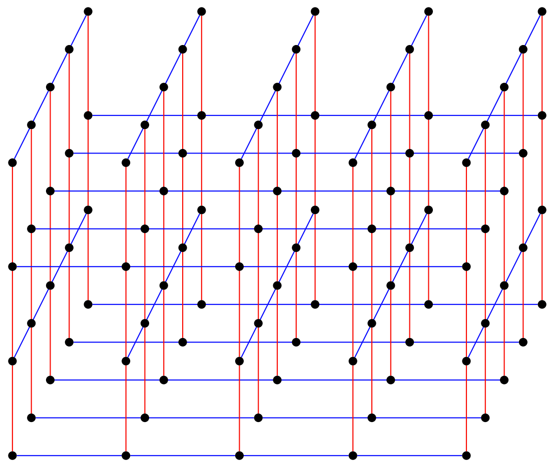

Figure 3.2. A portion of the Cayley graph of $\langle\rho\rangle \ltimes \mathbb{Z}[i]$ with respect to the generating set $\{\rho, 1+0i\}$. Edges labeled by ρ are red, while edges labeled by $1+0i$ are blue.

Γ and H. In general, there may be some limit shapes for stationary FPPs on H which are *not* attained by stationary FPPs on Γ. Consider the following example.

Let $\Gamma := \langle\rho\rangle \ltimes \mathbb{Z}[i]$, the semidirect product of the Gaussian integers with a cyclic group of order four, the generator of the cyclic group acting by multiplication by i. Γ contains the abelian (hence nilpotent) group $\mathbb{Z}[i] \cong \mathbb{Z}^2 =: H$ as a subgroup of index 4. We know from our work above (and from [20]) that any norm on $\mathbb{R}^2$ is attainable as a limit shape for H. However, we claim that the scaling limit of any invariant metric on Γ which is $\lesssim d$ (such as $\mathbb{E}T$ for a stationary FPP T with integrable weights) must be a norm on $\mathbb{R}^2$ which has $\frac{\pi}{4}$ rotational symmetry. Take any $(x+iy) \in \mathbb{Z}[i]$. Then

$$\begin{aligned}\mathbb{E}T(1, i(x+iy)) &= \mathbb{E}T(1, \rho^{-1}(x+iy)\rho) \\ &\leq \mathbb{E}T(1,\rho^{-1}) + \mathbb{E}T(\rho^{-1}, \rho^{-1}(x+iy)) + \mathbb{E}T(\rho^{-1}(x+iy), \rho^{-1}(x+iy)\rho) \\ &= \mathbb{E}T(1,\rho^{-1}) + \mathbb{E}T(1,(x+iy)) + \mathbb{E}T(1,\rho) \leq \mathbb{E}T(1,(x+iy)) + 2(\text{const.}).\end{aligned}$$

Iterating this inequality four times and taking a scaling limit gives

$$\lim_{n\to\infty} \frac{\mathbb{E}T(1, n(x+iy))}{n}$$
$$= \lim_{n\to\infty} \frac{\mathbb{E}T(1, ni(x+iy))}{n} = \lim_{n\to\infty} \frac{\mathbb{E}T(1, -n(x+iy))}{n} = \lim_{n\to\infty} \frac{\mathbb{E}T(1, -ni(x+iy))}{n},$$

which is precisely the statement that the limit norm has quarter-turn symmetry.

A similar restriction arises in any virtually nilpotent group. As in Section 3.1.2, let Γ be a finitely generated virtually nilpotent group, and let N be a torsion-free nilpotent normal subgroup of finite index (for the construction of such a subgroup see Section 3.9). The conjugation action of Γ on N induces an action of $\Gamma/N =: Q$ on N^{ab}_{free}. It will be convenient later to phrase things in terms of the right conjugation action, and so we think of the action as a homomorphism $\phi : Q \to \mathrm{Aut}(N^{ab}_{free})^{op}$. This further induces a right action of Q on $N^{ab} \otimes \mathbb{R} \cong N^{ab}_{free} \otimes \mathbb{R} \cong \mathfrak{g}^{ab}$, which, by abuse of notation, we also denote by $\phi : Q \to \mathrm{Aut}(\mathfrak{g}^{ab})^{op}$. We say that a norm on Φ on $\mathfrak{g}^{ab}$ is *conjugation-invariant* if it is ϕ-invariant, that is,

$$\Phi(x^{\phi(q)}) = \Phi(x)$$

for all $x \in N^{ab} \otimes \mathbb{R}, q \in Q$.

Proposition 3.4.2. *Let Γ, N, ϕ be as above. If T is a stationary integrable FPP on Γ such that the scaling limit of $\mathbb{E}T$ is a Carnot-Carathéodory metric on a nilpotent Lie group L_∞, then the norm on $\mathfrak{g}^{ab}$ associated to this metric is ϕ-invariant.*

Proof. The proof is very similar to our example. First, let $\tilde{Q}$ be a finite set of coset representatives of N, that is, a finite subset $\tilde{Q} \subset \Gamma$ such that the quotient map $\Gamma \to Q$ induces a bijection $\tilde{Q} \leftrightarrow Q$; Since $\tilde{Q}$ is finite and the FPP is integrable, there exists some constant $C < \infty$

such that $\mathbb{E}T(1,\tilde{q}), \mathbb{E}T(1,\tilde{q}^{-1}) \leq C$ for all $\tilde{q} \in \tilde{Q}$. Then, for any $x \in N$ and any $\tilde{q} \in \tilde{Q}$,

$$\mathbb{E}T(1, x^{\tilde{q}}) \leq \mathbb{E}T(1,\tilde{q}^{-1}) + \mathbb{E}T(1,x) + \mathbb{E}T(1,\tilde{q}) \leq \mathbb{E}T(1,x) + 2C$$

where we have used the fact that $\mathbb{E}T$ is left-invariant. Similarly, we have

$$\mathbb{E}T(1,x) = \mathbb{E}T(1,(x^{\tilde{q}})^{\tilde{q}^{-1}}) \leq \mathbb{E}T(1,x^{\tilde{q}}) + 2C,$$

and thus

$$|\mathbb{E}T(1,x) - \mathbb{E}T(1,x^{\tilde{q}})| \leq 2C.$$

Since ϕ respects the quotient map $N \to N^{ab}_{free}$, taking infima over $x \in N$ such that $x^{ab}_{free} = z$ for some fixed $z \in N^{ab}_{free}$ gives

$$|\tilde{T}(z) - \tilde{T}(z^{\phi(q)})| \leq 2C = o(z);$$

that is, $\tilde{T}$ is asymptotically equivalent to $\tilde{T}^{\phi(q)}$ for all $q \in Q$, and hence the norm Φ it induces on $\mathfrak{g}^{ab}$ is $\phi(q)$-invariant. Pansu's theorem [30] tells us that Φ is the norm in the Carnot-Carathéodory construction of the scaling limit of $(\Gamma, \mathbb{E}T)$, so we are done. □

Although there is certainly more work to be done in exploring necessary conditions for the existence of a limit shape, in all cases which we know how to prove ([5], [8]), the scaling limit of the random space (Γ, T) coincides with the scaling limit of its mean $(\Gamma, \mathbb{E}T)$, so this tells us that conjugation invariance is a necessary feature of a limit shape at least in all cases in which we can prove there is a scaling limit.

Theorem 3.1.1 then states that this is the *only* obstruction to a Carnot-Carathéodory metric on L_∞ being the limit shape of a stationary FPP on Γ; that is, as long as the Carnot-Carathéodory metric comes from a norm which is conjugation-invariant, it is the scaling limit of some FPP with stationary weights.

3.5. Construction of the edge weights in the virtually nilpotent case

Transferring our theorem to the general virtually nilpotent case is far from automatic, essentially since our Cayley graph may not be nice with respect to the the finite index subgroups we wish to pass to. Recall that N is a finite-index torsion-free nilpotent normal subgroup of Γ. Instead of keeping track of "displacements" of paths by looking at the projection to Γ^{ab}, we want to instead look at N^{ab}_{free}, and there is typically no nice homomorphism from Γ to N^{ab}_{free}. Nor is there a nice embedding $N^{ab} \to \Gamma^{ab}$; the natural map can have very large kernel (e.g. in our example $\Gamma := \langle\rho\rangle \ltimes \mathbb{Z}[i]$ above, Γ^{ab} is finite, while $N = N^{ab} = \mathbb{Z}[i]$). Ultimately, we resolve this by looking at a slightly nonabelian notion of "displacement" via the projection $\Gamma \to \Gamma/\widetilde{[N,N]}$, where we define $\widetilde{[N,N]}$ to be the kernel of the projection $N \to N^{ab}_{free}$. Note that $\widetilde{[N,N]}$ is indeed normal in Γ: an element $x \in \Gamma$ is in $\widetilde{[N,N]}$ if and only if $x \in N$ and for some $1 \le k < \infty$, $x^k \in [N,N]$; since N is normal in Γ, both these properties are preserved under conjugation by any element $g \in \Gamma$. Note also that $\Gamma/\widetilde{[N,N]}$ contains N^{ab}_{free} as a subgroup of finite index.

In spite of these complications, the spirit of the proof exactly the same. Heuristically, we want to ensure that every direction has the correct "speed" at large scales, and we do this by sprinkling long "fast" paths throughout the graph which travel at a certain speed in a certain direction; the rest of the edges are "slow" so that any long geodesic must largely avoid them.

It is clear from our above proof that the weight K of the slow edges can be as large as we like, as long as it is finite. We use the slowness of the edges to account for any error in the fast paths–that is, to guard against the fact that a subpath of a fast path might not go in exactly the right direction or exactly at the right speed.

In our first proof, we used the existence of nice paths (Proposition 3.2.1) which had the property that they (1) stayed close to the straight line through b_n, and (2) traveled "monotonically forward" along b_n. In the general case, we will want to find nice paths in $\Gamma/[\widetilde{N,N}]$ which satisfy these properties in a certain "coarse" sense to be described below.

Let us now go into more detail understanding the group $\Gamma/[\widetilde{N,N}]$, especially considering it as a finite extension of N^{ab}_{free}. First, take a finite set of coset representatives $\tilde{Q} \subset \Gamma/[\widetilde{N,N}]$ for $N/[\widetilde{N,N}]$; we assume for convenience that $\tilde{Q}$ contains the identity. The quotient map $\Gamma/[\widetilde{N,N}] \to Q := (\Gamma/[\widetilde{N,N}])/(N/[\widetilde{N,N}]) \cong \Gamma/N$ induces a bijection $\tilde{Q} \to Q$, and we denote its inverse by $s : Q \to \tilde{Q}$. If s were a homomorphism, we would have a semidirect product, but this is not always possible in general. In general, define a function $\eta : Q \times Q \to N^{ab}_{free}$ satisfying

$$s(q_1)s(q_2) = s(q_1q_2)\eta(q_1, q_2).$$

This then allows us to understand $\Gamma/[\widetilde{N,N}]$ more explicitly thus: note that $Q \times N^{ab}_{free} \to \Gamma/[\widetilde{N,N}]$, $(q, n) \mapsto s(q)n$ is a bijection. Pulling back the multiplication from $\Gamma/[\widetilde{N,N}]$ to the set $Q \times N^{ab}_{free}$ then gives the multiplication

$$(Q \times N^{ab}_{free}) \times (Q \times N^{ab}_{free}) \to Q \times N^{ab}_{free}$$

$$(q_1, n_1) \cdot (q_2, n_2) := (q_1q_2, \eta(q_1, q_2) + n_1^{\phi(q_2)} + n_2).$$

Thus, $\Gamma/[\widetilde{N,N}]$ looks like a semidirect product up to the "finite error" introduced by η.

Remark 3.5.1. *η is in fact a cocycle; the cocycle condition comes precisely from the associativity of the above multiplication. However, we will not use this fact. Rather, we will repeatedly use the simple fact that η is a map from the finite set $Q \times Q$, and thus has finite image and hence uniformly bounded image.*

Remark 3.5.2. *The cocycle η of course depends on our choice of $\tilde{Q}$, and the choice is non-unique.*

We will now introduce two modified notions of displacement which will be convenient for us. Let γ be a path in E (the Cayley graph of Γ) starting at $x \in \Gamma$ and ending at $y \in \Gamma$. We define

$$\tilde{D}(\gamma) := \bar{x}^{-1}\bar{y} \in \Gamma/[\widetilde{N,N}],$$

where $\bar{x}, \bar{y}$ are the images of x, y under the projection $\Gamma \to \Gamma/[\widetilde{N,N}]$. Note that $\tilde{D}$ is invariant with respect to the action of Γ on paths in E by left multiplication. Note also that for concatenations of paths $\gamma = \alpha * \beta$ we have

$$\tilde{D}(\gamma) = \tilde{D}(\alpha)\tilde{D}(\beta).$$

It will also be helpful for us to have a notion of displacement which lives in N^{ab}_{free} rather than $\Gamma/[\widetilde{N,N}]$; for this, we take a particular choice of point in N^{ab}_{free} nearby (in the Cayley graph of $\Gamma/[\widetilde{N,N}]$) to $\tilde{D}(\gamma)$:

$$D(\gamma) := \tilde{D}(\gamma)\tilde{q}(\gamma)^{-1} \in N^{ab}_{free},$$

where $\tilde{q}(\gamma)$ is the image of $\tilde{D}(\gamma)$ under the composition $\Gamma/[\widetilde{N,N}] \to Q \xrightarrow{s} \tilde{Q}$; put another way, using the identification $\Gamma/[\widetilde{N,N}] \leftrightarrow Q \times N^{ab}_{free}$, if $\tilde{D}(\gamma) = (q, n)$, then $D(\gamma) = (q, n)(q^{-1}, 0) = n^{\phi(q)^{-1}}$. Note also that if $\tilde{D}(\gamma) \in N^{ab}_{free}$, then $\tilde{D}(\gamma) = D(\gamma)$.

$D(\gamma)$ is convenient because it always lands in N^{ab}_{free}, the space we are trying to induce the correct norm on; however, instead of being additive on paths, using the definition and the concatenation property for $\tilde{D}$, we instead get the slightly more complicated equation

$$D(\alpha\beta) = D(\alpha) + D(\beta)^{\phi(\alpha)} + \eta(\alpha,\beta)^{\phi(\alpha\beta)^{-1}}, \tag{3.5.1}$$

where in an abuse of notation, we define $\eta(\alpha,\beta) := \eta(q(\alpha), q(\beta))$, $\phi(\alpha) := \phi(q(\alpha))$, where $q(\alpha)$ is the image of $\tilde{D}(\alpha)$ under the quotient map $\Gamma/[N,N] \to Q$. Iterating the above fact easily gives the following by induction:

Proposition 3.5.1. *For any paths $\alpha_1, ..., \alpha_N$ in E, we have*

$$D(\alpha_1 \cdots \alpha_N) = D(\alpha_1) + \sum_{i=1}^{N-1} \left(D(\alpha_{i+1}) + \eta(\alpha_1 \cdots \alpha_i, \alpha_{i+1})^{\phi(\alpha_{i+1})^{-1}} \right)^{\phi(\alpha_1 \cdots \alpha_i)^{-1}}$$

Thus, although the displacements do not add, besides the twisting of ϕ we only accumulated at most one uniformly bounded error term per path concatenated, which will end up being enough later.

From now on we fix an isomorphism $\mathfrak{g}^{ab} \cong \mathbb{R}^d$ such that N^{ab}_{free} is identified with $\mathbb{Z}^d \subset \mathbb{R}^d$ via the map $N^{ab}_{free} \to N^{ab}_{free} \otimes \mathbb{R} \cong \mathfrak{g}^{ab} \cong \mathbb{R}^d$. We will often thus identify $D(\gamma)$ with its image in $\mathbb{R}^d$.

We are now ready to state the properties we want for our "nice" paths in E (which will become "fast" paths).

Lemma 3.5.1. *There exists a constant $C_0' > 0$ depending only on Γ, S, N, and $\tilde{Q}$ such that, for any vector $u \in \mathbb{R}^d$ and any $n \in \mathbb{Z}_{\geq 0}$ there exists a simple path γ in E such that*

(1) γ *starts at* $1 \in \Gamma$ *and* $\|D(\gamma) - 2^n u\|_2 \leq C_0'$.

(2) $|\gamma| \lesssim |D(\gamma)| \lesssim 2^n \|u\|_2$.

(3) γ stays near the line through u: If α is a subpath of γ starting at 1, then $\|D(\alpha) - \mathrm{proj}_u D(\alpha)\|_2 \leq C_0'$.

(4) γ is a finite concatenation of paths β_i where for each i, $|\beta_i| \leq C_0'$, $\|D(\beta')^{\phi(q)}\|_2 \leq C_0'$ for all $q \in Q$ and every subpath β' of β_i, and

$$\left\langle D(\beta_0 \cdots \beta_{i+1}) - D(\beta_0 \cdots \beta_i), \frac{u}{\|u\|_2} \right\rangle \geq \frac{1}{C_0'},$$

that is, γ is "coarsely monotone."

We also assume that $\max_{q_1,q_2,q_3 \in Q} \|\eta(q_1, q_2)^{\phi(q_3)}\|_2 \leq C_0'$.

This lemma will be proven in Section 3.7.

For now, we define the edge weights, very similarly to the first construction. First, given a Carnot-Carathéodory metric with associated norm Φ on $\mathfrak{g}^{ab}$, let $B \subset \mathfrak{g}^{ab} \cong \mathbb{R}^d$ be the unit ball of Φ. Let $\{b_n\}_{n \geq 0}$ be a countable dense subset of the boundary of B. For each n, let γ_n be the path given in Lemma 3.5.1 associated to the vector b_n and the natural number n. Let E_n be the set of edges in E which share at least one vertex with the path γ_n.

Pick $h > 0$ small enough so that $B_2(0, h) \subset B$ and then choose $K > 0$ large enough so that

$$\max_{f \in S, q, q_1, q_2, q_3 \in Q} \frac{\|D(f)^{\phi(q)}\|_2 + \|\eta(q_1, q_2)^{\phi(q_3)}\|_2}{K - 9C_0' h^{-1}} \leq h.$$

Then define $\eta_n : E_n \to \mathbb{R}_+$ by

$$\eta_n(f) = \begin{cases} \frac{\left\langle D(\beta_0 \cdots \beta_i) - D(\beta_0 \cdots \beta_{i-1}), \frac{b_n}{\|b_n\|_2} \right\rangle}{\|b_n\|_2 |\beta_i|} & f \in \beta_i, \\ K, & \text{otherwise.} \end{cases}$$

where the β_i are the subpaths of $\gamma = \gamma_n$ alluded to in Lemma 3.5.1 (the dependence of β_i on n is suppressed in the notation).

Lastly, we superimpose randomly sprinkled translated copies of the η_n exactly as in the first construction; that is, define $\{Z_x\}_{x\in\Gamma}, \{Y_x\}_{x\in\Gamma}$, X_f, x_f, and n_f exactly as above and then define $w : E_n \to \mathbb{R}_+$

$$w(f) = \begin{cases} T_{x_f}\eta_{n_f}(f) & X_f \neq \emptyset \\ K & \text{otherwise.} \end{cases}$$

By the same arguments as above, these weights are well-defined, ergodic, and uniformly bounded above. Moreover, the monotonicity condition in Lemma 3.5.1 implies that each edge has weight at least

$$\min_{b\in B} \frac{1}{C_0'^2 \|b\|_2} > 0,$$

which is to say that T is bi-Lipschitz to the word metric, and we can apply Theorem 3.1.3.

3.6. Proof of Theorem 3.1.4 in the general case

Once again, the proof that the correct norm is induced on $\mathfrak{g}^{ab}$ can be reduced to showing the conditions in Proposition 3.3.1. The proof of the second condition is the same argument as in the simplified case. (We construct the desired paths by traveling along the center of N until we reach the first fast path that goes in the correct direction, and then we travel back along the center of N. We have the same volume growth estimates that we used above as long as we assume Γ is not virtually $\mathbb{Z}$. In the virtually $\mathbb{Z}$ case, our limit shapes are norms on $\mathbb{R}$, and since all norms on $\mathbb{R}$ are scalar multiples of each other, we can achieve any desired norm we like by appropriately scaling the weights of, say, the deterministic FPP which assigns weight 1 to each edge and gives $T = d$.)

For the first condition of Proposition 3.3.1, the spirit of the proof is the same, but we have to deal with more error terms.

First, we consider a fast subpath γ of E (that is, a path which does not contain any edges of length K), and again we note that it is (up to translation) a subpath of some γ_n. First consider the case that γ travels forward rather than backward along γ_n. Then we write

$$\gamma = \alpha\beta_j \cdots \beta_i\omega,$$

where the β_i are the subpaths alluded to in Lemma 3.5.1 and α and ω are subpaths of β_{j-1} and β_{i+1} respectively.

Now, by Equation (3.5.1), we know that

$$D(\beta_j \cdots \beta_i)^{\phi(\beta_0 \cdots \beta_{j-1})} = [D(\beta_0 \cdots \beta_i) - D(\beta_0 \cdots \beta_{j-1})] - \eta(\beta_0 \cdots \beta_{j-1}, \beta_j \cdots \beta_i)^{\phi(\beta_0 \cdots \beta_i)^{-1}}.$$

We can further decompose $[D(\beta_0 \cdots \beta_i) - D(\beta_0 \cdots \beta_{j-1})]$ into its components parallel to b_n and perpendicular to b_n:

$$[D(\beta_0 \cdots \beta_i) - D(\beta_0 \cdots \beta_{j-1})] = [D(\beta_0 \cdots \beta_i) - D(\beta_0 \cdots \beta_{j-1})]_{\|} + [D(\beta_0 \cdots \beta_i) - D(\beta_0 \cdots \beta_{j-1})]_{\perp}.$$

Now, by our definition of η_n we have

$$T(\beta_j \cdots \beta_i) = \frac{1}{\|b_n\|_2} \left\langle D(\beta_0 \cdots \beta_i) - D(\beta_0 \cdots \beta_{j-1}), \frac{b_n}{\|b_n\|_2} \right\rangle,$$

where we have used coarse monotonicity of γ. Thus, we have

$$\frac{[D(\beta_0 \cdots \beta_i) - D(\beta_0 \cdots \beta_{j-1})]_{\|}}{T(\beta_j \cdots \beta_i)} = b_n \in B.$$

Moreover, since γ stays near to the line through b_n we have

$$\frac{[D(\beta_0\cdots\beta_i)-D(\beta_0\cdots\beta_{j-1})]_\perp}{2C_0'h^{-1}}\in B_2(0,h)\subset B,$$

and by assumptions on C_0' we have

$$\frac{-\eta(\beta_0\cdots\beta_{j-1},\beta_j\cdots\beta_i)^{\phi(\beta_0\cdots\beta_i)}}{C_0'h^{-1}}\in B_2(0,h)\subset B.$$

Hence by Lemma 3.3.1

$$\frac{D(\beta_j\cdots\beta_i)^{\phi(\beta_0\cdots\beta_{j-1})}}{T(\beta_j\cdots\beta_i)+3C_0'h^{-1}}\in B,$$

and then by conjugation-invariance of B we have

$$\frac{D(\beta_j\cdots\beta_i)}{T(\beta_j\cdots\beta_i)+3C_0'h^{-1}}\in B.$$

Now, since α and ω are subpaths of β_{i-1} and β_{j+1}, we have

$$\frac{D(\alpha)}{C_0'h^{-1}},\frac{D(\omega)}{C_0'h^{-1}}\in B_2(0,h)\subset B,$$

and hence by Lemma 3.3.1

$$\frac{D(\alpha\beta_j\cdots\beta_i\omega)}{T(\beta_j\cdots\beta_i)+7C_0'h^{-1}}=\frac{D(\alpha)+D(\beta_j\cdots\beta_i)^{\phi(\cdot)}+\eta(\cdot,\cdot)^{\phi(\cdot)}+D(\omega)^{\phi(\cdot)}+\eta(\cdot,\cdot)^{\phi(\cdot)}}{C_0'h^{-1}+T(\beta_j\cdots\beta_i)+3C_0'h^{-1}+C_0'h^{-1}+C_0'h^{-1}+C_0'h^{-1}}\in B,$$

where we have again used conjugation-invariance of B.

Now, if γ travels backwards rather than forwards along γ_n, we apply the above argument to $\overline{\gamma}$ (the reverse of γ) to obtain

$$\frac{D(\overline{\gamma})}{T(\overline{\gamma}) + 7C_0'h^{-1}} = \frac{D(\overline{\gamma})}{T(\gamma) + 7C_0'h^{-1}} \in B.$$

Since we chose $\tilde{Q}$ to contain 1, $D(\gamma\overline{\gamma}) = 0$ and so by Equation (3.5.1) we have that

$$D(\gamma) = -D(\overline{\gamma})^{\phi(\gamma)} - \eta(\gamma, \overline{\gamma}).$$

So again using symmetry and conjugation invariance of B, together with assumptions on C_0' and Lemma 3.3.1, we conclude

$$\frac{D(\gamma)}{T(\gamma) + 8C_0'h^{-1}} \in B.$$

Now, for slow edges f, by choice of K we have

$$\frac{D(f) + \eta(\cdot,\cdot)^{\phi(\cdot)}}{T(f) - 9C_0'h^{-1}} \in B_2(0, h) \subset B.$$

Writing an arbitrary path γ as a concatenation of fast paths and slow edges and using Propositon 3.5.1 gives

$$D(\gamma) = \sum_{f \text{ slow edges}} (D(f) + \eta(\cdot,\cdot)^{\phi(\cdot)})^{\phi(\cdot)} + \sum_{\gamma' \text{ fast paths}} (D(\gamma') + \eta(\cdot,\cdot)^{\phi(\cdot)})^{\phi(\cdot)},$$

and so using the above and Lemma 3.3.1 gives

$$\frac{D(\gamma)}{\sum_{f \text{ slow edges}}(T(f) - 9C_0'h^{-1}) + \sum_{\gamma' \text{ fast paths}}(T(\gamma') + 9C_0'h^{-1})} \in B,$$

and since there is at most one more fast path than there are slow edges, we conclude

$$\frac{D(\gamma)}{T(\gamma) + 9C_0'h^{-1}} \in B.$$

The rest of the proof is just as in the above argument.

3.7. Proof of Lemma 3.5.1

To prove the existence of "nice paths" we want to approximate the nice paths in $\mathbb{Z}^d \cong N_{free}^{ab}$ from Proposition 3.2.1 and prove that our approximation retains the nice properties "coarsely".

First, we prove a lemma which will help control error terms:

Lemma 3.7.1. *There exists a constant K' such that for any paths α, β in E, we have*

$$||D(\alpha\beta) - D(\alpha)||_2 \leq K'|\beta|.$$

Proof. By Equation (3.5.1), we know that

$$D(\alpha\beta) - D(\alpha) = D(\beta)^{\phi(\alpha)} + \eta(\alpha,\beta)^{\phi(\alpha\beta)^{-1}}.$$

First, since the image of Q in $\mathrm{Aut}(N_{free}^{ab}) \cong SL_d^{\pm}(\mathbb{Z})$ is a finite family of bounded operators on $\mathbb{R}^d$, there is some constant $M < \infty$ such that

$$||v^{\phi(q)}||_2 \leq M||v||_2$$

for all $q \in Q, v \in \mathbb{R}^d$. Thus we have $||D(\beta)^{\phi(\alpha)}||_2 \leq M||D(\beta)||_2$.

Next, since N_{free}^{ab} is finite index in $\Gamma/[\widetilde{N,N}]$, it is undistorted, which is to say that any word metric on N_{free}^{ab} is bi-Lipschitz to the restriction to N_{free}^{ab} of any word metric on $\Gamma/[\widetilde{N,N}]$. (This can be seen using Schreier generators for N^{ab}_{free}, see e.g. Theorem 14.3.1 in [25]). In particular, this means that the Euclidean norm $\|\cdot\|_2$ on N^{ab}_{free} is bi-Lipschitz to the metric induced by the Cayley graph on $\Gamma/[\widetilde{N,N}]$. Hence

$$\|D(\beta)\|_2 \leq K''|D(\beta)| = K''|\tilde{D}(\beta)\tilde{q}(\beta)^{-1}| \leq K''(|\beta| + \max_{\tilde{q}\in\tilde{Q}} |\tilde{q}|).$$

Lastly, since Q is finite, we have a uniform bound on the norm of the second term, that is,

$$\max_{q_1,q_2,q_3\in Q} \|\eta(q_1,q_2)^{\phi(q_3)}\|_2 < \infty.$$

Putting everything together gives

$$\|D(\alpha\beta) - D(\alpha)\|_2 \leq MK''|\beta| + \text{const.},$$

and since every nonempty β has $|\beta| \geq 1$ we can easily adjust to get a finite K' which satisfies the desired inequality. □

Now, we construct the paths. Given u and n, first consider the path γ_n in $\mathbb{Z}^d \cong N^{ab}_{free}$ using the standard generators e_i of $\mathbb{Z}^d$ given by Proposition 3.2.1. Next, for each edge e of the path in the standard generators, choose a path β' in the Cayley graph for $\Gamma/[\widetilde{N,N}]$ induced by the image of S which starts one vertex of e and ends at the other; pick these paths to satisfy

$$|\beta'| \leq \max_{i=1,\ldots,d} d'(1,e_i) =: C \tag{3.7.1}$$

where d' is the word metric on $\Gamma/[\widetilde{N,N}]$ induced by the image of S. We then lift to a path $\tilde{\beta}'_0 \cdots \tilde{\beta}'_{N-1}$ in E. Note that by the properties guaranteed by Proposition 3.2.1 we have that:

$$\|D(\tilde{\beta}'_0 \cdots \tilde{\beta}'_{N-1}) - 2^n u)\|_2 \leq \frac{\sqrt{d}}{2}, \tag{3.7.2}$$

$$|\tilde{\beta}'_0 \cdots \tilde{\beta}'_{N-1}| \lesssim 2^n \|u\|_2, \tag{3.7.3}$$

and

$$\|D(\tilde{\beta}'_0 \cdots \tilde{\beta}'_i) - \mathrm{proj}_u D(\tilde{\beta}'_0 \cdots \tilde{\beta}'_i)\|_2 \leq C_0 \tag{3.7.4}$$

for all i. If α is a general subpath of $\tilde{\beta}'_0 \cdots \tilde{\beta}'_{N-1}$ starting at 1, it is of the form $\alpha = \tilde{\beta}'_0 \cdots \tilde{\beta}'_i \alpha'$ where α' is a subpath of $\tilde{\beta}'_{i+1}$, and hence combining Lemma 3.7.1 together with Equations (3.7.1) and (3.7.4) gives

$$\|D(\alpha) - \mathrm{proj}_u D(\alpha)\|_2 \leq C_0 + K'C. \tag{3.7.5}$$

Thus, $\tilde{\beta}'_0 \cdots \tilde{\beta}'_{N-1}$ satisfies many of the properties we desire. However, it may contain loops, and it may not satisfy coarse monotonicity. So first erase loops to get a simple path $\tilde{\beta}_0 \cdots \tilde{\beta}_{N'-1}$. The particular manner in which loops are erased does not matter, so long as the resulting path is a simple path with the same starting and ending point which is obtained from the original path by deleting subpaths. If entire segments $\tilde{\beta}'_i$ are deleted, the number N' of new segments $\tilde{\beta}_0, ..., \tilde{\beta}_{N'-1}$ need not be the same as N the number of original segments, and some reindexing may be required so that we don't skip indices; however, every $\tilde{\beta}_i$ is composed of subpaths of a

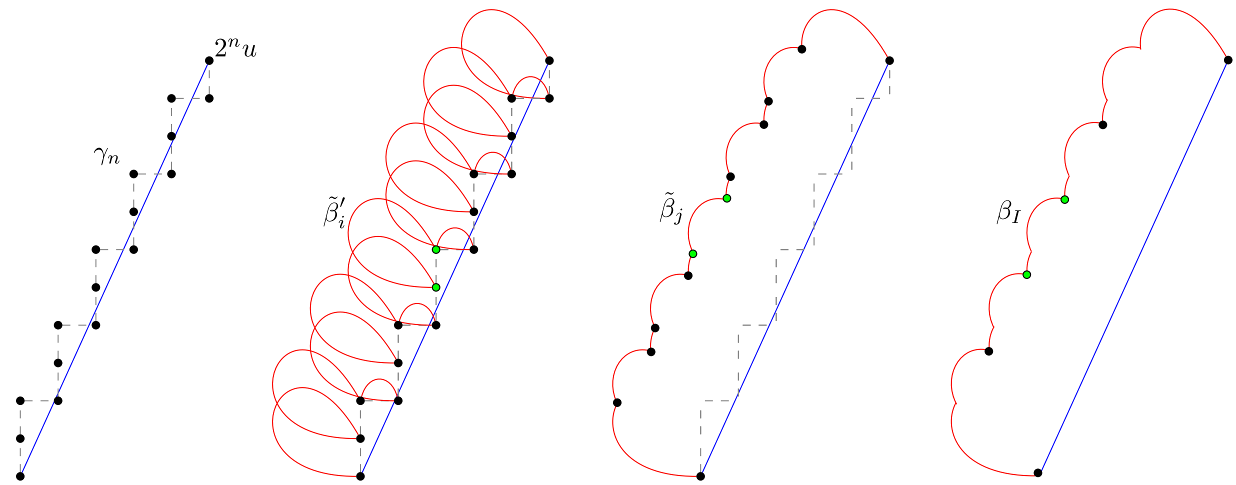

Figure 3.3. Construction of "nice paths". (The "lifting" step is omitted here to aid visualization).

single $\tilde{\beta}'_j$, j depending on i. Thus, each segment $\tilde{\beta}_i$ of the new path still consists of at most C edges.

Moreover, since the set of displacements of subpaths of the loop-erased path is a subset of the set of displacements of subpaths of the original path, Equation (3.7.5) holds for the new path as well. Equations (3.7.2) and (3.7.3) also clearly pass to the loop-erased path as well.

Now we obtain coarse monotonicity. First we prove the following version of coarse monotonicity for the original Euclidean paths:

Lemma 3.7.2. *There exists some $k' > 0$ and $M < \infty$ such that any subpath γ of γ_n (γ_n the path in the standard Cayley graph of $\mathbb{Z}^d$ from Proposition 3.2.1 associated to $u = b_n$) of length at least M satisfies*

$$\left\langle D(\gamma), \frac{u}{||u||_2} \right\rangle \geq k'|\gamma|.$$

Proof. First, we claim that there is a constant C depending only on d such that for any subpath of any γ_n of edge-length at least C, at least one edge f of the path satisfies

$$\left\langle D(f), \frac{u}{\|u\|_2} \right\rangle \geq \frac{1}{\sqrt{d}}.$$

Heuristically, this is because the path cannot travel too long in directions perpendicular to u while staying close to the line through 0 and u. More rigorously, for some coordinate $i_0 \in \{1, ..., d\}$ we have

$$|\pi_{i_0}(u)| \geq \frac{\|u\|_2}{\sqrt{d}}.$$

For notational convenience, let's replace some of the standard basis vectors with their opposites to ensure that $\langle u, e_i \rangle = |\pi_i(u)| \geq 0$ for all i, and further, let's reindex so that $e_1, ..., e_l$ satisfy $c_i := \left\langle e_i, \frac{u}{\|u\|_2} \right\rangle < \frac{1}{\sqrt{d}}$ and $e_{l+1}, ..., e_d$ satisfy $c_i \geq \frac{1}{\sqrt{d}}$ for some $0 \leq l < d$.

Now let γ be a subpath of γ_n starting at $x \in \mathbb{Z}^d$ and ending at $y \in \mathbb{Z}^d$, and assume that for every edge f in γ,

$$\left\langle D(f), \frac{u}{\|u\|_2} \right\rangle < \frac{1}{\sqrt{d}}.$$

By Proposition 3.2.1, x and y must be within Euclidean distance C_0 of the line L passing through 0 and b_n in $\mathbb{R}^n$. Moreover, since we only travel in directions with low weights, we have $y = x + n_1 e_1 + \cdots + n_l e_l$ for some positive integers n_i. Now, the distance from y to L is

$$\begin{aligned}
\operatorname{dist}(y, L) &= \left\| x + n_1 e_1 + \cdots + n_l e_l - \left\langle x + n_1 e_1 + \cdots + n_l e_l, \frac{u}{\|u\|_2} \right\rangle \frac{u}{\|u\|_2} \right\|_2 \\
&\geq \left\| n_1 e_1 + \cdots + n_l e_l - \left\langle n_1 e_1 + \cdots + n_l e_l, \frac{u}{\|u\|_2} \right\rangle \frac{u}{\|u\|_2} \right\|_2 - \operatorname{dist}(x, L),
\end{aligned}$$

so, since both distances are less than C_0, we have

$$\begin{aligned}
2C_0 &\geq \left\| n_1 e_1 + \cdots + n_l e_l - \left\langle n_1 e_1 + \cdots + n_l e_l, \frac{u}{\|u\|_2} \right\rangle \frac{u}{\|u\|_2} \right\|_2 \\
&= \left\| n_1 e_1 + \cdots n_l e_l - (n_1 c_1 + \cdots n_l c_l) \left(\sum_{i=1}^{d} c_i e_i \right) \right\|_2 \\
&\geq \sqrt{\sum_{i=1}^{l} (n_i - (n_1 c_1 + \cdots + n_l c_l) c_i)^2} \\
&\geq C' \sum_{i=1}^{l} (n_i - (n_1 c_1 + \cdots + n_l c_l) c_i) \\
&\geq C' \sum_{i=1}^{l} \left(n_i - (n_1 + \cdots + n_l) \frac{1}{d} \right) = C' \left(1 - \frac{l}{d} \right) \sum_{i=1}^{l} n_i \\
&\geq \frac{C'}{d} \sum_{i=1}^{l} n_i = \frac{C'}{d} |\gamma|.
\end{aligned}$$

To go from the third to the fourth line, we used that the Euclidean norm is equivalent to the ℓ_1 norm on $\mathbb{R}^d$, to go from the fourth to the fifth line, we used that $0 < c_i < \frac{1}{\sqrt{d}}$ for $i = 1, ...l$, and to get to the final line we used that $l \leq d - 1$.

Thus, any subpath of γ_n which consists of at least $C := \lfloor \frac{2C_0 d}{C'} \rfloor + 1$ edges contains at least one edge with displacement at least $1/\sqrt{d}$ in the u direction.

Finally, this implies that, for any subpath γ of γ_n with length at least $2C$ we have

$$\left\langle D(\gamma), \frac{u}{\|u\|_2} \right\rangle \geq k \lfloor \frac{|\gamma|}{C} \rfloor \geq \frac{k}{2C} |\gamma|.$$

That is, we have the lemma with $M = 2C$ and $k' = \frac{k}{2C}$.

□

Now take $M' = \max(M, \lceil \frac{2K'C+1}{k'} \rceil)$. We then define a new segmentation $\beta_0, ..., \beta_{\lfloor N'/M' \rfloor - 1}$ of the path by

$$\beta_i = \tilde{\beta}_{M'i}\tilde{\beta}_{M'i+1} \cdots \tilde{\beta}_{M'i+(M'-1)}$$

if $i < \lfloor N'/M' \rfloor - 1$ and

$$\beta_i = \tilde{\beta}_{M'i} \cdots \tilde{\beta}_{N'-1}$$

if $i = \lfloor N'/M' \rfloor - 1$. Note that we have

$$|\beta_i| \leq 2M'C.$$

To show that this segmentation of the path gives coarse monotonicity, we have to compare with the original path before erasing loops. To this end, for a given $i < \lfloor N'/M' \rfloor - 1$, let I be such that $\tilde{\beta}_{(M'+1)i}$ is a subpath of $\tilde{\beta}'_I$; that is, the index such that the next edge in $\beta_1 \cdots \beta_{\lfloor N'/M' \rfloor - 1}$ after the segment β_i lies in $\tilde{\beta}'_I$. For $i = \lfloor N'/M' \rfloor - 1$, we set $I = N$. We also set J to be such that the last edge in the path β_{i-1} lies in $\tilde{\beta}'_J$; that is, $\tilde{\beta}_{(i-1)M'-1}$ is a subpath of $\tilde{\beta}'_I$. If $i = 0$, we set $J = 0$.

Now note that there exists some (possibly empty) subpath α of $\tilde{\beta}'_J$ such that

$$D(\beta_0 \cdots \beta_{i-1}\alpha) = D(\tilde{\beta}'_0 \cdots \tilde{\beta}'_J)$$

and there exists some subpath ω of $\tilde{\beta}'_I$ such that

$$D(\beta_0 \cdots \beta_i\omega) = D(\tilde{\beta}'_0 \cdots \tilde{\beta}'_I).$$

Hence, by Lemma 3.7.1 and Equation (3.7.1), we have that

$$(3.7.6) \qquad ||D(\beta_0 \cdots \beta_i) - D(\tilde{\beta}'_0 \cdots \tilde{\beta}'_I)||_2, ||D(\beta_0 \cdots \beta_{i-1}) - D(\tilde{\beta}'_0 \cdots \tilde{\beta}'_J)||_2 \leq K'C,$$

which then implies that

$$\left\langle D(\beta_0\cdots\beta_i)-D(\beta_0\cdots\beta_{i-1}),\frac{u}{\|u\|}\right\rangle \geq \left\langle D(\tilde{\beta}'_0\cdots\tilde{\beta}'_I)-D(\tilde{\beta}'_0\cdots\tilde{\beta}'_J),\frac{u}{\|u\|_2}\right\rangle - 2K'C.$$

Now, by construction each $\tilde{D}(\tilde{\beta}'_0\cdots\tilde{\beta}'_i)\in N^{ab}_{free}$, and hence we have

$$D(\tilde{\beta}'{}_0\cdots\tilde{\beta}'{}_I)-D(\tilde{\beta}'_0\cdots\tilde{\beta}'_J)=D(\tilde{\beta}'_{J+1}\cdots\tilde{\beta}'_I),$$

and then since $D(\tilde{\beta}'_{J+1}\cdots\tilde{\beta}_I)$ is the displacement of a subpath of the path γ_n (in the *standard* Cayley graph of $\mathbb{Z}^d$) with edge length at least $I-(J+1)\geq M'\geq M$, Lemma 3.7.2 then gives

$$\left\langle D(\tilde{\beta}'_0\cdots\tilde{\beta}'_I)-D(\tilde{\beta}'_0\cdots\tilde{\beta}'_J),\frac{u}{\|u\|_2}\right\rangle \geq k'M'\geq 2K'C+1,$$

and so combining with Equation (3.7.6) gives

$$\left\langle D(\beta_0\cdots\beta_i)-D(\beta_0\cdots\beta_{i-1}),\frac{u}{\|u\|}\right\rangle \geq 2K'C+1-2K'C=1.$$

Thus, taking

$$C'_0 := \max\left(\sqrt{d}/2, C_0+K'C, 2M'C, 1, \max_{q_1,q_2,q_3\in Q}\eta(q_1,q_2)^{\phi(q_3)}\right)$$

and $\gamma := \beta_0\cdots\beta_{\lfloor N'/M'\rfloor-1}$ gives the Lemma as desired. □

3.8. Carnot-Carathéodory metrics and the associated graded Lie group

In this section we explain the construction needed to describe continuum limits of nilpotent groups, i.e. the associated graded nilpotent Lie group associated to a finitely generated virtually nilpotent group, and Carnot-Carathéodory metrics on this group. As above, let Γ be a finitely generated virtually nilpotent group, and let N be a torsion-free nilpotent group of finite index. A theorem of Mal'cev ([27], see also Theorem 2.18 in [31]) says that there exists a simply connected nilpotent Lie group L such that N is (isomorphic to) a cocompact lattice in L. Let $\mathfrak{g}$ be the Lie algebra of L. Let $\mathfrak{g}_\infty$ be the associated graded nilpotent Lie algebra, that is

$$\mathfrak{g}_\infty := \bigoplus_{i\geq 1} \mathfrak{g}^i/\mathfrak{g}^{i+1},$$

where $\mathfrak{g}^1 := \mathfrak{g}$, $\mathfrak{g}^{i+1} := [\mathfrak{g}^i, \mathfrak{g}]$ is the descending central series for $\mathfrak{g}$. Let L_∞ be the unique simply connected Lie group which has $\mathfrak{g}_\infty$ as its Lie algebra. We will refer to L_∞ as the *graded nilpotent Lie group associated to* Γ.

The map

$$N \hookrightarrow L \to L/[L, L] \cong \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}] =: \mathfrak{g}^{ab}$$

induces an inclusion $N^{ab}_{free} \to \mathfrak{g}^{ab}$ and an isomorphism $N^{ab} \otimes \mathbb{R} \to \mathfrak{g}^{ab}$. Now consider a norm Ψ on $N^{ab} \otimes \mathbb{R} \cong \mathfrak{g}^{ab}$. Note that $\mathfrak{g}^{ab} = \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}^1/\mathfrak{g}^2$ is a vector subspace of $\mathfrak{g}_\infty$. By left translation in L_∞, the subspace $\mathfrak{g}^{ab} \subset \mathfrak{g}_\infty$ gives a left-invariant distribution on TL_∞, and we can extend the norm to any vector in the distribution. Let us call a path $\xi : [a, b] \to L_\infty$ *admissible* if it is differentiable a.e. and a.e. ξ' belongs to the support of the distribution. We can then define the Ψ-length of ξ to be

$$\Psi(\xi) := \int_a^b \Psi(\xi'(t))dt,$$

and this gives a metric on L_∞ by

$$d_\Psi(x, y) := \inf\{\Psi(\xi) : \xi \text{ is an admissible path from } x \text{ to } y\}. \tag{3.8.1}$$

The metric d_Ψ is called the Carnot-Carathéodory metric on L_∞ associated to Ψ. Since $\mathfrak{g}^{ab}$ generates $\mathfrak{g}_\infty$ as a Lie algebra, by Chow's theorem [17], the topology induced on L_∞ by d_Ψ coincides with the usual topology on L_∞.

The above information is sufficient to understand the statement of the main theorem. The following further data is required to understand Section 3.10. The Lie algebra $\mathfrak{g}_\infty$ has a one-parameter family of automorphisms $\delta_t : \mathfrak{g}_\infty \to \mathfrak{g}_\infty, t > 0$ given by setting

$$\delta_t(X) = t^i X$$

if $X \in \mathfrak{g}^i/\mathfrak{g}^{i+1}$ and extending by linearity. This of course integrates to a 1-parameter family of automorphisms of L_∞, which we also denote by δ_t. We refer to δ_t as *dilations*.

Note that d_Ψ is *homogeneous* in the sense that $d_\Psi(\delta_t(x), \delta_t(y)) = td_\Psi$. In the abelian case, $\Gamma = \mathbb{Z}^d$, $L_\infty = \mathbb{R}^d$, the dilations are scalar multiplication by t, and d_Ψ is the usual metric induced by the norm Ψ on $\mathbb{R}^d$.

We now describe a sequence of maps $\Gamma \to L_\infty$ which will be Gromov-Hausdorff approximations (see Section 3.10) when Γ and L_∞ are endowed with the appropriate metrics. First, choose a collection of linear subspaces $V_1, ..., V_k$ of $\mathfrak{g}$ such that for each i

$$\mathfrak{g} = V_1 \oplus \cdots \oplus V_i \oplus \mathfrak{g}^{i+1}.$$

Note that for each i, $V_i \subset \mathfrak{g}^i$ and the natural map $V_i \to \mathfrak{g}^i/\mathfrak{g}^{i+1}$ is in isomorphism of vector spaces. Let

$$M : \mathfrak{g} = V_1 \oplus \cdots \oplus V_k \to \oplus_{i=1}^k \mathfrak{g}^i/\mathfrak{g}^{i+1} = \mathfrak{g}_\infty$$

be the associated linear isomorphism. Then we define a family of maps

$$\mathrm{scl}_t : \Gamma \hookrightarrow L \xrightarrow{\log} \mathfrak{g} \xrightarrow{M} \mathfrak{g}_\infty \xrightarrow{\delta_t} \mathfrak{g}_\infty \xrightarrow{\exp} L_\infty.$$

(Here log is the inverse of $\exp : \mathfrak{g} \to L$, which is a diffeomorphism, since L is a simply connected nilpotent Lie group).

3.9. Understanding the limit norm Φ via N^{ab}_{free}

Our description of the construction of the limit norm Φ on $\mathfrak{g}^{ab}$ differs slightly from the description in [8]. The two descriptions certainly coincide in the case that $\Gamma = N$ is a torsion-free finitely generated nilpotent group with torsion-free abelianization. However, it's not immediately obvious that their description matches ours in the general virtually nilpotent case. This section is primarily intended to show how our statement of Theorem 3.1.3 follows from the following:

Theorem 3.9.1. *[8] Let H be a finitely generated nilpotent group which is torsion-free and has torsion-free abelianization. Let T be a stationary random metric on H which is* inner *(see below) and bi-Lipschitz to a word metric on H. Let d_Φ be the Carnot-Carathéodory metric on L_∞ associated to the metric $\mathbb{E}T$, as in Section 3.1.2 (with $\Gamma = N = H$). Then almost surely*

$$(H, \frac{1}{n}T, 1) \xrightarrow[n\to\infty]{} (L_\infty, d_\Phi, 1)$$

is the sense of pointed Gromov-Hausdorff convergence.

First let us construct relevant finite-index subgroups.

Proposition 3.9.1. *Let Γ be a finitely generated virtually nilpotent group. Then there exists a finite index subgroup H of Γ which is nilpotent, torsion-free,* and *has torsion-free abelianization.*

Proof. By definition, Γ contains a nilpotent subgroup Γ' of finite index, and this is also finitely generated by Schreier's lemma (see e.g. [25] Theorem 14.3.1). Thus Γ' contains a *torsion-free* subgroup Γ'' of finite index (see [25], Theorem 17.2.2).

Take N to be the kernel of the map $\Gamma \to \mathrm{Sym}(\Gamma/\Gamma'')$ given by the action of Γ on the cosets of Γ'' by left multiplication. Since $N \leq \Gamma''$, N is nilpotent and torsion free, and since N is the kernel of a map to a finite subgroup, it is a finite index normal subgroup of Γ.

Now we extract a finite index subgroup H of Γ which is nilpotent, torsion-free, *and* has torsion-free abelianization as follows. One explicit construction is given by Yves Cornulier in the MathOverflow post [24]; this construction also has the advantage that that the natural map $H^{ab} \to N^{ab}$ induced by the inclusion $H \hookrightarrow N$ is itself an inclusion (also of finite index).

Here is the construction: recall that we have a projection map $N \to N^{ab} \to N^{ab}/N^{ab}_{tor} =: N^{ab}_{free}$. Take a basis of d generators $e_1, ..., e_d$ for $\mathbb{Z}^d \cong N^{ab}_{free}$, and lift them to $s_1, ..., s_d \in N$; then we claim that $H := \langle s_1, ..., s_d \rangle \leq N$ is a finite index subgroup with torsion free abelianization.

To see that H has torsion-free abelianization, consider the natural map $H^{ab} \to N^{ab}_{free}$ induced by the map $H \hookrightarrow N \to N^{ab}_{free}$. We claim this is an injection. For if $n_1\bar{s}_1 + \cdots + n_d\bar{s}_d$ is in the kernel of this map, by the choice of $s_1, ..., s_d$ this means that $n_1e_1 + \cdots + n_de_d = 0$, which implies that $n_1, ..., n_d = 0$, since $e_1, ..., e_d$ is a basis. The map is also clearly surjective by construction, so $H^{ab} \cong N^{ab}_{free}$ and so H has torsion-free abelianization.

To see that H is finite index and finish the proof of Proposition 3.9.1, first note that, from the above, $H^{ab} \leq N^{ab}$ is finite index. We then use the following lemma below; the proof is taken from Cornulier's argument in [24].

□

Lemma 3.9.1. *Let N be a finitely generated nilpotent group, and let H be subgroup of N such that $H[N,N]$ is finite index in N (equivalently, $H^{ab} \to N^{ab}$ has finite-index image in N^{ab}). Then H is finite index in N.*

Proof. We proceed by induction on the nilpotency degree of N. If N is abelian, then the statement is immediate.

Suppose the statement holds for all nilpotent groups of degree $k-1$, and suppose N is degree k. Let N^k be the k^{th} subgroup in the descending central series for N. By our inductive hypothesis applied to N/N^k, HN^k is a finite index subgroup of N. So all that remains is to show that H is finite index in HN^k.

For this, first note that since all $(k+1)$-fold commutators vanish, the k-fold commutator map $N \times \cdots \times N \to N^k$ is "multilinear" in the sense that

$$[a_1, \cdots, xy, \cdots, a_k] = [a_1, \cdots, x, \cdots, a_k] \cdot [a_1, \cdots, y, \cdots, a_k];$$

we also see that the output only depends on the abelianizations of $a_1, ..., a_k$, and thus the k-fold commutator map induces a surjective homomorphism from the tensor product $N^{ab} \otimes \cdots \otimes N^{ab} \to N^k$. We claim that the map $\bigotimes^k H^{ab} \to \bigotimes^k N^{ab}$ induced by the finite index inclusion $H \to N$ has image which is finite index in $\bigotimes^k N^{ab}$. Once we know this, since H^k is precisely the composition of the map $\bigotimes^k H^{ab} \to \bigotimes^k N^{ab} \to N^k$, H^k is finite index in N^k, and hence H is finite index in HN^k.

Now, to see that the image of $\bigotimes^k H^{ab} \to \bigotimes^k N^{ab}$ is finite index, we use the following general fact: If A is a finitely generated abelian group and $B \leq A$ is a subgroup of finite index, then for any $i \geq 1$, $\bigotimes^i B \leq \bigotimes^i A$ is finite index. For $i = 1$, this is immediate. Now, inductively assume T' is a finite set such that $T' + \bigotimes^i B = \bigotimes^i A$, and let S' be a finite generating set for $\bigotimes^i B$. Also let T be a finite set such that $T + B = A$ and let S be a finite generating set for B. We claim that the set

$$\left\{ \sum_{\sigma \in S'} t_\sigma \otimes \sigma + \sum_{\tau \in T'} \left(t_\tau \otimes \tau + \sum_{s \in S} s \otimes t'_{s,\tau} \right) : t_\sigma, t_\tau \in T, t'_{s,\tau} \in T' \right\}$$

forms a finite set of coset representatives for $\bigotimes^{i+1} B$ in $\bigotimes^{i+1} A$.

To see this, first consider a general element of $\bigotimes^{i+1} A$. It is a sum of elements of the form

$$(\sum_{s \in S} m_s s + t) \otimes (\sum_{s' \in S'} m_{s'} s' + t')$$

where $t \in T, t' \in T'$, $m_s, m_{s'} \in \mathbb{Z}$, and hence, by expansion, equal to

$$\sum_{\sigma \in S'} (\sum_{s \in S} m_{\sigma,s} s + t_\sigma) \otimes \sigma + \sum_{\tau \in T'} (\sum_{s \in S} m_{\tau,s} s + t_\tau) \otimes \tau$$

for some $m_{\sigma,s}, m_{\tau,s} \in \mathbb{Z}$, $t_\sigma, t_\tau \in T$. Since every $s \otimes \sigma \in \otimes^{k+1} B$, the element

$$\sum_{\sigma \in S'} t_\sigma \otimes \sigma + \sum_{\tau \in T'} \left(t_\tau \otimes \tau + \sum_{s \in S} s \otimes m_{\tau,s} \tau \right)$$

represents the same coset of $\otimes^{k+1} B$. For each s, τ, by the inductive hypothesis, we have

$$s \otimes m_{\tau,s} \tau = s \otimes \left(\sum_{s' \in S'} n_{s',\tau} s' + t'_{s,\tau} \right)$$

for some $n_{s',\tau} \in \mathbb{Z}$ and $t'_{s,\tau} \in T'$, and this is equivalent *modulo* $\otimes^k B$ to

$$\sum_{s' \in S'} s \otimes t'_{s,\tau}.$$

That is, an arbitrary element is equivalent to one in the set provided, as desired. □

In sum, by Proposition 3.9.1, we have $H \leq N \trianglelefteq \Gamma$ finite index inclusions, where N is torsion-free and H is torsion-free with torsion-free abelianization.

Now, let T be a stationary random metric on Γ which is almost surely inner and bi-Lipschitz to a word metric on Γ. Recall that a metric space is called *inner* if for all $\epsilon > 0$, there exists $0 < R < \infty$ such that for any $x, y \in \Gamma$, there exists an (ϵ, R)-coarse geodesic from x to y, that is, a sequence $x = p_0, p_1, ..., p_M = y$ in Γ such that each $d(x_{i-1}, x_i) \leq R$ and

$$\sum_{i=1}^{M} d(p_{i-1}, p_i) \leq (1 + \epsilon) d(x, y).$$

(Note that, in Chapter 3, we consider T an FPP with edge weights w uniformly bounded above; such T is automatically inner). We want to show that

$$(\Gamma, \frac{1}{n} T) \to (L_\infty, d_\Phi).$$

By Proposition 3.4.1, it suffices to show that

$$(H, \frac{1}{n} T|_H) \to (L_\infty, d_\Phi).$$

Thus, we want to apply Theorem 3.1.3 to H, so first we must check that the hypotheses are satisfied.

Proposition 3.9.2. *Let Γ, H, T be as above. Then $T|_H$ is bi-Lipschitz to a word metric on H and $T|_H$ is inner.*

Proof. $T|_H$ is bi-Lipschitz to $d|_H$, and since $H \leq \Gamma$ is finite index, any word metric on H is bi-Lipschitz to $d|_H$ (this can be seen using Schreier generators for H, see e.g. Theorem 14.3.1 in [25]), so we have the first claim.

Next, we show innerness. Let $\epsilon > 0$. First, using the innerness of T on Γ, choose $r > 0$ so that any $x, y \in \Gamma$ can be joined by an $(\frac{\epsilon}{2}, r)$-coarse geodesic. Next, note that since $H \leq \Gamma$ is finite index and $T \leq Kd$ a.s. for some $K < \infty$, we have

$$\max_{g \in \Gamma} T(g, H) \leq K \max_{g \in \Gamma} d(g, H) =: C$$

for some non-random constant $0 < C < \infty$. Now choose $0 < R < \infty$ sufficiently large so that $0 < \frac{4C}{R-r} \leq \frac{\epsilon}{2}$. We claim that any $h, h' \in H$ can be joined by an $(\epsilon, R + 2C)$-coarse geodesic in H.

To construct such a coarse geodesic, first take an $(\frac{\epsilon}{2}, r)$-coarse geodesic $h = p'_0, p'_1, ..., p'_{M'} = h'$ in Γ. By deleting points, we can construct a $(\frac{\epsilon}{2}, R)$-coarse geodesic $h = p_0, ..., p_M = h'$ with

$$M \leq \left\lceil \frac{T(h, h')}{R - r} \right\rceil \leq \frac{2T(h, h')}{R - r},$$

where the last inequality only holds for $T(h, h') \geq R - r$, but if $T(h, h') \leq R + 2C$ then $p_0 = h, p_1 = h'$ trivially gives an $(\epsilon, R+2C)$-coarse geodesic, so we may assume this inequality holds.

Lastly, for each p_i, choose $q_i \in H$ with $T(p_i, q_i) \leq C$ (and of course $q_0 = p_0 = h, q_M = p_M = h'$). Then each $T(q_{i-1}, q_i) \leq T(p_{i-1}, p_i) + 2C \leq R + 2C$ and

$$\begin{aligned}\sum_{i=1}^{M} T(q_{i-1}, q_i) &\leq \sum_{i=1}^{M} T(p_{i-1}, p_i) + 2CM \leq (1 + \frac{\epsilon}{2})T(h, h') + 2CM \\ &\leq (1 + \frac{\epsilon}{2})T(h, h') + 2C \cdot \frac{2T(h, h')}{R - r} \\ &\leq (1 + \epsilon)T(h, h'),\end{aligned}$$

so $q_0, ..., q_M$ is an $(\epsilon, R + 2C)$-coarse geodesic in H, as desired. □

Now, note that the Malcev completions of H and N coincide; if N is a cocompact lattice in G, then as a finite-index subgroup of N, H is also cocompact in G. Therefore H and N have the same associated graded nilpotent Lie group L_∞ as well. Thus, Theorem 3.1.3 tells us that

$$(H, \frac{1}{n} T|_H) \to (L_\infty, d_{\Phi_H}),$$

where we define Φ_H to be the unique norm on $\mathfrak{g}^{ab}$ asymptotically equivalent to the subadditive function

$$\tilde{T}_H(h) := \inf_{t \in H : t^{ab} = h} \mathbb{E} T(1, t)$$

on H^{ab}. (Recall that we can relate functions on H^{ab} and $\mathfrak{g}^{ab}$, since we have a map $H^{ab} \to \mathfrak{g}^{ab}$ and an isomorphism $H^{ab} \otimes \mathbb{R} \cong \mathfrak{g}^{ab}$ induced by the composition

$$H \hookrightarrow G \to G/[G, G] \cong \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}] =: \mathfrak{g}^{ab}.)$$

Thus, to deduce our statement of Theorem 3.1.3, it only remains to show that $\Phi_H = \Phi$, where recall that we define Φ to be the unique norm on $\mathfrak{g}^{ab}$ which is asymptotically equivalent to the subadditive function

$$\tilde{T}(n) := \inf_{t \in N : t^{ab}_{free} = n} \mathbb{E}(1, t)$$

on N^{ab}_{free}.

Proposition 3.9.3. $\Phi_H = \Phi$.

Proof. Note that H^{ab} and N^{ab}_{free} are identified with the same subgroup of $\mathfrak{g}^{ab}$ since the inclusion $H^{ab} \to \mathfrak{g}^{ab}$ is exactly equal to the composition of the isomorphism $H^{ab} \cong N^{ab}_{free}$ and the inclusion $N^{ab}_{free} \to \mathfrak{g}^{ab}$. Using the isomorphism $H^{ab} \cong N^{ab}_{free}$ to consider $\tilde{T}_H$ as a subadditive function on N^{ab}_{free}, we have

$$\tilde{T}_H(n) = \inf_{t \in H : t^{ab}_{free} = n} \mathbb{E}T(1, t).$$

From this it is clear that $\tilde{T} \leq \tilde{T}_H$.

To show a lower bound, first note that since H is finite index in N, $H \cap [\widetilde{N, N}]$ is finite-index in $[\widetilde{N, N}]$. Let R be a finite set of right coset representatives for $H \cap [\widetilde{N, N}]$ in $[\widetilde{N, N}]$, that is, $N \cap [\widetilde{N, N}] = \bigcup_{r \in R} H \cap [\widetilde{N, N}]r$. Set $C := \max_{r \in R} |r|$, where $| \cdot | = d(1, \cdot)$ is, as always, the word length in Γ with respect to the generating set S. Then we have

$$\tilde{T}(n) = \inf_{t \in H, r \in R : t^{ab}_{free} = n} \mathbb{E}T(1, tr) \geq \inf_{t \in H, r \in R : t^{ab}_{free} = n} \mathbb{E}T(1, t) - \mathbb{E}T(1, r) \geq \Phi_H(n) - KC,$$

where we have used that $T \leq Kd$. Thus $|\tilde{T}(n) - \tilde{T}_H(n)| \leq KC = o(n)$ and $\Phi = \Phi_H$, as desired. □

3.10. Gromov-Hausdorff convergence to the limit shape

Recall the notion of pointed Gromov-Hausdorff convergence ([18]). There are many equivalent conditions for this convergence, but here we use a particular sufficient condition. Let $(X_n, d_n, o_n), (X_0, d_0, o_0)$ be metric spaces with distinguished basepoints o_n, o_0. A sequence of maps $f_n : X_n \to X_0$ is called a *sequence of of pointed Gromov-Hausdorff approximations* if for every $\epsilon > 0$, for all sufficiently large n we have

(1) $d_0(f_n(o_n, o_0)) < \epsilon$,

(2) every point of $B(o_0, 1/\epsilon)$ is within distance ϵ of $f_n(B(o_n, 1/\epsilon))$,

(3) $(1-\epsilon)d_n(x,y) - \epsilon \le d_0(f_n(x), f_n(y)) \le (1+\epsilon)d_n(x,y)_\epsilon$ for all $x, y \in B(o_n, 1/\epsilon)$.

If $f_n : X_n \to X_0$ is a sequence of pointed Gromov-Hausdorff approximations, then X_n pointed Gromov-Hausdorff converges to X_0. Here, our metric spaces are groups with various metrics, and the basepoint will always be the identity element.

In [8], Section 4.4, Cantrell and Furman prove the following: for any fixed $g, g' \in G^\infty$, almost surely

(3.10.1)
$$\lim_{\epsilon\to 0}\limsup_{t\to\infty}\sup\left\{\frac{1}{t}|T(\gamma,\gamma') - d_\Phi(g,g')| : \gamma,\gamma' \in \Gamma, d_{\|\cdot\|}(\mathrm{scl}_{\frac{1}{t}}\gamma, g), d_{\|\cdot\|}(\mathrm{scl}_{\frac{1}{t}}\gamma', g') < \epsilon\right\} = 0,$$

where $\Gamma, L_\infty, T, d_\Phi, d_{\|\cdot\|}$ are all as defined in Section 3.1.2, and the maps $\mathrm{scl}_{\frac{1}{t}} : N \to L_\infty$ are as defined in Section 3.8. In particular, $(L_\infty, d_{\|\cdot\|})$ is the scaling limit of Γ endowed with the word metric as given by Pansu's theorem:

Theorem 3.10.1. *(Pansu, [30])*

$$\mathrm{scl}_{\frac{1}{t}} : (\Gamma, \frac{1}{t}d) \to (L_\infty, d_{\|\cdot\|})$$

is a sequence of Gromov-Hausdorff approximations.

To prove that $\mathrm{scl}_{\frac{1}{t}} : (\Gamma, \frac{1}{t}T) \to (L_\infty, d_\Phi)$ is a sequence of Gromov-Hausdorff approximations, by homogeneity of the norm d_Φ, it suffices to show that, for any $\epsilon > 0$, there exists $R > 0$ such that for any $|\gamma|, |\gamma'| \geq R$,

$$|T(\gamma, \gamma') - d_\Phi(\mathrm{scl}_1(\gamma), \mathrm{scl}_1(\gamma'))| \leq \epsilon \max(|\gamma|, |\gamma'|).$$

The rest of this section is devoted to proving this fact.

Remark 3.10.1. *In [8], it is shown that the event of failure of Gromov-Hausdorff convergence is contained in an uncountable union of null-sets. More specifically, they show that failure of Gromov-Hausdorff convergence entails the existence of some pair $g, g' \in L_\infty$ for which Equation* (3.10.1) *fails, but a priori (g, g') ranges over the uncountable set $L_\infty \times L_\infty$. It is necessary to show that it is contained in a* countable *union of null-sets.*

Now, let $\{(g_n, g'_n)\}$ be a countable dense subset of $L_\infty \times L_\infty$. With probability 1, Equation (3.10.1) holds for all (g_n, g'_n) simultaneously. We show that on this probability 1 subset Gromov-Hausdorff convergence holds.

Suppose that Gromov-Hausdorff convergence fails, that is, there exists $\epsilon_0 > 0$ and some sequence $(\gamma_n, \gamma'_n) \in \Gamma \times \Gamma$ with $\min(|\gamma_n|, |\gamma'_n|) \to \infty$ such that

$$\frac{1}{t_n}|T(\gamma, \gamma') - d_\Phi(\mathrm{scl}_1(\gamma), \mathrm{scl}_1(\gamma')| \geq \epsilon_0,$$

where we define $t_n := \max(|\gamma_n|, |\gamma'_n|)$. By homogeneity of d_Φ, this is equivalent to

$$\left|\frac{1}{t_n}T(\gamma, \gamma') - d_\Phi(\mathrm{scl}_{\frac{1}{t_n}}\gamma_n, \mathrm{scl}_{\frac{1}{t_n}}\gamma'_n)\right| \geq \epsilon_0. \tag{3.10.2}$$

Since the sequence $(\mathrm{scl}_{\frac{1}{t_n}}\gamma_n, \mathrm{scl}_{\frac{1}{t_n}}\gamma'_n)$ lies in the product of the unit $d_{\|\cdot\|}$ balls of L_∞, by compactness we may pass to a subsequence and assume that

$$(\mathrm{scl}_{\frac{1}{t_n}}\gamma_n, \mathrm{scl}_{\frac{1}{t_n}}\gamma'_n) \to (g_0, g'_0)$$

for some $(g_0, g'_0) \in L_\infty \times L_\infty$. Convergence holds in the $d_{\|\cdot\|}$ metric as well as the d_Φ metric.

Now choose N sufficiently large so that

$$|d_\Phi(\mathrm{scl}_{\frac{1}{t_n}}\gamma_n, \mathrm{scl}_{\frac{1}{t_n}}\gamma'_n) - d_\Phi(g_0, g'_0)| \leq \frac{\epsilon_0}{2} \tag{3.10.3}$$

for all $n \geq N$. Combining Equations (3.10.2) and (3.10.3) gives

$$|\frac{1}{t_n}T(\gamma_n, \gamma'_n) - d_\Phi(g_0, g'_0)| \geq \frac{\epsilon_0}{2}. \tag{3.10.4}$$

Fix $\delta' > 0$ (to be chosen later). Now choose (g_{m_0}, g'_{m_0}) from our countable dense set such that

$$\max(d_{\|\cdot\|}(g_{m_0}, g_0), d_{\|\cdot\|}(g'_{m_0}, g'_0), d_\Phi(g_{m_0}, g_0), d_\Phi(g'_{m_0}, g'_0)) \leq \delta'.$$

For each $k \geq 1$ define $\gamma^k_{m_0}$ to be the $\gamma \in \Gamma$ such that $\mathrm{scl}_{\frac{1}{k}}$ has minimal distance to g_{m_0}, and similarly define $\gamma'^k_{m_0}$. Then by Equation (3.10.1) we have

$$\left|\frac{1}{k}T(\gamma^k_{m_0}, \gamma'^k_{m_0}) - d_\Phi(g_{m_0}, g'_{m_0})\right| \xrightarrow[k\to\infty]{} 0,$$

and so we can choose N also sufficiently large that for all $n \geq N$,

$$\left|\frac{1}{t_n}T(\gamma^{t_n}_{m_0}, \gamma'^{t_n}_{m_0}) - d_\Phi(g_{m_0}, g'_{m_0})\right| \leq \delta'.$$

By Theorem 3.10.1 we can also choose N so that for all $n \geq N$,

$$\left|\frac{1}{t_n} d(\gamma_n, \gamma_{m_0}^{t_n}) - d_{\|\cdot\|}(g_0, g_{m_0})\right| \leq \delta',$$

$$\left|\frac{1}{t_n} d(\gamma'_n, \gamma'^{t_n}_{m_0}) - d_{\|\cdot\|}(g'_0, g'_{m_0})\right| \leq \delta'.$$

Thus we have (again taking $k = \max(|\gamma_n|, |\gamma'_n|)$)

$$\begin{aligned}\left|\frac{1}{t_n} T(\gamma_n, \gamma'_n) - d_\Phi(g_0, g'_0)\right| \leq & \left|\frac{T(\gamma_n, \gamma'_n) - T(\gamma_{m_0}^{t_n}, \gamma'^{t_n}_{m_0})}{t_n}\right| \\ & + \left|\frac{1}{t_n} T(\gamma_{m_0}^{t_n}, \gamma'^{t_n}_{m_0}) - d_\Phi(g_{m_0}, g'_{m_0})\right| \\ & + |d_\Phi(g_{m_0}, g'_{m_0}) - d_\Phi(g_0, g'_0)|.\end{aligned}$$

By our choice of (g_{m_0}, g'_{m_0}), we have that the last term is bounded by 2δ. If $n \geq N$, we have that the second term is bounded by δ. To bound the first term, recall that by assumption, $T \leq Kd$ and hence

$$|T(\gamma_n, \gamma'_n) - T(\gamma_{m_0}^{t_n}, \gamma'^{t_n}_{m_0})| \leq T(\gamma_n, \gamma_{m_0}^{t_n}) + T(\gamma'_n, \gamma'^{t_n}_{m_0}) \leq K(d(\gamma_n, \gamma_{m_0}^{t_n}) + d(\gamma'_n, \gamma'^{t_n}_{m_0})),$$

and so

$$\begin{aligned}\left|\frac{T(\gamma_n, \gamma'_n) - T(\gamma_{m_0}^{t_n}, \gamma'^{t_n}_{m_0})}{t_n}\right| & \leq K\left(\frac{1}{t_n} d(\gamma_n, \gamma_{m_0}^{t_n}) + \frac{1}{t_n} d(\gamma'_n, \gamma'^{t_n}_{m_0})\right) \\ & \leq K\left(d_{\|\cdot\|}(g_0, g_{m_0}) + \delta + d_{\|\cdot\|}(g'_0, g'_{m_0}) + \delta\right) \leq 4K\delta.\end{aligned}$$

All in all we have

$$\left|\frac{1}{t_n} T(\gamma_n, \gamma'_n) - d_\Phi(g_0, g'_0)\right| \leq 4K\delta + 3\delta,$$

and for a sufficiently small choice of δ, this contradicts Equation (3.10.4), and so we are done.

CHAPTER 4

Strict monotonicity for independent first passage percolation

4.1. Introduction

Recall from the last chapter that if T is an independent FPP metric on $G = (V, E)$ a Cayley graph of a virtually nilpotent group, then under mild assumptions, the random metric space (V, T) has a deterministic scaling limit (L_∞, d_Φ), which coincides with the scaling limit of the deterministic metric space $(V, \mathbb{E}T)$. However, it is very difficult to understand the relationship between the weight measure ν and the limit metric d_Φ for nontrivial ν, since the proofs of existence of such scaling limits depend on ergodic theorems and are not constructive. Therefore, instead of trying to determine d_Φ explicitly for a given ν, in this chapter we try to understand the map $\nu \mapsto d_\Phi$ by asking whether it is in some sense *strictly monotonic*.

In the classical case of the standard Cayley graph of $\mathbb{Z}^d$, recall that the limiting space is $(\mathbb{R}^d, \mu)$, where μ is a norm defined by the "time constants"

$$\mu_v := \lim_{n\to\infty} \frac{T(0, nv)}{n} = \lim_{n\to\infty} \frac{\mathbb{E}T(0, nv)}{n}.$$

In this setting, van den Berg and Kesten [36] proved a strict monotonicity theorem. That is, they proved that if a probability measure $\tilde{\nu}$ is "strictly more variable" than a probability measure ν, both have finite mean, and ν is subcritical (in a sense to be described later), then for G the standard Cayley graph of $\mathbb{Z}^d$, $d \geq 2$, we have a strict inequality of time constants $\tilde{\mu}_v < \mu_v$ for all

$\nu \neq 0$. In fact, their proof shows that

$$\liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y) - \mathbb{E}\tilde{T}(x,y)}{d(x,y)} > 0,$$

where $d(x,y) := \sum_{i=1}^{d} |x_i - y_i|$ is the graph distance between the vertices $x, y \in \mathbb{Z}^d$ in the standard Cayley graph of $\mathbb{Z}^d$. If one showed this inequality above for Cayley graphs of virtually nilpotent groups, it would translate to an analogous strict inequality of CC-metrics $d_{\tilde{\Phi}}(\xi) < d_{\Phi}(\xi)$ for all $\xi \neq 1$. But note that this inequality also makes sense for any graph; one does not require the existence of any scaling limits or even "time constants."[1] We will often abbreviate the above inequality as $\mathbb{E}\tilde{T} \ll \mathbb{E}T$. This naturally raises the question: for which other graphs G does the same conclusion hold?

More explicitly, we call a measure ν *exponential-subcritical* if the mass it assigns to the infimum of its support is less than a certain threshold, as defined in Section 4.2.[2] A measure $\tilde{\nu}$ is *strictly more variable* than ν if $\tilde{\nu} \neq \nu$ and $\int f d\tilde{\nu} \leq \int f d\nu$ for any concave nondecreasing f such that both integrals converge absolutely. We say that a graph *has the van den Berg-Kesten (vdBK) property* if the following holds: For any $\nu, \tilde{\nu}$ with finite mean, if ν is exponential subcritical and $\tilde{\nu}$ is strictly more variable than ν, then $\mathbb{E}\tilde{T} \ll \mathbb{E}T$. The main theorem of [36] is precisely the statement that the standard Cayley graph of $\mathbb{Z}^d, d \geq 2$ is vdBK. In this chapter we explore the question of which other graphs are vdBK.

The first main theorem of this chapter extends the strict monotonicity theorem of van den Berg and Kesten to the setting of Cayley graphs of virtually nilpotent groups:

---

[1]Note for future reference that $\liminf_{d(x,y)\to\infty} = \lim_{R\to\infty} \inf_{(x,y)\in V^2 : d(x,y)\geq R}$. In particular, neither x nor y is fixed, which is a relevant point when considering inhomogeneous graphs.

[2]The extent to which such an assumption on ν is necessary for strict monotonicity is discussed in Section 4.2.1 as well as Chapter 5.

Theorem 4.1.1. *Let G be any Cayley graph of a finitely generated virtually nilpotent group which is not isomorphic as a graph to the standard Cayley graph of $\mathbb{Z}$. Then G has the vdBK property.*

A way of rephrasing this result (using Gromov's theorem [16]) is that any Cayley graph which has polynomial growth is either isomorphic to the standard Cayley graph of $\mathbb{Z}$ or has the vdBK property. Moreover, a theorem of Trofimov [35] establishes a close relationship between transitive graphs (graphs G such that $\mathrm{Aut}(G)$ acts transitively on the vertex set V) of polynomial growth and virtually nilpotent Cayley graphs, so this result goes a long way towards resolving the question of which transitive graphs of polynomial growth are vdBK.

One might wonder if *all* graphs have the vdBK property. This is not true; the easiest counterexample is when the graph G is a tree. In this case, since there is only one self-avoiding path between any two points, we have that $\frac{\mathbb{E}T(x,y)}{d(x,y)}$ is a *constant* equal to the mean of ν. It is easy to produce two different probability measures $\nu, \tilde{\nu}$ with the same mean such that $\tilde{\nu}$ is more variable than ν (see the proof of Theorem 4.4.1). This is, of course, why the standard Cayley graph of $\mathbb{Z}$ had to be excluded from the above theorem.

However, trees are not the only counterexample. Consider the Cayley graph of the free group $F(a, b)$ on the two letters a, b which is associated to the [redundant] generating set $\{a, b, ab\}$ (see Figure 4.1). It is not hard to see that, although there is more than one self-avoiding path between any two points, each self-avoiding path between two points must pass through every vertex of the edge-geodesic path between those two points, and that each step, one only has the choice to travel along the edge lying in the geodesic, or to take a particular path of length two. Hence, in this case $\frac{\mathbb{E}T(x,y)}{d(x,y)}$ is a constant given by $\mathbb{E}\min(w_1, w_2 + w_3)$, where w_1, w_2, w_3 are independent variables with distribution ν. One can again produce two distinct distributions, one

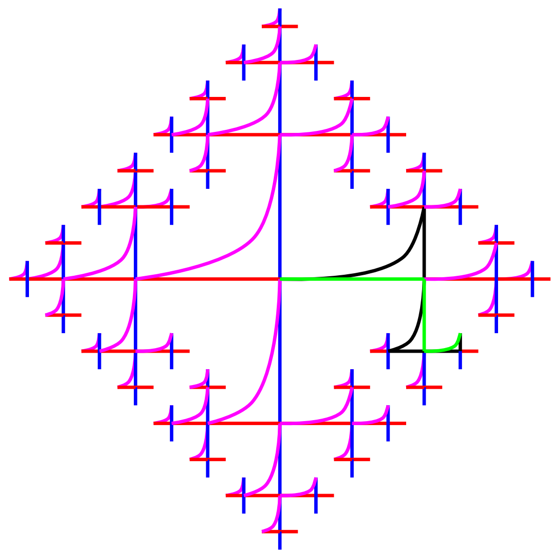

Figure 4.1. The Cayley graph of the free group $F(a,b)$ with respect to the generating set {a,b,ab}. In green is the unique edge-geodesic path from 1 to $ab^{-1}ab$. Every self-avoiding path from 1 to $ab^{-1}ab$ in this graph must visit all the vertices of the green path and can only use green or black edges.

more variable than the other, such that their "time constants" are equal, contradicting the vdBK property.

It turns out that the crucial property for determining whether a graph is vdBK is a property of the graph which we call "admitting detours" (defined in Section 4.4). If a bounded degree graph does not admit detours, then it is not vdBK, by Theorem 4.4.1 below. (One should note, however, that if the graph in question has few symmetries, the vdBK property is quite a strong property; this is discussed in more detail in Section 4.4.1). On the other hand, our main theorems prove that given one of two quite different "large scale" assumptions on the geometry of G, admitting detours *implies* the vdBK property.

For the first coarse-geometric setting, we say that a graph has *strict polynomial growth* if the number of vertices in a ball of radius R in the graph metric is bounded above and below by polynomials in R of the same degree. We have:

Theorem 4.1.2. *Let G be a graph of strict polynomial growth. Then G is vdBK if and only if G admits detours.*

To explain the second setting, we say that a (set) map $f : X \to Y$ between metric spaces is a *quasi-isometry* if there exists $0 < C < \infty$ such that $\frac{1}{C}d_X(a,b) - C \leq d_Y(f(a), f(b)) \leq Cd_X(a,b) + C$ for all $a, b \in X$ and $d_Y(y, f(X)) \leq C$ for all $y \in Y$. If a quasi-isometry $f : X \to Y$ exists, then X and Y are said to be *quasi-isometric*; this is an equivalence relation on metric spaces. When we say that two graphs are quasi-isometric, we mean that the graph metrics on their vertex sets are. We then have:

Theorem 4.1.3. *Let G be a bounded degree graph which is quasi-isometric to a tree. Then G is vdBK if and only if G admits detours. In fact, if G admits detours, then even if ν is not exponential-subcritical, whenever $\tilde{\nu}$ is strictly more variable than ν, we have $\mathbb{E}\tilde{T} \ll \mathbb{E}T$.*

Theorem 4.1.2 is used to prove Theorem 4.1.1: the latter follows from the former once we prove that all Cayley graphs of virtually nilpotent groups which are not isomorphic to the standard Cayley graph of $\mathbb{Z}$ admit detours, which is proven in Section 4.5. Theorem 4.1.3 can be combined with results proved in Section 4.5 to give many examples of Cayley graphs which are quasi-isometric to trees and are vdBK:

Theorem 4.1.4. *Let F_k be the free group on $k \geq 1$ letters, and let F be a nontrivial finite group. If Γ contains $F \times F_k$ as a finite-index subgroup or if Γ is a semidirect product $F \rtimes F_k$, then any Cayley graph of Γ is vdBK. More generally, if Γ has a finite-index free subgroup, is not isomorphic to $\mathbb{Z}$ or $\mathbb{Z}/2 * \mathbb{Z}/2$, and either contains a finite index subgroup with nontrivial center or contains a nontrivial finite normal subgroup, then any Cayley graph of Γ is vdBK.*

Finally, as noted in [36], the question of strict monotonicity with respect to *stochastic domination* is related to "absolute continuity with respect to the expected empirical measure." What precisely we mean by this is explained in Section 4.8; note that this condition does not imply *existence* of a limiting expected empirical measure. In any case, the methods of our chapter easily prove absolute continuity of the weight distribution with respect to the expected empirical measure under the same "large-scale" assumptions:

Theorem 4.1.5. *Let G be a bounded degree graph which is quasi-isometric to a tree. Then for any probability measure ν on $[0, \infty)$ with finite mean, ν is absolutely continuous with respect to the expected empirical measure of the associated first passage percolation T. Moreover, if ν strictly stochastically dominates a measure $\tilde{\nu}$, then $\mathbb{E}\tilde{T} \ll \mathbb{E}T$.*

Theorem 4.1.6. *Let G be a graph of strict polynomial growth. Suppose that ν has finite mean and is exponential-subcritical. Then ν is absolutely continuous with respect to the expected empirical measure of the associated first passage percolation T. Moreover, if ν strictly stochastically dominates a measure $\tilde{\nu}$, then $\mathbb{E}\tilde{T} \ll \mathbb{E}T$.*

Note that these strict monotonicity theorems for stochastic domination hold whether or not the graph in question admits detours.

The layout of the chapter is as follows: in Section 4.2 we establish definitions and notations. In Section 4.3 we collect various lemmata, most of which are essentially proven in [36], which we will need to prove our main theorems. The key conclusion of this section is that, in order to prove that $\mathbb{E}\tilde{T} \ll \mathbb{E}T$, it suffices to show that the expected number of times the T-geodesic between two points x and y passes through a certain type of configuration called a "feasible pair" is *linear* in $d(x, y)$. In Section 4.4 we introduce the concept of "admitting detours" and show that this is a necessary condition for a graph to be vdBK (although we note that the vdBK property is quite a strong one in the inhomogeneous case). We also give examples of graphs which admit detours. In Section 4.5 we prove sufficient conditions for Cayley graphs to admit detours, particularly those that we will need to prove Theorem 4.1.1. This section is entirely group-theoretical and combinatorial, so some readers may choose to skip it on first reading. In Section 4.6 we prove Theorem 4.1.3, which will follow almost immediately from the results of Section 4.3 combined with a characterization of graphs quasi-isometric to trees. Because paths are so constrained in this setting, it is not hard to produce local events which imply that the T-geodesic from x to y passes through a feasible pair, and this makes the proof quite simple. We also prove Theorem 4.1.4 as corollary.

In Section 4.7 we prove Theorem 4.1.2, which is much more involved. The three key components are a Peierls-type lemma, a resampling argument, and a "geometric construction" (a construction of a set of weights suitable for use in the resampling argument). Although this general strategy is the same as in [36], the methods given here apply to general graphs of strict polynomial growth which are not necessarily almost-transitive. The geometric constructions in particular are rather different from those of [36] and are quite intricate, since we are given the task of manipulating the geodesic while remaining largely agnostic to the fine geometry of the graph. By far, these geometric constructions are the most involved part of the proof of Theorem 4.1.2. At the end of this section we give some examples of graphs which are not almost-transitive to which our results apply.

Lastly, in Section 4.8 we prove absolute continuity with respect to the expected empirical measure for graphs of strict polynomial growth and for graphs quasi-isometric to trees, which implies a strict monotonicity theorem with respect to stochastic domination, regardless of whether the graph in question admits detours. The proofs are just easier versions of the proofs of the main theorems of this chapter.

4.2. Variability, subcriticality, and the van den Berg–Kesten property

Let ν and $\tilde{\nu}$ be two probability measures on $[0, \infty)$. We say that $\tilde{\nu}$ is *more variable* than ν if for every concave nondecreasing function $f : \mathbb{R} \to \mathbb{R}$ we have

$$\int f d\tilde{\nu} \leq \int f d\nu$$

as long as both integrals converge absolutely. We say that $\tilde{\nu}$ is *strictly more variable* than ν if $\tilde{\nu}$ is more variable than ν and $\tilde{\nu} \neq \nu$.

We now define some percolation thresholds associated to a graph G. For $p \in [0, 1]$, denote by G_p the random subgraph of G given by including each edge $e \in E(G)$ in G_p independently with probability p, excluding with probability $1 - p$. We define the *exponential percolation threshold* for G to be

$$\underline{p_c} := \sup \left\{ p \in [0, 1] : \limsup_{R \to \infty} \sup_{o \in V} \frac{1}{R} \log \mathbb{P}(G_p \text{ contains an edge path from } o \text{ to } B_G(o, R)^c) < 0 \right\}$$

and we define the *exponential geodesic percolation threshold* for G to be

$$\underline{\vec{p_c}} := \sup \left\{ p \in [0, 1] : \limsup_{R \to \infty} \sup_{o \in V} \frac{1}{R} \log \mathbb{P} \left(\begin{array}{c} G_p \text{ contains an edge path from } o \text{ to} \\ B_G(o, R)^c \text{ which is edge-geodesic in } G \end{array} \right) < 0 \right\}.$$

Below the exponential percolation thresholds, we have uniform exponential upper bounds on connection events. We call a measure ν *exponential-subcritical* if either $\inf := \inf \operatorname{supp} \nu = 0$ and $\nu(\{0\}) < \underline{p_c}$ or $\inf > 0$ and $\nu(\{\inf\}) < \underline{\vec{p_c}}$.

Definition 1. *We say that an infinite graph G* has the van den Berg–Kesten (vdBK) property *if for every $\nu, \tilde{\nu}$ with finite mean such that ν is exponential-subcritical and $\tilde{\nu}$ is strictly more variable than ν, we have*

$$\liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y) - \mathbb{E}\tilde{T}(x,y)}{d(x,y)} > 0. \tag{4.2.1}$$

We will often abbreviate the "asymptotic strict inequality" (4.2.1) as $\mathbb{E}\tilde{T} \ll \mathbb{E}T$. The main theorems of this chapter give sufficient or necessary conditions for a graph to be vdBK.

4.2.1. Remarks on the condition of exponential subcriticality

Here are some remarks which, while not necessary to the proofs below, are worth noting, on the condition of exponential subcriticality, its relationship to various other percolation thresholds, and the extent to which it is a necessary assumption to get a strict monotonicity result.

First, it is clear from the definitions that for any graph, $\underline{\vec{p_c}} \geq \underline{p_c}$, and a simple union bound (counting self-avoiding paths from a fixed vertex) shows that if G has degree at most D, then $\underline{p_c} \geq 1/D > 0$. It is also clear that for any connected graph, $\underline{p_c} \leq p_c$, where p_c is the percolation threshold as usually defined:

$$p_c := \inf\{p \in [0,1] : \mathbb{P}(G_p \text{ contains an infinite edge path from } o) > 0\}.$$

For almost-transitive graphs, the sharpness of the percolation threshold [12] shown by Duminil-Copin and Tassion implies that $\underline{p_c} = p_c$. (The proof of sharpness in [12] is stated for transitive graphs, but is not hard to generalize to almost-transitive graphs; here by *almost-transitive graph* we mean a graph G such that the action of $\mathrm{Aut}(G)$ on V has finitely many orbits.)

Furthermore, on amenable almost-transitive graphs (in particular graphs of polynomial growth), the original argument of Burton-Keane ([7], see also [19] for an explicitly general proof) shows that $p_c = p_u$, where p_u is the *uniqueness threshold*

$$p_u := \inf\{p \in [0,1] : \mathbb{P}(G_p \text{ contains a unique infinite connected component }) = 1\}.$$

If $\nu(\{0\}) \geq p_u$, then one expects that $\lim_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y)}{d(x,y)} = 0$ (although this has only been proven in certain cases, see e.g. Theorem 6.1 of [26]). In that case, it is impossible that $\mathbb{E}\tilde{T} \ll \mathbb{E}T$, so for polynomial growth almost-transitive graphs, the assumption on the atom at 0 is really as weak as one could hope for.

In the case of the standard Cayley graph of $\mathbb{Z}^d$, $\underline{\vec{p_c}}$ is the classical oriented percolation threshold $\vec{p_c}$; this is because of the nature of edge-geodesics in this graph, combined with the sharpness results of Aizenman and Barsky [1]. In fact, if G is the standard Cayley graph of $\mathbb{Z}^d$, the condition here of being exponential-subcritical is precisely the condition of being "useful" in [36]. Furthermore, in this case, if $\inf \operatorname{supp} \nu =: \inf > 0$ and $\nu(\{\inf\}) \geq \vec{p_c}$, then $\lim_{n\to\infty} \frac{\mathbb{E}T(0,(n,\ldots,n))}{dn} = \inf$ [13, 29]. So if $\nu \neq \delta_{\inf}$, we can take $\tilde{\nu} = \delta_{\inf}$ to get $\tilde{\nu}$ strictly more variable than ν but $\mathbb{E}\tilde{T} \not\ll \mathbb{E}T$. Thus, the assumption $\nu(\{\inf\}) < \underline{\vec{p_c}}$ in the definition of the vdBK property is also necessary at least in this setting.[3]

In fact, similar behavior happens more generally, i.e. if G is transitive graph of polynomial growth, one has $\liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y)}{d(x,y)} = \inf$ whenever $\inf > 0$ and $\nu(\{\inf\}) > \underline{\vec{p_c}}$. This shows that the subcriticality assumption on ν is needed for strict monotonicity in the exact same way as above. This is proven in Chapter 5.

[3]Of course, this is because $\mathbb{E}\tilde{T} \ll \mathbb{E}T$ is equivalent to a strict inequality of time constants *in all directions simultaneously*; if one instead only cares about strict inequality of a time constant in a fixed direction, the assumption $\nu(a) < \underline{\vec{p_c}}$ may not be necessary; for instance Marchand [29] proved that for G the standard Cayley graph of $\mathbb{Z}^2$, we get strict inequality in the e_1 direction without that assumption (as long as still $\nu(\{0\}) < p_c$).

On the other hand, for graphs quasi-isometric to a tree, we will see in Theorem 4.1.3 that exponential subcriticality as defined here is not necessary at all. In fact, in this setting, if G admits detours (see Section 4.4), then $\mathbb{E}\tilde{T} \ll \mathbb{E}T$ whenever $\tilde{\nu}$ is strictly more variable than ν, with no further assumptions needed on either measure. This is consistent with the perspective that generally the *uniqueness* threshold, rather than p_c, is the correct threshold to consider for the atom at 0, since almost-transitive graphs quasi-isometric to trees can have $p_c < 1$ but always have $p_u = 1$ (since they have more than one end, see page 86 of [21]). The proper "uniqueness" analogue of $\underline{\vec{p_c}}$ outside of the amenable case is unclear.

Finally, percolation on graphs which are not almost-transitive is poorly understood, and so it is entirely unclear how close exponential subcriticality is to the "right" condition on ν to consider in this general setting. However, if G has degree at most D then the inequalities $\underline{\vec{p_c}} \geq \underline{p_c} \geq 1/D > 0$ tell us that our main theorems are never vacuous for bounded degree graphs; in particular, we get sufficient conditions to conclude strict monotonicity (e.g. that ν is atomless), even if the parameters $\underline{\vec{p_c}}$ and $\underline{p_c}$ are quite mysterious.

4.3. Reduction to a lower bound on expected number of traversed "feasible pairs"

In this section, we reduce the task of deducing a strict inequality $\mathbb{E}\tilde{T} \ll \mathbb{E}T$ to the task of showing that the T-geodesic traverses linearly many "feasible pairs" in expectation. Most of the argument from this section can be transferred directly from [36], with the main difference being that we define a weaker notion of "feasible pair." Proofs are given where the necessary modifications from [36] are not obvious. Note that the arguments of this section allow us to stop considering $\tilde{w}$ or $\tilde{T}$ and simply focus on understanding the T-geodesic.

First, note that we have the following theorem of van den Berg and Kesten:

Theorem 4.3.1 ([36], Theorem 2.9a)**.** *Let ν and $\tilde{\nu}$ be probability measures on* $[0, \infty)$ *with finite mean such that $\tilde{\nu}$ is more variable than ν. Then for all $x, y \in V$*

$$\mathbb{E}\tilde{T}(x, y) \leq \mathbb{E}T(x, y).$$

Although the proof in [36] is stated only for $G = \mathbb{Z}^d$, it easily extends to all locally finite graphs.

We also have

Theorem 4.3.2 ([32, 37])**.** *Let ν and $\tilde{\nu}$ be probability measures on* $[0, \infty)$ *with finite mean such that $\tilde{\nu}$ is strictly more variable than ν. Then there exists a coupling* $(w(e), \tilde{w}(e))$ *such that $w(e)$ is ν-distributed, $\tilde{w}(e)$ is $\tilde{\nu}$-distributed, and*

$$\mathbb{E}[\tilde{w}(e)|w(e)] \leq w(e)$$

almost surely.

Another lemma from [36] which we will need is the following:

Lemma 4.3.1 ([36], Lemma 4.5)**.** *We may assume without loss of generality that in our coupling*

$$\mathbb{P}(\tilde{w}(e) > w(e)) > 0. \tag{4.3.1}$$

Explicitly, either this holds, or there exists some $\bar{w}(e)$ such that $\mathbb{E}\tilde{T}(x, y) \leq \mathbb{E}\bar{T}(x, y)$ for all $x, y \in V$, the distribution of $\bar{w}(e)$ is strictly more variable than ν, and such that (4.3.1) *holds with $\tilde{w}$ replaced by $\bar{w}$, i.e. $\mathbb{P}(\bar{w}(e) > w(e)) > 0$.*

A key technical lemma we will use is the following:

Lemma 4.3.2 ([36], Lemma 4.8). *Let ν, $\tilde{\nu}$ be probability measures with finite mean such that $\tilde{\nu}$ is more variable than ν and such that* (4.3.1) *holds. Then there exist $\epsilon > 0$, $a > 0$, $b > 0$, $g > 0$, and a bounded Borel set $I_0 \subset [0, \infty)$ and $y_0 \in I_0$ with the following properties:*

- *For all $\delta > 0$, $\nu(I_0 \cap (y_0 - \delta, y_0 + \delta)) > 0$.*
- *For all $y \in I_0$,*

$$\mathbb{P}(\tilde{w}(e) > y + a | w(e) = y) \geq b.$$

- *For any $k \geq 1$ and any $y_1, ..., y_k, y'_1, ..., y'_{\lfloor (1+\epsilon)k \rfloor} \in I_0$, we have*

$$\sum_{i=1}^{k} (y_i + a) - \sum_{i=1}^{\lfloor (1+\epsilon)k \rfloor} y'_i > kg \geq g.$$

Proof. The proof is essentially the same as that given in [36], but simpler since we do not actually need as many conditions. By (4.3.1), for some sufficiently small $a, b > 0$ there is some Borel set $B \subset [0, \infty)$ such that $\nu(B) > 0$ and for all $y \in B$,

$$\mathbb{P}(\tilde{w}(e) > y + a | w(e) = y) \geq b.$$

Let y_0 be a point of support for B, that is, a point such that $\nu(B \cap (y_0 - \delta, y_0 + \delta)) > 0$ for all $\delta > 0$. Choose $\epsilon > 0$ sufficiently small such that $\epsilon y_0 < a$. Then choose $\delta_0 > 0$ sufficiently small that

$$\epsilon y_0 + 2\delta_0 + \epsilon\delta_0 < a,$$

and choose

$$0 < g < a - (\epsilon y_0 + 2\delta_0 + \epsilon\delta_0).$$

Then we can take $I_0 := B \cap (y - \delta_0, y + \delta_0)$. The first two conditions clearly hold by construction; let us show the last condition:

$$\sum_{i=1}^{k}(y_i + a) - \sum_{i=1}^{\lfloor(1+\epsilon)k\rfloor} y_i' \geq k(y_0 - \delta_0 + a) - k(1 + \epsilon)(y_0 + \delta_0)$$
$$= k(a - 2\delta_0 - \epsilon y_0 - \epsilon\delta_0) > kg \geq g.$$

□

Definition 2. *Let π, π' be a pair of paths with the same starting and ending point such that $\pi \neq \pi'$ (as edge sets). We say that π' is a ϵ-*detour *for π if*

$$|\pi' \setminus \pi| \leq (1 + \epsilon)|\pi \setminus \pi'|.$$

Here, π' and π are identified with the sets of edges they contain, $\setminus$ denotes set difference, and $|\cdot|$ is the cardinality of a set.

Note that the condition that $\pi' \neq \pi$ implies that $|\pi \setminus \pi'| \geq 1$. For otherwise we would have that $|\pi' \setminus \pi| \leq (1 + \epsilon)|\pi \setminus \pi'| = (1 + \epsilon) \cdot 0 = 0$, that is, $|\pi' \setminus \pi| = |\pi \setminus \pi'| = 0$ and hence $\pi' = \pi$. Intuitively, one should think of π' as an "alternate path" which misses a fair number of edges of the original path but is not much longer than the original path. A simple but useful observation is:

Proposition 4.3.1. *π' is an ϵ-detour for π if and only if π and π' have the same endpoints, $\pi \neq \pi'$, and*

$$|\pi'| - |\pi| \leq \epsilon|\pi \setminus \pi'|.$$

Proof. This follows immediately from the facts that $|\pi \setminus \pi'| = |\pi| - |\pi \cap \pi'|$ and $|\pi' \setminus \pi| = |\pi'| - |\pi \cap \pi'|$, together with some algebraic manipulation. □

Definition 3. *Let ϵ and I_0 all be as in Lemma 4.3.2, and let C be a constant. We call (α, γ) a* feasible pair *with respect to the T-geodesic*[4] *π from x to y if both α and γ are self-avoiding, γ is an ϵ-detour for α of length at most $C(1+\epsilon)$, α is a subpath of π, and for all $e \in (\alpha \cup \gamma) \setminus (\alpha \cap \gamma)$, $w(e) \in I_0$.*

This notion of course depends on C, ϵ, and I_0 even though this is suppressed in the notation. Here C is an unspecified constant, but in practice there will be one particular $C = C(\epsilon)$ that we end up using. These detours turn out to be key to proving the strict inequalities we want to show, as we shall see in the next lemma.

The following lemma is essentially contained within Lemma 5.19 and the proof of Theorem 2.9(b) from Proposition 5.22 in [36], but we write it here to be explicit about what the necessary modifications are.

Lemma 4.3.3. *Let $\nu, \tilde{\nu}$ have finite mean and be such that $\tilde{\nu}$ is strictly more variable than ν and* (4.3.1) *holds. Let ϵ and I_0 be given as in Lemma 4.3.2 and let C be fixed. Then there exists some constant $c_0 > 0$ such that if $G = (V, E)$ is a graph and $\{B_i\}_{i \in I} \subset E$ is a family of disjoint*

[4] A T-geodesic from x to y is a path $\pi : x \to y$ with $T(\pi) = T(x, y)$. In general there may be more than one T-geodesic; we implicitly fix an arbitrary well-ordering on self-avoiding paths in G and define "the" T-geodesic π from x to y to be the T-geodesic which is least in this ordering. On the other hand, it is not a priori obvious that a T-geodesic exists. If $\nu(\{0\}) < p_c(G)$ and G is locally finite, then it is easily shown (see Proposition 4.4 in [2]) that all pairs of points $x, y \in V$ admit a T-geodesic; in particular, if ν is exponential-subcritical, T-geodesics exist. In Section 4.6 we do not assume that ν is exponential-subcritical, but the arguments are easily modified to avoid the assumption that T-geodesics exist, by considering paths $\pi : x \to y$ with $T(\pi) \leq T(x, y) + \epsilon$ and letting $\epsilon \to 0$.

subsets, for any $x, y \in V$ we have

$$\mathbb{E}T(x,y) - \mathbb{E}\tilde{T}(x,y) \geq c_0 \sum_{i \in I} \mathbb{P}(B_i \text{ contains a feasible pair for the } T\text{-geodesic } \pi : x \to y).$$

Proof. As in [36], let $\hat{w} : E \to [0,\infty)$ be given by

$$\hat{w}(e) := \mathbb{1}_{\xi(e)=0} w(e) + \mathbb{1}_{\xi(e)=1} \tilde{w}(e),$$

where $\{\xi(e)\}_{e \in E}$ is a family of i.i.d. Unif($\{0,1\}$) variables, also independent of $w, \tilde{w}$. As shown in Lemma 5.19 of [36], $\mathbb{E}\tilde{T}(x,y) \leq \mathbb{E}\hat{T}(x,y)$ for all $x, y \in V$, so it suffices to show the desired inequality with $\tilde{T}$ replaced by $\hat{T}$.

We will call a pair (α, γ) *advantageous* for the T-geodesic $\pi : x \to y$ if it is feasible for π and furthermore $\xi(e) = 0$ for all $e \in \gamma \setminus \alpha$, $\xi(e) = 1$ for all e in a subset $S \subset \alpha \setminus \gamma$ of size at least $|S| \geq \frac{1}{1+\epsilon}|\gamma \setminus \alpha|$, and if for all $e \in S$ we have $\tilde{w}(e) > w(e) + a$, where $a > 0$ is as given in Lemma 4.3.2. Note that, by Lemma 4.3.2, if (α, γ) is advantageous then

$$\hat{T}(\alpha) - \hat{T}(\gamma) = \hat{T}(\alpha \setminus \gamma) - \hat{T}(\gamma \setminus \alpha) \geq \tilde{T}(S) - T(\gamma \setminus \alpha) \geq g,$$

where $g > 0$ is as in the lemma. Furthermore, for any pair (α, γ) we have

$$\mathbb{E}\left[\mathbb{1}_{\{(\alpha,\gamma) \text{ is advantageous}\}} \middle| w\right] \geq 2^{-(C(1+\epsilon)+C)} b^C \mathbb{1}_{\{(\alpha,\gamma) \text{ is feasible}\}}.$$

(Here we have used that $|\gamma| \leq C(1+\epsilon)$).

Therefore, consider a T-geodesic π from x to y. Construct another (random) path π' by starting with π and, for each B_i, if B_i contains an advantageous pair (α_i, γ_i) for π, replacing the subsegment α_i with γ_i. (If B_i contains more than one advantageous pair, choose the least one in some arbitrary ordering). We then have

$$\hat{T}(\pi) - \hat{T}(\pi') \geq g \sum_i \mathbb{1}_{\{B_i \text{ contains an advantageous pair for } \pi\}}.$$

Since, as shown in Lemma 5.19 of [36], $\mathbb{E}T(\pi) \geq \mathbb{E}\hat{T}(\pi)$, we have

$$\mathbb{E}T(x,y) - \mathbb{E}\hat{T}(x,y) \geq \mathbb{E}\hat{T}(\pi) - \mathbb{E}\hat{T}(\pi') \geq g \sum_{i \in I} \mathbb{P}(B_i \text{ contains an advantageous pair for } \pi).$$

But (again using some fixed ordering on pairs inside B_i) we have

$$\begin{aligned}
&\mathbb{P}(B_i \text{ contains an advantageous pair for } \pi) \\
\geq &\sum_{(\alpha,\gamma)\subset B_i} \mathbb{E}\left[\mathbb{1}_{\{(\alpha,\gamma) \text{ is the least pair in } B_i \text{ which is feasible}\}} \mathbb{1}_{\{(\alpha,\gamma) \text{ is advantageous}\}}\right] \\
= &\sum_{(\alpha,\gamma)\subset B_i} \mathbb{E}\left[\mathbb{1}_{\{(\alpha,\gamma) \text{ is the least pair in } B_i \text{ which is feasible}\}} \mathbb{E}[\mathbb{1}_{\{(\alpha,\gamma) \text{ is advantageous}\}} | w]\right] \\
\geq &2^{-(C(1+\epsilon)+C)} b^C \sum_{(\alpha,\gamma)\subset B_i} \mathbb{E}\left[\mathbb{1}_{\{(\alpha,\gamma) \text{ is the least pair in } B_i \text{ which is feasible}\}}\right] \\
= &2^{-(C(1+\epsilon)+C)} b^C \mathbb{P}(B_i \text{ contains a feasible pair for } \pi).
\end{aligned}$$

Thus we have the lemma with $c_0 := g \cdot 2^{-(C(1+\epsilon)+C)} b^C > 0$. □

Inequalities up to a constant factor will appear many times in this chapter, so from here we fix the following notation. For two functions f and g of a parameter t, we will write $f(t) \lesssim g(t)$ or $g(t) \gtrsim f(t)$ if there is some constant $c > 0$ and $t_0 < \infty$ such that $f(t) \leq cg(t)$ for all $t \geq t_0$. In this chapter our parameter t is typically either $d(x,y)$ or R, and which it is should be clear from context.

Finally, we can weaken the disjointness assumption on the $\{B_i\}$ in Lemma 4.3.3 and obtain the following key lemma (where the same hypotheses on ν and $\tilde{\nu}$ are assumed as in Lemma 4.3.3):

Lemma 4.3.4. *Let $\{B_i\}_{i\in I}$ be a family of subgraphs of G and suppose that*

$$\sup_{i\in I} \#\{j \in I : B_j \cap B_i \neq \emptyset\} < \infty.$$

Then, if

$$\sum_{i\in I} \mathbb{P}(B_i \textit{ contains a feasible pair for the } T\textit{-geodesic } \pi : x \to y) \gtrsim d(x,y)$$

for all $x, y \in V$ with $d(x,y)$ sufficiently large, then

$$\liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y) - \mathbb{E}\tilde{T}(x,y)}{d(x,y)} > 0.$$

Proof. First, consider the graph whose vertex set is I and whose edges are $\{i, j\}$ such that $B_i \cap B_j \neq \emptyset$. Our first assumption states precisely that this graph has degree bounded by some constant, let's call it $D' < \infty$. Then this graph can be colored by $D' + 1$ colors using a greedy coloring. Hence we get a decomposition $I = \bigsqcup_{\ell=1}^{D'+1} I_\ell$ such that for each fixed ℓ, for all $i, j \in I_\ell$, if $i \neq j$ then $B_i \cap B_j = \emptyset$. Moreover, we have

$$\begin{aligned} &\max_{\ell\in\{1,\dots,D'+1\}} \sum_{i\in I_\ell} \mathbb{P}(B_i \text{ contains a feasible pair for the geodesic } \pi : x \to y) \\ \geq &\frac{1}{D'+1} \sum_{i\in I} \mathbb{P}(B_i \text{ contains a feasible pair for the geodesic } \pi : x \to y) \gtrsim d(x,y). \end{aligned}$$

Thus we will have our lemma once we show that for each ℓ

$$\mathbb{E}T(x,y) - \mathbb{E}\tilde{T}(x,y) \gtrsim \sum_{i\in I_\ell} \mathbb{P}(B_i \text{ contains a feasible pair for the geodesic } \pi : x \to y).$$

But since the $\{B_i\}_{i\in I_\ell}$ are disjoint families, this follows immediately from Lemma 4.3.3. □

In light of the previous lemma, our strategy for proving our main theorems will be to find suitable subgraphs B_i of G and then prove that the expected number of B_i containing a feasible pair for the T-geodesic from x to y is at least a constant times $d(x,y)$.

4.4. Graphs that admit detours

Here we introduce and prove facts about the key fine-geometric condition on our graphs. Recall (Definition 2) that we call π' an *ϵ-detour* for π if $\pi \neq \pi'$, π and π' have the same endpoints, and $|\pi' \setminus \pi| \leq (1+\epsilon)|\pi \setminus \pi'|$.

Definition 4. *We say that a graph G* admits detours *if for every $\epsilon > 0$, there exists some C such that for every self-avoiding path π in G of length C, there exists a self-avoiding ϵ-detour π' for π.*

Note that this is equivalent to the (*a priori* stronger) condition that any self-avoiding path of length *at least* C admits a self-avoiding ϵ-detour; given a self-avoiding path π of length greater than C, replace some subpath π_1 of length C with an ϵ-detour π_1' for π_1, and we obtain an ϵ-detour π' for π. π' can be made into a self-avoiding ϵ-detour for π simply by loop-erasing if necessary.

Our main theorems say that at least in certain coarse-geometric settings, this fine-geometric condition on the graph is equivalent to the vdBK property. We first give an equivalent condition and show that this condition is *necessary* for a graph to have the vdBK property. We then discuss a subtlety in interpreting the negation of the vdBK property for inhomogeneous graphs. We then give examples of graphs which admit detours.

4.4.1. vdBK graphs admit detours

Recall that a path π from v to w is called an *edge-geodesic* if for all paths π' from v to w, $|\pi| \le |\pi'|$. π is called a *unique (edge)-geodesic* if for all $\pi' \neq \pi$ from v to w, $|\pi| < |\pi'|$.

Proposition 4.4.1. *G admits detours if and only if G* admits detours along unique geodesics *in the following sense: for all $\epsilon > 0$, there exists $C < \infty$ such that for every unique edge-geodesic π of length C, there exists an ϵ-detour π' for π.*

Proof. The forward implication is clear. Now assume that G admits detours along unique geodesics. Note that the ϵ-detour π' for a unique geodesic π can be made self-avoiding simply by loop erasing; the resulting path is still an ϵ-detour for π because the process of loop erasing cannot increase $|\pi' \setminus \pi|$ and cannot decrease $|\pi \setminus \pi'| \ge 1$. So it only remains to construct self-avoiding ϵ-detours for self-avoiding paths which are not unique geodesics. Let π be a self-avoiding path which is not a unique geodesic, and let π' be an edge-geodesic connecting the endpoints of π which is not equal to π. Since π' is a geodesic, we have

$$0 \le |\pi| - |\pi'| = |\pi \setminus \pi'| - |\pi' \setminus \pi|,$$

so that

$$|\pi' \setminus \pi| \le |\pi \setminus \pi'| \le (1+\epsilon)|\pi \setminus \pi'|$$

for all $\epsilon > 0$. □

We can now easily prove that admitting detours is a *necessary* condition for a graph to be vdBK.

Theorem 4.4.1. *Let G be a graph which does not admit detours. Then there exists a sequence of pairs $(x_n, y_n) \in V^2$ with $d(x_n, y_n) \xrightarrow[n\to\infty]{} \infty$ and a pair of atomless measures $\nu, \tilde{\nu}$ with finite mean which are supported away from* 0 *and such that $\tilde{\nu}$ is strictly more variable than ν but*

$$\mathbb{E}T(x_n, y_n) = \mathbb{E}\tilde{T}(x_n, y_n)$$

for all n. In particular, if G has bounded degree, then G does not satisfy the vdBK property.

Proof. Since G does not admit detours, by Proposition 4.4.1 there exists $\epsilon_0 > 0$ such that for each n we have a unique geodesic π_n of length n which does not admit a ϵ_0-detour, which is to say that, if x_n and y_n are the endpoints of π_n, then any other self-avoiding π_n' from x_n to y_n satisfies

$$|\pi_n' \setminus \pi_n| \geq (1 + \epsilon_0)|\pi_n \setminus \pi_n'|.$$

Note that, canceling a term of $T(\pi_n \cap \pi_n')$, we always have

$$T(\pi_n') - T(\pi_n) = T(\pi_n' \setminus \pi_n) - T(\pi_n \setminus \pi_n').$$

Now assume ν is supported on $[1, 1 + \epsilon_0]$. Then for any $\pi_n' \neq \pi_n$ we have

$$\begin{aligned} T(\pi_n' \setminus \pi_n) - T(\pi_n \setminus \pi_n') &\geq 1 \cdot |\pi_n' \setminus \pi_n| - (1 + \epsilon_0)|\pi_n \setminus \pi_n'| \\ &\geq (1 + \epsilon_0)|\pi_n \setminus \pi_n'| - (1 + \epsilon_0)|\pi_n \setminus \pi_n'| = 0. \end{aligned}$$

That is, when ν is supported on $[1, 1+\epsilon_0]$, π_n is almost surely has optimal passage time, that is,

$$T(x_n, y_n) = T(\pi_n) \text{ a.s.}$$

But then

$$\mathbb{E}T(x_n, y_n) = \mathbb{E}T(\pi_n) = (\mathbb{E}w)d(x_n, y_n).$$

In particular, if both ν and $\tilde{\nu}$ are supported on $[1, 1+\epsilon_0]$ and $\mathbb{E}w = \mathbb{E}\tilde{w}$, we get

$$\mathbb{E}T(x_n, y_n) = (\mathbb{E}w)d(x_n, y_n) = (\mathbb{E}\tilde{w})d(x_n, y_n) = \mathbb{E}\tilde{T}(x_n, y_n),$$

so to complete our proof we just need to find two such $\nu, \tilde{\nu}$ such that $\tilde{\nu}$ is strictly more variable than ν. For example, we can take $\tilde{\nu}$ to be the uniform measure on $[1, 1+\epsilon_0]$ and ν to be the uniform measure on $[1+(\epsilon_0/4), 1+(3\epsilon_0/4)]$ (see Example 2.17 in [36]). Finally, if G has bounded degree, then $\nu(\{1+(\epsilon_0/4)\}) = 0 < 1/D \leq \underline{\vec{p_c}}$, so ν is exponential-subcritical and the pair $\nu, \tilde{\nu}$ contradicts the vdBK property. □

Although this theorem applies to any bounded degree graph, one should note that the vdBK is quite a strong property to ask of an inhomogeneous graph. That is, the condition $\mathbb{E}\tilde{T} \ll \mathbb{E}T$, or

$$\liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y) - \mathbb{E}\tilde{T}(x,y)}{d(x,y)} > 0,$$

requires that we have a discrepancy of linear order for *any* pair of points $x, y \in V$ which are sufficiently far apart. This implies, but is in general stronger than, another natural condition: fixing a basepoint $o \in V$, we may ask that

$$\liminf_{x\to\infty} \frac{\mathbb{E}T(o,x) - \mathbb{E}\tilde{T}(o,x)}{d(o,x)} > 0.$$

Note that by the triangle inequality, when $\nu, \tilde{\nu}$ have finite mean, the above limit is actually independent of the choice of basepoint o. This is the relevant strict inequality if one is considering scaling limits based at o.

The difference in these two notions means that the negation of the vdBK property is not as strong of a property as we would like it to be. For instance (see Section 4.7.6), supercritical percolation clusters in $\mathbb{Z}^d$ almost surely do *not* satisfy the vdBK property, but it seems very likely that they should satisfy "the vdBK property with one endpoint fixed."

On the other hand, it should be emphasized that if G is almost-transitive (i.e. the action of $\mathrm{Aut}(G)$ on V has finitely many orbits), then the two above notions of strict inequality are equivalent. This is because if we take a finite fundamental domain $F \subset V$ for the action of $\mathrm{Aut}(G)$ on V, using $\mathrm{Aut}(G)$-invariance of $\mathbb{E}T, \mathbb{E}\tilde{T}$, and d, we have that

$$\begin{aligned}
\liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y) - \mathbb{E}\tilde{T}(x,y)}{d(x,y)} &= \lim_{R\to\infty} \min_{u\in F} \inf_{\substack{x\in V \\ d(u,x)\geq R}} \frac{\mathbb{E}T(u,x) - \mathbb{E}\tilde{T}(u,x)}{d(u,x)} \\
&= \min_{u\in F} \liminf_{x\to\infty} \frac{\mathbb{E}T(u,x) - \mathbb{E}\tilde{T}(u,x)}{d(u,x)} \\
&= \liminf_{x\to\infty} \frac{\mathbb{E}T(o,x) - \mathbb{E}\tilde{T}(o,x)}{d(o,x)}.
\end{aligned}$$

Thus, while the vdBK property may be "too strong" for general inhomogeneous graphs, it is "the right" property for almost-transitive graphs. Moreover, there are several inhomogeneous graphs which are nonetheless vdBK; see Section 4.7.6.

4.4.2. Examples of graphs which admit detours

Proposition 4.4.1 gives us an easy way to produce graphs which admit detours, namely by "doubling" edges. Simply take any graph G and create a new graph G' by taking the edge set of G and adding an extra edge between each $v, w \in V$ which are connected by an edge in G. Since every edge has a "parallel" edge, G' has no unique geodesics, and hence by Proposition 4.4.1 admits detours.

This is a rather "cheap" way to get a graph that admits detours, especially since in first-passage percolation often the graphs one is interested in are simple, i.e. contain no parallel edges. However, this is a simple way to see that the property of admitting detours is not a quasi-isometry invariant; every graph G is quasi-isometric to a graph G' which admits detours, so admitting detours is a "fine" rather than a "coarse" geometric property.

The property is not group-theoretic either; that is, for some groups Γ, some Cayley graphs of Γ admit detours and others do not. This can be seen using the same technique as above if one allows Cayley graphs to have double edges (for discussion on what exactly is meant by "Cayley graph" see Section 4.5). But even if one restricts to simple Cayley graphs, there are counterexamples. The standard Cayley graph of $\mathbb{Z}$ does not admit detours (since it is a tree), but every Cayley graph of $\mathbb{Z}$ not isomorphic to this one does. Similarly, Cayley graphs of $\mathbb{Z}/2 * \mathbb{Z}/2$ which are isomorphic to the standard Cayley graph of $\mathbb{Z}$ do not admit detours, but all others do. This is proven in Section 4.5.

On the other hand, there are several properties of groups which ensure that *all* of their Cayley graphs admit detours. For instance, we have

Proposition 4.5.1. *Let Γ be a finitely generated group, and suppose that Γ contains $F \trianglelefteq \Gamma$ a nontrivial finite normal subgroup. Then any Cayley graph of Γ admits detours.*

Proposition 4.5.4. *Let Γ be a finitely generated group not isomorphic to $\mathbb{Z}$ or $\mathbb{Z}/2 * \mathbb{Z}/2 \cong \mathbb{Z} \rtimes \mathbb{Z}/2$ with a finite index subgroup H such that H has nontrivial center. Then any Cayley graph G of Γ admits detours.*

These allow us to conclude:

Theorem 4.5.1. *Let G be a Cayley graph of a virtually nilpotent group. If G is not isomorphic as a graph to the standard Cayley graph of $\mathbb{Z}$, then G admits detours.*

The proofs of all of these facts are entirely combinatorial and group-theoretic and are given in Section 4.5. There may be many weaker group-theoretic conditions which ensure that every Cayley graph of a group admits detours; the ones proven here were mostly chosen in order to prove Theorem 4.5.1, since this is needed to prove Theorem 4.1.1. Of course, they readily apply to many groups which are not virtually nilpotent.

4.5. Cayley graphs which admit detours

In this section we prove sufficient conditions for a Cayley graph of a group to admit detours, and in particular Theorem 4.5.1, which will allow us to conclude Theorem 4.1.1 from Theorem 4.1.2. The arguments here involve no probability, only combinatorics and group theory. They are independent of the rest of the chapter, so some readers may choose to skip this section upon first reading.

4.5.1. Paths in Cayley graphs

We begin by defining two slightly different notions of Cayley graph. Both definitions are natural, and the distinction between the two does not make a difference to geometric group theorists (i.e., the two constructions produce quasi-isometric graphs, and the metric induced on the group is the same) but it will make a slight difference in the study of FPP.

Let Γ be a finitely generated group, S a finite set, and $f : S \to \Gamma \setminus \{1\}$ a map whose image generates Γ. We define the *unreduced Cayley graph* associated to (Γ, S) to be the graph $G = (V, E)$ with vertex set $V = \Gamma$ and edge set $E = \Gamma \times S$, where the boundary of the edge $e = (g, s)$ is $\{g, gf(s)\} \subset V$. We define the *reduced Cayley graph* associated to (Γ, S) to be the graph $G = (V, E)$ whose vertex set is $V = \Gamma$, whose edge set is the set $E = \{\{x, y\} \subset \Gamma : x \neq y, \exists s \in S \cup S^{-1} \text{ s.t. } y = xf(s)\}$ and the boundary map is the natural inclusion. The reduced Cayley graph is the simple graph obtained from the unreduced Cayley graph by deleting parallel edges. Typically the two graphs are isomorphic unless some $f(s) \in \Gamma$ is of order two, but the two graphs may also be nonisomorphic if one takes $f : S \to \Gamma$ to be a "redundant" generating set, containing repeated elements or inverses of elements already included. We call a graph a *Cayley graph* it is a reduced or unreduced Cayley graph.

Note that the action of Γ on itself by left multiplication induces an action of Γ on G by graph isomorphisms whenever G is a Cayley graph. So Γ acts on the set of vertices of G, the set of edges of G, the set of geodesics of G, the set of unique geodesics of G, etc. The action of Γ on the vertex set $V = \Gamma$ is transitive, and so any path and in particular any unique geodesic is the translate of some path (unique geodesic) starting at $1 \in \Gamma = V$.

Finite paths in the unreduced Cayley graph associated to (Γ, S) starting from the identity vertex are in one-to-one correspondence with finite words in $S \sqcup S^{-1}$. (Here, S is considered as an abstract set, and S^{-1} consists of symbols of the form s^{-1} where $s \in S$). In the case of a *reduced* Cayley graph, there is a bijection between finite paths starting at the identity and elements of $(f(S) \cup f(S)^{-1})^*$ (i.e. finite words in the subset $f(S) \cup f(S)^{-1} \subset \Gamma$). We use A to denote $S \sqcup S^{-1}$ in the case that G is the unreduced Cayley graph associated to (Γ, S) and we use A to denote $f(S) \cup f(S)^{-1} \subset \Gamma$ in the case that G is the reduced Cayley graph associated to (Γ, S). So in either case, paths starting from $1 \in V = \Gamma$ correspond to finite words in A; we denote the set of finite words in A by A^*.

A^* together with the operation of concatenation is the free monoid on A. Since A has a natural "formal inverse" map $A \to A, a \mapsto a^{-1}$ in both cases, this induces a "formal inverse" map on A^* given by $(a_1 \cdots a_n)^{-1} := a_n^{-1} \cdots a_1^{-1}$. Note that α^{-1} is not a true inverse of α; the word $\alpha\alpha^{-1} \in A^*$ represents the concatenation of the path represented by α with its reverse, which is in particular not equal to the trivial path (consisting of no edges) represented by the empty word in A^*.

We have an "evaluation map" $\rho : (S \sqcup S^{-1})^* \to \Gamma$ induced by $f : S \to \Gamma$ given by

$$\rho(s_1^{\epsilon_1} \cdots s_k^{\epsilon_k}) := f(s_1)^{\epsilon_1} \cdots f(s_k)^{\epsilon_k},$$

where the $s_i \in S$, $\epsilon_i \in \{+1, -1\}$. We also have an evaluation map $\rho : (f(S) \cup f(S)^{-1})^* \to \Gamma$ just given by group multiplication in Γ; $\rho(g_1 \cdots g_k)$ is the product $g_1 * \cdots * g_k \in \Gamma$. (Typically the group multiplication in Γ is denoted in the same way as concatenation in A^*; we only write it with $*$ here to emphasize that $\rho : A^* \to \Gamma$ is not the identity map).

Note that in either case $\rho : A^* \to \Gamma$ is a homomorphism of monoids which respects the inverse operation, i.e. $\rho(\alpha\beta) = \rho(\alpha)\rho(\beta)$, $\rho(\alpha^{-1}) = \rho(\alpha)^{-1}$ for all $\alpha, \beta \in A^*$. Geometrically, if $\alpha \in A^*$ represents a path in G starting from $1 \in V = \Gamma$, $\rho(\alpha) \in \Gamma = V$ is the endpoint of that path. If $\alpha, \beta \in A^*$, then the path represented by $\alpha\beta$ is the concatenation of the path represented by α with the left-translate by $\rho(\alpha)$ of the path represented by β. The path represented by α^{-1} is the left-translate by $\rho(\alpha)^{-1}$ of the reverse of the path represented by α.

The condition that $\pi \in A^*$ corresponds to a geodesic in G is exactly the condition that for any $\pi' \in A^*$ such that $\rho(\pi') = \rho(\pi)$ we have $|\pi'| \geq |\pi|$ (where here $|\cdot|$ is the length of a word). The condition that $\pi \in A^*$ corresponds to a unique geodesic is precisely the condition that π corresponds to a geodesic, and for any π' such that $\rho(\pi') = \rho(\pi)$ and $|\pi'| = |\pi|$ we have $\pi = \pi'$. Recall also that the properties of being geodesic and uniquely geodesic pass to subpaths and are invariant under translations, and so the above properties pass to subwords. In what follows we will use the same symbol to denote a word in A^* and the path in G starting from the identity which it corresponds to. For instance, for $\alpha, \beta \in A^*$, $\rho(\alpha)\beta$ is the left-translate by $\rho(\alpha) \in \Gamma$ of the path in G represented by β; $\rho(\alpha)\beta$ is a path in G starting at $\rho(\alpha)$ and ending at $\rho(\alpha\beta)$, and the path $\alpha\beta$ is the concatenation of the paths α and $\rho(\alpha)\beta$.

4.5.2. Sufficient conditions for a Cayley graph to admit detours

In many groups, there are relatively explicit constructions for constructing an ϵ-detour for a given unique geodesic. All the constructions given in this section are very similar, and come from intuition from $\mathbb{Z}^d$, $d \geq 2$. In the standard Cayley graph of $\mathbb{Z}^d$, unique geodesics are just long straight lines; one can construct a path with the same endpoints which is totally edge-disjoint from the original and only two edges longer by simply first taking a step perpendicular to the original path, following a translated version of the original path, and then taking a step back. Our basic strategy will be: given a unique geodesic with some nice properties, find two words of bounded length which when appended to the beginning and end of this path give a path which has the same endpoints but misses a positive proportion of the edges in the original path; this constructs detours for "nice" unique geodesics. Then, if every unique geodesic contains a subpath with nice properties with length at least a positive proportion of the original length, we can now construct a detour for a general unique geodesic simply by replacing the nice subpath with its detour.

The first large class of groups for which we prove that all Cayley graphs admit detours is the following:

Proposition 4.5.1. *Let* Γ *be a finitely generated group, and suppose that* Γ *contains* $F \trianglelefteq \Gamma$ *a nontrivial finite normal subgroup. Then any Cayley graph of* Γ *admits detours.*

Proof. Set $\ell := \max_{f\in F} |f|$ (where $|f|$ is the minimum length of a word $w \in A^*$ with $\rho(w) = f$). Let $\pi \in A^*$ be a uniquely geodesic word. We first claim that there exists a subword π' of π with $|\pi'| \geq \frac{|\pi|-\ell}{\ell+1}$ such that no nonempty subword v of π' has $\rho(v) \in F$. To see this, consider the path in the Cayley graph of Γ/F corresponding to π. A subword v of π with $\rho(v) \in F$ corresponds exactly to a loop in the path in Γ/F. We then loop-erase; that is, we can find a collection of disjoint subwords $\beta_1, ..., \beta_k$ of π, corresponding to loops in the path in Γ/F, such that if we take the word π'' obtained from π by removing these subwords, we get a path in Γ/F with the same endpoints (that is, $\rho(\pi'')$ and $\rho(\pi)$ have the same image under the map $\Gamma \to \Gamma/F$) and this path does not have any loops (that is, for all nonempty subwords v of π'', we have $\rho(v) \notin F$). Note that since π represents a geodesic path in the Cayley graph of Γ, we have

$$|\rho(\pi)| = |\pi| = |\pi''| + \sum_{i=1}^{k} |\beta_i|,$$

but on the other hand, since $\rho(\pi)$ and $\rho(\pi'')$ have the same image in Γ/F, there exists $f \in F$ such that $f\rho(\pi'') = \rho(\pi)$, and therefore

$$|\rho(\pi)| \leq |f| + |\rho(\pi'')| \leq \ell + |\pi''|,$$

whence we conclude that

$$\sum_{i=1}^{k} |\beta_i| \leq \ell$$

and hence (assuming without loss of generality that the β_i are nonempty) also that $k \leq \ell$. Since π'' was obtained from π by deleting k subwords from π, it is equal to the concatenation of at most $k+1$ subwords of π. Thus, by the pigeonhole principle, there exists a subword π' of π with

$$|\pi'| \geq \frac{|\pi''|}{k+1} \geq \frac{|\pi| - \ell}{\ell + 1},$$

and such that no nonempty subword v of π' has $\rho(v) \in F$, as desired.

Now, we construct a detour for π'. Let $f \in F\backslash\{1\}$. Since F is normal, $f^{\rho(\pi')} := (\rho(\pi'))^{-1} f \rho(\pi') \in F$. Taking geodesics $\gamma_1 \in A^*$ from 1 to f and $\gamma_2 \in A^*$ from 1 to $(f^{\rho(\pi')})^{-1}$ and setting

$$\tilde{\pi}' := \gamma_1 \pi' \gamma_2$$

gives a path with $\rho(\tilde{\pi}') = f\rho(\pi')(f^{\rho(\pi')})^{-1} = \rho(\pi')$ (that is, $\tilde{\pi}'$ has the same endpoints as π'), and

$$|\tilde{\pi}'| = |\pi'| + |\gamma_1| + |\gamma_2| \leq |\pi'| + 2\ell.$$

We now show that if $\tilde{\pi}'$ intersects π' it only does so in the first ℓ or the last ℓ edges of π'. Note that it suffices to show that the vertex sets $V(\pi')$ and $V(\tilde{\pi}')$ only intersect in the first $\ell + 1$ or the last $\ell + 1$ vertices of $V(\pi')$. (Recall that if $S \subset E$, then $V(S) \subset V$ is the set of vertices which are endpoints of some edge $e \in S$.) Note also that intersections of $V(\pi')$ and $V(\tilde{\pi}')$ correspond to initial subwords of π' and $\tilde{\pi}'$ which have the same image under $\rho : A^* \to \Gamma$.

So first, suppose that for some initial subword v of γ_1 and some initial subword π_1 of π' we have that

$$\rho(v) = \rho(\pi_1);$$

geodesicity of π' implies that $|\pi_1| \leq |v| \leq \ell$; that is, the intersection of $V(\pi')$ and $V(\tilde{\pi}')$ has happened in the first $\ell + 1$ vertices of $V(\pi')$. Similarly, if for some initial subword v of γ_2 and some initial subword π_1 of π' we have

$$\rho(\gamma_1 \pi' v) = \rho(\pi_1);$$

taking the v' to be the subword of γ_2 such that $\gamma_2 = vv'$ we get

$$\rho(\pi') = \rho(\gamma_1 \pi' \gamma_2) = \rho(\pi_1)\rho(v'),$$

and taking π_2 to be the subword of π' such that $\pi' = \pi_1\pi_2$ we get

$$\rho(\pi_1)\rho(\pi_2) = \rho(\pi') = \rho(\pi_1)\rho(v') \Rightarrow \rho(\pi_2) = \rho(v')$$

by cancellation, and so by geodesicity of π' we conclude that $|\pi_2| \leq |v'| \leq \ell$; that is, the intersection of $V(\pi)'$ and $V(\tilde{\pi}')$ has happened in the last $\ell + 1$ vertices. The only remaining case is that there exist initial subwords π_1 and π_2 of π such that

$$\rho(\gamma_1 \pi_1) = \rho(\pi_2),$$

and hence

$$f\rho(\pi_1) = \rho(\pi_1)f^{\rho(\pi_1)} = \rho(\pi_2) \Rightarrow f^{\rho(\pi_1)} = \rho(\pi_1)^{-1}\rho(\pi_2) \in F \setminus \{1\}.$$

But since π_1 and π_2 are both initial subwords of π', there exists some subword π_3 of π such that $\rho(\pi_1)^{-1}\rho(\pi_2) = \rho(\pi_3)^{\pm}$, but then π_3 is a subword of π' with $\rho(\pi_3) \in F \setminus \{1\}$, which contradicts our construction of π', so this case cannot occur. Thus we have proven that $V(\pi')$ and $V(\tilde{\pi}')$ only intersect in the first $\ell + 1$ or the last $\ell + 1$ vertices of $V(\pi')$, and so in particular $|\pi' \setminus \tilde{\pi}'| \geq |\pi'| - 2\ell$.

Finally, if $\pi = \alpha\pi'\omega$, set $\tilde{\pi} = \alpha\tilde{\pi}'\omega$. We then have that $\rho(\pi) = \rho(\tilde{\pi})$, that is, π and $\tilde{\pi}$ have the same endpoints, and

$$|\tilde{\pi}| - |\pi| = |\tilde{\pi}'| - |\pi'| \leq 2\ell,$$

while

$$|\pi \setminus \tilde{\pi}| \geq |\pi' \setminus \tilde{\pi}'| \geq |\pi'| - 2\ell \geq \frac{|\pi| - \ell}{\ell + 1} - 2\ell.$$

Thus, given $\epsilon > 0$, if we choose C large enough that $(2\ell)\left(\frac{C-\ell}{\ell+1} - 2\ell\right)^{-1} \leq \epsilon$, then for any unique geodesic π with $|\pi| \geq C$, we have that

$$|\tilde{\pi}| - |\pi| \leq \epsilon|\pi \setminus \tilde{\pi}|,$$

that is, $\tilde{\pi}$ is an ϵ-detour for π. □

For other classes of groups, we will need to treat two cases separately: either we can construct a detour for our unique geodesic with a simple construction, or our unique geodesic has a very special form. This is made precise in the following lemma, which we use several times below:

Lemma 4.5.1. *Let Γ be a finitely generated group and let $H \leq \Gamma$ be a subgroup of finite index. Let G be a Cayley graph associated to a generating set S for Γ, A be the relevant alphabet. Suppose that there exists some $z \in H \setminus \{1\}$ such that for all $h \in H$*

$$|z^h| := |h^{-1}zh| = |z|.$$

Fix $w \in A^*$ a geodesic from 1 to z, and for each H-conjugate z^h of z fix $w^h \in A^*$ a geodesic from 1 to z^h (note the property of z ensures that $|w| = |w^h|$ for all $h \in H$). Let $\pi \in A^*$ be a unique geodesic in G such that $\rho(\pi) \in H$. Then:

(1) If there exist $\alpha, \omega \in A^*$, $h \in H$ such that $\rho(\alpha), \rho(\omega) \in H$ and $\pi = \alpha w^h \omega$, then π is an initial subpath of $(w^{h'})^N$ for some $h' \in H$, $N < \infty$. If $\pi = \alpha(w^h)^{-1}\omega$ then π is an initial subpath of $((w^{h'})^{-1})^N$ for some $h' \in H, N < \infty$.

(2) Suppose that the condition above fails, that is, for all choices of $\alpha, \omega \in A^*$ and $h \in H$ with $\rho(\alpha), \rho(\omega) \in H$, we have $\pi \neq \alpha w^h \omega$ and $\pi \neq \alpha(w^h)^{-1}\omega$. Then the path

$$\pi' := w\pi(w^{\rho(\pi)})^{-1}$$

has the same endpoints as π and satisfies

$$|V(\pi') \cap V(\pi) \cap H| \leq 2(|w| + 1),$$

and hence

$$|V(\pi) \setminus V(\pi')| \geq |(V(\pi) \setminus V(\pi')) \cap H| = |V(\pi) \cap H| - |V(\pi' \cap \pi \cap H)|$$

$$\geq |V(\pi) \cap H| - 2(|w| + 1)$$

so that

$$|\pi \setminus \pi'| \geq \frac{1}{2}|V(\pi) \setminus V(\pi')| \geq \frac{1}{2}|V(\pi) \cap H| - |w| - 1.$$

Before proving the lemma, we give context by proving the implications we need it for.

Proposition 4.5.2. *Let G be a Cayley graph of $\mathbb{Z}$. Either $A = \{a, a^{-1}\}$ for some $a \in A$, in which case G is isomorphic to the standard Cayley graph of $\mathbb{Z}$, or G admits detours.*

Proof. If $A = \{a, a^{-1}\}$, then clearly G is isomorphic to the standard Cayley graph of $\mathbb{Z}$. Assume that there exists $s \in A \setminus \{a, a^{-1}\}$ for some $a \in A$. Letting $H = \Gamma = \mathbb{Z}$, $w = a$, we see that the assumptions of the lemma are satisfied, since $\mathbb{Z}$ is abelian and hence the conjugation action is trivial (we also set $w^h = a$ for all $h \in H$). Therefore, if π is any unique geodesic in Γ, either $\pi = a^N, \pi = (a^{-1})^N$, or π does not contain a or a^{-1}. By part (2) of the lemma, in the latter case, setting $\pi' := a\pi a^{-1}$ gives a path with the same endpoints such that[5]

$$|\pi \setminus \pi'| \geq \frac{1}{2}|\pi| - 2.$$

Now consider the former case, i.e. $\pi = a^N$ or $\pi = (a^{-1})^N$. Note that taking $H = \Gamma = \mathbb{Z}$ and $w = s$ also satisfies the hypotheses of the lemma, and since s, s^{-1} do not appear in π, by part (2) of the lemma, taking $\pi' := s\pi s^{-1}$ gives a path with the same endpoints as π and

$$|\pi \setminus \pi'| \geq \frac{1}{2}|\pi| - 2.$$

[5]By looking more closely at the paths, one can see that the prefactor of $\frac{1}{2}$ is not necessary; the crude bound $|\pi \setminus \pi'| \geq \frac{1}{2}|V(\pi) \setminus V(\pi')|$ from which it comes could be avoided by a more careful argument, but this is not necessary for our purposes.

Thus, for any $\epsilon > 0$, if C is sufficiently large that $2 \leq \epsilon(\frac{C}{2} - 2)$ then we have that if π is a unique geodesic with $|\pi| \geq C$ then with the above construction of π' we have

$$|\pi'| - |\pi| = 2 \leq \epsilon|\pi \setminus \pi'|,$$

as desired. □

Proposition 4.5.3. *Let G be a Cayley graph of $\Gamma := \mathbb{Z}/2 * \mathbb{Z}/2$. Either G is a reduced Cayley graph associated to a generating set S such that $f(S)$ consists of exactly two elements of order 2, in which case G is isomorphic to the standard Cayley graph of $\mathbb{Z}$, or G admits detours.*

Proof. If G is a reduced Cayley graph with $A = f(S)$ consisting of exactly two elements of order 2, then G is a connected regular infinite graph of degree 2, and so it is isomorphic to the standard Cayley graph of $\mathbb{Z}$. We want to show that in any other case, G admits detours.

Recall that all nonidentity elements of $\Gamma = \mathbb{Z}/2 * \mathbb{Z}/2 \cong \mathbb{Z} \rtimes \mathbb{Z}/2$ have either infinite order or order two. Hence, the remaining cases to check are (1): $f(S)$ contains an element of infinite order; (2): $f(S)$ contains only elements of order 2 but G is an unreduced Cayley graph; (3): $f(S)$ contains only elements of order 2, G is a reduced Cayley graph, and $|f(S)| \geq 3$. Also recall that any product of two distinct elements of order 2 is an element of infinite order, and that for any element $x \in \Gamma$ of infinite order, for any $g \in \Gamma$ we have $g^{-1}xg = x^{-1}$ if g has order 2 and $g^{-1}xg = x$ otherwise.

So assume that for some $z \in A$, $\rho(z)$ has infinite order. Then for any $g \in \Gamma$ we have $\rho(z)^g = \rho(z)$ or $\rho(z)^g = \rho(z)^{-1} = \rho(z^{-1})$. Therefore if we take $H = \Gamma$ and $w = z$, the hypotheses of the lemma are satisfied, and for each $h \in H$ we can choose w^h to be either z or z^{-1}. Let π be a unique geodesic in G. By the lemma, either $\pi = z^N$, $\pi = (z^{-1})^N$ or π does not contain z or z^{-1}.

In the latter case, by the lemma, taking $\pi' := z\pi z^{-1}$ gives a path with the same endpoints and $|\pi \setminus \pi'| \geq \frac{1}{2}|\pi| - 2$. In the former cases, take $a \in A$ with $\rho(a)$ of order 2 (any generating set of Γ must contain such an element). Taking $\pi' := a\pi^{-1}a$ then gives a path with the same endpoints as π, and $V(\pi) \cap V(\pi')$ is equal to the endpoints of π and π', since for any $0 \leq n, m \leq N$, $\rho(az^{\pm n}) \neq \rho(z^{\pm m})$, since the two elements lie in distinct cosets of $\langle \rho(z) \rangle \leq \Gamma$. Hence in fact $|\pi \setminus \pi'| = |\pi|$.

Next, assume that $f(S)$ contains only elements order 2, but G is an unreduced Cayley graph. In this case, every edge of G is a double edge, so there are no unique geodesics, and so G admits detours.

Lastly, assume that $f(S)$ contains only elements of order 2, G is a reduced Cayley graph, and $|f(S)| \geq 3$. Pick $a, b, c \in A$ such that $\rho(a), \rho(b), \rho(c)$ are all distinct. Since $\rho(ab)$ has infinite order, again taking $w = ab$, $H = \Gamma$ satisfies the hypotheses of the lemma, and we can choose for each $h \in H$ either $w^h = ab$ or $w^h = ba$. Hence, letting π be a unique geodesic in G, we have that either π does not contain ab or ba as a subword, or that π is a subword of $(ab)^N$ for some N. In the first case, by the lemma, we have that taking $\pi' := (ab)\pi(ba)^{(-1)^{|\pi|}}$ gives a path with the same endpoints as π with $|\pi \setminus \pi'| \geq \frac{1}{2}|\pi| - 3$. In the second case, take $\pi' := (ac)\pi(ca)^{(-1)^{|\pi|}}$. Then π and π' have the same endpoints, and we claim that $|V(\pi') \cap V(\pi)|$ is contained in the union of the first two and last two vertices of $V(\pi')$, implying that $|\pi \setminus \pi'| \geq |\pi| - 3$. To see this, suppose to the contrary that for some initial subpaths π_1, π_2 of π we had

$$\rho((ac)\pi_1) = \rho(\pi_2).$$

We have $\rho((ac)\pi_1) = \rho(\pi_1(ac)^{(-1)^{|\pi_1|}})$. Moreover, there is some subpath $\tilde{\pi}$ of π such that either $\pi_2\tilde{\pi} = \pi_1$ or $\pi_1\tilde{\pi} = \pi_2$. Then in the first case, by cancellation we have

$$\rho((ac)^{(-1)^{|\pi_1|}}) = \rho(\tilde{\pi}).$$

As a subpath of π, $\tilde{\pi}$ is uniquely geodesic, but (ac) and (ca) are both geodesics, since $f(S)$ only contains elements of order 2 and hence any path to an element of infinite order has length at least 2. Hence we have that $\tilde{\pi} = (ac)^{(-1)^{|\pi_1|}}$, contradicting our assumption that π is a subpath of $(ab)^N$. In the case that $\pi_1\tilde{\pi} = \pi_2$, we similarly conclude that $\tilde{\pi} = (ac)^{(-1)^{|\pi_1|+1}}$, which is similarly a contradiction.

Thus, in any of these three cases, given a unique geodesic π, we can produce a path π' with the same endpoints such that $|\pi \setminus \pi'| \geq \frac{1}{2}|\pi| - 3$ and $|\pi'| \leq |\pi| + 4$. Thus, given $\epsilon > 0$, if we choose C sufficiently large that $4 \leq \epsilon(\frac{1}{2}C - 3)$, then for any self avoiding path π of length at least C have have π' with

$$|\pi'| - |\pi| \leq 4 \leq \epsilon|\pi \setminus \pi'|,$$

as desired. □

Proposition 4.5.4. *Let Γ be a finitely generated group not isomorphic to $\mathbb{Z}$ or $\mathbb{Z}/2 * \mathbb{Z}/2 \cong \mathbb{Z} \rtimes \mathbb{Z}/2$ with a finite index subgroup H such that H has nontrivial center. Then any Cayley graph G of Γ admits detours.*

Proof. First, we show that we can assume without loss of generality that H is not cyclic. Suppose that H were cyclic; then Γ would be virtually $\mathbb{Z}$, and therefore (see Lemma 11.4 on page 102 of [23]) there is a finite normal subgroup $F \trianglelefteq \Gamma$ such that Γ/F is isomorphic to either $\mathbb{Z}$ or $\mathbb{Z}/2 * \mathbb{Z}/2$. If F is not trivial, then Proposition 4.5.1 tells us that any Cayey graph of Γ admits detours. If F is trivial, then Γ itself is isomorphic to either $\mathbb{Z}$ or $\mathbb{Z}/2 * \mathbb{Z}/2$, and these cases are excluded by assumption.

So assume that H is a non-cyclic finite index subgroup of Γ with nontrivial center. Fix $z \neq 1$ a nontrivial central element of H of minimal distance to 1 and consider a geodesic path $w \in A^*$ from 1 to z. Since by definition $z^h = z$ for all $h \in H$, this choice of w and H satisfy the conditions of the lemma; we also set $w^h = w$ for all $h \in H$.

Take a unique geodesic $\tilde{\pi}$ in G. By the pigeon-hole principle, there is some $t \in \Gamma$ such that at least $\frac{1}{[\Gamma:H]}|V(\pi)|$ of the vertices in $V(\pi)$ lie in the coset tH. Denote by η_i the subpath of $\tilde{\pi}$ starting at the i^{th} such vertex and ending at the $(i+1)^{th}$ such vertex. Set $\pi := \eta_1 \cdots \eta_M$, the subpath of $\tilde{\pi}$ from the first such vertex to the last such. Note that each $\rho(\eta_i) \in H$, $\rho(\pi) \in H$, and that the η_i are minimal in the sense that if α is a proper initial or final subword of some η_i, then $\rho(\alpha) \notin H$.

Now, by the lemma, either π is an initial subpath of w^N or $(w^{-1})^N$ for some $N < \infty$ or there are no $\alpha, \omega \in A^*$ with $\rho(\alpha), \rho(\omega) \in H$ and either $\pi = \alpha w \omega$ or $\pi = \alpha w^{-1} \omega$. In the latter case, the lemma also tells us that taking $\pi' := w \pi w^{-1}$ gives a path with the same endpoints as π and

$$|\pi \setminus \pi'| \geq \frac{1}{2}|V(\pi) \cap H| - |w| - 1.$$

So consider the case that π is an initial subpath of w^N (the case that π is an initial subpath of $(w^{-1})^N$ is exactly analogous). First note that as long as $|\pi| > |w|$, this implies that all the η_i are equal. For suppose that $\eta_1 \cdots \eta_l = w$; then we have

$$\rho(\pi) = \rho(w \eta_{l+1} \cdots \eta_M) = \rho(\eta_{l+1} w \eta_{l+2} \cdots \eta_M) = \rho(\eta_{l+1} \eta_1 \cdots \eta_l \eta_{l+2} \cdots \eta_M),$$

so unique geodesicity, together with the fact that the decomposition $w = \eta_1 \cdots \eta_M$ is uniquely specified by w and H, implies that $\eta_i = \eta_{i+1}$ for all $i = 1, ..., l-1$, so $w = \eta_1^l$, and hence $\pi = \eta_1^M$.

Now take a geodesic path α from 1 to some element of $H \setminus \langle\rho(\eta_1)\rangle$ (which is possible since H is not cyclic) with minimal distance to 1. Let $0 \leq r < l$ be minimal such that $M + r$ is a multiple of l. Then we set

$$\pi' := \alpha\eta_1^{M+r}\alpha^{-1}\eta_1^{-r}.$$

Since $\rho(\eta_1^{M+r})$ is a power of $\rho(\eta_1^l) = \rho(w) = z$, it is central in H, and hence we see that π and π' have the same endpoints. We further claim that

$$V(\pi') \cap V(\pi) \cap H \subset V(\alpha) \cup \rho(\alpha\eta_1^{M+r})V(\alpha^{-1}\eta_1^{-r}),$$

which implies that

$$|V(\pi') \cap V(\pi) \cap H| \leq 2|\alpha| + |w| + 2,$$

hence $|V(\pi \setminus \pi')| \geq |V(\pi) \cap H| - (2|\alpha| + |w| + 2)$ and

$$|\pi \setminus \pi'| \geq \frac{1}{2}|V(\pi) \setminus V(\pi')| \geq \frac{1}{2}|V(\pi) \cap H| - |\alpha| - |w| - 1.$$

To see this, suppose to the contrary that for some $0 \leq i \leq M + r, 0 \leq j \leq M$ we had

$$\rho(\alpha\eta_1^i) = \rho(\eta_1^j).$$

Then cancellation gives us

$$\rho(\alpha) = \rho(\eta_1)^{j-i} \in \langle\rho(\eta_1)\rangle,$$

contradicting our choice of α.

Thus, in either case, given π, we get π' with the same endpoints satisfying

$$|\pi'| \leq |\pi| + 2(|\alpha| + |w|)$$

and

$$|\pi \setminus \pi'| \geq \frac{1}{2}|V(\pi) \cap H| - |\alpha| - |w| - 1.$$

If $\tilde{\pi} = \pi_1 \pi \pi_2$, set $\tilde{\pi}' := \pi_1 \pi' \pi_2$. Then we again have

$$|\tilde{\pi}'| - |\tilde{\pi}| = |\pi'| - |\pi| \leq 2(|\alpha| + |w|)$$

and

$$\begin{aligned} |\tilde{\pi} \setminus \tilde{\pi}'| = |\pi \setminus \pi'| &\geq \frac{1}{2}|V(\pi) \cap H| - |\alpha| - |w| - 1 \\ &= \frac{1}{2}|V(\tilde{\pi}) \cap tH| - |\alpha| - |w| - 1 \geq \frac{1}{2[\Gamma : H]}|V(\tilde{\pi})| - |\alpha| - |w| - 1. \end{aligned}$$

Note that $|w| = |z|$ is a constant independent of the path $\tilde{\pi}$. Moreover,

$$|\alpha| \leq \sup_{h \in H} \inf_{h' \in H \setminus \langle h \rangle} |h'| =: K < \infty,$$

where K is a constant also independent of the path $\tilde{\pi}$. To see that K is finite, take a finite generating set S' for H (H is finitely generated by Schreier's lemma, as a finite index subgroup of the finitely generated group Γ.) Since H is not cyclic, for every $h \in H$, there is some $s \in S'$ such that $s \notin \langle h \rangle$, and so $K \leq \sup_{s \in S'} |s| < \infty$, as desired.

Thus, given $\epsilon > 0$, if we choose $C > |w|$ sufficiently large that $2(K + |w|) \leq \epsilon\left(\frac{1}{2[\Gamma:H]}C - K - |w| - 1\right)$, then for any unique geodesic $\tilde{\pi}$ in G of length at least C we have $\tilde{\pi}'$ with the same endpoints such that

$$|\tilde{\pi}'| - |\tilde{\pi}| \leq \epsilon|\tilde{\pi} \setminus \tilde{\pi}'|,$$

as desired. □

Remark 4.5.1. *The preceding three propositions are in some sense as far as we can push the lemma. If there exists $z \neq 1$ such $|z^h| = |z|$ for all $h \in H$, in particular the H-orbit of z under the conjugation action is finite; hence by the orbit stabilizer theorem,* $\mathrm{Stab}_H(z) \leq H$ *is a finite index subgroup of H with nontrivial center (the center contains z). Thus if H is itself finite index in Γ, Γ has a finite index subgroup with nontrivial center, and one of the preceding propositions apply. The lemma still holds if H is not finite index, but in that case it is not clear how to use it to construct detours for G.*

Now to prove the lemma.

Proof of Lemma 4.5.1. Let π be a unique geodesic with $\rho(\pi) \in H$ and suppose that for some $\alpha, \omega \in A^*, h \in H$, we have $\rho(\alpha), \rho(\omega) \in H$ and $\pi = \alpha w^h \omega$. (The case that $\pi = \alpha (w^h)^{-1} \omega$ is exactly analogous). Note that since $\rho(\alpha), \rho(w^h), \rho(\omega) \in H$, we have decompositions

$$\alpha = \iota_1 \cdots \iota_j,$$

$$w^h = \eta_1 \cdots \eta_l$$

$$\omega = \kappa_1 \cdots \kappa_m,$$

where each $\rho(\iota_i), \rho(\eta_i), \rho(\kappa_i) \in H$ and if γ is an initial or final proper subpath of some ι_i, η_i, or κ_i, then $\rho(\gamma) \neq H$. Note that this property also uniquely specifies the decomposition.

Now we will show that α is a final subword of some $(w^h)^N$ and that ω is an initial subword of some $(w^h)^N$. First, suppose that $j \leq l$. We have that

$$\rho(\alpha \eta_1 \cdots \eta_l) = \rho(\alpha w^h) = \rho(w^{h\rho(\alpha)^{-1}} \alpha) = \rho(w^{h\rho(\alpha)^{-1}} \iota_1 \cdots \iota_j)$$

where $|w^{h\rho(\alpha)}| = |w^h|$; hence by unique geodesicity we have that the above paths are equal. By the uniqueness of the decomposition, we then have that $\iota_i = \eta_{i+l-j}$ for $i = 1, ..., j$, that is, α is a final subword of w If $j > l$, we have

$$\rho(\iota_1 \cdots \iota_j w^h) = \rho(\iota_1 \cdots \iota_{j-l} w^{h\rho(\iota_{j-l+1}\cdots\iota_j)^{-1}} \iota_{j-l+1} \cdots \iota_j),$$

so that by unique geodesicity, $\iota_{j-l+1} \cdots \iota_j = w^h$, that is $\alpha = \alpha' w$ for some α' of shorter length. Thus, by induction, α is a final subword of $(w^h)^N$ for some N. The argument that ω is an initial subword of some $(w^h)^N$ is exactly analogous.

Thus, $\pi = \alpha w \omega$ is a subword of some $(w^h)^N$; even more than that, α starts with η_i for some i, so we see that π is an initial subword of some w'^N, $w' := \eta_i \eta_{i+1} \cdots \eta_l \eta_1 \cdots \eta_{i-1}$. Note that $\rho(w') = \rho(w^h)^{\rho(\eta_1 \cdots \eta_{i-1})} = \rho(w^{h\rho(\eta_1 \cdots \eta_{i-1})})$, and w' is a subword of π, so by unique geodesicity $w' = w^{h\rho(\eta_1 \cdots \eta_{i-1})}$, so we have proven (1).

Now to prove (2). Let π be a unique geodesic with $\rho(\pi) \in H$ and such that no decomposition of the form $\pi = \alpha(w^h)^{\pm 1}\omega$ holds for any $\rho(\alpha), \rho(\omega), h \in H$. Then take $\pi' := w\pi w^{-1}$. We claim that

$$V(\pi') \cap V(\pi) \cap H \subset V(w) \cup \rho(w\pi)V(w^{-1}),$$

which then clearly implies that $|V(\pi') \cap V(\pi) \cap H| \leq 2(|w| + 1)$.

To prove our claim, suppose that to the contrary there were two initial subpaths π_1, π_2 of π with $\rho(\pi_1), \rho(\pi_2) \in H$ such that

$$\rho(w\pi_1) = \rho(\pi_2).$$

Then we have

$$\rho(w\pi_1) = \rho(\pi_1 w^{\rho(\pi_1)}) = \rho(\pi_2).$$

There is some subpath $\tilde{\pi}$ of π such that $\rho(\tilde{\pi}) \in H$ and either $\pi_1 = \pi_2\tilde{\pi}$ or $\pi_2 = \pi_1\tilde{\pi}$. In the second case, cancellation and unique geodesicity give

$$\rho(w^{\rho(\pi_1)}) = \rho(\tilde{\pi}) \Rightarrow w^{\rho(\pi_1)} = \tilde{\pi}$$

which contradicts our assumption on π, since $\pi = \pi_1\tilde{\pi}\omega$ for some $\rho(\omega) \in H$ and the above equation says that $\tilde{\pi}$ is a geodesic from 1 to $z^{\rho(\pi_1)}$. The first case gives $\rho(\tilde{\pi}) = \rho(w^{\rho(\tilde{\pi})})^{-1}$, which is similarly a contradiction. So we are done.

The final inequalities are consequences of straightforward algebraic manipulations, together with the inequality

$$|\pi \setminus \pi'| \geq \frac{1}{2}|V(\pi) \setminus V(\pi')|,$$

which follows from the fact that we can construct a map $V(\pi) \setminus V(\pi') \to \pi \setminus \pi'$ with fibers of size at most 2 as follows: take each $v \in V(\pi) \setminus V(\pi')$ and associate to it an edge $e \in \pi$ such that v is an endpoint of e; such an edge exists since $v \in V(\pi)$ and $e \notin \pi'$ because $v \notin V(\pi')$. □

Theorem 4.5.1. *Let G be a Cayley graph of a virtually nilpotent group. If G is not isomorphic as a graph to the standard Cayley graph of $\mathbb{Z}$, then G admits detours.*

Proof. If G is a Cayley graph of a group which is not isomorphic to $\mathbb{Z}$ or $\mathbb{Z}/2 * \mathbb{Z}/2$, this follows from Proposition 4.5.4, since nilpotent groups have nontrivial center. If G is a Cayley graph of $\mathbb{Z}$, this follows from Proposition 4.5.2. If G is a Cayley graph of $\mathbb{Z}/2 * \mathbb{Z}/2$, this follows from Proposition 4.5.3. □

4.6. Proof of Theorem 4.1.3

In this section, we prove the following:

Theorem 4.1.3. *Let G be a bounded degree graph which is quasi-isometric to a tree. Then G is vdBK if and only if G admits detours. In fact, if G admits detours, then even if ν is not exponential-subcritical, whenever $\tilde{\nu}$ is strictly more variable than ν, we have $\mathbb{E}\tilde{T} \ll \mathbb{E}T$.*

A metric space which is quasi-isometric to a tree (where the tree is given the usual graph metric) is called a *quasi-tree*. The following is a well-known equivalent condition for a geodesic metric space to be a quasi-tree (the original, slightly weaker condition is due to Manning [28]; the following extension is a well-known consequence, see e.g. [6]):

Theorem 4.6.1 (Manning's bottleneck criterion)**.** *A geodesic metric space X is a quasi-tree if and only if there exists some $\Delta < \infty$ such that for every $x, y \in X$, for every geodesic $[x, y]$ from x to y, for every $z \in [x, y]$, any path π from x to y intersects $B(z, \Delta)$.*

Corollary 4.6.1. *Let $G = (V, E)$ be a graph which is a quasi-tree. Then there exists $R < \infty$ such that for any $x, y \in V$, for any edge geodesic $[x, y]$ from x to y and any $z \in V([x, y])$, every path π from x to y intersects $E(B(z, R))$.*

Here (and later in this chapter), if $S \subset V$, then $E(S) \subset E$ is defined to be the set of edges of G which have both endpoints lying in S, and if $S \subset E$, then $V(S) \subset V$ is defined to be the set of vertices which are an endpoint of an edge in S.

Proof. (V, d) is naturally a subspace of the geodesic metric space (G, d) given by the geometric realization of G (i.e. the 1-dimensional metric cell complex where each $e \in E$ corresponds to 1-cell in G isometric to $[0, 1]$, joining 0-cells corresponding to the endpoints of e). The combinatorial edge-geodesics we study here correspond to geodesics in (G, d), and one quickly sees that the corollary holds with $R = \Delta + 1$. □

Proof of Theorem 4.1.3. We only need to prove that, if G admits detours, then we have $\mathbb{E}\tilde{T} \ll \mathbb{E}T$ whenever $\tilde{\nu}$ is strictly more variable than ν, since the other direction is given by Theorem 4.4.1.

To this end, let $\nu, \tilde{\nu}$ have finite mean with $\tilde{\nu}$ strictly more variable than ν and first assume that (4.3.1) holds. Then let $\epsilon > 0, I_0$ be given as in Lemma 4.3.2. Since G admits detours, there is some C such that every self-avoiding path π of length C admits a self-avoiding ϵ-detour (which is necessarily of length at most $(1+\epsilon)C$). Since G is a quasi-tree, take $R < \infty$ such that for all $x, y \in V$, for any geodesic $[x, y]$ from x to y, any path $\pi : x \to y$ intersects $E(B(z, R))$ for all $z \in V([x, y])$.

Now, define the family $\{B_v := B(v, R + C(2+\epsilon)) : v \in V\}$. First we claim that for any v,

$$\begin{aligned}&\mathbb{P}(B_v \text{ contains a feasible pair for the geodesic } \pi : x \to y)\\ \geq &\mathbb{P}(\pi \text{ visits } B(v, R) \text{ and leaves } B_v, w(e) \in I_0 \text{ for all } e \in E(B_v)).\end{aligned}$$

To see this, note that if π visits $B(v, R)$ and exits B_v, there is a segment α of π of length at least C contained in $B(v, R + C)$; this segment admits a self-avoiding ϵ-detour γ contained in $B(v, R + C(2+\epsilon))$. Then, if also $w(e) \in I_0$ for all $e \in E(B_v)$, (α, γ) forms a feasible pair.

Next note that if $v \in V([x, y]) \setminus B_y$ then *any* path from x to y visits $B(v, R)$ and exits B_v, and so for such v we have

$$\begin{aligned}&\mathbb{P}(B_v \text{ contains a feasible pair for the geodesic } \pi : x \to y)\\ \geq &\mathbb{P}(w(e) \in I_0 \text{ for all } e \in E(B_v)) = \nu(I_0)^{|E(B_v)|} \geq (\nu(I_0))^{(D+1)^{R+C(2+\epsilon)+1}} =: c\end{aligned}$$

where D is the maximum degree of G. Therefore for $d(x, y)$ sufficiently large we have

$$\begin{aligned}
&\sum_{v \in V} \mathbb{P}(B_v \text{ contains a feasible pair for the geodesic } \pi : x \to y) \\
\geq &\sum_{v \in V([x,y]) \setminus B_y} \mathbb{P}(w(e) \in I_0 \text{ for all } e \in E(B_v)) \\
\geq &|V([x, y]) \setminus B_y| c \geq cd(x, y) - c(D + 1)^{R+C(2+\epsilon)} \gtrsim d(x, y).
\end{aligned}$$

Moreover, we have that

$$\sup_v \#\{w : B_w \cap B_v \neq \emptyset\} \leq \sup_v |B(v, 2(R + C(2 + \epsilon)))| \leq (D + 1)^{2(R+C(2+\epsilon))} < \infty,$$

and so by Lemma 4.3.4 we have that

$$\liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x, y) - \mathbb{E}\tilde{T}(x, y)}{d(x, y)} > 0,$$

as desired.

On the other hand, if w and $\tilde{w}$ do not satisfy (4.3.1), then take $\bar{w}$ as in Lemma 4.3.1; applying our above argument to $\bar{w}$ gives

$$\liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x, y) - \mathbb{E}\tilde{T}(x, y)}{d(x, y)} \geq \liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x, y) - \mathbb{E}\bar{T}(x, y)}{d(x, y)} > 0,$$

and so we are done. □

As a corollary we also obtain Theorem 4.1.4:

Proof of Theorem 4.1.4. Let Γ be a virtually free group (i.e. Γ contains a finite index free subgroup). Since free groups have Cayley graphs which are regular trees, any Cayley graph of

Γ is quasi-isometric to a regular tree, and so by Theorem 4.1.3 a Cayley graph of Γ is vdBK if and only if it admits detours. If Γ has a finite index subgroup with nontrivial center and is not isomorphic to $\mathbb{Z}$ or $\mathbb{Z}/2 * \mathbb{Z}/2$, then by Proposition 4.5.4, all its Cayley graphs admit detours. If Γ has a finite normal subgroup, then by Proposition 4.5.1, all its Cayley graphs admit detours; hence under either condition all Cayley graphs of Γ are vdBK.

If Γ is a semidirect product $F \rtimes F_k$ then F is a nontrivial finite normal subgroup, and hence its Cayley graphs are vdBK by the above. If Γ contains $F \times F_k$ as a finite index subgroup, let $A \leq F$ be a nontrivial cyclic (hence abelian) subgroup; then $A \times F_k$ has nontrivial center (containing $A \times \{1\}$) and $A \times F_k$ is finite index in Γ, and so the Cayley graphs of Γ are vdBK by above. □

4.7. Proof of Theorem 4.1.2

We say that a graph has *strict polynomial growth* if there exists some $0 < d < \infty$ and some $0 < c_1 \leq C_1 < \infty$ such that for all $R \geq 1$,

$$c_1 R^d \leq \inf_{v \in V} |B(v, R)| \leq \sup_{v \in V} |B(v, R)| \leq C_1 R^d.$$

(Note that this in particular entails that G has bounded degree). It is well known that Cayley graphs of finitely generated virtually nilpotent groups are of strict polynomial growth; in fact, $|B(R)|/R^d$ converges to a constant as $R \to \infty$ [30]. Moreover, it is clear that half-planes, sectors, and many other subgraphs of the standard Cayley graph of $\mathbb{Z}^d$ have strict polynomial growth.

The goal of this section is to prove the following:

Theorem 4.1.2. *Let G be a graph of strict polynomial growth. Then G is vdBK if and only if G admits detours.*

Given Theorem 4.1.2, we can quickly prove Theorem 4.1.1 as follows:

Proof of Theorem 4.1.1 given Theorem 4.1.2. Let Γ be a virtually nilpotent group. Then any Cayley graph for Γ has strict polynomial growth [30]. If a Cayley graph G of Γ is not isomorphic as a graph to the standard Cayley graph of $\mathbb{Z}$, then G admits detours, by Theorem 4.5.1. So Theorem 4.1.2 implies that G is vdBK. □

4.7.1. A Peierls argument for graphs of strict polynomial growth

In the previous case, where G was quasi-isometric to a tree, we benefitted from the fact that we could find areas which the geodesic visited with probability 1, and hence for such an area A and any event C, $\mathbb{P}(\{\pi \text{ visits } A\} \cap C) = \mathbb{P}(C)$. Here we have to deal with graphs which may have many paths with mostly disjoint support between each pair of points, and so a generic event C will not have $\mathbb{P}(\{\pi \text{ visits } A\} \cap C) = \mathbb{P}(C)$. In fact, there are very few events C for which we can get nontrivial inequalities for $\mathbb{P}(\{\pi \text{ visits } A\} \cap C)$. Therefore, a key tool will be Lemma 4.7.1 below, which ensures that, for certain types of local events, the geodesic will in expectation visit $\gtrsim d(x, y)$ many regions where the prescribed events hold. This is a Peierls-type argument, similar to the one used in [36], but with a different choice of coarse-graining.

Before we can state the lemma, we must describe a specific coarse-graining construction which has desirable properties if our graph G has strict polynomial growth. First, for each R, choose a maximal subset $\{o_i^R\}_i \subset V$ which is R-separated, that is, such that if $i \neq j$, then $d(o_i^R, o_j^R) \geq R$. (From here on we will suppress the dependence on R and write o_i for o_i^R). Maximality implies that for each vertex $v \in V$, there exists some i such that $d(o_i, v) \leq R$. Also fix an arbitrary well-ordering on the indices i, and for each i, let B_i^R be the "Voronoi tile" containing o_i, that is, set

$$B_i^R := \{v \in V : d(o_i, v) < d(o_j, v) \text{ for all } j < i, d(o_i, v) \leq d(o_j, v) \text{ for all } j \geq i\}.$$

(B_i^R consists of the vertices which are closer to o_i than any other o_j, but we "break ties" when v is equidistant from o_i and o_j using the ordering on indices). We see that $V = \bigsqcup_i B_i^R$ and that $\sup_i \operatorname{diam} B_i^R \leq 2R$ (since each $B_i^R \subset B(o_i, R)$). That is, $\{B_i^R\}$ forms a partition of V, and the partition elements have bounded diameter.

Next, we fix $0 < \Sigma < \infty$ (a scaling parameter that will be chosen to suit our separate constructions below). The following proposition says that the vertex sets B_i^R and the edge sets $E(B(o_i, \Sigma R))$ have "degree" uniformly bounded in R if G has strict polynomial growth:

Proposition 4.7.1. *Suppose that G has strict polynomial growth. Then there exists $D < \infty$ independent of R such that the following holds. For each R let G^R be the simple graph whose vertex set is $\{o_i\}$ and is such that $o_i \sim o_j$ if and only if there is an edge in G with one endpoint in B_i^R and the other endpoint in B_j^R. Let $\tilde{G}^R$ be the simple graph whose vertex set is $\{o_i\}$ and such that $o_i \sim o_j$ if and only if $E(B(o_i, \Sigma R)) \cap E(B(o_j, \Sigma R)) \neq \emptyset$. Then for all sufficiently large R, both G^R and $\tilde{G}^R$ have degree at most D.*

Proof. Fixing some o_i, we have

$$\{j : o_j \sim o_i \text{ in } G^R\} \subset \{j : d(o_i, o_j) \leq 2R + 1\} \subset \{j : B_j^R \subset B(o_i, 3R + 1)\},$$

as well as

$$\{j : o_j \sim o_i \text{ in } \tilde{G}^R\} \subset \{j : d(o_i, o_j) \leq 2\Sigma R\} \subset \{j : B_j^R \subset B(o_i, (2\Sigma + 1)R\}.$$

Thus in order to bound both degrees it suffices to show that, given any constant Σ', the quantity

$$\#\{j : B_j^R \subset B(o_i, \Sigma' R)\}$$

is uniformly bounded in both i and R. To this end, note that, since o_i is R-separated, it follows that $B(o_i, (R/2) - 1) \subset B_i^R$. So using our volume bounds and the fact that the B_j^R are disjoint we have

$$\#\{j : B_j^R \subset B(o_i, \Sigma' R)\} c_1 \left[\frac{R}{2} - 1\right]^d \leq \sum_{B_j^R \subset B(o_i, \Sigma' R)} |B_j^R| \leq |B(o_i, \Sigma' R)| \leq C_1 (\Sigma' R)^d,$$

so that

$$\#\{j : B_j^R \subset B(o_i, \Sigma' R)\} \leq \frac{C_1 (\Sigma' R)^d}{c_1 [(R/2) - 1]^d} \xrightarrow[R \to \infty]{} \frac{C_1}{c_1} (2\Sigma')^d,$$

so we are done. □

Remark 4.7.1. *This is actually the only point in the proof where we use strict polynomial growth. In every other part of the proof, we will only use that G has a uniform strictly subexponential volume bound and bounded degree (which is equivalent to a uniform bound on $|B(v, 1)|$). If one could find a suitable coarse-graining for more general subexponential growth graphs, the methods in this chapter would immediately show that such graphs are vdBK if and only if they admit detours. However, constructing such a coarse-graining would take some ingenuity; if for instance we attempt to do the Voronoi construction for a graph with growth of order $e^{\sqrt{R}}$, the degree bounds we get from the above analysis are superpolynomial in R, and the proof of Lemma 4.7.1 does not go through.*

Now we can state and prove the Peierls lemma:

Lemma 4.7.1. *Let $G = (V, E)$ be a graph of strict polynomial growth. Suppose that for each sufficiently large $R < \infty$ we have a family of events $\{A_i^R\}$ which depend only on the edges in $E(B(o_i, \Sigma R))$. Suppose also that*

$$\rho(R) := \sup_i \mathbb{P}\left((A_i^R)^c\right) \xrightarrow[R\to\infty]{} 0.$$

Then, for all sufficiently large R, there exist $c_2(R), \epsilon_2(R) > 0$ such that for all $x, y \in V$ with $d(x, y)$ sufficiently large,

$$\mathbb{P}\begin{pmatrix} \exists\gamma : x \to y \text{ visiting at most } c_2 d(x,y) \text{ distinct} \\ B_i^R \text{ such that } A_i^R \text{ holds} \end{pmatrix} \le e^{-\epsilon_2 d(x,y)}$$

Proof. Recall the graphs G^R and $\tilde{G}^R$ defined in Proposition 4.7.1. Let γ be a path from x to y in G. This induces a path $\tilde{\gamma}$ in G^R in a natural way: $\tilde{\gamma}$ starts at the o_{i_1} corresponding to the unique $B_{i_1}^R$ containing x, and each time γ crosses an edge from a vertex in some B_i^R to a vertex in some distinct $B_{i'}^R$, $\tilde{\gamma}$ crosses an edge from o_i to $o_{i'}$. Note that since the diameter of the B_i^R is bounded uniformly in i, there exists $\eta(R) > 0$ such that if $\gamma : x \to y$, then $\tilde{\gamma}$ visits at least $\eta d(x, y)$ distinct B_i^R. We want to bound the probability that (for some $c_2(R)$ to be chosen later) some such $\tilde{\gamma}$ visits at most $c_2 d(x, y)$ o_i such that A_i^R holds. First, note that if $\tilde{\gamma}$ visits at most $c_2 d(x, y)$ o_i such that A_i^R holds, a self-avoiding path obtained from $\tilde{\gamma}$ from erasing loops has the same property. So if $B_{i_1}^R$ is the unique tile containing x and $B_{i_2}^R$ the unique tile containing y, it suffices to bound the probability that some self-avoiding path $\tilde{\gamma}$ in G^R which starts at o_{i_1} and ends at o_{i_2} visits at most $c_2 d(x, y)$ o_i such that A_i^R holds; to reduce clutter, let us write A_i instead of A_i^R.

Now, for a fixed self-avoiding path $\tilde{\gamma}$ visiting k distinct o_i, we have

$$\mathbb{P}\left(\begin{array}{c}\tilde{\gamma} \text{ visits at most } c_2 d(x,y) \\ o_i \text{ such that } A_i \text{ holds}\end{array}\right) \leq \sum_{S\subset V(\tilde{\gamma}),|S|=k-c_2 d(x,y)} \mathbb{P}\left(\bigcap_{o_i\in S} A_i^c\right).$$

Since $\tilde{G}^R$ has degree bounded by D (where D is as in Proposition 4.7.1) each such $S \subset \{o_i\}_i$ contains a subset S' which is independent in $\tilde{G}^R$ (that is, no two elements of S' are joined by an edge of $\tilde{G}^R$) and which has size at least $|S'| \geq \frac{1}{D+1}|S|$. From the definition of $\tilde{G}^R$ we see that if S' is an independent set in $\tilde{G}^R$ then the collection of events $\{A_i^R\}_{o_i\in S'}$ is independent. Hence the above is bounded by

$$\sum_{\substack{S\subset V(\tilde{\gamma}),\\ |S|=k-c_2 d(x,y)}} \mathbb{P}\left(\bigcap_{o_i\in S'} A_i^c\right) = \sum_{\substack{S\subset V(\tilde{\gamma}),\\ |S|=k-c_2 d(x,y)}} \prod_{o_i\in S'} \mathbb{P}(A_i^c) \leq \binom{k}{c_2 d(x,y)} \rho^{\frac{k-c_2 d(x,y)}{D+1}}.$$

On the other hand, the number of self-avoiding paths of length k in G^R starting at o_{i_1} is at most D^k (since G^R has degree at most D). Thus we have

$$\begin{aligned}\mathbb{P}\left(\begin{array}{c}\exists\gamma : x\to y \text{ visiting at most } c_2 d(x,y) \text{ distinct} \\ B_i^R \text{ such that } A_i^R \text{ holds}\end{array}\right) &\leq \sum_{k=\lceil\eta d(x,y)\rceil}^{\infty} D^k \binom{k}{c_2 d(x,y)} \left(\rho^{\frac{1}{D+1}}\right)^{k-c_2 d(x,y)} \\ &\leq \left(\rho^{\frac{1}{D+1}}\right)^{-c_2 d(x,y)} \sum_{k=\lceil\eta d(x,y)\rceil}^{\infty} \left(2D\rho^{\frac{1}{D+1}}\right)^k;\end{aligned}$$

for R sufficiently large we have $2D\rho^{\frac{1}{D+1}} < 1$, and then the right hand side above is equal to

$$\left(\rho^{\frac{1}{D+1}}\right)^{-c_2 d(x,y)} \left(2D\rho^{\frac{1}{D+1}}\right)^{\lceil\eta d(x,y)\rceil} \cdot \frac{1}{1-2D\rho^{\frac{1}{D+1}}}.$$

If we choose $c_2 > 0$ sufficiently small that

$$-c_2 \log(\rho^{1/(D+1)}) + \eta \log(2D\rho^{1/(D+1)}) > 0$$

then our upper bound decays exponentially in $d(x, y)$, and so we are done. □

Remark 4.7.2. *If we assume that $\rho(R) \leq Ce^{-cR}$ (as will be the case in the proof of Theorem 4.1.2), then Lemma 4.7.1 holds for slightly larger class of graphs, i.e. those which have*

$$R^{d'} \lesssim \inf_v |B(v, R)| \leq \sup_v |B(v, R)| \lesssim R^d$$

for some d, d' with $d - d' < 1$. This is because for such graphs, the proof of Proposition 4.7.1 shows that G^R and $\tilde{G}^R$ have degree bounded by $D(R) = o(R)$, and then the proof of Lemma 4.7.1 goes through. This allows us to extend Theorem 4.1.2 to such graphs. However, it is difficult to come up with a natural example of a graph which has such a property but is not already of strict polynomial growth.

Lastly, let us use Lemma 4.7.1 to prove the following, which will be very important to our later constructions.

Lemma 4.7.2. *Let G be a graph with strict polynomial growth, and suppose that ν is exponential-subcritical. Then, there exist $q, c > 0$ such that, for all $x, y \in V$ with $d(x, y)$ sufficiently large,*

$$\mathbb{P}(T(x, y) < (\inf +q)d(x, y)) \leq e^{-cd(x,y)}.$$

Remark 4.7.3. *The conclusion of the above lemma also holds for* any *graph G of degree at most D if one assumes $\nu(\{\inf\}) < 1/D$; this is proved in the course of proving Lemma A.1 in [34].*

Proof. First, suppose $\inf = 0$; since ν is exponential-subcritical, $\nu(\{0\}) < \underline{p_c}$, and we can pick $q' > 0$ sufficiently small that if $\nu([\inf, \inf + q']) < \underline{p_c}$. Then, by the definition of $\underline{p_c}$, there is some $c' > 0$ such that for any R sufficiently large, for any $v \in V$,

$$\mathbb{P}(v \text{ is connected to } B(v,R)^c \text{ by a path of edges which each have weight} < \inf + q') \leq e^{-c'R}.$$

In particular, for any $\Sigma \geq 2$, we have that

$$\begin{aligned}
&\mathbb{P}(\exists p \in S(v, \Sigma R), x \in B(v,R), \text{ path } \alpha : p \to x \text{ in } B(v, \Sigma R) \text{ s.t. } w(e) < \inf + q' \text{ for all } e \in \alpha) \\
&\quad \leq \mathbb{P}(\exists x \in B(v,R), p' \in S(x, (\Sigma - 1)R), \text{ path } \alpha : x \to p' \text{ s.t. } w(e) < \inf + q' \text{ for all } e \in \alpha) \\
&\quad \leq |B(v,R)| e^{-c'(\Sigma-1)R} \leq C_1 R^d e^{-c'(\Sigma-1)R} \xrightarrow[R\to\infty]{} 0.
\end{aligned}$$

In particular,

$$\inf_{v\in V} \mathbb{P}\begin{pmatrix} \text{all paths from } S(v, \Sigma R) \text{ to } B(v,R) \\ \text{contain at least one edge of weight } \geq \inf + q' \end{pmatrix} \xrightarrow[R\to\infty]{} 1,$$

and so by Lemma 4.7.1, for all sufficiently large R, there exist $c_2(R) > 0, \epsilon_2(R) > 0$ such that for all sufficiently large $d(x,y)$,

$$\mathbb{P}\left(\begin{array}{l}\exists\gamma : x \to y \text{ visiting at most } c_2 d(x,y) \text{ distinct } B_i \text{ such that} \\ \text{all paths from } S(o_i, \Sigma R) \text{ to } B(o_i, R) \text{ contain at least one} \\ \text{edge of weight } \geq \inf + q'\end{array}\right) \leq e^{-\epsilon_2 d(x,y)}.$$

Now, each $B(o_i, \Sigma R)$ intersects at most D' other $B(o_i', \Sigma R)$ by Proposition 4.7.1, and so if a path γ visits at least $c_2 d(x,y)$ B_i such that all paths from $S(o_i, \Sigma R)$ to $B(o_i, R)$ contain at least edge with weight at least $\inf + q'$, there is some collection of at least $\frac{1}{D'+1} c_2 d(x,y)$ *disjoint* $B(o_i, \Sigma R)$ with this property such that γ visits B_i. If $x \notin B(o_i, \Sigma R)$ then in particular γ starts outside of $B(o_i, \Sigma R)$ and so since γ visits $B_i \subset B(o_i, R)$, some subpath of γ joins $S(o_i, \Sigma R)$ to $B(o_i, R)$ and so some edge of $\gamma \cap E(B(o_i, \Sigma R))$ has weight at least $\inf + q'$. So by disjointness we conclude that γ has at least $\frac{c_2}{D'+1} d(x,y) - 1$ edges of length at least $\inf + q'$, and so

$$T(\gamma) \geq (\inf) d(x,y) + q' \left(\frac{c_2}{D'+1} d(x,y) - 1 \right)$$

in this case. So taking $q := q' c_2 / (2(D'+1))$, we see that whenever $d(x,y) \geq 2(D'+1)/c_2$ we have

$$\mathbb{P}(T(\gamma) < (\inf + q) d(x,y)) \leq e^{-\epsilon_2 d(x,y)},$$

and the lemma follows.

Now, suppose that $\inf > 0$. Then choose $q' > 0$ such that $\nu([\inf, \inf + q']) < \underline{\vec{p_c}}$ to obtain $c' > 0$ such that for any R sufficiently large, for any $v \in V$ we have

$$\mathbb{P}\left(\begin{array}{c} v \text{ is connected to } B(v,R)^c \text{ by an edge-geodesic path} \\ \text{of edges which each have weight} < \inf + q' \end{array} \right) \leq e^{-c'R}.$$

Then arguing similarly as above, by Lemma 4.7.1, for all sufficiently large R, there exist $c_2(R) > 0, \epsilon_2(R) > 0$ such that for all sufficiently large $d(x, y)$,

$$\mathbb{P}\left(\begin{array}{l} \exists \gamma : x \to y \text{ visiting at most } c_2 d(x,y) \text{ distinct } B_i \text{ such that} \\ \text{all edge-geodesic paths from } S(o_i, \Sigma R) \text{ to } B(o_i, R) \text{ contain} \\ \text{at least one edge of weight } \geq \inf + q' \end{array} \right) \leq e^{-\epsilon_2 d(x,y)}.$$

Similar to above, we then see that (except on an exponentially small event) every path γ from x to y contains at least $\frac{c_2}{D'+1} d(x, y) - 1$ disjoint subpaths which are either not edge-geodesic, or contain an edge of weight at least $\inf + q'$. Each such subpath γ_i has passage time $T(\gamma_i) \geq (\inf)|\gamma_i| + \min(\inf, q')$. So taking $q := \min(q', \inf) c_2/(2(D' + 1)) > 0$ and $c = \epsilon_2$ gives the lemma.

□

Remark 4.7.4. *This is the only part of the proof where we use the exponential subcriticality of ν.*

Remark 4.7.5. *This lemma implies in particular that if G is a Cayley graph of a finitely generated virtually nilpotent group and if $\nu(\{0\}) < p_c$, then there exists $a > 0$ such that for all $x, y \in V$, $\mathbb{E}T(x, y) \geq a d(x, y)$. This means, for instance, that the results of [5] giving the existence of a scaling limit apply when ν has an exponential moment and $\nu(\{0\}) < p_c$ (a weaker condition than the condition $\nu(\{0\}) < 1/D$ quoted in that paper).*

4.7.2. Proof strategy: a resampling scheme

Note that if we have any family of events A_i^R as in Lemma 4.7.1, for fixed R sufficiently large, we have that in particular

$$\begin{aligned}
&\sum_i \mathbb{P}(\{\text{the geodesic } \pi : x \to y \text{ visits } B_i^R\} \cap A_i^R) \\
=&\mathbb{E}[\# B_i^R \text{ such that } \pi \text{ visits } B_i^R \text{ and } A_i^R \text{holds}] \\
\geq&(c_2 d(x,y))\mathbb{P}(\pi \text{ visits at least } c_2 d(x,y)\ B_i^R \text{ such that } A_i^R \text{ holds}) \\
\geq&(c_2 d(x,y))(1 - e^{-\epsilon_2 d(x,y)}) \gtrsim d(x,y).
\end{aligned}$$

We will say that π *crosses* $B(o_i, \Sigma R)$ if π starts at a vertex outside $B(o_i, \Sigma R)$, ends at a vertex outside $B(o_i, \Sigma R)$, and visits B_i. Since the number of o_i such that $x \in B(o_i, \Sigma R)$ or $y \in B(o_i, \Sigma R)$ is bounded independent of x and y, we see also from the above that

$$\sum_i \mathbb{P}(\{\text{the geodesic } \pi : x \to y \text{ crosses } B(o_i, \Sigma R)\} \cap A_i^R) \gtrsim d(x,y).$$

Thus, if we find a family of events $\{A_i^R\}$ such that for each i, $\mathbb{P}(B(o_i, \Sigma R) \text{ contains a feasible pair})$ is at least a positive constant (independent of x, y, i, but possibly depending on R) times $\mathbb{P}(\{\pi \text{ crosses } B(o_i, \Sigma R)\} \cap A_i^R)$, we will have

$$\sum_i \mathbb{P}(B(o_i, \Sigma R) \text{ contains a feasible pair for } \pi : x \to y) \gtrsim d(x,y),$$

and hence by Lemma 4.3.4 we will have our theorem. (Note that here the role of the B_i from Lemma 4.3.4 is played by $B(o_i, \Sigma R)$, not the Voronoi tiles B_i^R we defined in the last section).

We will obtain a bound of the form

$$\mathbb{P}(B(o_i, \Sigma R) \text{ contains a feasible pair}) \geq c(R)\mathbb{P}(\{\pi \text{ crosses } B(o_i, \Sigma R)\} \cap A_i^R)$$

by introducing a resampling scheme, as in [36]. Explicitly, fix some o_i; throughout the rest of the chapter, we abbreviate $B(s) := B(o_i, s)$. Define new random weights $w^* : E \to [0, \infty)$ as follows: $w^*|_{E(B(\Sigma R)^c} = w|_{E(B(\Sigma R)^c}$, but the $w^*(e), e \in E(B(\Sigma R)$ are i.i.d. ν-distributed random variables, also independent of w. (Recall that for $S \subset V$, we define $E(S) \subset E$ to be the set of edges of G with endpoints lying in S). Note that w and w^* are equal in distribution. For each R we will define a w-measurable random set of configurations $E_w \subset [0, \infty)^{E(B(\Sigma R))}$ such that

$$\{\pi \text{ crosses } B(\Sigma R)\} \cap A_i^R \cap \{w^*|_{E(B(\Sigma R))} \in E_w\} \subset \{B(\Sigma R) \text{ contains a feasible pair for } \pi^*\}, \tag{4.7.1}$$

where π is the T-geodesic from x to y and π^* is the T^*-geodesic from x to y. To reduce clutter, let us abbreviate the event $\{w^*|_{E(B(\Sigma R)} \in E_w\}$ by $\{w^* \in E_w\}$. If in addition we ensure that the conditional probability $\mathbb{P}(w^* \in E_w | w) \geq c(R) > 0$ on the event $\{\pi \text{ crosses } B(\Sigma R)\} \cap A_i^R$ (where $c(R)$ is some non-random constant), we get

$$\begin{aligned}
\mathbb{P}(B(\Sigma R) \text{ contains a feasible pair for } \pi) &= \mathbb{P}(B(\Sigma R) \text{ contains a feasible pair for } \pi^*) \\
&\geq \mathbb{P}(\{\pi \text{ crosses } B(o_i, \Sigma R)\} \cap A_i^R \cap \{w^* \in E_w\}) \\
&= \mathbb{E}\left[\mathbb{1}_{\{\pi \text{ crosses } B(o_i, \Sigma R)\} \cap A_i^R\}} \mathbb{E}[\mathbb{1}_{\{w^* \in E_w\}} | w]\right] \\
&\geq c(R)\mathbb{P}(\{\pi \text{ crosses } B(o_i, \Sigma R)\} \cap A_i^R),
\end{aligned}$$

as desired. The discussion in this section is summarized in following proposition:

Proposition 4.7.2. *Suppose there exist w-measurable events A_i^R satisfying the conditions of Lemma 4.7.1 and w-measurable random sets of configurations E_w such that for sufficiently large R (4.7.1) holds and $\mathbb{P}(w^* \in E_w|w) \geq c(R)$ on the event $\{\pi$ crosses $B(\Sigma R)\} \cap A_i^R$, where $c(R) > 0$ is a constant depending only on $R, \nu, \tilde{\nu}$, and G. Then*

$$\liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y) - \mathbb{E}\tilde{T}(x,y)}{d(x,y)} > 0.$$

Thus the meat of the proof of Theorem 4.1.2 consists of performing a "geometric" construction to obtain suitable A_i^R and E_w.

4.7.3. Geometric construction: bounded case

First, suppose that ν has bounded support. We want to construct A_i^R and E_w satisfying the hypotheses of Proposition 4.7.2. Denote by inf the infimum of the support of ν and denote by sup the supremum of the support of ν. Assume that (4.3.1) holds, and then choose $\epsilon > 0, y_0, I_0$ as in Lemma 4.3.2. Assuming that G admits detours, let C be such that every self-avoiding path of length C admits a self-avoiding ϵ-detour. Set $C' := C(3 + 2\epsilon)$. Assume that ν is exponential-subcritical, and then let $q > 0$ be the parameter given by Lemma 4.7.2. Denote by D the maximum degree of G.

First, let us consider the case that $y_0 = \sup$; in fact this allows us to do a much simpler construction. In this case, choose $\Sigma > 2$ large enough that $(\inf + q)\left(1 - \frac{1}{\Sigma}\right) > \inf$ and choose $\delta > 0$ such that $\inf + \delta < (\inf + q)\left(1 - \frac{1}{\Sigma}\right)$. Choose a sequence $\delta_{\sup}(R) \xrightarrow[R\to\infty]{} 0$ such that for each R $\nu([\sup - \delta_{\sup}(R), \sup]) > 0$ but $\lim_{R\to\infty} \nu([\inf, \sup - \delta_{\sup}(R)])^{DC_1(\Sigma R)^d} = 1$. Then let $A_i^R := A_1 \cap A_2$

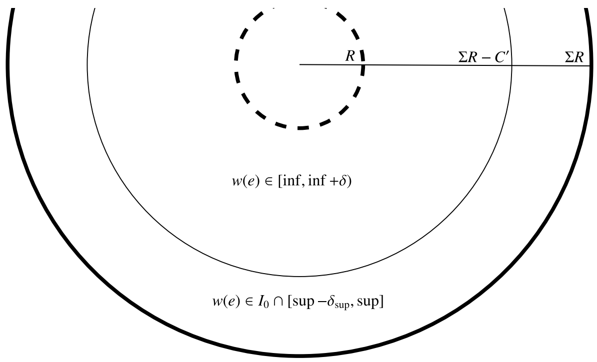

Figure 4.2. A schematic diagram of the prescribed set of configurations E_w in the case that ν has bounded support and $y_0 = \sup$.

where A_1 and A_2 are as follows:

$$A_1 := \left\{ \begin{array}{c} \text{For all vertices } v, w \in B(\Sigma R) \text{ with } d(v,w) \geq R, \\ \text{all paths } \gamma \text{ from } v \text{ to } w \text{ in } B(\Sigma R) \text{ have } T(\gamma) \geq (\inf + q) d(v,w) \end{array} \right\}.$$

$$A_2 := \left\{ w(e) \leq \sup - \delta_{\sup} \text{ for all } e \in E(B(\Sigma R)) \right\}.$$

We see that both events only depend on the weights of edges in $B(\Sigma R)$, by choice of $\delta_{\sup}(R)$ we have $\mathbb{P}(A_2) \xrightarrow[R\to\infty]{} 1$, and by Lemma 4.7.2 we have that for sufficiently large R

$$\mathbb{P}(A_1^c) \leq \sum_{\substack{v,w \in B(\Sigma R), \\ d(v,w) \geq R}} \mathbb{P}(T(v,w) < (\inf + q) d(v,w)) \leq (C_1 R^d)^2 e^{-\epsilon_2 R} \xrightarrow[R\to\infty]{} 0$$

uniformly in i, so the hypotheses of Lemma 4.7.1 are satisfied.

Now in this case set of configurations E_w does not actually depend on w; we simply set

$$E_w := \left\{ \omega \in [0,\infty)^{E(B(\Sigma R))} : \omega(e) \in \begin{array}{ll} [\inf, \inf + \delta) & \text{if } e \in E(B(\Sigma R - C')), \\ [\sup, \sup - \delta_{\sup}] \cap I_0 & \text{otherwise} \end{array} \right\}.$$

Let us show that, for sufficiently large R, on the event $\{\pi \text{ crosses } B(\Sigma R)\} \cap A_1 \cap A_2 \cap \{w^* \in E_w\}$, $B(\Sigma R)$ contains a feasible pair for any T^*-geodesic.

For a subset $S \subset E$, denote by $T_S(p,q)$ the infimal weight of a path from p to q which only uses edges lying in S. First, let a and b be points of $S(\Sigma R)$ such that $T_{E(B(\Sigma R))^c}(x,a)$ and $T_{E(B(\Sigma R))^c}(b,y)$ are infimal. Fix a T-geodesic $\alpha \subset E(B(\Sigma R))^c$ from x to a, an edge geodesic $[a, o_i]$ from a to o_i, an edge-geodesic $[o_i, b]$ from o_i to b, and a T-geodesic $\beta \subset E(B(\Sigma R))^c$ from b to y, and define $\pi' := \alpha * [a, o_i] * [o_i, b] * \beta$. We claim that $T^*(\pi') < T(\pi)$ when R is sufficiently large. To see this, first note that, if v and w are the first and last vertices of π lying on $S(\Sigma R)$, we have

$$T^*(\pi'_{x,a}) + T^*(\pi'_{b,y}) = T(\pi'_{x,a}) + T(\pi'_{b,y}) \leq T(\pi_{x,v}) + T(\pi_{w,y}),$$

where here and elsewhere, for a path γ and vertices $p, q \in V(\gamma)$, $\gamma_{p,q}$ denotes the subpath of γ starting at p and ending at q.

Next, since π crosses B_i, $\pi_{v,w}$ contains at least two subsegments connecting $S(\Sigma R)$ and $S(R)$, and so since A_1 holds we have

$$T(\pi_{v,w}) \geq 2(\inf + q)(\Sigma - 1)R = (\inf + q)\left(1 - \frac{1}{\Sigma}\right)2\Sigma R,$$

while if $w^* \in E_w$, we have

$$T^*(\pi'_{a,b}) \leq 2\Sigma R(\inf + \delta) + (\sup)C'.$$

Since by construction $\inf + \delta < (\inf + q)\left(1 - \frac{1}{\Sigma}\right)$ and $(\sup)C' = o(R)$, for sufficiently large R we have $T^*(\pi') < T(\pi)$.

Now, consider a T^*-geodesic π^* from x to y. On our event, we have $w^* \geq w$ on $E(B(\Sigma R - C'))^c$, so if π^* did not intersect $E(B(\Sigma R - C'))$, we would have $T^*(\pi^*) \geq T(\pi) > T^*(\pi')$, a contradiction. Thus, π^* must visit $B(\Sigma R - C')$. In particular, it contains a subpath connecting $S(\Sigma R)$ and $S(\Sigma R - C')$, and so a subpath connecting $S(\Sigma R - C(1+\epsilon))$ and $S(\Sigma R - C' + C(1+\epsilon))$, which must have length at least $C' - 2C(1+\epsilon) = C$. Choose a self-avoiding ϵ-detour γ for such a segment. Since γ has length at most $C(1+\epsilon)$, it is contained in $E(B(\Sigma R)) \setminus E(B(\Sigma R - C'))$. But since $w^* \in E_w$, this means that all the edges of both γ and the subsegment of π^* have weights in I_0. Hence $B(\Sigma R)$ contains a feasible pair for π^*.

Furthermore, since $y_0 = \sup$, by the construction of I_0 we have

$$\mathbb{P}(w^* \in E_w) \geq \min\Big(\nu([\inf, \inf + \delta)), \nu([\sup - \delta_{\sup}, \sup] \cap I_0)\Big)^{DC_1(\Sigma R)^d} > 0$$

independent of o_i, so both hypotheses of Proposition 4.7.2 hold.

Now we suppose that $y_0 < \sup$ and do a different construction of the A_i^R and E_w. Again take $\epsilon, y_0, I_0, C, C', q$ as above. Then take some large $\Sigma_0 > 2$ such that

$$\inf < \left(1 - \frac{1}{\Sigma_0}\right)(\inf + q) < \sup;$$

then take some $\delta_0 > 0$ sufficiently small that

$$\inf + \delta_0 < \left(1 - \frac{1}{\Sigma_0}\right)(\inf + q) < \sup,$$

$$\sup - \mathbb{E}w - 2\delta_0 > 0,$$

and

$$\sup - y_0 - 2\delta_0 > 0.$$

(Note that $\mathbb{E}w < \sup$ since in the case that ν is Dirac, $y_0 = \sup$). Next, fix some $0 < s < \left(1 - \frac{1}{\Sigma_0}\right)\frac{(\inf + q)}{\sup}$ such that

$$(\inf + \delta_0) + s \sup < \left(1 - \frac{1}{\Sigma_0}\right)(\inf + q).$$

Then fix some $\Sigma \geq \Sigma_0$ such that $s\Sigma > 1$. Also fix some some $0 < \kappa < \frac{\sup - \delta_0 - \mathbb{E}w}{\sup - \inf} s$.

The event A_i^R will be defined as the intersection of three events $A_1 \cap A_2 \cap A_3$. We set

$$A_1 := \left\{ \begin{array}{c} \text{For all vertices } v, w \in B(\Sigma R) \text{ with } d(v, w) \geq R, \\ \text{all paths } \gamma \text{ from } v \text{ to } w \text{ in } B(\Sigma R) \text{ have } T(\gamma) \geq (\inf + q) d(v, w) \end{array} \right\},$$

just as in the first case. We set

$$A_2 := \left\{ \begin{array}{c} \text{For all vertices } v, w \in B(\Sigma R) \text{ with } d_{E(B(\Sigma R))}(v, w) \geq R, \\ T_{E(B(\Sigma R))}(v, w) \leq (\mathbb{E}w + \delta_0) d_{E(B(\Sigma R))}(v, w) \end{array} \right\}.$$

For this, note that for each fixed pair of points v, w with $d_{E(B(\Sigma R))}(v, w) \geq R$, fixing an edge-minimal path $\gamma : v \to w$ in $B(\Sigma R)$, we have

$$\mathbb{P}(T_{E(B(\Sigma R))}(v, w) > (\mathbb{E}w + \delta_0) d_{E(B(\Sigma R))}(v, w)) \leq \mathbb{P}(T(\gamma) > (\mathbb{E}w + \delta_0)|\gamma|),$$

which, since $T(\gamma)$ is just a sum of $|\gamma|$ i.i.d. ν-distributed random variables, decays exponentially in $|\gamma|$, (hence R), by a standard Chernoff bound (ν has bounded support and hence exponential moments). Since the number of pairs of such (v, w) is strictly subexponential in R, we have $\mathbb{P}(A_2^c) \xrightarrow[R \to \infty]{} 0$, as desired. Clearly also A_2 only depends on the weights of edges in $B(\Sigma R)$.

Lastly we choose for each R some $0 \leq \delta_{\sup}(R) < \delta_0$ such that $\nu([\sup - \delta_{\sup}, \sup]) > 0$ and $\nu([\inf, \sup - \delta_{\sup}(R)])^{DC_1(\Sigma R)^d} \xrightarrow[R\to\infty]{} 1$, and then set

$$A_3 := \left\{ w(e) \leq \sup - \delta_{\sup} \text{ for all } e \in E(B(\Sigma R)) \right\}.$$

Clearly A_3 only depends on edges in $B(\Sigma R)$ and by our construction of $\delta_{\sup}(R)$, we have $\mathbb{P}(A_3) \xrightarrow[R\to\infty]{}$ 1 uniformly in i, as desired.

Now, let $a', b' \in S(\Sigma R)$ be such that $T_{E(B(\Sigma R))^c}(x, a')$ and $T_{E(B(\Sigma R))^c}(b', y)$ are minimal. Choose edge geodesics $[a', o_i]$ and $[b', o_i]$. Let $a \in V([a', o_i]), b \in V([b', o_i])$ be the unique vertices such that $d(a, a'), d(b, b') = \lceil s\Sigma R \rceil$. Moreover, for each $t \in [0, \Sigma R - \lceil s\Sigma R \rceil] \cap \mathbb{Z}$, let $a_t \in V([a, o_i]), b_t \in V([b, o_i])$ be the unique vertices such that $d(a, a_t), d(b, b_t) = t$. Now, let $t_a \geq 0$ be minimal such that

$$d(a_{t_a+1}, [b, o_i]) \leq 2C',$$

and let $t_b \geq 0$ be minimal such that

$$d(b_{t_b+1}, [a, o_i]) \leq 2C',$$

and set $c := a_{t_a}$, $d := b_{t_b}$. Note that minimality implies that for all $0 \leq t \leq t_a$ we have $d(a_t, [b, o_i]) \geq 2C' + 1$ and for all $0 \leq t \leq t_b$ we have $d(b_t, [a, o_i]) \geq 2C' + 1$. Here we have tacitly used the fact that $d(a, b) \geq d(a', b') - 2\lceil s\Sigma R \rceil \gtrsim R$ is strictly larger than $2C'$ for sufficiently large R. To see the bound $d(a', b') - 2\lceil s\Sigma R \rceil \gtrsim R$, let v and w be the entry and exit points from $B(\Sigma R)$ of the T-geodesic $\gamma : x \to y$, and note that

$$d(a', b') \geq \frac{1}{\sup} T(a', b') \geq \frac{1}{\sup} T(v, w) \geq \frac{\inf + q}{\sup} \left(1 - \frac{1}{\Sigma}\right) 2\Sigma R,$$

so

$$d(a,b) \geq d(a',b') - 2\lceil s\Sigma R\rceil \geq \left[\frac{\inf + q}{\sup}\left(1 - \frac{1}{\Sigma}\right) - s\right](2\Sigma R) - 1,$$

which is $\gtrsim R$ by choice of s. The bound $T(a',b') \geq T(v,w)$ comes from the fact that, since v, w lie on the T-geodesic from x to y, $T(x,y) = T(x,v) + T(v,w) + T(w,y) \leq T(x,a') + T(a',b') + T(b',y)$, and by definition of a', b' we have $T(x,v) + T(w,y) \geq T(x,a') + T(b',y)$. The bound $T(v,w) \geq (\inf + q)\left(1 - \frac{1}{\Sigma}\right) 2\Sigma R$ comes from the fact that π crosses B_i and hence contains at least two paths connecting $S(\Sigma R)$ and $S(R)$, which, since A_1 holds, have total passage time at least $2(\inf + q)(\Sigma - 1)R$.

Now consider the sets of integers

$$S_n(C', \kappa) := \{n\lfloor \kappa\Sigma R\rfloor + j : j \in [0, C'] \cap \mathbb{Z}\} \subset \mathbb{Z}$$

and

$$S'_n(C', \kappa) := \{n\lfloor \kappa\Sigma R\rfloor + j : j \in [C(1+\epsilon), C(2+\epsilon)] \cap \mathbb{Z}\} \subset \mathbb{Z},$$

where $n \geq 0, n \in \mathbb{Z}$. Then let α_n and β_n respectively be the subpaths of $[a,c]$ and $[b,d]$ respectively induced by the vertex sets $\{a_t : t \in S_n\}$ and $\{b_t : t \in S_n\}$. Similarly let α'_n and β'_n be induced by $\{a_t : t \in S'_n\}$ and $\{b_t : t \in S'_n\}$. For each $n \geq 0$ with $(n+1)\lfloor \kappa\Sigma R\rfloor \leq t_a, t_b$, fix a self-avoiding ϵ-detour γ_n for α'_n and a self-avoiding ϵ-detour δ_n for β'_n. Note that by construction each $\alpha_n \cup \gamma_n$ is disjoint from $[b,d]$ and all $\beta_m \cup \delta_m$, and vice versa. Moreover, $\alpha_n \cup \gamma_n$ is disjoint from $\alpha_m \cup \gamma_m$ for $n \neq m$, and the same is true for the $\beta_n \cup \delta_n$.

Finally, define

$$S_I := \bigcup_{\substack{n \geq 0, \\ (n+1)\lfloor \kappa\Sigma R\rfloor \leq t_a, t_b}} (\alpha_n \cup \gamma_n) \cup (\beta_n \cup \delta_n),$$

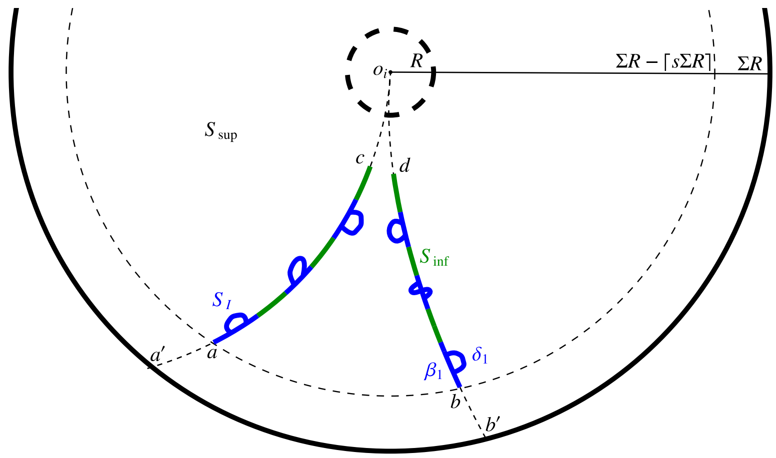

Figure 4.3. A schematic diagram of the prescribed set of configurations E_w in the case that ν has bounded support and $y_0 < \sup$.

define

$$S_{\inf} := ([a,c] \cup [b,d]) \setminus S_I,$$

and set $S_{\sup} := E(B(\Sigma R)) \setminus (S_{\inf} \cup S_I)$. For each R we choose $0 < \delta_{\inf}(R) < \delta_0$ sufficiently small that $(DC_1 R^d + 2)\delta_{\inf} < \sup - \delta_0 - y_0$. We finally define our random set of configurations by

$$E_w := \left\{ \omega \in [0,\infty)^{E(B(\Sigma R))} : \omega(e) \in \begin{array}{ll} I_0 \cap (y_0 - \frac{\delta_{\inf}}{2}, y_0 + \frac{\delta_{\inf}}{2}) & e \in S_I \\ [\inf, \inf + \delta_{\inf}] & e \in S_{\inf} \\ [\sup - \delta_{\sup}, \sup] & e \in S_{\sup} \end{array} \right\}.$$

Now let us prove that $A_1 \cap A_2 \cap A_3 \cap \{\pi \text{ crosses } B_i\} \cap \{w^* \in E_w\}$ is contained in the event that $B(\Sigma R)$ contains a feasible pair with respect to T^*.

First, define a path π' by taking a T-geodesic from x to a' in $B(\Sigma R)^c$, then taking the path $[a', c]$, taking an edge-geodesic from c to d, taking $[d, b']$ and then taking a T-geodesic from b' to y in $B(\Sigma R)^c$. For all sufficiently large R, on the event $\{\pi \text{ crosses } B_i\} \cap A_i^R \cap \{w^* \in E_w\}$, we have that $T^*(\pi') < T(\pi)$. To see this, first note that by definition of a', b', if v, w are the first entrance and last exit of π from $B(\Sigma R)$ then we have $T(\pi_{x,v}) + T(\pi_{w,y}) = T(x,v) + T(w,y) \geq T(x,a') + T(b',y) = T^*(\pi'_{x,a'}) + T^*(\pi'_{b',y})$. Thus it suffices to show that $T(\pi_{v,w}) > T^*(\pi'_{a',b'})$ for sufficiently large R. Since π visits $B_i \subset B(R)$ and since A_1 holds we have

$$T(\pi_{v,w}) \geq (\inf + q)2(\Sigma - 1)R = (\inf + q)\left(1 - \frac{1}{\Sigma}\right)2\Sigma R,$$

whereas

$$\begin{aligned} T(\pi'_{a',b'}) &\leq \left[(\inf + \delta_{\inf}) + \frac{C'}{\lfloor \kappa\Sigma R\rfloor}(y_0 + \delta_0)\right](d(a,c) + d(b,d)) + (\sup)(2s\Sigma R + d(c,d)) \\ &\leq (\inf + \delta_{\inf} + (\sup)s)2\Sigma R + o(R), \end{aligned}$$

so this follows from our choice to ensure $\inf + \delta_{\inf} + (\sup)s < (\inf + q)\left(1 - \frac{1}{\Sigma}\right)$. (We get the bound $d(c,d) = o(R)$ as follows: assume that $t_a \leq t_b$; in the opposite case the argument is analogous. By definition there exists some $t' \geq t_b + 1$ such that $d(a_{t_a+1}, b_{t'}) \leq 2C'$. But then

$$|t' - (t_a + 1)| = |d(o_i, b_{t'}) - d(o_i, a_{t_a+1})| \leq d(b_{t'}, a_{t_a+1}) \leq 2C',$$

that is, $t' \leq t_a + 1 + 2C' \leq t_b + 1 + 2C'$, and so

$$d(c,d) \leq d(c, a_{t_a+1}) + d(a_{t_a+1}, b_{t'}) + d(b_{t'}, b_{t_b}) \leq 4C' + 2 = O(C') = o(R).)$$

Now, let π^* be a T^*-geodesic from x to y. We show that π^* traverses a feasible pair.

We first show that if $p, q \in V(\pi^*) \cap V(S_{\inf})$ with p and q lying in the same connected component of $S_{\inf} \cup S_I$, then $\pi^*_{p,q} \subset S_{\inf} \cup S_I$. To see this, note that, when $w^* \in E_w$, if e is an edge in $[a, c]$ or $[b, d]$ with one endpoint in $S(t)$ and one in $S(t+1)$, then

$$w^*(e) \leq \inf\{w^*(e') : e' \text{ has one endpoint in } S(t) \text{ and the other in } S(t+1)\} + \delta_{\text{inf}}.$$

This is because, if $e \in S_{\text{inf}}$, then $e' \in S_{\text{inf}}$ or $e' \in S_{\text{sup}}$ and if $e \in S_I$ then $e' \in S_I$ or $e' \in S_{\text{sup}}$.

Since every path from p to q must have at least one edge connecting $S(t)$ to $S(t+1)$ for all $t, t+1$ between $d(a, p)$ and $d(a, q)$, we see that

$$T^*([p, q]) \leq T^*(\alpha) + \delta_{\text{inf}}|\alpha|$$

for any path α from p to q. If furthermore α leaves $S_{\text{inf}} \cup S_I$, then it contains at least one edge of weight at least $\sup -\delta_{\text{sup}}$; such an edge has weight at least $\sup -\delta_{\text{sup}} - y_0 - \delta_{\text{inf}}$ greater than any edge in $[p, q]$. Hence in this case we get the bound

$$T^*([p, q]) + \sup -\delta_{\text{sup}} - y_0 - \delta_{\text{inf}} \leq T^*(\alpha) + \delta_{\text{inf}}(|\alpha| - 1).$$

But applying our assumption on $\delta_{\text{inf}}(R)$ we get

$$T^*(\alpha) - T^*([p, q]) \geq \sup -\delta_{\text{sup}} - y_0 - (|\alpha| + 2)\delta_{\text{inf}} \geq \sup -\delta_0 - y_0 - (|B(\Sigma R)| + 2)\delta_{\text{inf}} > 0.$$

That is, such an α is not optimal, and hence an optimal T^*-path $\pi^*_{p,q}$ must lie in $S_{\text{inf}} \cup S_I$.

Hence, if we can show that $V(\pi^*)$ contains some p and q which lie in the same connected component of $S_{\text{inf}} \cup S_I$ but lie in different components of S_{inf}, then we can apply the previous argument to deduce that π^* passes through some $\alpha_n \cup \gamma_n$ or $\beta_n \cup \delta_n$, and then use the following proposition to conclude that $\pi^*_{p,q}$ contains a feasible pair:

Proposition 4.7.3. *Let ξ be an edge geodesic in G, and let γ be a self-avoiding ϵ-detour for a subpath of ξ. Suppose that $w^*(e) \in I_0$ for all $e \in \xi \cup \gamma$. Let π^* be a T^*-geodesic, and suppose that some subpath of π^* has the same endpoints as ξ and that this subpath is contained in $\xi \cup \gamma$. Then $\xi \cup \gamma$ contains a feasible pair for π^*.*[6]

Proof of Proposition. Let ξ' be the subpath of ξ such that γ is an ϵ-detour for ξ', and write $\xi = \xi_1 * \xi' * \xi_2$. Let us also abuse notation and denote by π^* the subpath of π^* contained in $\xi \cup \gamma$ which has the same endpoints as ξ. If $\pi^* = \xi$, then $\xi_1 * \gamma * \xi_2$ is an ϵ-detour for π^*; loop-erasing then gives a *self-avoiding* ϵ-detour γ' for ξ (see the proof of Proposition 4.4.1), so (π^*, γ') forms a feasible pair. If $\pi^* \neq \xi$, then since ξ is an edge geodesic, ξ is a self-avoiding ϵ-detour for π^* (see the proof of Proposition 4.4.1), and hence (π^*, ξ) forms a feasible pair. □

So it only remains to find such p and q. The idea is that, in order to make up for the slow edges π^* runs over when it enters and exits $B(\Sigma R)$, π^* must visit many fast edges; we will then use the pigeonhole principle to conclude that it must contain suitable p and q.

Explicitly, first note that since $T^*(\pi^*) \leq T^*(\pi') < T(\pi) \leq T(\pi^*)$, we have $T(\pi^*) - T^*(\pi^*) > 0$. Since $w^* \geq w$ on $E(B(\Sigma R))^c \cup S_{\sup}$, π^* must therefore contain some edges in $S_I \cup S_{\inf}$. But note that by construction, any path connecting $S(\Sigma R)$ and $S_I \cup S_{\inf}$ contains a a subpath which lies in $S_{\sup}$ and connects two points in $B(\Sigma R)$ of distance at least $s\Sigma R > R$. Since π^* starts and ends outside of $B(\Sigma R)$ and visits $S_{\inf} \cup S_I$, it contains at least two such subpaths, α and β. We then have

$$T^*(\alpha) \geq (\sup - \delta_{\sup})|\alpha|, T^*(\beta) \geq (\sup - \delta_{\sup})|\beta|$$

[6]Technically we should include assumptions controlling the lengths of these paths to satisfy our definition of a feasible pair; in our applications of this proposition it is easy to see that the length of the detour is at most $C'(1+\epsilon)$.

and

$$T(\alpha) \leq (\mathbb{E}w)|\alpha|, T(\beta) \leq (\mathbb{E}w)|\beta|$$

(since A_2 holds). Since $w^* \geq w$ on S_{sup} we then have

$$T^*(\pi^* \cap S_{\text{sup}}) - T(\pi^* \cap S_{\text{sup}}) \geq T^*(\alpha \cup \beta) - T(\alpha \cup \beta) \geq (\sup - \mathbb{E}w - \delta_{\text{sup}}) s(2\Sigma R).$$

Since $T^*(\pi^* \cap E(B(\Sigma R))^c) - T(\pi^* \cap E(B(\Sigma R))^c) = 0$, in order to ensure that $T^*(\pi^*) - T(\pi^*) < 0$, it must be the case that

$$T(\pi^* \cap (S_{\text{inf}} \cup S_I)) - T^*(\pi^* \cap (S_{\text{inf}} \cup S_I)) > (\sup - \mathbb{E}w - \delta_{\text{sup}}) s(2\Sigma R).$$

Since each edge e admits savings at most $w(e) - w^*(e) \leq \sup - \inf$, this gives

$$|\pi^* \cap (S_{\text{inf}} \cup S_I)| > \frac{\sup - \mathbb{E}w - \delta_{\text{sup}}}{\sup - \inf} s(2\Sigma R).$$

Moreover, since each component of S_I is composed of less than $2C'$ edges

$$|S_I| \leq 2C' \frac{2\Sigma R}{\lfloor \kappa \Sigma R \rfloor} = O(C') = o(R),$$

and so

$$|\pi^* \cap S_{\text{inf}}| \geq \frac{\sup - \mathbb{E}w - \delta_{\text{sup}}}{\sup - \inf} s(2\Sigma R) - o(R);$$

since by assumption $\kappa < \frac{\sup - \delta_0 - \mathbb{E}w}{\sup - \inf} s$, for sufficiently large R we have in particular

$$|\pi^* \cap S_{\text{inf}}| > 2\kappa \Sigma R.$$

Since $S_{\inf} \cup S_I$ has two connected components, at least one of the components contains more than $\kappa\Sigma R$ edges of $\pi^* \cap S_{\inf}$. But each connected component of $S_{\inf}$ contains at most $\lfloor \kappa\Sigma R \rfloor - C'$ edges, so $V(\pi^*)$ must contain some pair of points p, q which lie in different connected components of $S_{\inf}$ but in the same connected component of $S_{\inf} \cup S_I$, as desired. Thus, this construction satisfies (4.7.1).

To see that the construction satisfies the other hypothesis of Proposition 4.7.2, note that

$$\begin{aligned}\mathbb{P}(w^* \in E_w | w) &= \nu([\inf, \inf + \delta_{\inf}))^{|S_{\inf}|} \nu(I_0 \cap (y_0 - \frac{\delta_{\inf}}{2}, y_0 + \frac{\delta_{\inf}}{2}))^{|S_I|} \nu([\sup - \delta_{\sup}, \sup])^{|S_{\sup}|} \\ &\geq \min\left(\nu([\inf, \inf + \delta_{\inf})), \nu(I_0 \cap (y_0 - \frac{\delta_{\inf}}{2}, y_0 + \frac{\delta_{\inf}}{2})), \nu([\sup - \delta_{\inf}, \sup])\right)^{DC_1(\Sigma R)^d}\end{aligned}$$

is bounded away from 0 independently of o_i and x, y, as desired.

4.7.4. Geometric construction: unbounded case

Now, suppose ν has unbounded support. We construct the relevant events A_i^R and configurations E_w and show that they satisfy (4.7.1). The main challenge for the case that ν has unbounded support is in ensuring that the beginning and end of our prescribed path are far enough away from each other that we "have enough room" to make a segment and a detour which don't collide with the rest of the path. Once we construct our prescribed path it will not be hard to force the resampled geodesic π^* to take it, since we can resample the prescribed path to have very small passage time and resample the surrounding edges to have arbitrarily large passage time.

Again assume that (4.3.1) holds, and then choose $\epsilon > 0, y_0, I_0$ as in Lemma 4.3.2. Assume that ν is exponential-subcritical and let $q > 0$ be the parameter from Lemma 4.7.2. Then fix

$\sigma > \max(2, \frac{2(\inf + q)}{q})$ and $\Sigma > \sigma$. The event A_i^R will be constructed as the intersection of five events $A_i^R := A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5$. The first event is

$$A_1 := \{\text{every path } \gamma : v \to w \text{ in } B(\Sigma R) \text{ with } d(v,w) \geq R \text{ satisfies } T(\gamma) \geq (\inf + q)d(v,w)\}.$$

This evidently only depends on the edges in $E(B(\Sigma R))$. Moreover, by Lemma 4.7.2, for all sufficiently large R we have

$$\begin{aligned} \mathbb{P}(A_1^c) &\leq \sum_{v,w \in B(o_i, \Sigma R), d(v,w) \geq R} \mathbb{P}(T(v,w) < (\inf + q)d(v,w)) \\ &\leq |B(o_i, \Sigma R)| e^{-cR} \leq C_1 R^d e^{-cR} \xrightarrow[R \to \infty]{} 0. \end{aligned}$$

For the next event we choose $\delta_{\inf}(R) \xrightarrow[R \to \infty]{} 0$ such that $\nu([\inf, \inf + \delta_{\inf}]) > 0$ and $DC_1(\Sigma R)^d \delta_{\inf}(R) \leq 1$ for all R, and $(\nu([\inf + \delta_{\inf}(R), \infty)))^{DC_1(\Sigma R)^d} \xrightarrow[R \to \infty]{} 1$. (Note that if there is an atom at inf, then eventually we will have $\delta_{\inf}(R) = 0$, but $\delta_{\inf}(R) \geq 0$ always). Note that the second condition implies in particular that $|E(B(\Sigma R))| \delta_{\inf}(R) \leq 1$. We define

$$A_2 := \{w(e) \geq \inf + \delta_{\inf} \text{ for all } e \in E(B(\Sigma R))\}.$$

This clearly only depends on the weights in $E(B(\Sigma R))$ and the third condition on $\delta_{\inf}(R)$ implies that

$$\mathbb{P}(A_2) = \nu([\inf + \delta_{\inf}, \infty))^{|E(B(\Sigma R)|} \geq \nu([\inf + \delta_{\inf}, \infty))^{DC_1(\Sigma R)^d} \xrightarrow[R \to \infty]{} 1.$$

For the third event, we choose $M(R) \xrightarrow[R \to \infty]{} \infty$ such that $\nu^{*DC_1(\Sigma R)^d}([0, M(R)]) \xrightarrow[R \to \infty]{} 1$. We set

$$A_3 := \left\{ \sum_{e \in E(B(\Sigma R))} w(e) \leq M \right\}.$$

It is clear by the choice of $M(R)$ that $\mathbb{P}(A_3) \xrightarrow[R\to\infty]{} 1$. Also note that since ν is assumed to have infinite support, $\nu((M(R),\infty)) > 0$ for all R.

Let us call a value $p \in \operatorname{supp}\nu$ (δ,η)-*resamplable* if $\nu([p, p+\delta)) \geq \eta$. Set $\delta_{sim}(R) := (DC_1R^d)^{-1}$. Then, using Proposition 4.7.5 below, choose $\eta(R) > 0$ such that

$$\nu(\{p : p \text{ is } (\delta_{sim}(R), \eta(R))\text{-resamplable}\})^{DC_1R^d} \geq 1 - e^{-R}.$$

Set

$$A_4 := \{w(e) \text{ is } (\delta_{sim}, \eta)\text{-resamplable for all } e \in E(B(\Sigma R))\}.$$

Clearly A_4 only depends on weights of edges in $E(B(\Sigma R))$, and by our choice of $\eta(R)$ we have

$$\mathbb{P}(A_4) \geq 1 - e^{-R} \xrightarrow[R\to\infty]{} 1.$$

The event A_5 is more complicated to describe, so we delay its description and the proof that $\mathbb{P}(A_5) \xrightarrow[R\to\infty]{} 1$ until the end of the section.

Next we describe the construction of E_w. Denote by π the geodesic from x to y, and denote by v and w the first vertex of π which lies in $B(\Sigma R)$ and the last vertex of π which lies in $B(\Sigma R)$, respectively. As will be proved in Lemma 4.7.3 at the end of the section, the event A_5 implies that, for some $\Sigma R \geq r \geq \sigma R$, we have disjoint self-avoiding paths α and β with the following properties:

(1) α starts at x and ends at a point $v' \in S(r)$; moreover $V(\alpha) \cap B(r-1) = \emptyset$.

(2) β starts at a point $w' \in S(r)$ and ends at y; moreover $V(\beta) \cap B(r-1) = \emptyset$.

(3) $d_{E(B(r))}(v', w') > K := 4C(1+\epsilon)$.

(4) α coincides with π until its last entrance into $B(\Sigma R)$ and β coincides with π after its first exit from $B(\Sigma R)$. Explicitly, Choose $\tilde{v}$ to be the last entrance of α into $B(\Sigma R)$, so that $\alpha_{\tilde{v},v'}$ is the connected component of $E(B(\Sigma R)) \cap \alpha$ containing v'. Similarly choose $\tilde{w}$ to be the first exit of β from $B(\Sigma R)$, so that $\beta_{w',\tilde{w}}$ is the connected component of $E(B(\Sigma R) \cap \beta$ containing w'. We have that $\tilde{v}, \tilde{w} \in V(\pi)$, and $\pi_{x,\tilde{v}} = \alpha_{x,\tilde{v}}$ and $\pi_{\tilde{w},y} = \beta_{\tilde{w},y}$.

(5) Let $v_r \in S(r)$ be the vertex of π immediately preceding the first vertex of π which lies in $B(r-1)$, and let $w_r \in S(r)$ be the vertex of π immediately following the last vertex of π which lies in $B(r-1)$. Then $|\alpha_{\tilde{v},v'}| \leq |\pi_{\tilde{v},v_r}|$ and $|\beta_{w',\tilde{w}}| \leq |\pi_{w_r,\tilde{w}}|$.

Now choose edge-geodesics $[v', o_i]$ from v' to o_i and $[o_i, w']$ from o_i to w'. Again let C be such that every self-avoiding path of length C admits a self-avoiding ϵ-detour. Let a be the vertex of $[v', o_i]$ which is distance $C(1+\epsilon)$ from v'. Let b be the vertex of $[v', o_i]$ which is distance $C(2+\epsilon)$ from v'. Then $[a, b] := [v', o_i]_{a,b}$ is a self-avoiding path of length C, and hence it admits a self-avoiding ϵ-detour γ.

Proposition 4.7.4. *γ is contained in $B(r-1)$ and $V(\gamma) \cap V([o_i, w']) = \emptyset$.*

Proof. The first claim follows from the fact that γ has length at most $C(1+\epsilon)$; To see the second claim, suppose to the contrary that there was some $z \in V(\gamma) \cap V([o_i, w'])$. Since $d_{B(r)}(z, a) \leq C(1+\epsilon)$ and z and a both lie on edge-geodesics to o_i, we have that

$$\begin{aligned}|d_{B(r)}(v', a) - d_{B(r)}(w', z)| &= |[d_{B(r)}(v', o_i) - d_{B(r)}(a, o_i)] - [d_{B(r)}(w', o_i) - d_{B(r)}(z, o_i)]| \\ &= |d_{B(r)}(a, o_i) - d_{B(r)}(z, o_i)| \leq d_{B(r)}(a, z) \leq C(1+\epsilon),\end{aligned}$$

and therefore

$$d_{B(r)}(w', z) \leq d_{B(r)}(v', a) + C(1+\epsilon) = 2C(1+\epsilon),$$

hence

$$d_{B(r)}(v', w') \leq d_{B(r)}(v', a) + d_{B(r)}(a, z) + d_{B(r)}(z, w') \leq 2C(1+\epsilon) + 2C(1+\epsilon) = K,$$

contradicting the fact that $d(v', w') > K$. □

Set b' to be the vertex in $V([v', o_i])$ which has distance $C' = C(3+2\epsilon)$ from v'. Set o' to be the first intersection of $V([v', o_i])$ with $V([o_i, w'])$; the previous proposition shows that o' is strictly closer to o_i than b'. Define as usual $[v', o'] := [v', o_i]_{v', o'}$, $[o', w'] := [o_i, w']_{o', w'}$, $[v', b'] := [v', o_i]_{v', b'}$. We now define the following subsets of $E(B(\Sigma R))$:

$$\begin{aligned} S_I &:= [v', b'] \cup \gamma, \\ S_{\inf} &:= (\alpha_{\tilde{v}, v'} * [v', o'] * [o', w'] * \beta_{w', \tilde{w}}) \setminus S_I, \\ S_{sim} &:= (\alpha \cup \beta) \cap E(B(\Sigma R)) \setminus S_{\inf} \\ S_M &:= E(B(\Sigma_R)) \setminus (S_I \cup S_{\inf} \cup S_{sim}). \end{aligned}$$

Note that these sets are all pairwise disjoint and cover $E(B(\Sigma R))$. Now we can finally define our set of configurations E_w:

$$E_w := \left\{ \omega \in [0,\infty)^{E(B(\Sigma R))} : \omega(e) \in \begin{array}{ll} I_0 & e \in S_I \\ [\inf, \inf + \delta_{\inf}] & e \in S_{\inf} \\ [w(e), w(e) + \delta_{sim}) & e \in S_{sim} \\ [M, \infty) & e \in S_M \end{array} \right\}.$$

(We have used the assumption that $w \in A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5$ to construct E_w, and this is really the only case we care about; off of this event we may define $E_w = \emptyset$). We now show that this

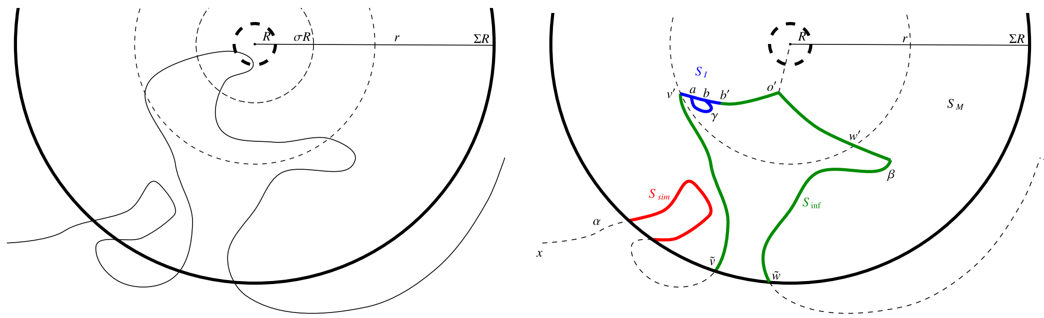

Figure 4.4. If the T-geodesic from x to y is as in the diagram on the left, the prescribed weights E_w might be given by the diagram on the right.

choice of A_i and E_w satisfies (4.7.1). Set

$$\pi' := \alpha * [v', o'] * [o', w'] * \beta.$$

First we show that $T^*(\pi') < T(\pi)$. By construction we have $\pi'_{x,\tilde{v}} = \pi_{x,\tilde{v}}$ and $\pi'_{\tilde{w},y} = \pi_{\tilde{w},y}$. Moreover, each edge in either of those paths is also by construction either in $E(B(\Sigma R))^c$ or S_{sim}, and hence when $w^* \in E_w$,

$$T^*(\pi'_{x,\tilde{v}} \sqcup \pi'_{\tilde{w},y}) \leq T(\pi'_{x,\tilde{v}} \sqcup \pi'_{\tilde{w},y}) + |E(B(\Sigma R))|\delta_{sim}$$

$$\leq T(\pi_{x,\tilde{v}} \sqcup \pi_{\tilde{w},y}) + 1.$$

Next, since $|\alpha_{\tilde{v},v'}| \leq |\pi_{\tilde{v},v_r}|$, $|\beta_{w',\tilde{w}}| \leq |\pi_{w_r,\tilde{w}}|$, since A_2 holds, and since $\alpha_{\tilde{v},v'}, \beta_{w',\tilde{w}} \subset S_{\inf}$, we have

$$T^*(\alpha_{\tilde{v},v'} \sqcup \beta_{w',\tilde{w}}) \leq (\inf + \delta_{\inf})|\alpha_{\tilde{v},v'} \sqcup \beta_{w',\tilde{w}}| \leq (\inf + \delta_{\inf})|\pi_{\tilde{v},v_r} \sqcup \pi_{w_r,\tilde{w}}| \leq T(\pi_{\tilde{v},v_r} \sqcup \pi_{w_r,\tilde{w}}).$$

Now, since $\pi'_{v',w'} \setminus [v', b'] \subset S_{\inf}$ and $[v', b'] \subset S_I$, we have

$$T^*(\pi'_{v',w'}) \leq (\inf + \delta_{\inf})2r + (\sup I_0)C(3 + 2\epsilon),$$

while since A_1 holds and π_{v_r,w_r} starts and ends at $S(r)$ ($r \geq \sigma R > 2R$) and visits $S(R)$, we have

$$T(\pi_{v_r,w_r}) \geq 2(\inf + q)(r - R),$$

so that

$$\begin{aligned} T(\pi_{v_r,w_r}) - T^*(\pi'_{v',w'}) &\geq 2R[(q - \delta_{\inf})\frac{r}{R} - (\inf + q)] - (\sup I_0)C(3 + 2\epsilon) \\ &\geq 2R[(q - \delta_{\inf})\sigma - (\inf + q)] - (\sup I_0)C(3 + 2\epsilon), \end{aligned} \tag{4.7.2}$$

For R sufficiently large we have $\delta_{\inf} < q/2$ so that

$$(q - \delta_{\inf})\sigma - (\inf + q) > (q/2)\sigma - (\inf + q) > 0,$$

that is, the coefficient of R (4.7.2) is strictly positive. Altogether we have

$$T(\pi) - T^*(\pi') \geq 2R[(q/2)\sigma - (\inf + q)] - (\sup I_0)C(3 + 2\epsilon) - 1 \gtrsim R,$$

so in particular $T^*(\pi') < T(\pi)$ for all sufficiently large R.

From this, we can conclude that the T^*-geodesic π^* must contain some edges in $S_{\inf} \cup S_I$. For suppose it did not; since $w^* \geq w$ on $(S_{\inf} \cup S_I)^c$, we would have

$$T^*(\pi^*) \geq T(\pi^*) \geq T(\pi) > T^*(\pi'),$$

contradicting T^*-geodesicity of π^*.

Next, we know that π^* contains no edge in S_M. For suppose that it did; then, since A_3 holds,

$$\begin{aligned} T^*(\pi^*) &\geq T^*(\pi^* \cap E(B(\Sigma R))^c) + M \\ &\geq T(\pi^* \cap E(B(\Sigma R))^c) + \sum_{e \in E(B(\Sigma R))} w(e) \\ &\geq T(\pi^*) \geq T(\pi) > T^*(\pi'), \end{aligned}$$

again contradicting T^*-geodesicity of π^*.

Note that $S_{\inf} \cup S_I$ and S_{sim} by construction share no vertices in common, and so we see that π^*, as a self-avoiding path which enters $S_{\inf}$, does not intersect S_M and eventually exits $B(\Sigma R)$, must contain $\pi'_{\tilde{v},v'}$ and $\pi'_{b',\tilde{w}}$ or their reverses as a subpath. In particular, some subpath of π^* has endpoints v', b' and is restricted to $S_I = [v', b'] \cup \gamma$ and hence by Proposition 4.7.3 S_I contains a feasible pair for π^*, and we are done showing that (4.7.1) is satisfied.

To complete the proof that the A_i^R, E_w satisfy the hypotheses of Proposition 4.7.2, it remains to prove the "resampling lemma" relevant to A_4, to describe and prove the relevant properties of A_5, and to give a lower bound on the conditional probability of $\{w^* \in E_w\}$.

Proposition 4.7.5. *For any fixed $\delta > 0$, we have*

$$\lim_{\eta \to 0} \nu(\{p : p \text{ is } (\delta, \eta)\text{-resamplable}\}) = 1.$$

Proof. By continuity of measure we have that

$$\lim_{\eta \to 0} \nu(\{p : \nu([p, p+\delta)) > \eta) = \nu(\{p : \nu([p, p+\delta)) > 0),$$

so it will suffice to show that

$$\nu(\{p : \nu([p, p + \delta)) = 0\}) = 0.$$

Set $N := \{p : \nu([p, p + \delta)) = 0\}$. We claim that there is a countable subset $X \subset N$ such that

$$N \subset \bigcup_{p \in X} [p, p + \delta).$$

Once we know this, the proposition follows, since then

$$\nu(N) \leq \nu\left(\bigcup_{p \in X} [p, p + \delta)\right) \leq \sum_{p \in X} \nu([p, p + \delta)) = 0.$$

To construct X, first set $X_0 := \emptyset$. For each $i \geq 0$, consider $n_{i+1} := \inf N \setminus \left(\bigcup_{p \in X_i} [p, p + \delta)\right)$. If $n_{i+1} \in N$, then set $X_{i+1} := X_i \cup \{n_{i+1}\}$. Otherwise choose a (countable) sequence S_{i+1} of points of N approaching n_{i+1} and set $X_{i+1} := X_i \cup S_{i+1}$. It is simple to inductively check that each X_i is countable and that $\bigcup_{p \in X_i} [p, p + \delta)$ covers at least $N \cap [0, i\delta)$, so $X := \cup_{i=1}^{\infty} X_i$ is a countable subset of N with $\bigcup_{p \in X} [p, p + \delta) \supset N$, as desired. □

Now we describe the event A_5 and its properties. The intuition is as follows: considering the T-geodesic $\pi : x \to y$, for each ball $B(r)$, if the first entrance of π into that ball is far from the last exit of π from that ball, then we have "enough room" to do our construction, that is, we have paths satisfying (1)-(5) above. So we want to bound the probability that, to the contrary, for all radii r, the first entry and last exit are close. In fact, an even weaker event gives us "enough room," and we bound the probability of the failure of this event by showing that it would entail that π is constrained to a "narrow" subgraph as it crosses $B(o_i, \Sigma R)$, making it unlikely that the geodesic would enter so deep into $B(o_i, \Sigma R)$ before turning around.

For the formal construction of the event, first, given a pair of points $p, q \in S(\Sigma R)$, take edge-geodesics $[p, o_i]$ and $[q, o_i]$ from p and q respectively to o_i. For each $\Sigma R \geq r \geq \sigma R$, let p^r and q^r be the unique elements of $V([p, o_i]) \cap S(r)$ and $V([q, o_i]) \cap S(r)$ respectively. We then define

$$S_0^r(p, q) := (B_{E(B(r))}(p^r, 3K) \cup B_{E(B(r))}(q^r, 3K)) \cap S(r),$$

where $K := 4C(1 + \epsilon)$. Then for each $\ell \geq 0$ we define

$$S_\ell^r(p, q) := \{z \in B(\Sigma R) \setminus B(r-1) : d_{E(B(\Sigma R) \setminus B(r-1))}(z, S_0^r(p, q)) = \ell\}.$$

Lastly, for $\Sigma R - 3K \geq r \geq \sigma R$, set

$$S^r(p, q) := \bigsqcup_{\ell=0}^{3K} S_\ell^r(p, q)$$

and define the event

$$C^r(p, q) := \left\{ \begin{array}{c} \text{there exist paths } \gamma_1, \gamma_2 \text{ in } S^r(p, q) \text{ such that} \\ \text{the endpoints } a_1, b_1 \text{ of } \gamma_1 \text{ lie in } S_{2K}^r \text{ and} \\ |\gamma_1| \leq K, \text{ one endpoint of } \gamma_2 \text{ lies in} \\ S_{2K}^r \text{ and the other lies in } S_0^r, \text{ and } T(\gamma_2) \leq T(\gamma_1) \end{array} \right\}.$$

We now define the event A_5 by

$$A_5 := \bigcap_{\substack{p,q \in S(\Sigma R), \\ d_{B(\Sigma R)}(p,q) \leq K}} \left(\bigcap_{r=\sigma R}^{\Sigma R - 3K} C^r(p, q) \right)^c,$$

that is, A_5 is the event that for each pair p, q of close points on $S(\Sigma R)$, $C^r(p, q)$ fails for at least some r. Note that A_5 only depends on the weights of edges in $E(B(\Sigma R))$.

Proposition 4.7.6. *There exists some constant $\rho < 1$ depending only on the degree D of G, ν, and K such that*

$$\mathbb{P}(C^r(p,q)) \leq \rho$$

for all R, p, q, r.

Proof. First note that, since each S^r_0 lies in the union of two balls of radius $3K$, S^r_0 contains at most $2(D+1)^{3K}$ vertices. Since the entirety of S^r lies within distance $3K$ of S^r_0, we further have that

$$|S^r| \leq |S^r_0|(D+1)^{3K} \leq 2(D+1)^{6K}.$$

That is, we have a uniform bound on the possible number of vertices in S^r, and so it is not hard to see that the subgraph induced by $S^r(p,q)$ can only take on finitely many isomorphism types as all parameters except D and K vary. Hence, to show our claim, it suffices to show that for each fixed isomorphism type, $\mathbb{P}(C^r(p,q)) < 1$. (Here "isomorphism type" includes the relevant extra data of which subsets correspond to S^r_0 and S^r_{2K}, but even with this extra data it is easy to see that a bound on the number of vertices implies a bound on the number of possible isomorphism types).

To this end, fix an isomorphism type, and let E' be the set of edges in S^r which lie in some path in S^r of length at most K joining two vertices of S^r_{2K}. Since ν is assumed to have unbounded support (in particular it is not Dirac), there is some $a > 0$ such that $\nu([0,a)) > 0$ and $\nu([a,\infty)) > 0$. Then the event

$$\{w(e) < a \text{ for all } e \in E', w(e) \geq a \text{ for all } e \notin E'\}$$

has nonzero probability. Moreover, this event entails the failure of $C^r(p,q)$, since on it all candidates for γ_1 necessarily have edges in E' and hence have $T(\gamma_1) < aK$, while all candidates for γ_2 must have at least K edges lying in E'^c, and hence $T(\gamma_2) \geq aK > T(\gamma_1)$. □

Proposition 4.7.7. $\mathbb{P}(A_5) \xrightarrow[R\to\infty]{} 1$.

Proof. For each fixed p,q, note that whenever $S \subset [\sigma R, \Sigma R - 3K] \cap \mathbb{Z}$ is such that each element has distance at least $3K$ from every other element, the subgraphs $\{S^r(p,q) : r \in S\}$ are all disjoint and hence the events $\{C^r(p,q) : r \in S\}$ are all independent. Since K, σ, Σ are constants fixed independent of R, it is easy to see that there is some $c_3 > 0$ such that for all large R we can pick such an S with $|S| \geq c_3 R$, and so

$$\mathbb{P}\left(\bigcap_{r=\sigma R}^{\Sigma R - 3K} C^r(p,q)\right) \leq \mathbb{P}\left(\bigcap_{r\in S} C^r(p,q)\right)$$

$$= \prod_{r\in S} \mathbb{P}(C^r(p,q)) \leq \rho^{c_3 R},$$

where $\rho < 1$ is provided by the previous proposition. But then we have

$$\mathbb{P}(A_4^c) \leq \sum_{p,q\in S(\Sigma R)} \mathbb{P}\left(\bigcap_{r=\sigma R}^{\Sigma R-3K} C^r(p,q)\right) \leq (C_1(\Sigma R)^d)^2 \rho^{c_3 R} \xrightarrow[R\to\infty]{} 0,$$

as desired. □

Now we prove the key property of A_5.

Lemma 4.7.3. *On the event $A_5 \cap \{x, y \notin B(\Sigma R), \pi$ visits $B_i\}$, for some $\Sigma R \geq r \geq \sigma R$, there exist paths α and β satisfying conditions 1 through 5 above.*

Proof. Denote by v and w respectively the first and last vertices of π which lie in $B(\Sigma R)$. Now, for each $\Sigma R \geq r \geq \sigma R$, define $v_r \in S(r)$ to be the vertex of π immediately preceding the first vertex of π lying in $B(r-1)$, and define $w_r \in S(r)$ to be the vertex of π immediately following the last vertex of π lying in $B(r-1)$. All of these are well defined, since π starts and ends outside of $B(\Sigma R)$ and visits $B_i \subset B(R) \subset B(\sigma R)$. Then define $\alpha_r := \pi_{x,v_r}$, $\beta_r := \pi_{w_r,y}$. If for some r, $d_{E(B(r))}(v_r, w_r) > K$, then we can just take $\alpha = \alpha_r$ and $\beta = \beta_r$ and we are done. So from here on assume that $d_{E(B(r))}(v_r, w_r) \leq K$ for all $\Sigma R \geq r \geq \sigma R$.

Next, for each r we define the set

$$\tilde{S}^r := \left(\bigcup_{p \in V(\alpha_r \cup \beta_r) \cap B(\Sigma R)} \bigcup_{\substack{\gamma: p \to o_i \\ \text{edge geodesic}}} V(\gamma) \right) \cap S(r).$$

Suppose that for some $\Sigma R \geq r \geq \sigma R$, there is some $z \in \tilde{S}^r$ with $d_{E(B(r))}(z, v_r), d_{E(B(r))}(z, w_r) > K$. Then we can construct α and β as follows. z by definition lies on some edge-geodesic γ from some point $p \in V(\alpha_r \cup \beta_r)$ to o_i. Consider the last vertex of $V(\gamma) \cap (V(\alpha_r \cup \beta_r))$ (that is, the nearest vertex to o_i), and call it z'. If $z' \in V(\alpha_r)$, set $\alpha := (\alpha_r)_{x,z'} * \gamma_{z',z}$ and $\beta := \beta_r$. If $z' \in V(\beta_r)$, set $\beta := \overline{\gamma}_{z,z'} * (\beta_r)_{z',y}$ and $\alpha := \alpha_r$ (here an overline denotes the reverse of a path). In either case, α and β give disjoint self-avoiding paths because the original paths were disjoint and self-avoiding and because by construction $V(\gamma_{z',z})$ only intersects $V(\alpha_r \cup \beta_r)$ at z'. Conditions (1)-(3) are satisfied by choice of z, (4) is satisfied because α and $\overline{\beta}$ agree with α_r and $\overline{\beta_r}$ until one of them reaches z', and from that point the path follows γ; in particular, it stays inside $B(\Sigma R)$ until it reaches its endpoint. For (5), note that, since γ is an edge-geodesic from z' to o_i,

$$|\gamma_{z',z}| = d(z', o_i) - d(z, o_i) = d(z', o_i) - r.$$

Since $(\alpha_r)_{z',v_r}$ (or $(\overline{\beta}_r)_{z',w_r}$, if $z' \in V(\beta_r)$) is a path from z' to $S(r)$, it must have length at least $d(z', o_i) - r$ by the triangle inequality, and so we get (5).

Lastly, we show that, if both of the above conditions fail, i.e. for all $\Sigma R \geq r \geq \sigma R$ we have

$$d_{E(B(r))}(v_r, w_r) \leq K \tag{4.7.3}$$

and

$$\tilde{S}^r \subset (B_{E(B(r))}(v_r, K) \cup B_{E(B(r))}(w_r, K)) \cap S(r), \tag{4.7.4}$$

then the event A_5 fails.

For this, first note that, since every $V(\alpha_r \cup \beta_r) \cap B(\Sigma R)$ contains the entry and exit points v and w, every $\tilde{S}^r$ contains v^r, w^r (in the notation used in defining the set $S^r(p, q)$ in the case $(p, q) = (v, w)$). Then (4.7.4) implies that v^r and w^r are each distance at most K from either v_r or w_r. A general element $z \in \tilde{S}^r$ has the same property, and combining with (4.7.3) gives

$$d(z, v^r) \leq \min(d(z, v_r), d(z, w_r)) + d(v_r, w_r) + \min(d(v_r, v^r), d(w_r, v^r)) \leq 3K,$$

and similarly $d(z, w^r) \leq 3K$. Hence $\tilde{S}^r \subset S^r_0(v, w)$. Moreover we have

Claim 1. *If $\Sigma R \geq r + \ell, r \geq \sigma R$, then $v_{r+\ell}, w_{r+\ell} \in S^r_\ell(v, w)$.*

Proof of Claim. Since $v_{r+\ell}, w_{r+\ell} \in S(r + \ell)$ and $S^r_0(v, w) \subset S(r)$, we have that

$$d_{E(B(\Sigma R)\setminus B(r-1))}(v_{r+\ell}, S^r_0(v, w)), d_{E(B(\Sigma R)\setminus B(r-1))}(w_r, S^r_0(v, w)) \geq \ell,$$

so we only have to show the opposite inequality. For this, let $(v_{r+\ell})^r$ be as usual the intersection of $S(r)$ with an edge geodesic from $v_{r+\ell}$ to o_i. Since $v_{r+\ell} \in V(\alpha_{r+\ell}) \subset V(\alpha_r)$, we have that $(v_{r+\ell})^r \in \tilde{S}^r \subset S^r_0(v, w)$; moreover, the geodesic from $v_{r+\ell}$ to $(v_{r+\ell})^r$ is a path of length ℓ which lies in $B(\Sigma R) \setminus B(r-1)$, and so we have

$$d_{B(\Sigma R)\setminus B(r-1)}(v_{r+\ell}, S_0^r(v,w)) \leq d_{B(\Sigma R)\setminus B(r-1)}(v_{r+\ell}, (v_{r+\ell})^r) = l,$$

as desired. The argument for $w_{r+\ell}$ is the same. □

Finally, we contradict A_5. For each $\Sigma R - 3K \geq r \geq \sigma R$, consider $\gamma_3 := \pi_{v_{r+2K}, v_r}$. Since γ_3 by construction does not visit $B(r-1)$, and since it starts at a point with distance $d_{E(B(\Sigma R)\setminus B(r-1))}(v_{r+2K}, S_0^r(v,w)) = 2K$ and ends at a point $v_r \in S_0^r(v,w)$, some subpath γ_2 of γ_3 is contained in $S^r(v,w)$, starts at $S_{2K}^r(v,w)$ and ends at $S_0^r(v,w)$. On the other hand, let γ_1 be an edge-geodesic from v_{r+2K} to w_{r+2K}. By assumption, $d(v_{r+2K}, w_{r+2K}) \leq K$, so $|\gamma_1| \leq K$; therefore γ_1 does not intersect $B(r-1)$, and since the endpoints of γ_1 lie in $S_{2K}^r(v,w)$, γ_1 is totally contained in $S^r(v,w)$. But since π is a T-geodesic, we have

$$T(\gamma_1) \geq T(\pi_{v_{r+2K}, w_{r+2K}}) \geq T(\gamma_3) \geq T(\gamma_2),$$

and so $C^r(v,w)$ holds. But then $C^r(v,w)$ holds for all $\Sigma R - 3K \geq r \geq \sigma R$, so A_5 fails. □

To apply Proposition 4.7.2 it only remains to obtain a lower bound on $\mathbb{P}(w^* \in E_w|w)$ on the event A_i^R which is independent of o_i. But on A_i^R we have

$$\begin{aligned}\mathbb{P}(w^* \in E_w|w) &\geq \nu(I_0)^{|S_I|}\nu([\inf, \inf+\delta_{\inf}])^{|S_{\inf}|}\eta^{|S_{sim}|}\nu([M,\infty))^{|S_M|} \\ &\geq \min(\nu(I_0), \nu([\inf, \inf+\delta_{\inf}]), \eta, \nu([M,\infty)))^{DC_1(\Sigma R)^d} > 0,\end{aligned}$$

as desired.

4.7.5. Proof of Theorem 4.1.2

Let G be a graph of strict polynomial growth which admits detours. Let ν be an exponential-subcritical measure with finite mean, and let $\tilde{\nu}$ be a measure which has finite mean and is strictly more variable than ν. First assume (4.3.1). Then let $\epsilon > 0, I_0, y_0$ be as in Lemma 4.3.2. In case ν has bounded support, construct $B_i, B(o_i, \Sigma R), A_i^R$, and E_w as in Section 4.7.3. In case ν has unbounded support, construct $B_i, B(o_i, \Sigma R), A_i^R$, and E_w as in Section 4.7.4. In their respective sections, we prove that both constructions satisfy the hypotheses of Proposition 4.7.2, and so

$$\liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y) - \mathbb{E}\tilde{T}(x,y)}{d(x,y)} > 0.$$

Now, if $w, \tilde{w}$ do not satisfy (4.3.1), take $\bar{w}$ as in Lemma 4.3.1. Then we have

$$\liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y) - \mathbb{E}\tilde{T}(x,y)}{d(x,y)} \geq \liminf_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y) - \mathbb{E}\bar{T}(x,y)}{d(x,y)} > 0.$$

Thus G is vdBK. The reverse implication is given by Theorem 4.4.1.

4.7.6. Non-homogeneous graphs of polynomial growth

Theorem 4.1.2 does not require almost-transitivity (although assuming almost-transitivity does make both the hypotheses and the conclusions easier to interpret, see Sections 4.4.1 and 4.2.1). This means that the theorem applies to a very broad class of graphs, but it can be difficult to produce examples of non-transitive graphs which have *strict* polynomial growth and for which it is easy to check whether the graph admits detours. Here we give some examples and a counterexample.

First, the theorem can be applied to a broad range of subgraphs of the standard Cayley graph of $\mathbb{Z}^d$. For instance $G := \mathbb{Z}^{d_1}_{\geq 0} \times \mathbb{Z}^{d_2} \subset \mathbb{Z}^{d_1+d_2}$ will be vdBK whenever $d_1 + d_2 \geq 2$. These graphs have growth bounds $B_G(R) \leq B_{\mathbb{Z}^{d_1+d_2}}(R) \leq 2^{-d_1} B_G(R)$, from which we can deduce that G has strict polynomial growth. Moreover, the unique geodesics in G are all also unique geodesics in $\mathbb{Z}^d$, (that is, they are represented by words of the form e_i^k, where $\{e_i\}$ is the standard generating set), and when $d_1 + d_2 \geq 2$ one can easily see that these admit detours.

Moreover, we can apply the theorem to "sectors", that is, graphs $G_{\theta,\theta'}$ induced by the vertex subset

$$V_{\theta,\theta'} := \{(x, y) \in \mathbb{Z}^2 : \theta \leq \arctan(y/x) \leq \theta'\}$$

for fixed $\theta < \theta'$. Again we see that this is of strict polynomial growth. Moreover, the unique geodesics in this graph are either already unique geodesics in $\mathbb{Z}^2$, or they run along the "boundary" $\{(x, y) : \arctan(y/x) \approx \theta \text{ or } \theta'\}$ (in fact most geodesics along the boundary are also not unique). But again it is simple to check that these admit detours, and hence G is vdBK. Similar constructions can be done in higher dimensions, and in fact many more subgraphs of $\mathbb{Z}^d$ satisfy the hypotheses of the theorem.

Another example of an inhomogeneous graph which is vdBK can be constructed as follows. Start with the standard Cayley graph of $\mathbb{Z}^2$, and choose some subset S of the square faces (for instance, one can choose S randomly by independently including each face with probability p). For each face in S, add a vertex in the center, connecting it with edges to the four corners. It is not hard to show that the graph obtained has strict polynomial growth and admits detours, and so again by Theorem 4.1.2 it is vdBK. One can construct many examples of such "lattices with impurities" which are vdBK, but in general one must be a bit careful to be sure that the constructed graph admits detours (for instance, naïvely implementing the above construction in higher dimensions may produce new unique geodesics which do not admit detours).

Next, Theorem 4.1.2 likely applies to the famous "kites and darts" Penrose tilings, although rigorously writing out the details may take a bit of work. It is not hard to show strict polynomial growth, e.g. by comparing number of vertices in a graph-distance ball to the total area of the adjacent tiles. Showing rigorously that these tilings admit detours is trickier, but one should note that unique geodesics are very constrained in these tilings, and it seems likely that some version of the strategy used in $\mathbb{Z}^2$, i.e. "follow a translated version of the original path and then come back," should work (even though *exact* translates of paths typically do not exist).

Unfortunately, the most obvious candidate is in fact a counterexample. That is, consider the infinite cluster of a supercritical Bernoulli percolation on a Cayley graph of a virtually nilpotent group. In fact, this will not have strict polynomial growth as we have defined here, since we require *uniform* volume lower bounds. But beyond that, one can see that (for $p < 1$) almost surely the cluster does *not* admit detours, and hence by Theorem 4.4.1 it is *not* vdBK.

This can be seen by a simple "finite energy" type argument. For any $C < \infty$, choose a large radius $R \geq C$ such that the probability that $B(R)$ intersects the infinite cluster is positive; this event is actually independent of the configuration of edges inside $E(B(R))$, so chose a particular configuration in $E(B(R))$ such that all edges in contact with the vertex boundary of $B(R)$ are open, such that all these edges are connected to each other by open edges, and such that these open edges on the boundary are connected to an open path of length $\geq 3C$ which is otherwise surrounded by closed edges. The probability that the boundary of $B(R)$ is connected to infinity *and* that the restriction of the sampled configuration restricted to $E(B(R))$ is our prescribed configuration, is also positive. One quickly sees that on this event, the infinite cluster contains a self-avoiding path of length C which does not admit a detour of length at most, say, $(3/2)C$. The event that the infinite cluster contains a such a path is clearly a translation-invariant event, and so by ergodicity, since this event occurs with positive probability, it occurs with probability 1. Intersecting all these events for a countable collection $C \to \infty$ shows that almost surely the infinite cluster does not admit detours.

Of course, for graphs which are not almost-transitive, the vdBK condition is quite strong. One may ask for the following weaker condition (which is equivalent in the case of almost-transitive graphs, see Section 4.4.1): fix $o \in V$. If $\tilde{\nu}$ is strictly more variable than ν and ν is exponential-subcritical, is

$$\liminf_{x\to\infty} \frac{\mathbb{E}T(o,x) - \mathbb{E}\tilde{T}(o,x)}{d(o,x)} > 0?$$

It is conceivable that the answer might be “yes” in the case of supercritical percolation clusters on nilpotent Cayley graphs, since supercritical clusters “generally behave like their underlying graph” at large scales. Perhaps the proofs in this chapter could be adapted to this case, but it would require “large scale” and perhaps “statistical” weakenings of the geometric properties used, and the precise adaptation is not clear.

4.8. Absolute continuity with respect to the expected empirical measure

For a graph G and a probability measure ν on $[0, \infty)$, we say that the associated first passage percolation *has weight distribution absolutely continuous with respect to the expected empirical measure* if for any Borel set $A \subset [0, \infty)$ with $\nu(A) > 0$ we have

$$\liminf_{d(x,y)\to\infty} \frac{\mathbb{E}[\sum_{e\in\pi} \mathbb{1}_{w(e)\in A}]}{d(x,y)} > 0,$$

where π denotes the T-geodesic from x to y. Note that this does not imply or presuppose that a literal expected empirical measure, that is, a weak limit of the expected empirical measures $\frac{1}{d(x,y)}\mathbb{E}\sum_{e\in\pi}\delta_{w(e)}$, exists,[7] although it does imply that ν is absolutely continuous with respect to any subsequential weak limit of this collection of measures. As noted in [36], the above property implies strict monotonicity with respect to stochastic domination:

Proposition 4.8.1. *Suppose that ν is absolutely continuous with respect to the expected empirical measure. Then whenever ν strictly stochastically dominates $\tilde{\nu}$, that is, whenever $\tilde{\nu} \neq \nu$ and there exists some coupling $(\tilde{w}, w)$ of $\tilde{\nu}$ and ν such that $\tilde{w} \leq w$ almost surely, we have $\mathbb{E}\tilde{T} \ll \mathbb{E}T$.*

Proof. Fix a coupling $(\tilde{w}, w)$ with $\tilde{w} \leq w$; since $\tilde{\nu} \neq \nu$, $\mathbb{P}(\tilde{w} < w) > 0$, and so one can find sufficiently small $a > 0, b > 0$, and Borel set $A \subset [0, \infty)$ such that $\nu(A) > 0$ and such that for every $y \in A$,

$$\mathbb{P}(\tilde{w} < y - a | w = y) \geq b.$$

[7] In the $G = \mathbb{Z}^d$ case it was recently proven by Bates [4] that for "generic" ν, the sequence of random empirical measures $\frac{1}{d(0,n\nu)}\sum_{e\in\pi}\delta_{w(e)}$ in a fixed direction almost surely weakly converges to a deterministic limit measure, an even stronger result than the existence of an *expected* empirical measure in a particular direction.

We thus have

$$\begin{aligned}
\mathbb{E}T(x,y) - \mathbb{E}\tilde{T}(x,y) &\geq \mathbb{E}[T(\pi) - \tilde{T}(\pi)] \\
&= \mathbb{E}\sum_{e\in\pi}(w(e) - \tilde{w}(e)) \\
&= \mathbb{E}\left[\sum_{e\in\pi}\mathbb{E}[w(e) - \tilde{w}(e)|w(e)]\right] \\
&\geq \mathbb{E}\left[\sum_{e\in\pi} ab\mathbb{1}_{w(e)\in A}\right] \\
&\gtrsim d(x,y),
\end{aligned}$$

where the last inequality follows from the fact that ν is absolutely continuous with respect to the expected empirical measure. □

If ν strictly stochastically dominates $\tilde{\nu}$, then $\tilde{\nu}$ is strictly more variable than ν, so our theorems above already prove strict monotonicity with respect to stochastic domination for graphs which admit detours and are either quasi-trees or have strict polynomial growth (in the later case, on the condition that ν is also exponential-subcritical). However, we can prove absolute continuity with respect to the empirical measure—and hence strict monotonicity with respect to stochastic domination—whether or not G admits detours directly, by using essentially identical methods to those above.

Theorem 4.1.5. *Let G be a bounded degree graph which is quasi-isometric to a tree. Then for any probability measure ν on $[0,\infty)$ with finite mean, ν is absolutely continuous with respect to the expected empirical measure of the associated first passage percolation T. Moreover, if ν strictly stochastically dominates a measure $\tilde{\nu}$, then $\mathbb{E}\tilde{T} \ll \mathbb{E}T$.*

Proof sketch. Let $A \subset [0,\infty)$ be Borel with $\nu(A) > 0$. Set $I_0 = A$. Set $C = 1, \epsilon = 1$. Then do the same construction as in the proof of Theorem 4.1.3. Whereas Theorem 4.1.3 gives $\gtrsim d(x,y)$ detours in expectation, now this construction gives $\gtrsim d(x,y)$ edges of π which have weights in $I_0 = A$, as desired. More explicitly, the construction gives a family of subgraphs $\{B_i\}$ with

$$\sum_i \mathbb{P}(\text{the geodesic } \pi : x \to y \text{ contains an edge } e \text{ in } B_i \text{ of weight } w(e) \in A) \gtrsim d(x,y).$$

Then arguing similarly as in the proof of Lemma 4.3.4, one gets a *disjoint* family $\{B_i\}$ with this property, and so one concludes that π contains $\gtrsim d(x,y)$ edges with weight lying in A in expectation, as desired. Lastly, the stochastic domination statement follows from Proposition 4.8.1. □

Theorem 4.1.6. *Let G be a graph of strict polynomial growth. Suppose that ν has finite mean and is exponential-subcritical. Then ν is absolutely continuous with respect to the expected empirical measure of the associated first passage percolation T. Moreover, if ν strictly stochastically dominates a measure $\tilde{\nu}$, then $\mathbb{E}\tilde{T} \ll \mathbb{E}T$.*

Proof sketch. Let $A \subset [0,\infty)$ be Borel with $\nu(A) > 0$ (one may without loss of generality replace A with a bounded positive ν-measure subset). Assume ν is exponential-subcritical, and choose q, Σ, and δ as in the first and simplest construction in the proof of Theorem 4.1.2, that is, in the case that ν has bounded support and $y_0 = \sup$. Also define the event A_1 as in that construction. Then define $E_w = E$ by

$$E := \left\{\omega \in [0,\infty)^{E(B(\Sigma R))} : \omega(e) \in [\inf, \inf + \delta) \text{ if } e \in E(B(\Sigma R - 1)), \omega(e) \in A \text{ otherwise}\right\}.$$

Using this construction, we have that for sufficiently large R, whenever A_1 holds, the T-geodesic π crosses B_i, and $w^* \in E_w = E$, there is a path entering $B(o_i, \Sigma R)$ which has T^*-weight strictly smaller than any path not entering $B(o_i, \Sigma R)$; hence the T^* geodesic enters $B(o_i, \Sigma R)$ and therefore contains an edge with weight valued in A. (Note that this is much simpler than in the proof of Theorem 4.1.2 when $y_0 \neq \sup$; this is because in that proof, we had to ensure that the T^*-geodesic made a long excursion away from the boundary of $B(o_i, \Sigma R)$, whereas here we only need it to hit a single edge with weight in A).

The Peierls lemma (Lemma 4.7.1) gives, in expectation, $\gtrsim d(x, y)$ B_i such that the T-geodesic visits B_i and A_1 holds; combining this with resampling (similar to Proposition 4.7.2) thus gives in expectation at least $\gtrsim d(x, y)$ $B(o_i, \Sigma R)$ which contain an edge of the T-geodesic with weight lying in A. Again arguing similarly as in the proof of Lemma 4.3.4 then gives that, in expectation, the T-geodesic contains $\gtrsim d(x, y)$ edges with weight lying in A, as desired. The stochastic domination statement again follows from Proposition 4.8.1. □

CHAPTER 5

Percolation cone for virtually nilpotent groups and failure of monotonicity

In Chapter 4, we proved several strict monotonicity theorems, but always under the assumption that the dominating or less variable measure ν is exponential-subcritical. One might wonder whether this assumption is an artifact of our proof or whether it is really a necessary assumption. The purpose of this chapter is to show that the exponential-subcritical assumption on ν is in fact in some sense necessary to obtain strict monotonicity, at least in the setting of Cayley graphs of virtually nilpotent groups.

5.1. The case $\inf \operatorname{supp} \nu = 0$

Firstly, in the case that $\inf \operatorname{supp} \nu = 0$, the exponential percolation threshold $\underline{p_c}$ as defined in Section 4.2 is well-known to be equal to the usual percolation threshold for transitive graphs and in particular Cayley graphs ([1], see also [12]); that is, if $\nu(\{0\}) > \underline{p_c}$ then with probability 1 there exists an infinite connected subset of edges which all have weight 0. The following proposition is then an easy corollary of recent results in supercritical percolation on polynomial growth graphs [9]:

Proposition 5.1.1. *Let G be a transitive graph of polynomial growth, and let $\nu \in \operatorname{Prob}([0,\infty))$ with finite mean such that $\nu(\{0\}) > \underline{p_c}$. Then*

$$\lim_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y)}{d(x,y)} \to 0.$$

Proof. It suffices to show that for fixed o, $\lim_{x\to\infty} \frac{\mathbb{E}T(o,x)}{d(o,x)} = 0$. Since $\nu(\{0\}) > p_c$, the collection of edges in E such that $w(e) = 0$ has a unique[1] infinite component; denote the set of vertices in this component by C. Note that we have

$$T(o,x) \leq (\inf_{v\in C} T(o,v)) + (\inf_{w\in C} T(w,x))$$

Therefore, by transitivity,

$$\mathbb{E}T(o,x) \leq 2\mathbb{E}[T(o,C)]$$

so it suffices to show that $\mathbb{E}[T(o,C)] < \infty$. In fact it suffices to show that $\mathbb{E}[d(o,C)] < \infty$, since if γ is a path with a minimal number of edges from o to C, then $\mathbb{E}[T(o,C)] \leq \mathbb{E}[T(\gamma)]$, and conditioning gives that $\mathbb{E}[T(\gamma)] \leq (\mathbb{E}\hat{w})\mathbb{E}[|\gamma|] = (\mathbb{E}\hat{w})\mathbb{E}[d(o,C)]$, where $\hat{w}$ is distributed as $w|w > 0$ (so integrability of w implies integrability of $\hat{w}$). The fact that $\mathbb{E}[d(o,C)] < \infty$ follows from Proposition 1.3 of [9] as follows. Since $\nu(\{0\}) > p_c$, that proposition tells us that for all sufficiently large n,

$$\mathbb{P}\left(B(o,\frac{n}{10}) \leftrightarrow B(o,n)^c, U(\frac{n}{5},\frac{n}{2})\right) \geq 1 - e^{-\sqrt{n}}, \tag{5.1.1}$$

where $B(o,\frac{n}{10}) \leftrightarrow B(o,n)^c$ means that there is a path of edges of weight 0 from some vertex in $B(o,\frac{n}{10})$ to some vertex in $B(o,n)^c$, and $U(\frac{n}{5},\frac{n}{2})$ is the event that there is at most one connected component of edges of weight zero which intersects both $B(o,\frac{n}{5})$ and $B(o,\frac{n}{2})^c$. Note that the uniqueness event $U(\frac{n}{5},\frac{n}{2})$ allows us to "glue" the connections $\{B(o,\frac{n}{10}) \leftrightarrow B(o,n)\}$ at different

[1]Since G has subexponential growth, it is in particular amenable. The original argument of Burton-Keane [7], then shows that there is only one infinite cluster; see also [19] for an explicitly general proof.

scales n, so that

$$\bigcap_{n=N}^{\infty} \{B(o,n) \leftrightarrow B(o,10n)^c, U(2n,5n)\} \subset \{B(o,N) \leftrightarrow \infty\}.$$

Therefore, using the layer-cake formula for expectation and using a union bound on the complement, we get

$$\begin{aligned}
\mathbb{E}d(o,C) &= \sum_{N=0}^{\infty} \mathbb{P}(d(o,C) > N) \\
&= \sum_{N=0}^{\infty} \mathbb{P}(B(o,N) \not\leftrightarrow \infty) \\
&\le N_0 + \sum_{N=N_0}^{\infty} \mathbb{P}\left(\bigcup_{n=N}^{\infty} \{B(o,n) \leftrightarrow B(o,10n)^c, U(2n,5n)\}^c \right) \\
&\le N_0 + \sum_{N=N_0}^{\infty} \sum_{n=N}^{\infty} e^{-\sqrt{n}} < \infty,
\end{aligned}$$

where N_0 is such that (5.1.1) holds for all $10n \ge N$. □

Remark 5.1.1. *If one is willing to assume ν has an exponential moment, one can prove the above proposition using the same methods we will use below for the case* $\inf \operatorname{supp} \nu > 0$*. That is, one argues that the normalized point-to-sphere passage time tends to zero almost surely, and then uses Talagrand concentration, rather than by showing that $\mathbb{E}[d(o,C)]$ is finite. In fact, this gives a proof under the weaker assumption that the growth of G is subexponential, not necessarily polynomial.*

This in particular implies that strict monotonicity is impossible in this regime:

Corollary 5.1.1. *Let G, ν be as above such that $\nu \neq \delta_0$. Then there exists $\tilde{\nu}$ which is strictly stochastically dominated by ν but is such that*

$$\lim_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y) - \mathbb{E}\tilde{T}(x,y)}{d(x,y)} = 0.$$

Proof. Any $\tilde{\nu}$ which is strictly stochastically dominated by ν will suffice; stochastic domination implies that the limit in question is nonnegative, and it is at most equal to $\lim_{d(x,y)\to\infty} \frac{\mathbb{E}T(x,y)}{d(x,y)}$, which by the above is 0. To get such $\tilde{\nu}$, simply take a nontrivial convex combination of ν and δ_0. □

Note that we have shown that when $\nu(\{0\}) > p_c$, strict monotonicity fails *in all directions*. This will contrast the next case, where $\inf \operatorname{supp} \nu > 0$.

5.2. The case $\inf \operatorname{supp} \nu > 0$

Next, we cover the case that $a := \inf \operatorname{supp} \nu > 0$ and $\nu(\{a\}) > \underline{\vec{p_c}}$ (under an extra moment assumption on ν). This case will be a bit more subtle, since monotonicity will only fail in *some* directions. To get cleaner statements in terms of scaling limits, we assume that G is the Cayley graph of a virtually nilpotent group, but in principle the arguments should apply to any almost-transitive graph of polynomial growth.

As in the classical case of $\mathbb{Z}^d$, the reason for failure of monotonicity will be that, once $\nu(\{a\})$ exceeds a certain threshold, edges of weight exactly a are common enough that there are infinite edge-geodesics in some directions with all weight a. Then exactly as above we can produce a counterexample to monotonicity, since these paths already (up to error $o(d(o,x))$) have the smallest possible passage time under ν.

The method of proof is as follows: first, we define the geodesic percolation threshold $\vec{p_c}$ and show sharpness for transitive graphs (and hence equality with the exponential geodesic percolation threshold $\underline{\vec{p_c}}$ as defined above). Then it is not hard to show that the normalized point-to-sphere passage times $T(o, \partial B(R))/R$ converge to a almost surely. To then conclude that *in some direction* $x_n \to \xi$ we have $\lim_{n\to\infty} \mathbb{E}T(o, x_n)/d(o, x_n) = a$, we assume an exponential moment and use Talagrand concentration (as well as polynomial growth of our graph). This will give us

Theorem 5.2.1. *Let G be a Cayley graph of a finitely generated virtually nilpotent group, Let* L_∞ *denote the associated graded nilpotent Lie group associated to G, and let* d_∞ *denote that CC-metric on* L_∞ *which is the scaling limit of the graph metric d on V. If* $a := \inf \operatorname{supp} \nu > 0$*,* $\nu(\{a\}) > \underline{\vec{p_c}}(G)$*, and* $\int \exp(\alpha t) d\nu(t) < \infty$ *for some* $\alpha > 0$*, then there exists some* $\xi \in L_\infty$ *with* $d_\infty(1, \xi) = 1$ *and* $d_\Phi(1, \xi) = a$*, where* d_Φ *is the scaling limit of the FPP metric T associated to* ν*. Moreover, if* $\nu \neq \delta_a$*, there exists* $\tilde{\nu}$ *such that* ν *strictly stochastically dominates* $\tilde{\nu}$ *but* $d_{\tilde{\Phi}}(1, \xi) = d_\Phi(1, \xi) = a$*.*

This theorem shows that subcriticality is necessary for strict monotonicity, but it also has the strange feature that the direction ξ of non-monotonicity *a priori* depends on ν, while in principle it should only depend on $\nu(\{a\})$ (since it should be a direction in which edges of weight a are "geodesically percolating"). With a bit more care we can formalize the idea that for each $p > \vec{p_c}$ there are directions which are "geodesically percolating," and this gives a sort of "percolation cone" analogous to the $\mathbb{Z}^d$ case (see [29], Section 2.5.1 of [2]). This in particular allows us to get a direction ξ which does not depend on ν:

Theorem 5.2.2. *Let G, L_∞, d_∞ as above. Then there exists $\xi \in L_\infty$, $d_\infty(1,\xi) = 1$ with the following property. If* $\inf \operatorname{supp} \nu = a > 0$, $\nu(\{a\}) > \vec{p_c}$, *and* $\int \exp(\alpha t) d\nu(t) < \infty$ *for some* $\alpha > 0$, *then $d_\Phi(1,\xi) = a$. Moreover, if $\nu \neq \delta_a$, then there exists $\tilde{\nu}$ which is strictly stochastically dominated by ν such that $d_{\tilde{\Phi}}(1,\xi) = d_\Phi(1,\xi) = a$.*

The rest of the chapter is devoted to proving these two theorems.

5.3. Sharpness of the geodesic percolation phase transition

Let $G = (V, E)$ be a graph, $p \in [0,1]$, and let G_p be the random subgraph of G with vertex set V and edge set given by independently including each $e \in E$ with probability p and excluding it with probability $1-p$. We call the included edges *open* and the excluded edges *closed*. We are concerned not just with connectivity in G_p but with connectivity *by edge geodesics*. Recall that an edge path $\pi : x \to y$ in G is called *edge-geodesic* if $|\pi| = d(x,y)$, that is, π uses the least number of edges possible to connect x and y.

For two vertices $x, y \in V$, we write $x \underset{\text{geo}}{\longleftrightarrow} y$ if there exists an open edge-geodesic path from x to y. We write $x \underset{\text{geo}}{\longleftrightarrow} \infty$ if there exists an infinite open path π starting from x such that every finite subpath of π is edge-geodesic. We then define

$$\vec{p_c}(G) := \inf\left\{p \in [0,1] : \mathbb{P}_p\left(\bigcup_{x\in V}\{x \underset{\text{geo}}{\longleftrightarrow} \infty\}\right) > 0\right\}.$$

Note that for each $x \in V$, we have

$$\{x \underset{\text{geo}}{\longleftrightarrow} \infty\} = \bigcap_{R=1}^{\infty}\{x \underset{\text{geo}}{\longleftrightarrow} \partial B(R)\},$$

and that the right hand side of this event is an intersection of events only depending on finitely many edges, so the event defining $\vec{p_c}(G)$ is measurable. Note also that the event $\bigcup_{x\in V}\{x \underset{\text{geo}}{\longleftrightarrow} \infty\}$ that there exists an infinite open edge-geodesic is tail-measurable, and so for each p it has probability either 0 or 1 by the Kolmogorov 0-1 Law. On the other hand, if G is vertex-transitive, then in particular $\mathbb{P}(x \underset{\text{geo}}{\longleftrightarrow} \infty)$ does not depend on $x \in V$, and so a union bound shows that, fixing any basepoint $o \in V$, we have

$$\vec{p_c}(G) = \inf\left\{p \in [0,1] : \mathbb{P}_p\left(o \underset{\text{geo}}{\longleftrightarrow} \infty\right) > 0\right\}.$$

Recall our definition of $\underline{\vec{p_c}}(G)$ from Section 4.2. It is clear from this definition that $\underline{\vec{p_c}}(G) \leq \vec{p_c}$. The key theorem of this section shows that for transitive graphs, $\underline{\vec{p_c}}(G) = \vec{p_c}$.

Theorem 5.3.1 (Sharpness of geodesic percolation phase transition)**.** *Let G be a transitive graph with no parallel edges. Suppose that $p < \vec{p_c}(G)$. Then for some $C, c > 0$ (depending on p)*

$$\mathbb{P}_p(x \underset{\text{geo}}{\longleftrightarrow} y) \leq C\exp(-cd(x,y)).$$

In particular, $\vec{p_c} = \underline{\vec{p_c}}$.

The proof of Theorem 5.3.1 follows closely the proof of sharpness of the percolation phase transition by Hugo Duminil-Copin and Vincent Tassion [12]. Surprisingly, their method adapts almost immediately to the geodesic percolation setting, the only technical point being that not all edge-geodesics compose to form edge-geodesics. For completeness, we present the full argument here, emphasizing the places where the argument is different.

Remark 5.3.1. *The assumption that G has no parallel edges is only for simplicity of exposition. If G has parallel edges, we can take $\tilde{G}$ to be the graph which has the same vertex set as*

G and whose edge set consists of pairs of vertices which have at least one edge between them in G (so in particular $\tilde{G}$ has no parallel edges). If we set $\beta > 0$ such that $1 - p = e^{-\beta}$, then the probability that all of k parallel edges are closed in G_p is $e^{-k\beta}$. Therefore, if one defines $J_{x,y}$ to be the number of edges from x to y in G, then [geodesic] connectivity on G_p is equivalent to [geodesic] connectivity on the random graph $\tilde{G}_\beta$, where $\tilde{G}_\beta$ is obtained from $\tilde{G}$ by, for each $\{x, y\} \in E(\tilde{G})$, deleting the edge independently with probability $e^{-J_{x,y}\beta}$ and otherwise retaining it. In fact, the proof of sharpness for percolation in [12] is already presented in the setting that we have a priori different coupling constants $J_{x,y}$ for different edges. The modifications to go from percolation to geodesic percolation in this setting will be identical to the modifications presented here (essentially, a slightly different definition of the quantity $\varphi_\beta(S)$).

To deal with the fact that not every composition of edge-geodesics is an edge-geodesic, we introduce the following definition:

Definition 5. *For $x, y, z \in V$, we say that y is* between x and z *if there exist edge geodesics $\pi_1 : x \to y$, $\pi_2 : y \to z$ such that the composition $\pi_1 * \pi_2 : x \to z$ is also an edge geodesic. In this case we write* $[x - y - z]$.

Note that $[x - y - z]$ if and only if there exists an edge geodesic $\pi : x \to z$ such that $y \in V(\pi)$.

As an example, consider the standard Cayley graph $\mathbb{Z}^2$. For any $n > 0$, the points which are between $(0, 0)$ and $(n, 0)$ are precisely those of the form $(i, 0)$ with $0 \leq i \leq n$. For any $n, m > 0$ the points between $(0, 0)$ and (n, m) are those of the form (i, j) with $0 \leq i \leq n, 0 \leq j \leq m$.

Proposition 5.3.1. *The following are equivalent:*

(1) y is between x and z.

(2) $d(x,z) = d(x,y) + d(y,z)$.

(3) For any *edge geodesics* $\pi_1 : x \to y$ *and* $\pi_2 : y \to z$*, the composition* $\pi_1 * \pi_2 : x \to z$ *is an edge geodesic.*

Proof. (3) ⇒ (1) is immediate from the definition of $[x - y - z]$.

(1) ⇒ (2): Assume $\pi_1 : x \to y$, $\pi_2 : y \to z$ are edge-geodesics such that the composition $\pi_1 * \pi_2$ is edge-geodesic. Then

$$d(x,z) = |\pi_1 * \pi_2| = |\pi_1| + |\pi_2| = d(x,y) + d(y,z).$$

(2) ⇒ (3): Let $\pi_1 : x \to y$ and $\pi_2 : y \to z$ be edge-geodesics. Their composition $\pi_1 * \pi_2$ is a path from x to z of length $|\pi_1| + |\pi_2| = d(x,y) + d(y,z)$. If (2) holds then this is equal to $d(x,z)$ and so $\pi_1 * \pi_2$ is an edge-geodesic. □

Let us assume from now on that G is vertex-transitive and fix a basepoint $o \in V$. Now, for each finite set $S \subset V$ containing o we define

$$\varphi_p(S) := p \sum_{x \in S} \sum_{\substack{y \notin S \\ \{x,y\} \in E \\ [o - x - y]}} \mathbb{P}_p(o \overset{S}{\underset{\text{geo}}{\longleftrightarrow}} x),$$

where $\{o \overset{S}{\underset{\text{geo}}{\longleftrightarrow}} x\}$ is the event that there is an open edge-geodesic from o to x all of whose edges have both endpoints in S. Note that this quantity is the same as $\varphi_\beta(S)$ from [12], except that the sum is restricted to edges $\{x,y\}$ such that x is between o and y and the connection events are by geodesics (and there is a change of variables $p = 1 - e^{-\beta}$).

We now define a threshold $\tilde{p}_c$ by

$$\tilde{p}_c := \sup\{p : \varphi_p(S) < 1 \text{ for some finite } S \subset V \text{ containing } o\}.$$

Again as in [12], we will prove sharpness of the threshold $\tilde{p}_c$, which will imply that $\underline{\vec{p_c}} = \tilde{p}_c = \vec{p_c}$. That is, we prove Theorem 5.3.1 by proving the following lemma:

Lemma 5.3.1. *For $p > \tilde{p}_c$ we have*

$$\mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} \infty) \geq \frac{1}{p}(p - \tilde{p}_c),$$

and for each $p < \tilde{p}_c$ we have $c_p > 0, C_p < \infty$ such that

$$\mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} B(o, R)^c) \leq C_p e^{-c_p R}.$$

Proof of Theorem 5.3.1 given Lemma 5.3.1. By the first statement in Lemma 5.3.1, we see that $\vec{p_c} \leq \tilde{p}_c$; by the second statement in Lemma implies that $\tilde{p}_c \leq \underline{\vec{p_c}}$. Since we already know $\underline{\vec{p_c}} \leq \vec{p_c}$, this gives

$$\tilde{p}_c = \underline{\vec{p_c}} = \vec{p_c} = \tilde{p}_c.$$

The exponential decay statement in Theorem 5.3.1 is equivalent to the statement that $\underline{\vec{p_c}} = \vec{p_c}$. ◻

Now we prove Lemma 5.3.1.

Proof of Lemma 5.3.1. First, let us prove exponential decay of geodesic connection probabilities below the threshold. Fix $p < \tilde{p}_c$. We then have some finite set $S \subset V$ containing o such that $\varphi_p(S) < 1$. Choose $r < \infty$ sufficiently large that all edges having at least one endpoint in S have both endpoints in $B(o,r)$. Then for $R > r$, if o is connected to $B(o,R)^c$ by an open edge-geodesic, then there is an open edge-geodesic inside of S connecting o to the boundary of S, there is an open edge-geodesic in S connecting S to S^c, there is an open edge-geodesic connecting the other endpoint of that edge to $B(o,R)^c$, and the composition of all three of these paths is edge geodesic. Thus, by the BK inequality (see e.g. Section 2.3 of [15])

$$\begin{aligned}\mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c) &\le \sum_{x\in S} \sum_{\substack{y\notin S\\ \{x,y\}\in E\\ [o-x-y]}} \mathbb{P}(o \underset{\text{geo}}{\overset{S}{\longleftrightarrow}} x) p \mathbb{P}(y \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c)\\ &\le \varphi_p(S)\mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} B(o,R-L)^c),\end{aligned}$$

where in the last line we have used the fact that a geodesic connecting y to $B(o,R)^c$ must necessarily connect y to $B(y,R-L)^c$, together with vertex-transitivity. By induction we have

$$\mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c) \le (\varphi_p(S))^{\lfloor R/L \rfloor}$$

and therefore (since $\varphi_p(S) < 1$ and L is fixed as $R \to \infty$) we have the desired exponential decay in R.

Now let us prove the lower bound on connection to infinity above the threshold. We do this by proving the following differential inequality:

$$\frac{d}{dp}\mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c) \ge \frac{1}{p}\left(\inf_{o\in S\subset B(o,R)} \varphi_p(S)\right)\left(1 - \mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c)\right). \tag{5.3.1}$$

Given (5.3.1), for all $p > \tilde{p}_c$ we have

$$\frac{\frac{d}{dp}\mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c)}{1 - \mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c)} \ge \frac{1}{p};$$

so integrating over $p' \in (\tilde{p}_c, p]$ gives

$$\log\left(\frac{1-\mathbb{P}_{\tilde{p}_c}(o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c)}{1-\mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c)}\right) \geq \log\frac{p}{\tilde{p}_c}.$$

Exponentiating both sides and cross-multiplying then gives

$$p - p\mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c) \leq \tilde{p}_c - \tilde{p}_c\mathbb{P}_{\tilde{p}_c}(o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c) \leq \tilde{p}_c,$$

which can be rearranged to

$$\mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c) \geq \frac{p-\tilde{p}_c}{p},$$

and taking $R \to \infty$ gives the desired inequality.

So it only remains to prove (5.3.1). Since $A := \{o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c\}$ is an increasing event depending on the state of only finitely many edges, by Russo's formula (see e.g. Section 2.2 of [15]).

$$\begin{aligned}\frac{d}{dp}\mathbb{P}_p(o \underset{\text{geo}}{\longleftrightarrow} B(o,R)^c) &= \sum_{e\in E(B(o,R))}\mathbb{P}_p(e \text{ is pivotal for } A)\\ &\geq \sum_{e\in E(B(o,R))}\mathbb{P}_p(e \text{ is closed and pivotal for } A).\end{aligned}$$

We then note that an edge $e = \{x, y\}$ is closed and pivotal for A if and only if: o is connected to x by a geodesic, y is connected to some $b \in B(o,R)^c$ by a geodesic, with y between o and b, x is between o and y, and no geodesic connects o to $B(o,R)^c$ (or the same event is true with the roles of x and y reversed). We can fruitfully rephrase this by defining the following random set

$\mathcal{S} \subset V$:

$$\mathcal{S} := \{v \in V : \forall b \in B(o,R)^c \text{ such that } [o - v - b], v \underset{\text{geo}}{\not\leftrightarrow} b\}.$$

That is, $\mathcal{S}$ is the set of vertices which are not connected to the boundary of $B(o,R)$ by an open edge-geodesic which is composable with an edge-geodesic from the origin. Then $\{x,y\}$ is closed and pivotal for A if and only if $o \in \mathcal{S}$, $x \in \mathcal{S}$, $y \notin \mathcal{S}$, $[o - x - y]$, and $o \overset{\mathcal{S}}{\underset{\text{geo}}{\longleftrightarrow}} x$ (or the same event holds with the roles of x and y reversed). Summing over the different possibilities for $\mathcal{S}$, we have

$$\begin{aligned}\sum_{e \in E(B(o,R))} \mathbb{P}_p(e \text{ closed, pivotal for } A) &= \sum_{e \in E(B(o,R))} \sum_{o \in S \subset B(o,R)} \mathbb{P}_p(e \text{ closed, pivotal for } A, \mathcal{S} = S) \\ &= \sum_{o \in S \subset B(o,R)} \sum_{\substack{x \in S, y \notin S \\ [o-x-y]}} \mathbb{P}_p(\{x,y\} \text{ closed, pivotal for } A, \mathcal{S} = S).\end{aligned}$$

Clearly the event $\{o \overset{S}{\underset{\text{geo}}{\longleftrightarrow}} x\}$ depends only on edges with both endpoints in S. On the other hand, we claim that the event $\{\mathcal{S} = S\}$ only depends on edges with at least one endpoint in S^c. This is because for any $S \subset B(o,R)$,

$$\{\mathcal{S} = S\} = \left(\bigcap_{y \notin S} \bigcup_{\substack{b \in B(o,R)^c \\ [o-y-b]}} \{y \overset{S^c}{\underset{\text{geo}}{\longleftrightarrow}} b\} \right) \cap \left(\bigcap_{\substack{x \in S, y \notin S \\ \{x,y\} \in E \\ [o-x-y]}} \{\{x,y\} \text{ closed}\} \right).$$

Above we have used several times the fact that if $[o - y - b]$ and $[y - x - b]$ then $[o - x - b]$. For instance, for each $y \notin S$, the fact that $y \notin \mathcal{S}$ implies that y is connected to some $b \in B(o,R)^c$ with $[o - y - b]$ by an open edge-geodesic; but such an open edge-geodesic cannot use any vertex $x \in S$, or else we would have $[o - x - b]$ and $x \underset{\text{geo}}{\longleftrightarrow} b$, implying that $x \notin \mathcal{S}$.

Now, since $\{o \overset{S}{\underset{\text{geo}}{\longleftrightarrow}} x\}$ and $\{\mathcal{S} = S\}$ depend on disjoint edges, by independence we have that for S with $o, x \in S$ and $y \notin S$ that

$$\mathbb{P}_p(\{x,y\} \text{ closed, pivotal for } A, \mathcal{S} = S) = \mathbb{P}_p(o \overset{S}{\underset{\text{geo}}{\longleftrightarrow}} x, \mathcal{S} = S) = \mathbb{P}_p(o \overset{S}{\underset{\text{geo}}{\longleftrightarrow}} x)\mathbb{P}_p(\mathcal{S} = S).$$

Thus we have

$$\begin{aligned}
\sum_{e \in E(B(o,R))} \mathbb{P}_p(e \text{ closed, pivotal for } A) &\geq \sum_{o \in S \subset B(o,R)} \sum_{\substack{x \in S, y \notin S \\ [o-x-y]}} \mathbb{P}_p(o \overset{S}{\underset{\text{geo}}{\longleftrightarrow}} x)\mathbb{P}_p(\mathcal{S} = S) \\
&= \sum_{o \in S \subset B(o,R)} \mathbb{P}_p(\mathcal{S} = S) \frac{1}{p} \sum_{\substack{x \in S, y \notin S \\ [o-x-y]}} p\mathbb{P}_p(o \overset{S}{\underset{\text{geo}}{\longleftrightarrow}} x) \\
&\geq \frac{1}{p} \left(\inf_{o \in S \subset B(o,R)} \varphi_p(S) \right) \sum_{o \in S \subset B(o,R)} \mathbb{P}_p(\mathcal{S} = S) \\
&= \frac{1}{p} \left(\inf_{o \in S \subset B(o,R)} \varphi_p(S) \right) \mathbb{P}_p(o \in \mathcal{S}) \\
&= \frac{1}{p} \left(\inf_{o \in S \subset B(o,R)} \varphi_p(S) \right) \mathbb{P}_p(o \underset{\text{geo}}{\not\leftrightarrow} B(o,R)^c),
\end{aligned}$$

and so we have established (5.3.1), so we are done. □

5.4. Proof of Theorem 5.2.1

Given Theorem 5.3.1, it is almost immediate that when $\nu(\{a\}) > \underline{\vec{p_c}}$, the renormalized point-to-sphere passage times $T(o, \partial B(R))/R$ converge to a almost surely. Call an edge e *open* if $w(e) = a$, *closed* otherwise.

Proposition 5.4.1. *Suppose that $\nu(\{a\}) > \underline{\vec{p_c}}$. Then*

$$\lim_{R \to \infty} \frac{T(o, \partial B(R))}{R} = a.$$

Proof. Since inf supp $\nu = a$, the inferior limit of the quantity in question is at least a, so we wish to show that the superior limit is at most a. By Theorem 5.3.1, $\nu(\{a\}) > \vec{p}_c$, so that with probability 1 there exists an open infinite edge geodesic π. Choose y (random) to be a starting point of such an infinite open edge-geodesic; take y to be as close as possible to o in the graph metric d, breaking ties according to some arbitrary ordering on vertices. Then note that we have

$$T(o, \partial B(R)) \leq T(o, y) + T(y, \partial B(R)).$$

For each sufficiently large R, pick z_R to be the first intersection of the edge-geodesic out of y with $\partial B(R)$. Then we have $d(o, z_R) = R$, and following the open edge-geodesic from y to z_R we have

$$T(y, \partial B(R)) \leq ad(y, z_R) \leq a(d(o, y) + d(o, z_R)) = ad(o, y) + aR.$$

Therefore, for all R we have

$$\frac{T(o, \partial B(R))}{R} \leq \frac{T(o, y)}{R} + \frac{ad(o, y)}{R} + a.$$

Since $T(o, y)$ and $d(o, y)$ are almost surely finite and do not depend on R, we have that the superior limit of $T(R)/R$ is almost surely at most a, and so we are done. □

Now we want to show that if G has subexponential growth, we can relate the point-to-sphere passage times to expected passage times in a sequence of directions. For this we use the following concentration result:

Theorem 5.4.1 (Talagrand [33] Proposition 8.3)**.** *For a locally finite graph G and a measure ν with a finite exponential moment, and $\epsilon > 0$, there exist constants $c, C > 0$ such that we have*

$$\mathbb{P}\left(\frac{|T(x,y) - \mathbb{E}T(x,y)|}{d(x,y)} > \epsilon\right) \leq C\exp(-cd(x,y)).$$

Remark 5.4.1. *Talagrand's result is much stronger, but implies the large-deviations upper bound needed above. In fact, we will only use that deviations* below *the mean happen with probability which decays faster than any inverse polynomial in $d(x,y)$. It is reasonable to expect that such deviations have exponentially small probability for any graph (as is the case for all $\mathbb{Z}^d$, see Theorem 5.2 in [26]), but the proofs in the literature do not seem immediately adaptable to the general setting.*

Given this concentration, we have:

Proposition 5.4.2. *Let G be a transitive graph of subexponential growth, and suppose that ν has a finite exponential moment. Then almost surely*

$$\limsup_{R\to\infty} \frac{T(o, \partial B(R)}{R} \geq \limsup_{R\to\infty} \inf_{x:d(o,x)=R} \frac{\mathbb{E}T(o,x)}{R}.$$

Proof. Note that for each $\epsilon > 0$,

$$\begin{aligned}
&\sum_{R=1}^{\infty} \mathbb{P}\left(\frac{T(o,\partial B(R))}{R} < \inf_{y:d(o,y)=R} \frac{\mathbb{E}T(o,y)}{R} - \epsilon\right) \\
&\leq \sum_{R=1}^{\infty} \sum_{x:d(o,x)=R} \mathbb{P}\left(\frac{T(o,x)}{R} < \inf_{y:d(o,y)=R} \frac{\mathbb{E}T(o,y)}{R} - \epsilon\right) \\
&\leq \sum_{R=1}^{\infty} \sum_{x:d(o,x)=R} \mathbb{P}\left(\frac{T(o,x)}{R} < \frac{\mathbb{E}T(o,x)}{R} - \epsilon\right) \\
&\leq \sum_{R=1}^{\infty} Ce^{-cR}|B(R)| \\
&\leq \sum_{R=1}^{\infty} Ce^{-c'R} < \infty,
\end{aligned}$$

where in the second-to-last line we have used Theorem 5.4.1, and in the last line we have used subexponential growth of G. Thus, by Borel-Cantelli, for each $\epsilon > 0$ we have

$$\limsup_{R\to\infty} \frac{T(o,\partial B(R))}{R} \geq \limsup \inf_{x:d(o,x)=R} \frac{\mathbb{E}T(o,x)}{R} - \epsilon$$

almost surely, and taking $\epsilon \to 0$ gives the desired lower bound. □

Remark 5.4.2. *The subexponential growth assumption here is important; for example, if we take G to be the standard Cayley graph of F_2, then we have $\mathbb{E}[T(o,x)]/d(o,x) = \mathbb{E}w$ for all x, but $T(o,\partial B(R))/R$ will typically be much smaller. (See [4] for a characterization of* $\lim_{R\to\infty} T(o,\partial B(R))/R$ *for regular trees.)*

Now we can prove Theorem 5.2.1:

Proof of Theorem 5.2.1. By Proposition 5.4.1, we have that

$$\lim_{R\to\infty} \frac{T(o, \partial B(R))}{R} = a$$

almost surely. Moreover, since each $\mathbb{E}T(o, x)/d(o, x) \geq a$, Proposition 5.4.2 then gives us that

$$\lim_{R\to\infty} \inf_{x:d(o,x)=R} \frac{\mathbb{E}T(o, x)}{d(o, x)} = a.$$

This implies that there exists some sequence $x_R \in V$ with each $d(o, x_R) = R$ such that

$$\lim_{R\to\infty} \frac{\mathbb{E}T(o, x_R)}{R} = a.$$

By compactness of the unit sphere in L_∞, the sequence $\mathrm{scl}_{\frac{1}{R}} x_R$ has a limit point $\xi \in L_\infty$ with $d_\infty(1, \xi) = 1$, and the fact that $\mathrm{scl}_{\frac{1}{R}} : (V, \mathbb{E}T) \to (L_\infty, d_\Phi)$ is a sequence of Gromov-Hausdorff approximations then gives $d_\Phi(1, \xi) = a$, as desired.

To contradict monotonicity in direction ξ, simply take $\tilde{\nu}$ to be some nontrivial convex combination of ν and δ_a. □

5.5. The percolation cone

A theorem of Marchand ([29], see also the statement of Theorem 2.24 in [2]) says that for the standard Cayley graph of $\mathbb{Z}^2$, the directions ξ with smallest possible time constant ($\lim_{n\to\infty} \frac{T(0,n\xi)}{|n\xi|_1} = a$) are precisely those directions which lie in the "percolation cone"—roughly speaking, those directions ξ such that an oriented percolation model on $\mathbb{Z}^2$ with parameter $p > \nu(\{a\})$ almost surely has an infinite open *directed* path in direction ξ. The set of such directions is nonempty if and only if $\nu(\{a\}) \geq \vec{p}_c$, where $\vec{p}_c$ is the oriented percolation threshold on the standard Cayley graph of $\mathbb{Z}^2$. Note that for the case of the standard Cayley graph of $\mathbb{Z}^2$, this exactly coincides with the geodesic percolation threshold $\vec{p_c}$ as we defined it in Section 5.3, so our notation is consistent.

In this section, we begin building an analogous picture. That is, for G the Cayley graph of a virtually nilpotent group, we show that for each $p > \vec{p_c} = \underline{\vec{p_c}}$ we have a set of "directions" C_p in the unit ball of (L_∞, d_∞) such *geodesic* percolation with parameter p almost surely has infinite open edge geodesics in direction (nearly) ξ. We show that C_p is nonempty for all $p > \vec{p_c} = \underline{\vec{p_c}}$. We then show that if $\xi \in C_{\nu(\{a\})}$, then $d_\Phi(1,\xi) = a$, i.e. the "time constant" is the smallest possible. Finally we use this to show Theorem 5.2.2. One might conjecture a sort of converse, i.e. that if ξ has $d(1,\xi) = 1$ but $\xi \notin C_{\nu(\{a\})}$ then $d_\Phi(1,\xi) > a$—in other words, the percolation cone is the *only* place where we have the minimum possible time constants. Proving or refuting this is left for future work.

We want to establish that for each $p > \vec{p_c}$ there is a nonempty set of directions which have time constant a for *any* ν with $\nu(\{a\}) \geq p$. The intuition behind the argument is as follows. Since $p > \vec{p_c}$, there exists some infinite open edge-geodesic. Even if this edge-geodesic does not have a well-defined direction, by compactness of the scaling limit, there must be some set of directions which it visits infinitely often. The set of directions which are visited infinitely often by an infinite open edge-geodesic should be deterministic by the Kolmogorov 0-1 law, and these will be our C_p. Lastly (assuming an exponential moment for ν) establishing that $T(o, x_n) \leq ad(o, x_n)$ for some $\mathrm{scl}_{\frac{1}{d(1,x_n)}} x_n \to \xi$ establishes that $d_\Phi(\xi) = a$ by Gromov-Hausdorff convergence.

The main technical issue with the above argument is that the event that there exists an infinite open edge-geodesic which visits direction ξ infinitely often is not a priori measurable. Unlike the event that there exists an infinite open edge-geodesic, we cannot simply write it as an intersection of events involving the existence of finite open edge-geodesics with some property. For instance, if we assume that we have infinitely many finite open edge geodesics (say, starting from the same point) which visit direction ξ, if we try to take a subsequential limit to get an infinite open edge-geodesic path, this infinite path may not visit ξ at all. Therefore, we have to define a slightly different family of events. A second issue is making sense of what it means to "visit direction ξ"; roughly speaking, we want there to be infinitely many rescaled open edge-geodesics visiting any neighborhood of ξ.

To this end, let us define the following events. Let $U \subset L_\infty$ be a bounded open set which intersects the unit sphere $S := \{\xi \in L_\infty : d_\infty(1, \xi) = 1\}$. For each finite subset $F \subset V$ and each $R \in \mathbb{N}$, we define

$$A_{U,F,R} := \{\exists x \in F, y \in V, \gamma : x \to y \text{ open edge geodesic s.t. } d(o, y) \geq R, \mathrm{scl}_{\frac{1}{R}} y \in U\}.$$

(Recall the map $\mathrm{scl}_{1/R} : \Gamma \to L_\infty$ defined in Section 3.8 which approximately embeds a rescaled copy of Γ into L_∞). We interpret $A_{U,F,R}$ as the the event that there is an open edge-geodesic with starting point in F and endpoint in direction U at scale R. Note that since F and $\mathrm{scl}_{\frac{1}{R}}^{-1}(U)$ are finite, there are only finitely many edge-geodesics between them in G; that is, $A_{U,F,R}$ depends on only finitely many edges and hence is measurable.

Next define

$$A_{U,F} := \limsup_{R \to \infty} A_{U,F,R} := \bigcap_{r=1}^{\infty} \bigcup_{R=r}^{\infty} A_{U,F,R},$$

the event that $A_{U,F,R}$ happens for infinitely many R. That is, at infinitely many scales, there is an open edge-geodesic starting from F and landing in direction U.

Lastly, we set

$$A_U := \bigcup_{F \subset V \text{ finite}} A_{U,F},$$

the event that there exists some finite F such that $A_{U,F}$ holds. Note that A_U is measurable, since it was obtained from the events $A_{U,F,R}$ by taking countable unions and intersections.

Proposition 5.5.1. *The event A_U is tail-measurable. Thus for each p, $\mathbb{P}_p(A_U) \in \{0, 1\}$.*

Proof. Let $\omega \in \{0, 1\}^E$ be a configuration of edges, and suppose that $\omega \in A_U$. We show that if ω' agrees with ω off of a finite subset of E, then $\omega' \in A_U$.

Since $\omega \in A_U$, there exists some $F \subset V$ finite such that $\omega \in A_{U,F}$, and hence there exists an infinite collection $\{\gamma_R\}$ of edge-geodesics such that each γ_R is open in ω, has starting point in F, and has endpoint in $\mathrm{scl}_{\frac{1}{R}}^{-1}(U)$.

Now suppose that ω' differs from ω only on the finite set $F' \subset E$. Then choose $r < \infty$ sufficiently large that $F \subset B(o, r-1)$ and $F' \subset E(B(o, r-1))$. Define $F'' := S(o, r)$, the set of points in V at distance r from o (note that F'' is finite). For any $R > r$, since γ_R begins inside $B(o, r-1)$ and ends outside of $B(o, r)$, γ_R has a subpath γ'_R starting at a point of F'' which only uses edges in $E(B(o, r))^c \subset F'^c$ and still ends in $\mathrm{scl}_{\frac{1}{R}}^{-1}(U)$. Since each such γ'_R is open in ω and ω' coincides with ω on F'^c, we have that each such γ'_R is open in ω' as well; thus $\omega' \in A_{U,F''} \subset A_U$.

This also shows that if $\omega \in A_U^c$ and ω' agrees with ω off of a finite subset, then $\omega' \in A_U^c$. For if $\omega' \in A_U$, then the above argument shows that $\omega \in A_U$, a contradiction. Thus, A_U is tail-measurable, as desired. The Kolmogorov 0-1 Law then implies that $\mathbb{P}_p(A_U) \in \{0, 1\}$. □

Proposition 5.5.2. *For each $p > \vec{p}_c$, the set*

$$C_p := \{\xi \in S : \forall \epsilon > 0, \mathbb{P}_p(A_{B_{d_\infty}(\xi,\epsilon)}) = 1\} \tag{5.5.1}$$

is nonempty and compact. Moreover, C_p is increasing in the sense that $p < p'$ implies $C_p \subset C_{p'}$. We call C_p the percolation cone *at level p.*

Proof. First, let us show that $S \setminus C_p$ is open in S. First note that if $U \subset V$, then $A_U \subset A_V$. Now, if $\xi \in S \setminus C_p$, then there exists some $\epsilon > 0$ such that $\mathbb{P}_p(A_{B_{d_\infty}(\xi,\epsilon)}) = 0$ (since $A_{B_{d_\infty}(\xi,\epsilon)}$ is 0-1 by the previous proposition). If $\xi' \in B_{d_\infty}(\xi, \epsilon)$, then for some $\epsilon' > 0$, $B_{d_\infty}(\xi', \epsilon') \subset B_{d_\infty}(\xi, \epsilon)$, so $\mathbb{P}_p(A_{B_{d_\infty}(\xi',\epsilon)'}) \leq \mathbb{P}_p(A_{B_{d_\infty}(\xi,\epsilon)}) = 0$, meaning $\xi' \notin C_p$. Thus C_p is a closed subset of S, and hence compact (since S is).

Next, since each $\mathbb{P}_p(A_U)$ is increasing in p, we have that $p < p'$ implies $C_p \subset C_{p'}$.

Finally, let us show that C_p is nonempty for $p > \vec{p_c}$. Since $p > \vec{p_c}$, $\mathbb{P}_p(\exists$ infinite open edge-geodesic $) = 1$. Suppose to the contrary that we had $C_p = \emptyset$. Then for each $\xi \in S$, we have some $\epsilon > 0$ such that $\mathbb{P}_p(A_{B_{d_\infty}(\xi,\epsilon)}) = 0$; by compactness of S, we have a finite subcover $\{B_{d_\infty}(\xi_1, \epsilon_1), ..., B_{d_\infty}(\xi_N, \epsilon_N)\}$ of S consisting of such balls. Note that this finite subcover covers some ϵ'-neighborhood of S by the tube lemma.

We claim that

$$\{\exists \text{ infinite open edge-geodesic }\} \subset \bigcup_{i=1}^{N} A_{B_{d_\infty}(\xi_i,\epsilon_i)}.$$

Let $\omega \in \{\exists$ infinite open edge-geodesic$\}$. Then for some $x \in V$, there is an infinite open edge-geodesic starting at x. Therefore, for all sufficiently large R, there exists y_R such that x is connected to y_R by an open edge-geodesic. For all R sufficiently large, $\mathrm{scl}_{\frac{1}{R}} y_R$ lies in the ϵ'-neighborhood of S, and hence lies in some $B_{d_\infty}(\xi_i, \epsilon_i)$. Then by the pigeonhole principle, some $B_{d_\infty}(\xi_{i_0}, \epsilon_{i_0})$ contains $\mathrm{scl}_{\frac{1}{R}} y_R$ for infinitely many R. Therefore

$$\omega \in A_{B_{d_\infty}(\xi_{i_0},\epsilon_{i_0}),\{x\}} \subset A_{B_{d_\infty}(\xi_{i_0},\epsilon_{i_0})} \subset \bigcup_{i=1}^{N} A_{B_{d_\infty}(\xi_i,\epsilon_i)},$$

as claimed. But then $1 = \mathbb{P}_p(\exists \text{ infinite open edge-geodesic }) \leq \sum_{i=1}^{N} \mathbb{P}(A_{B_{d_\infty}(\xi_i,\epsilon_i)}) = 0$, a contradiction. So C_p is nonempty. □

Now we show that all directions in the percolation cone $C_{\nu(\{a\})}$ have time constant a.

Theorem 5.5.1. *Suppose that* $\inf \operatorname{supp} \nu = a$, $\nu(\{a\}) = p \geq \vec{p}_c$, *and* ν *has a finite exponential moment. We have* $d_\Phi(\xi) = a$ *for all* $\xi \in C_p$, *where* C_p *is the percolation cone as defined in* (5.5.1).

Proof. $d_\Phi(1,\xi) \geq a$ for *all* ξ with $d_\infty(1,\xi) = 1$, so it remains to to show that if $\xi \in C_p$, then $d_\Phi(1,\xi) \leq a$. Let $\epsilon > 0$. By continuity of d_Φ with respect to d_∞, we can then choose a bounded open neighborhood U of ξ such that for all $\xi' \in U$ we have $d_\Phi(1,\xi') \geq d_\Phi(1,\xi) - \epsilon$ and $d_\infty(1,\xi') \leq 1 + \epsilon$. Since $d_\Phi \leq (\mathbb{E}w)d_\infty$, for sufficiently large R we have that

$$\mathrm{scl}_{\frac{1}{R}}^{-1} U \subset B_d(o, 2R) \subset B_{\mathbb{E}T}(o, 2\mathbb{E}wR) \subset V.$$

Then, since $\mathrm{scl}_{\frac{1}{R}}$ is a sequence of pointed Gromov-Hausdorff approximations (see section 3.10) both from $(V, \frac{1}{R}\mathbb{E}T)$ to (L_∞, d_Φ) ([5]) and from $(V, \frac{1}{R}d)$ to (L_∞), there is some R_0 such that for all $R \geq R_0$ and all $x, y \in B(o, 2R) \subset B_{\mathbb{E}T}(o, 2(\mathbb{E}w)R)$ we have

$$(1+\epsilon)\frac{1}{R}\mathbb{E}T(x,y) + \epsilon \geq d_\Phi(\mathrm{scl}_{\frac{1}{R}} x, \mathrm{scl}_{\frac{1}{R}} y)$$

and

$$(1-\epsilon)\frac{1}{R}d(x,y) - \epsilon \leq d_\infty(\mathrm{scl}_{\frac{1}{R}} x, \mathrm{scl}_{\frac{1}{R}} y).$$

We also choose R_0 sufficiently large that for all $R \geq R_0$, $\mathrm{scl}_{\frac{1}{R}}^{-1} U \subset B_d(o, 2R)$. Then in particular, taking $x = o$ and $y \in \mathrm{scl}_{\frac{1}{R}}^{-1} U$ we get

$$(1+\epsilon)\frac{1}{R}\mathbb{E}T(o,y) + \epsilon \geq d_\Phi(1, \mathrm{scl}_{\frac{1}{R}} y) \geq d_\Phi(1, \xi) - \epsilon,$$

so

$$\frac{1}{R}\mathbb{E}T(o,y) \geq \frac{d_\Phi(1,\xi) - 2\epsilon}{1+\epsilon}, \tag{5.5.2}$$

and also

$$(1-\epsilon)\frac{1}{R}d(o,y) - \epsilon \leq d_\infty(1, \mathrm{scl}_{\frac{1}{R}} y) \leq 1 + \epsilon$$

so that

$$\frac{d(o,y)}{R} \leq \frac{1+2\epsilon}{1-\epsilon}. \tag{5.5.3}$$

Now, by Talagrand concentration (Theorem 5.4.1) we have that

$$\sum_{R=1}^{\infty} \mathbb{P}(\exists x \text{ s.t. } d(o,x) = R, T(o,x) < \mathbb{E}T(o,x) - \epsilon d(o,x))$$
$$\leq \sum_{R=1}^{\infty} |B(o,R)| C e^{-cR} < \infty,$$

where the sum is finite because G has subexponential growth. So by Borel-Cantelli, the event

$$\Omega' := \{\omega : \exists R_1(\omega) \text{ s.t. } d(o,x) \geq R_1 \Rightarrow T(o,x) \geq \mathbb{E}T(o,x) - \eta d(o,x)\}$$

has probability 1. Next, since $\xi \in C_p$ and $\nu(\{a\}) \geq p$, if we call an edge e *open* whenever it has weight a, we have $\mathbb{P}(A_U) = 1$. Since $\mathbb{P}(\Omega' \cap A_U) = 1$, it is in particular nonempty.

So let $\omega \in A_U \cap \Omega'$; we then have $R_1(\omega) < \infty$ as in the definition of Ω' and a finite $F \subset V$ such that $\omega \in A_{U,F}$. By definition of $A_{U,F}$, there exist infinitely many edge-geodesics γ_R which have weight a in ω and which start at a point in F and end at a point in $y_R \in \mathrm{scl}_{\frac{1}{R}}^{-1} U$ with $d(o, y_R) \geq R$. For each such R we then have

$$\begin{aligned} T(o, y_R) &\leq \left[\max_{v \in F} T(o, v)\right] + T(\gamma_R) = \left[\max_{v \in F} T(o, v)\right] + a|\gamma_R| \\ &\leq \left[\max_{v \in F} T(o, v)\right] + a\left(d(o, y_R) + \max_{v \in F} d(o, v)\right). \end{aligned}$$

where we have used that γ_R is an edge-geodesic between some point of F and y_R. For all such R which are at least $\max(R_0, R_1)$, we further have

$$\begin{aligned} \frac{\mathbb{E}T(o, y_R)}{d(o, y_R)} - \epsilon \leq \frac{T(o, y_R)}{d(o, y_R)} &\leq a + \frac{[\max_{v \in F} T(o, v)] + a\,[\max_{v \in F} d(o, v)]}{d(o, y_R)} \\ &\leq a + \frac{[\max_{v \in F} T(o, v)] + a\,[\max_{v \in F} d(o, v)]}{R} \end{aligned}$$

Since $\max_{v \in F} T(o, v)$ and $\max_{v \in F} d(o, v)$ are finite and independent of R, for all sufficiently large such R we then have

$$\frac{\mathbb{E}T(o, y_R)}{d(o, y_R)} - \epsilon \leq a + \epsilon;$$

rearranging and using (5.5.2) and (5.5.3), we then get

$$a + 2\epsilon \geq \frac{\mathbb{E}T(o, y_R)}{R} \cdot \frac{R}{d(o, y_R)} \geq \left(\frac{d_\Phi(1, \xi) - 2\epsilon}{1 + \epsilon}\right)\left(\frac{1 - \epsilon}{1 + 2\epsilon}\right).$$

Finally, taking the limit as $\epsilon \to 0$ gives $d_\Phi(1, \xi) \leq a$, as desired. □

Finally, we can use this to prove Theorem 5.2.2, that is, that there is a direction ξ with $d_\infty(1,\xi) = 1$ such that $d_\Phi(1,\xi) = a$ for *any* ν with a finite exponential moment such that $\inf \operatorname{supp} \nu = a$ and $\nu(\{a\}) > \vec{p_c}$.

Proof of Theorem 5.2.2. Since the sets C_p for $p > \vec{p_c}$ are all nested, nonempty, and compact, it follows that $C := \bigcap_{\vec{p_c}<p\leq 1} C_p$ is nonempty and compact. But any $\xi \in C$ satisfies the desired property, by Theorem 5.5.1. As always, to get failure of monotonicity in direction ξ, simply take a nontrivial convex combination of ν and δ_a. □

References

[1] M. Aizenman and D. J. Barsky. Sharpness of the phase transition in percolation models. *Communications in Mathematical Physics*, 108(3):489 – 526, 1987.

[2] A. Auffinger, M. Damron, and J. Hanson. *50 years of first-passage percolation*, volume 68 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2017.

[3] A. Auffinger and C. Gorski. Asymptotic shapes for stationary first passage percolation on virtually nilpotent groups. *Probab. Theory Relat. Fields*, 186:285–326, 2023.

[4] E. Bates. Empirical measures, geodesic lengths, and a variational formula in first-passage percolation. *Mem. Amer. Math. Soc.* To appear.

[5] I. Benjamini and R. Tessera. First passage percolation on nilpotent cayley graphs and beyond. *Electronic Journal of Probability*, 20:1–20, 2015.

[6] M. Bestvina, K. Bromberg, and K. Fujiwara. Constructing group actions on quasi-trees and applications to mapping class groups. *Publ.math.IHES*, (122):1 – 64, 2015.

[7] R. M. Burton and M. Keane. Density and uniqueness in percolation. *Commun. Math. Phys.*, 121(3):501 – 505, 1989.

[8] M. Cantrell and A. Furman. Asymptotic shapes for ergodic families of metrics on nilpotent groups. *Groups Geom. Dyn.*, 11(4):1307–1345, 17.

[9] D. Contreras, S. Martineau, and V. Tassion. Supercritical percolation on graphs of polynomial growth, 2023.

[10] J. T. Cox and R. Durrett. Some limit theorems for percolation processes with necessary and sufficient conditions. *Ann. Probab.*, 9(4):583–603, 1981.

[11] C. Drutu and M. Kapovich. *Geometric Group Theory*.

[12] H. Duminil-Copin and V. Tassion. A New Proof of the Sharpness of the Phase Transition for Bernoulli Percolation and the Ising Model. *Commun. Math. Phys.*, 343(2):725 – 745, 2016.

[13] R. Durrett and T. M. Liggett. The Shape of the Limit Set in Richardson's Growth Model. *The Annals of Probability*, 9(2):186 – 193, 1981.

[14] C. Gorski. Strict monotonicity for first passage percolation on graphs of polynomial growth and quasi-trees, 2022.

[15] G. Grimmett. *Percolation*. Springer-Verlag, 1999.

[16] M. Gromov. Groups of polynomial growth and expanding maps. *Publications Mathématiques de L'Institut des Hautes Scientifiques*, 53:53–78, 1981.

[17] M. Gromov. Carnot-carathéodory spaces seen from within. *Progr. Math.*, 144, 1996.

[18] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1999. Based on the 1981 French original [MR0682063 (85e:53051)], With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.

[19] O. Häggström and J. Jonasson. Uniqueness and non-uniqueness in percolation theory. *Probability Surveys*, 3, 01 2007.

[20] O. Häggström and R. Meester. Asymptotic shapes for stationary first passage percolation. *Annals of Probability*, 23(4):1511–1522, 1995.

[21] O. Häggström, Y. Peres, and R. H. Schonmann. *Percolation on Transitive Graphs as a Coalescent Process: Relentless Merging Followed by Simultaneous Uniqueness*, pages 69–90. Birkhäuser Boston, Boston, MA, 1999.

[22] J. M. Hammersley and D. J. A. Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif*, pages 61–110. Springer-Verlag, New York, 1965.

[23] J. Hempel. *3-Manifolds*. American Mathematical Society, 2004.

[24] YCor (https://mathoverflow.net/users/14094/ycor). Finite index subgroup with free abelianization. MathOverflow. URL:https://mathoverflow.net/q/360855 (version: 2020-05-20).

[25] M.I. Kargopolov and Ju.I. Merzljakov. *Fundamentals of the Theory of Groups*. Springer-Verlag, 1979.

[26] H. Kesten. Aspects of first passage percolation. In *École d'été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Math.*, pages 125–264. Springer, Berlin, 1986.

[27] A. Malcev. On a class of homogeneous spaces. *Amer. Math. Soc. Transl.*, (39), 1951.

[28] J. F. Manning. Geometry of pseudocharacters. *Geometry & Topology*, 9(2):1147 – 1185, 2005.

[29] R. Marchand. Strict inequalities for the time constant in first passage percolation. *The Annals of Applied Probability*, 12(3):1001 – 1038, 2002.

[30] P. Pansu. Croissance des boules et de géodésiques fermées dans les nilvariétés. *Ergod. Th. & Dynam. Sys.*, 3:415–445, 1983.

[31] M.S. Raghunathan. *Discrete Subgroups of Lie groups*. Springer-Verlag, 1972.

[32] V. Strassen. The Existence of Probability Measures with Given Marginals. *The Annals of Mathematical Statistics*, 36(2):423 – 439, 1965.

[33] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathématiques de l'Institut des Hautes Scientifiques*, 81:73–205, 1995.

[34] R. Tessera. Speed of convergence in first passage percolation and geodesicity of the average distance. *Ann. Inst. H. Poincaré Probab. Statist.*, 54(2):569 – 586, 2018.

[35] V. I. Trofimov. Graphs of polynomial growth. *Math. USSR-Sb.*, 51:405–417, 1985.

[36] J. van den Berg and H. Kesten. Inequalities for the Time Constant in First-Passage Percolation. *The Annals of Applied Probability*, 3(1):56 – 80, 1993.

[37] W. Whitt. The effect of variability in the GI/G/s queue. *Journal of Applied Probability*, 17(4):1062–1071, 1980.