N.J.Towner

Coarse geometry and finitely generated groups

Bachelor thesis, July 2013

Supervisor: Dr. L.D.J. Taelman

Mathematisch Instituut, Universiteit Leiden

Contents

0 Conventions

1. We write $\mathbb{N}$ for the set of all integers greater than or equal to 0.
2. For sets A and B, $A \subset B$ means that every element of A is also an element of B.
3. We use the words "map" and "function" synonymously, with their set-theoretic meaning.
4. For subsets $A, B \subset \mathbb{R}$, a map $f : A \to B$ is called increasing if for all $x, y \in A$, $x > y \Rightarrow f(x) > f(y)$.
5. An edge of a graph is a set containing two distinct vertices of the graph.
6. For elements x, y of a group, the commutator $[x, y]$ is defined as $xyx^{-1}y^{-1}$.
7. For metric spaces (X, d_X), (Y, d_Y), we call a map $f : X \to Y$ an isometric embedding if for all $x, x' \in X$, $d_Y(f(x), f(x')) = d_X(x, x')$. We call f an isometry if it is a surjective isometric embedding.

1 Introduction

The topology of a metric space cannot tell us much about its large-scale, long-range structure. Indeed, if we take a metric space (X, d) and for each $x, y \in X$ we let $d'(x, y) = \min\{d(x, y), 1\}$ then d' is a metric on X equivalent to d. So instead of looking at homeomorphisms and the properties preserved by them, we consider a different type of map. There are various ways of formulating what we need, but they turn out to be equivalent for well-behaved spaces.

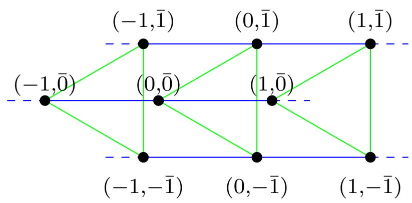
Let X and Y be metric spaces and f a map from X to Y. Our first requirement is that f should change sufficiently large distances by at most a constant factor: there should exist $a \geq 1$ and $b \geq 0$ such that for all $x, x' \in X$ the following inequality holds:

$$\frac{1}{a}d(x, x') - b \leq d(f(x), f(x')) \leq ad(x, x') + b.$$

Our second requirement is that f be nearly surjective: there must be a constant $c \geq 0$ such that for every element y of Y there is an element $x \in X$ with $d(f(x), y) \leq c$. Such a map is called a *quasi-isometry*. A simple example is the inclusion map from $\mathbb{Z}$ to $\mathbb{R}$. Another example is any map from $\mathbb{R}$ to $\mathbb{Z}$ that sends each real number to a nearest integer (note that quasi-isometries need not be continuous). We say X is *quasi-isometric* with Y if a quasi-isometry from X to Y exists, and it is not hard to see that this is an equivalence relation.

Now let G be a group and S a subset of G. For convenience, let us assume that S contains the inverses of all its elements, and that S does not contain 1. We can picture G as a graph with a vertex for each element of the group and an edge between vertices g and h if (and only if) $gs = h$ for some $s \in S$. For example, if we take $G = \mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$ and for S the set $\{(\pm 1, \bar{0}), (0, \pm\bar{1})\}$, we get the graph shown in figure 1. This is known as

Figure 1: A Cayley graph of $\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$.



the *Cayley graph* of G with respect to S. If the set S generates the whole group then the Cayley graph is connected, and we can define a metric on G by taking the distance between two elements of the group to be the length of a shortest path between them in the Cayley graph. Different generating sets can yield different metrics, of course, but if G is finitely generated (and we restrict ourselves to finite sets generating it) then all the metric spaces we obtain are quasi-isometric. For finite groups, the result is trivial (let S contain all non-trivial elements of G) but turning an infinite group into a metric space in this way helps us to understand its structure. In the mid-twentieth century, Schwarz and Milnor independently discovered the following far-reaching result: if a group G acts on a metric space X then, under certain broad conditions (see theorem 3.8 for details), the group G must be finitely generated and is quasi-isometric with X. This can be used to show, for example, that the fundamental group of any compact Riemannian manifold is finitely generated and quasi-isometric with its universal covering space.

A particularly important quasi-isometric invariant of finitely generated groups arises as follows. Let G be a group generated by a finite subset S and, for each $n \in \mathbb{N}$, let $v(n)$ be the number of elements of G that can be written as a product of at most n elements of S or their inverses. This is also the cardinality of the closed ball with centre 1 and radius n in the Cayley graph of G (with respect to S), and the function $v : \mathbb{N} \to \mathbb{N}$ is known as the *growth function* of G with respect to S. For example, the growth function of the free abelian group on 2 generators a and b is given by

$$\begin{aligned} v(n) &= \#\{ia + jb : i, j \in \mathbb{Z}, |i| + |j| \leq n\} \\ &= 2(1 + 3 + 5 + \ldots + (2n-1)) + 2n + 1 = 2n^2 + 2n + 1 \end{aligned}$$

while the growth function of the free group on 2 generators a and b is

$$\begin{aligned} v(n) &= \#\big\{s_1 s_2 \ldots s_m : m \leq n, s_1, s_2, \ldots, s_m \in \{a, b, a^{-1}, b^{-1}\} \\ &\qquad \text{where, for each } i \in \{1, 2, \ldots, m-1\}, s_{i+1} \neq s_i^{-1}\big\} \\ &= 1 + 4(3^0 + 3^1 + \ldots + 3^{n-1}) = 2 \cdot 3^n - 1. \end{aligned}$$

We say G has *polynomial growth* if its growth function is bounded above by a polynomial function (as in the first example above), or *exponential growth* if its growth function is bounded below by an exponential function (as in the second example). The precise growth function of G depends on the generators chosen, but whether that growth is polynomial, exponential or lies in between is a quasi-isometric invariant. Not only can we deduce the growth type of a group from information about its structure, but the converse is also possible, as Gromov proved in 1981:

Theorem 1.1 (Gromov[1]). *Let G be a finitely generated group. Then G has polynomial growth if and only if it has a nilpotent subgroup with finite index in G.*

[1] see [Gromov]

We do not prove Gromov's theorem in its entirety here, but cover the case of soluble groups, as established by Milnor and Wolf.

This thesis is organized as follows: chapter 2 contains the necessary background on length spaces and geodesic spaces, and the geometric realization of graphs. In chapter 3, we look at coarse equivalence and quasi-isometry and demonstrate that they coincide for length spaces. We show that the metric space obtained from a finitely generated group via the word metric is unique up to quasi-isometry and prove the Schwarz-Milnor theorem. As an application, we demonstrate that any finite index subgroup of a finitely generated group is itself finitely generated and quasi-isometric with the group. In chapter 4, we characterize the polycyclic groups, introduce the Hirsch length, and use it to prove that any subgroup of a virtually polycyclic group is equal to the intersection of all the finite index subgroups which contain it. In chapter 5, we define polynomial and exponential growth for finitely generated groups and show that they are quasi-isometric invariants. We prove that virtually nilpotent groups have polynomial growth, that finitely generated soluble groups have exponential growth or otherwise are polycyclic, and that polycyclic groups have exponential growth or otherwise are virtually nilpotent.

2 Length spaces and geodesic spaces

In this chapter, we look at what the lengths of paths in a metric space tell us about the space. The main results are the theorem of Arzelà and Ascoli for length spaces and the theorem of Hopf and Rinow. Further results can be found in [Bridson & Haefliger] and [Roe].

Definition. Let (X, d) be a metric space. For any $x, y \in X$, a *path* in X from x to y is a continuous map $p : [a, b] \to X$ (where $a, b \in \mathbb{R}$ with $a \leq b$) such that $p(a) = x$ and $p(b) = y$. We call p *rectifiable* if the set L_p defined by

$$L_p = \left\{ \sum_{i=1}^{n} d(p(t_{i-1}), p(t_i)) : n \in \mathbb{N}, a = t_0 < t_1 < \ldots < t_n = b \right\} \subset \mathbb{R}$$

is bounded above, and the *length* $l(p)$ of the path p is then defined as $\sup(L_p)$. We call (X, d) a *length space* if for all $x, y \in X$ there exists a rectifiable path in X from x to y and

$$d(x, y) = \inf\{l(p) : p \text{ a rectifiable path in } X \text{ from } x \text{ to } y\}.$$

A path $p : [a, b] \to X$ is called a *geodesic path* if it is an isometric embedding, i.e. for all $t, t' \in [a, b]$, $d(p(t), p(t')) = |t - t'|$. We call the metric space (X, d) a *geodesic space* if for all $x, y \in X$ there exists a geodesic path in X from x to y.

Examples. Any normed vector space is a geodesic space. Any geodesic space is a length space. The punctured plane $\mathbb{R}^2 \setminus \{(0, 0)\}$ is a length space but not a geodesic space.

We begin by checking that reversing the direction of a path or changing its parametrization does not alter the length, and that if we concatenate two paths then the length of the new path is the sum of the lengths of the old ones.

Lemma 2.1. *Let (X, d) be a metric space and $p : [a, b] \to X$ a rectifiable path.*

(i) If $f : [a, b] \to [a, b]$ is the map $t \mapsto a + b - t$ then $p \circ f$ is rectifiable and $l(p \circ f) = l(p)$.

(ii) If $f : [c, d] \to [a, b]$ is non-decreasing and surjective then $p \circ f$ is rectifiable and $l(p \circ f) = l(p)$.

(iii) If $q : [b, c] \to X$ is a rectifiable path with $q(b) = p(b)$ and we define

$$r : [a, c] \to X, t \mapsto \begin{cases} p(t) & \text{if } a \leq t \leq b \\ q(t) & \text{if } b \leq t \leq c \end{cases}$$

then r is rectifiable and $l(r) = l(p) + l(q)$.

Proof. (i) The map f is continuous so $p \circ f$ is a path. For any $n \in \mathbb{N}$ and any $t_0, t_1, \ldots, t_n \in [a, b]$ with $a = t_0 < t_1 < \ldots < t_n = b$, if we put $s_i = a+b-t_{n-i}$ for each $i \in \{0, 1, \ldots, n\}$ then $a = s_0 < s_1 < \ldots < s_n = b$ and $\sum_{i=1}^{n} d(p \circ f(s_{i-1}), p \circ f(s_i)) = \sum_{i=1}^{n} d(p(t_{i-1}), p(t_i))$ so $L_p \subset L_{p \circ f}$. Similarly $L_{p \circ f} \subset L_p$ so they are equal, making $L_{p \circ f}$ rectifiable with $l(p \circ f) = l(p)$.

(ii) Suppose f is not continuous at some point $x_0 \in [c, d]$. Then there exists $\varepsilon > 0$ such that for all $n \in \mathbb{N}_{\geq 1}$ there is an $x_n \in [c, d]$ satisfying $|x_n - x_0| < \frac{1}{n}$ but $|f(x_n) - f(x_0)| \geq \varepsilon$. If infinitely many terms of the sequence (x_n) are less than x_0 then $f(x_0) - \varepsilon \geq a$ and f is surjective so $f(x) = f(x_0) - \frac{\varepsilon}{2}$ for some $x \in [c, x_0]$. But there exists $n \in \mathbb{N}_{\geq 1}$ such that $x < x_n < x_0$ and f is non-decreasing so $f(x) \leq f(x_n) \leq f(x_0) - \varepsilon$, a contradiction. Similarly if infinitely many of the x_n are greater than x_0, so it follows that f is continuous.

For any $\lambda \in L_p$, $\lambda = \sum_{i=1}^{n} d(p(t_{i-1}), p(t_i))$ for some $n \in \mathbb{N}$ and some $t_0, t_1, \ldots, t_n \in [a, b]$ with $a = t_0 < t_1 < \ldots < t_n = b$. The map f is surjective so for each $i \in \{0, 1, \ldots, n\}$ there exists $s_i \in [c, d]$ such that $f(s_i) = t_i$, and f is non-decreasing so $c = s_0 < s_1 < \ldots < s_n = d$. Thus $\lambda = \sum_{i=1}^{n} d(p \circ f(s_{i-1}), p \circ f(s_i)) \in L_{p \circ f}$, giving $L_p \subset L_{p \circ f}$.

Conversely, for any $\lambda = \sum_{i=1}^{m} d(p \circ f(s_{i-1}), p \circ f(s_i)) \in L_{p \circ f}$ where $c = s_0 < s_1 < \ldots < s_m = d$, let $t_0, t_1, \ldots, t_n$ be $f(s_0), f(s_1), \ldots, f(s_m)$ with any duplicates omitted. Then we have $a = t_0 < t_1 < \ldots < t_n = b$ and $\lambda = \sum_{i=1}^{n} d(p(t_{i-1}), p(t_i)) \in L_p$, giving $L_{p \circ f} \subset L_p$ as well, and hence $L_{p \circ f} = L_p$.

(iii) The map r is well-defined as $p(b) = q(b)$ and is continuous by the Gluing Lemma. For any $z \in L_r$ there exist $x \in L_p$ and $y \in L_q$ such that $z \leq x + y$ (add b to the partition and split it) so $l(p) + l(q)$ is an upper bound for L_r. Thus r is rectifiable and $l(r) \leq l(p) + l(q)$. Conversely, for any $\varepsilon > 0$ there exist $x \in L_p$ and $y \in L_q$ with $l(p) - x < \frac{\varepsilon}{2}$ and $l(q) - y < \frac{\varepsilon}{2}$ so $l(p) + l(q) < x + y + \varepsilon \leq l(r) + \varepsilon$ (since $x + y \in L_r$) and this is for all $\varepsilon > 0$ so $l(r) \geq l(p) + l(q)$ as well, hence $l(r) = l(p) + l(q)$. □

Definition. Let (X, d) be a metric space and $p : [0, L] \to X$ a rectifiable path. We say p is *parametrized by arc length* if for all $t \in [0, L]$, $l(p|_{[0,t]}) = t$.

We can adjust the parametrization of any rectifiable path to give a path parametrized by arc length.

Proposition 2.2. *Let (X, d) be a metric space and $p : [a, b] \to X$ a rectifiable path. Let $L \in \mathbb{R}_{\geq 0}$ be the length of p and define*

$$\theta : [a, b] \to [0, L], t \mapsto l(p|_{[a,t]}).$$

Then θ is non-decreasing, uniformly continuous and surjective; there exists a unique path $\tilde{p} : [0, L] \to X$ such that $\tilde{p} \circ \theta = p$; and $\tilde{p}$ is parametrized by arc length.

Proof. We claim that for any $\varepsilon > 0$ there exists a partition (i.e. a dissection) $a = t_0 < t_1 < \ldots < t_n = b$ of $[a, b]$ such that for each $i \in \{1, 2, \ldots, n\}$ we have $l(p|_{[t_{i-1},t_i]}) < \varepsilon$. Take any $\varepsilon > 0$. The map p is uniformly continuous (as $[a, b]$ is compact) so there exists $\delta > 0$ such that for all $t, t' \in [a, b]$ with $|t - t'| < \delta$ we have $d(p(t), p(t')) < \frac{\varepsilon}{2}$. A partition $a = t_0 < t_1 < \ldots < t_n = b$ of $[a, b]$ exists such that $l(p) - \sum_{i=1}^{n} d(p(t_{i-1}), p(t_i)) < \frac{\varepsilon}{2}$ and (by refining it if necessary) such that for each $i \in \{1, 2, \ldots, n\}$, $t_i - t_{i-1} < \delta$. By (2.1iii), $l(p) = \sum_{i=1}^{n} l(p|_{[t_{i-1},t_i]})$, so

$$\sum_{i=1}^{n} \left(l(p|_{[t_{i-1},t_i]}) - d(p(t_{i-1}), p(t_i)) \right) < \frac{\varepsilon}{2}.$$

And for all $i \in \{1, 2, \ldots, n\}$, $l(p|_{[t_{i-1},t_i]}) \geq d(p(t_{i-1}), p(t_i))$ so for each i individually, $l(p|_{[t_{i-1},t_i]}) - d(p(t_{i-1}), p(t_i)) < \frac{\varepsilon}{2}$. Also $t_i - t_{i-1} < \delta$ and so $d(p(t_{i-1}), p(t_i)) < \frac{\varepsilon}{2}$, hence $l(p|_{[t_{i-1},t_i]}) < \varepsilon$, proving the claim.

The map p is rectifiable so θ is well-defined, and for any $t, t' \in [a, b]$ with $t < t'$ we have $\theta(t') = \theta(t) + l(p|_{[t,t']})$ by (2.1iii) so $\theta(t') \geq \theta(t)$, making θ non-decreasing. We now show that θ is uniformly continuous: take any $\varepsilon > 0$. By the claim, there exists a partition $a = t_0 < t_1 < \ldots < t_n = b$ of $[a, b]$ such that for each $i \in \{1, 2, \ldots, n\}$, $l(p|_{[t_{i-1},t_i]}) < \frac{\varepsilon}{2}$. Let

$$\delta = \min\{t_i - t_{i-1} : 1 \leq i \leq n\} > 0$$

and take any $s, t \in [a, b]$ with $|t - s| < \delta$. Without loss of generality, we may assume $t \geq s$. Then either $s, t \in [t_{i-1}, t_i]$ for some $i \in \{1, 2, \ldots, n\}$, in which case $l(p|_{[s,t]}) \leq l(p|_{[t_{i-1},t_i]}) < \frac{\varepsilon}{2} < \varepsilon$, or there exists $i \in \{1, 2, \ldots, n - 1\}$ such that $t_{i-1} < s < t_i < t < t_{i+1}$, in which case

$$l(p|_{[s,t]}) = l(p|_{[s,t_i]}) + l(p|_{[t_i,t]}) \leq l(p|_{[t_{i-1},t_i]}) + l(p|_{[t_i,t_{i+1}]}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus in both cases $|\theta(t) - \theta(s)| = l(p|_{[a,t]}) - l(p|_{[a,s]}) = l(p|_{[s,t]}) < \varepsilon$ and θ is therefore uniformly continuous.

As $\theta(a) = 0$ and $\theta(b) = L$, it follows that θ is also surjective (Intermediate Value Theorem). Moreover, for any $t, t' \in [a, b]$ with $t \leq t'$, if $\theta(t) = \theta(t')$ then $l(p|_{[t,t']}) = 0$ so $p|_{[t,t']}$ is a constant map and $p(t) = p(t')$. Thus for each $s \in [0, L]$, we can define $\tilde{p}(s) = p(t)$ where $\theta(t) = s$. This makes $\tilde{p}$ well-defined, and it is the unique map from $[0, L]$ to X satisfying $\tilde{p} \circ \theta = p$.

For any $s, s' \in [0, L]$ with $s \geq s'$, there exist $t, t' \in [a, b]$ with $t \geq t'$ such that $\theta(t) = s$ and $\theta(t') = s'$, and

$$\begin{aligned} d(\tilde{p}(s), \tilde{p}(s')) &= d(p(t), p(t')) \\ &\leq l(p|_{[t',t]}) = l(p|_{[a,t]}) - l(p|_{[a,t']}) = \theta(t) - \theta(t') = s - s'. \end{aligned}$$

It follows that $\tilde{p}$ is continuous (take $\delta = \varepsilon$) and hence a path in X. It remains to show that $\tilde{p}$ is rectifiable and parametrized by arc length. For any partition $0 = s_0 < s_1 < \ldots < s_n = L$ of $[0, L]$, there exists a partition $a = t_0 < t_1 < \ldots < t_n = b$ of $[a, b]$ such that for each $i \in \{0, 1, \ldots, n\}$ we have $\theta(t_i) = s_i$, so $\sum_{i=1}^n d(\tilde{p}(s_{i-1}), \tilde{p}(s_i)) = \sum_{i=1}^n d(p(t_{i-1}), p(t_i))$ which is in L_p, thus $L_{\tilde{p}}$ is bounded above and $\tilde{p}$ is rectifiable. Finally, for any $s \in [0, L]$ there exists $t \in [a, b]$ with $\theta(t) = s$, and $\tilde{p}|_{[0,s]} \circ \theta|_{[a,t]} = p|_{[a,t]}$ since θ is non-decreasing. As $\theta|_{[a,t]}$ with codomain $[0, s]$ is surjective, it follows from (2.1ii) that

$$l(\tilde{p}|_{[0,s]}) = l(\tilde{p}|_{[0,s]} \circ \theta|_{[a,t]}) = l(p|_{[a,t]}) = \theta(t) = s.$$

□

Lemma 2.3. *Let (X, d) be a length space, $x, z \in X$ and $\lambda, \mu \in \mathbb{R}_{>0}$ such that $d(x, z) < \lambda + \mu$. Then there exists $y \in X$ such that $d(x, y) < \lambda$ and $d(y, z) < \mu$.*

Proof. If $d(x, z) < \lambda$ then $y = z$ suffices, and if $d(x, z) < \mu$ then $y = x$ suffices. Suppose $d(x, z) \geq \lambda$ and $d(x, z) \geq \mu$. The metric space X is a length space, so there exists a rectifiable path $p : [a, b] \to X$ from x to z such that $l(p) < \lambda + \mu$. Let $\varepsilon = \lambda + \mu - l(p) > 0$. Note that $\lambda - \varepsilon \geq 0$ (since $l(p) \geq d(x, z) \geq \mu$) and $\mu - \varepsilon \geq 0$ (similarly). By (2.2) there exists $t \in [a, b]$ such that $l(p|_{[a,t]}) = \lambda - \frac{\varepsilon}{2}$ and, by (2.1iii), $l(p|_{[t,b]}) = \mu - \frac{\varepsilon}{2}$ so putting $y = p(t)$ gives $d(x, y) < \lambda$ and $d(y, z) < \mu$. □

Lemma 2.4. *Let (X, d_X), (Y, d_Y) be metric spaces, $A \subset X$ dense in X and $f : A \to Y$ a map. If Y is complete and f is uniformly continuous, then there exists a unique continuous map $\tilde{f} : X \to Y$ extending f, and $\tilde{f}$ is also uniformly continuous.*

Proof. Suppose Y is complete and f is uniformly continuous, and take any $x \in X$. Then $x \in \overline{A}$ (the topological closure) so for every $n \in \mathbb{N}_{\geq 1}$ there exists $a_n \in A$ such that $d_X(a_n, x) < \frac{1}{n}$. We show that $(f(a_n))$ is a Cauchy sequence: take any $\varepsilon > 0$. The map f is uniformly continuous, so there exists $\delta > 0$ such that for all $a, a' \in A$, $d_X(a, a') < \delta \Rightarrow d_Y(f(a), f(a')) < \varepsilon$. Choose $N \in \mathbb{N}_{\geq 1}$ with $N \geq \frac{2}{\delta}$. Then for all $m, n \in \mathbb{N}_{\geq 1}$ with $m, n \geq N$,

$$d_X(a_m, a_n) \leq d_X(a_m, x) + d_X(x, a_n) < \frac{1}{m} + \frac{1}{n} \leq \frac{2}{N} \leq \delta$$

so $d_Y(f(a_m), f(a_n)) < \varepsilon$. Thus $(f(a_n))$ is a Cauchy sequence, and Y is complete so $(f(a_n))$ converges to a limit $l_x \in Y$.

Now take any sequence (a'_n) in A such that $a'_n \to x$ as $n \to \infty$, and any $\varepsilon > 0$. The map f is uniformly continuous, so there exists $\delta > 0$ such that for all $a, a' \in A$ with $d_X(a, a') < \delta$ we have $d_Y(f(a), f(a')) < \frac{\varepsilon}{2}$. Also $a_n \to x$, $f(a_n) \to l_x$ and $a'_n \to x$ as $n \to \infty$ so there exists $N \in \mathbb{N}_{\geq 1}$ such that for all $n \geq N$, all of the following hold: $d_X(a_n, x) < \frac{\delta}{2}$, $d_Y(f(a_n), l_x) < \frac{\varepsilon}{2}$ and

$d_X(a'_n, x) < \frac{\delta}{2}$. Thus for all $n \geq N$, $d_X(a'_n, a_n) \leq d_X(a'_n, x) + d_X(x, a_n) < \delta$ and so

$$d_Y(f(a'_n), l_x) \leq d_Y(f(a'_n), f(a_n)) + d(f(a_n), l_x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

so $f(a'_n) \to l_x$ as $n \to \infty$ as well. Conclusion: for each $x \in X$, $l_x \in Y$ is the limit of $(f(a_n))$ for all sequences (a_n) in A such that $a_n \to x$ as $n \to \infty$. The only candidate for a continuous map on X extending f is therefore the map $\tilde{f} : X \to Y, x \mapsto l_x$. It is clear that $\tilde{f}|_A = f$, so it remains to show that $\tilde{f}$ is uniformly continuous. Take any $\varepsilon > 0$. The map f is uniformly continuous, so there exists $\delta_f > 0$ such that

$$\forall a, a' \in A, d_X(a, a') < \delta_f \Rightarrow d_Y(f(a), f(a')) < \frac{\varepsilon}{3}. \tag{1}$$

Let $\delta = \frac{\delta_f}{3} > 0$ and take any $x, x' \in X$ with $d_X(x, x') < \delta$. As above, there exist sequences (a_n) and (a'_n) in A such that $a_n \to x$ and $a'_n \to x'$ as $n \to \infty$, and by definition $f(a_n) \to \tilde{f}(x)$ and $f(a'_n) \to \tilde{f}(x')$ as $n \to \infty$. So there exists $N \in \mathbb{N}_{\geq 1}$ such that for all $n \geq N$, all of the following hold:

$$\begin{array}{rcl} d_X(a_n, x) < \dfrac{\delta_f}{3}, & & d_X(a'_n, x') < \dfrac{\delta_f}{3}, \\ d_Y(f(a_n), \tilde{f}(x)) < \dfrac{\varepsilon}{3} & \text{and} & d_Y(f(a'_n), \tilde{f}(x')) < \dfrac{\varepsilon}{3}. \end{array}$$

Thus

$$\begin{aligned} d_X(a_N, a'_N) &\leq d_X(a_N, x) + d_X(x, x') + d_X(x', a'_N) \\ &< \frac{\delta_f}{3} + \frac{\delta_f}{3} + \frac{\delta_f}{3} = \delta_f \end{aligned}$$

so $d_Y(f(a_N), f(a'_N)) < \frac{\varepsilon}{3}$ by (1) and hence

$$\begin{aligned} d_Y(\tilde{f}(x), \tilde{f}(x')) &\leq d_Y(\tilde{f}(x), f(a_N)) + d_Y(f(a_N), f(a'_N)) \\ &\quad + d_Y(f(a'_N), \tilde{f}(x')) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

□

Definition. A metric space (X, d) is called *separable* if it contains a countable subset A such that A is dense in X.

For example, $\mathbb{R}^n$ is separable for any $n \in \mathbb{N}$.

Definition. Let (X, d_X) and (Y, d_Y) be metric spaces and F a set of maps from X to Y. We call F *uniformly equicontinuous* if

$$\forall \varepsilon > 0, \exists \delta > 0, \forall f \in F, \forall x, x' \in X, d_X(x, x') < \delta \Rightarrow d_Y(f(x), f(x')) < \varepsilon.$$

Note that some authors call this equicontinuous.

Theorem 2.5 (Arzelà, Ascoli[2]). *Let X be a separable metric space, Y a compact metric space and $(f_n)_{n\in\mathbb{N}}$ a sequence of maps from X to Y such that $\{f_n : n \in \mathbb{N}\}$ is uniformly equicontinuous. Then there exists a subsequence $(f_{h(n)})$ of (f_n) and a uniformly continuous map $\tilde{f} : X \to Y$ such that for any compact subset $C \subset X$, $f_{h(n)}|_C$ converges uniformly to $\tilde{f}|_C$.*

Proof. The metric space X is separable, so it has a countable subset

$$A = \{a_0, a_1, a_2, \ldots\}$$

which is dense in X. The metric space Y is compact, so the sequence $(f_n(a_0))$ in Y has a convergent subsequence $(f_{g_0(n)}(a_0))$ where $g_0 : \mathbb{N} \to \mathbb{N}$ is increasing. Similarly, the sequence $(f_{g_0(n)}(a_1))$ in Y has a convergent subsequence $(f_{g_1(n)}(a_1))$ where $g_1 : \mathbb{N} \to \mathrm{im}(g_0)$ is increasing. Continuing in this way, we obtain a sequence of increasing maps (g_i) with domain $\mathbb{N}$ and $\mathbb{N} \supset \mathrm{im}(g_0) \supset \mathrm{im}(g_1) \supset \ldots$ with the property that for each $k \in \mathbb{N}$ and each $j \in \{0, 1, \ldots, k\}$, the sequence $f_{g_k(n)}(a_j)$ converges as $n \to \infty$. Define $h : \mathbb{N} \to \mathbb{N}$ by $h(n) = g_n(n)$, also an increasing map. Then for each $k \in \mathbb{N}$, $h(n)$ is in $\mathrm{im}(g_k)$ for all $n \geq k$ so $f_{h(n)}(a_k)$ converges as $n \to \infty$. Define

$$f : A \to Y, a \mapsto \lim_{n\to\infty} f_{h(n)}(a).$$

We show that f is uniformly continuous: take any $\varepsilon > 0$. By assumption, $\{f_n : n \in \mathbb{N}\}$ is uniformly equicontinuous, so there exists $\delta > 0$ such that for all $n \in \mathbb{N}$ and all $x, x' \in X$ with $d(x, x') < \delta$, $d(f_n(x), f_n(x')) < \varepsilon$. Taking the limit as $n \to \infty$, it follows that for all $a, a' \in A$ with $d(a, a') < \delta$, $d(f(a), f(a')) \leq \varepsilon$, so f is uniformly continuous. Moreover Y is complete (since it is a compact metric space), so by (2.4) there exists a uniformly continuous map $\tilde{f} : X \to Y$ extending f. Take any compact $C \subset X$. We shall show that $f_{h(n)}$ converges uniformly to $\tilde{f}$ on the domain C. Take any $\varepsilon > 0$. The maps f_n are uniformly equicontinuous so there exists $\delta > 0$ such that for all $x, x' \in X$ and all $n \in \mathbb{N}$,

$$d(x, x') < \delta \Rightarrow d(f_n(x), f_n(x')) < \frac{\varepsilon}{4}. \tag{2}$$

The subset C is compact so it is totally bounded and can be covered by a finite numer $m \in \mathbb{N}$ of (non-empty) open balls $B_1, B_2, \ldots, B_m \subset X$ of radius $\frac{\delta}{2}$. And A is dense in X so for each $i \in \{1, 2, \ldots, m\}$ there exists $z_i \in \mathbb{N}$ such that a_{z_i} is in B_i. For each $a \in A$, $f_{h(n)}(a) \to f(a)$ as $n \to \infty$ so there exists $N \in \mathbb{N}$ such that for each i in the finite set $\{1, 2, \ldots, m\}$ and for all $n \geq N$,

$$d(f_{h(n)}(a_{z_i}), f(a_{z_i})) < \frac{\varepsilon}{4}. \tag{3}$$

Take any $x \in C$ and any $n \geq N$. Then x is in B_i for some $i \in \{1, 2, \ldots, m\}$ and $d(x, a_{z_i}) < \delta$. The subset A is dense in X so there exists a sequence (b_j)

[2] see [Rudin]

in A such that $b_j \to x$ as $j \to \infty$, and $\tilde{f}$ is continuous so $\tilde{f}(b_j) \to \tilde{f}(x)$ as $j \to \infty$ too. And as $d(a_{z_i}, x) < \delta$ it follows that there exists $J \in \mathbb{N}$ such that for all $j \geq J$ we have both $d(a_{z_i}, b_j) < \delta$ and $d(\tilde{f}(b_j), \tilde{f}(x)) < \frac{\varepsilon}{4}$. Moreover $\tilde{f}(b_J) = f(b_J)$ (since $b_J \in A$) so

$$d(f(b_J), \tilde{f}(x)) < \frac{\varepsilon}{4}. \tag{4}$$

Now $d(a_{z_i}, b_J) < \delta$ so by (2) we have $d(f_{h(n)}(a_{z_i}), f_{h(n)}(b_J)) < \frac{\varepsilon}{4}$ and, taking the limit as $n \to \infty$,

$$d(f(a_{z_i}), f(b_J)) \leq \frac{\varepsilon}{4}. \tag{5}$$

Similarly, $d(x, a_{z_i}) < \delta$ so again by (2) we have

$$d(f_{h(n)}(x), f_{h(n)}(a_{z_i})) < \frac{\varepsilon}{4}. \tag{6}$$

Using (6), (3), (5) and then (4), it follows that

$$\begin{aligned} d(f_{h(n)}(x), \tilde{f}(x)) &\leq d(f_{h(n)}(x), f_{h(n)}(a_{z_i})) + d(f_{h(n)}(a_{z_i}), f(a_{z_i})) \\ &\quad + d(f(a_{z_i}), f(b_J)) + d(f(b_J), \tilde{f}(x)) \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned}$$

□

Example. Let X be a non-empty compact (and hence separable) metric space, and $\mathcal{C}(X)$ the set of all continuous functions from X to $\mathbb{R}$, which is a normed vector space under the uniform norm

$$\|f\|_\infty = \sup\{|f(x)| : x \in X\}$$

and complete as a metric space under the corresponding metric. Take any $F \subset \mathcal{C}(X)$ closed, bounded and uniformly equicontinuous. Using the Arzelà-Ascoli theorem, we can show that F is compact: take any sequence (f_n) in F. As F is bounded, there exists $M \geq 0$ such that for all $n \in \mathbb{N}$, $\|f_n\|_\infty \leq M$. Let $Y = \{x \in \mathbb{R} : |x| \leq M\}$. Then Y is closed (in $\mathbb{R}$) and bounded so Y is compact (Heine-Borel), and we can restrict the co-domain of f_n to Y for each $n \in \mathbb{N}$. By the Arzelà-Ascoli theorem (2.5), there is a subsequence of (f_n) which converges under the uniform norm to a continuous function $f \in \mathcal{C}(X)$, and F is closed in $\mathcal{C}(X)$ so f is in F, making F compact. In this way, the concept of uniform equicontinuity allows us to extend the Heine-Borel theorem for $\mathbb{R}^n$ to $\mathcal{C}(X)$, which is useful in functional analysis.

Definition. A metric space (X, d) is called *proper* if every closed and bounded (i.e. finite diameter) subset of X is compact.

A map $f : Y \to Z$ between topological spaces Y and Z is called *proper* if for every compact $C \subset Z$, $f^{-1}(C)$ is also compact.

Example. This example shows the link between the two definitions given above. Let (X, d) be a metric space, $x_0 \in X$ and define $f : X \to \mathbb{R}$ by $x \mapsto d(x_0, x)$. It follows that X is proper as a metric space if and only if the map f is proper: f is continuous so for any closed $B \subset \mathbb{R}$, $f^{-1}(B)$ is closed in X, and if $B \subset \mathbb{R}$ is bounded then $f^{-1}(B)$ is bounded (by definition of f), so if X is proper then f is proper. Conversely, take any closed and bounded $A \subset X$, let $B = f(A)$ and let $\overline{B}$ be the topological closure of B in $\mathbb{R}$. The set $\overline{B}$ is closed and bounded so if f is proper then $f^{-1}(B)$ is compact, and $A \subset f^{-1}(B)$ is closed so A is also compact, making X proper.

Example. Let X be an infinite set with the discrete metric, i.e. the distance between any pair of distinct points is 1. Then X itself is closed and bounded but not compact, so X is not proper.

Lemma 2.6. *Let (X, d) be a metric space and $p, p_0, p_1, p_2, \ldots : [a, b] \to X$ rectifiable paths such that p_n converges uniformly to p as $n \to \infty$. Then for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $l(p_n) > l(p) - \varepsilon$.*

Proof. Take any $\varepsilon > 0$. A partition $a = t_0 < t_1 < \ldots < t_m = b$ of $[a, b]$ exists such that $l(p) - \sum_{i=1}^{m} d(p(t_{i-1}), p(t_i)) < \frac{\varepsilon}{2}$, and p_n converges uniformly to p so there exists $N \in \mathbb{N}$ such that for all $n \geq N$ and all $t \in [a, b]$ we have $d(p_n(t), p(t)) < \frac{\varepsilon}{4m+1}$. Take any $n \in \mathbb{N}$ with $n \geq N$. For each $i \in \{1, 2, \ldots, m\}$,

$$\begin{aligned} d(p(t_{i-1}), p(t_i)) &\leq d(p(t_{i-1}), p_n(t_{i-1})) + d(p_n(t_{i-1}), p_n(t_i)) \\ &\quad + d(p_n(t_i), p(t_i)) \\ &< d(p_n(t_{i-1}), p_n(t_i)) + \frac{2\varepsilon}{4m+1}. \end{aligned}$$

Hence

$$\begin{aligned} l(p) &< \sum_{i=1}^{m} d(p(t_{i-1}), p(t_i)) + \frac{\varepsilon}{2} \leq \sum_{i=1}^{m} d(p_n(t_{i-1}), p_n(t_i)) + m\frac{2\varepsilon}{4m+1} + \frac{\varepsilon}{2} \\ &< \sum_{i=1}^{m} d(p_n(t_{i-1}), p_n(t_i)) + \varepsilon \leq l(p_n) + \varepsilon. \end{aligned}$$

□

Theorem 2.7 (Hopf, Rinow[3]). *Let (X, d) be a length space.*

(i) X is proper if and only if X is complete and locally compact.

(ii) If X is proper then it is a geodesic space.

[3] see [Bridson & Haefliger]

Proof. (i) If X is proper then it follows immediately that X is locally compact and complete. Suppose instead that X is locally compact and complete. If $X = \varnothing$ then X is trivially proper. Otherwise there exists $a \in X$. For each $r \in \mathbb{R}_{\geq 0}$ let $B_r[a] = \{x \in X : d(x, a) \leq r\}$ and define

$$I = \{r \in \mathbb{R}_{\geq 0} : B_r[a] \text{ is compact}\}.$$

Then $0 \in I$ and, for any $s, t \in \mathbb{R}_{\geq 0}$, if $t \in I$ and $s < t$ then $s \in I$. We shall show that I is both open and closed in $\mathbb{R}_{\geq 0}$. Take any $r \in I$. The metric space X is locally compact, so for each $x \in B_r[a]$ there exists $r_x \in \mathbb{R}_{>0}$ such that $B_{r_x}[x]$ is compact. In the case $r = 0$, it follows that $r_a \in I$ with $r_a > 0 = r$ so $\{s \in \mathbb{R}_{\geq 0} : s < r_a\}$ (an open ball in $\mathbb{R}_{\geq 0}$) is a subset of I. Suppose now that $r > 0$. The open balls $\{B_{r_x}(x) : x \in B_r[a]\}$ corresponding with the compact balls $\{B_{r_x}[x] : x \in B_r[a]\}$ form an open cover of $B_r[a]$, and $B_r[a]$ is compact so there exists a finite subcover:

$$B_r[a] \subset B_{r_1}(x_1) \cup B_{r_2}(x_2) \cup \ldots \cup B_{r_n}(x_n)$$

(with $n \in \mathbb{N}_{\geq 1}$). Let $F = X \setminus \bigcup_{i=1}^n B_{r_i}(x_i)$. Then $B_r[a]$ is compact, F is closed in X and $B_r[a] \cap F = \varnothing$ so there exists $\varepsilon > 0$ such that for all $x \in B_r[a]$ and all $y \in F$, $d(x, y) \geq \varepsilon$. Take any $y \in B_{r+\frac{\varepsilon}{2}}[a]$. Then $d(a, y) < r + \varepsilon$ so there exists $x \in X$ such that $d(a, x) < r$ and $d(x, y) < \varepsilon$ by (2.3). As $x \in B_r[a]$ and $d(x, y) < \varepsilon$, it follows that $y \notin F$, i.e. $y \in \bigcup_{i=1}^n B_{r_i}(x_i)$. Thus $B_{r+\frac{\varepsilon}{2}}[a]$ is a closed subset of $\bigcup_{i=1}^n B_{r_i}[x_i]$, which is compact (finite union of compact subsets) and so $B_{r+\frac{\varepsilon}{2}}[a]$ is compact and $r + \frac{\varepsilon}{2} \in I$. Hence $\{s \in \mathbb{R}_{\geq 0} : |s - r| < \frac{\varepsilon}{2}\}$ (an open ball in $\mathbb{R}_{\geq 0}$) is a subset of I. It follows that I is open in $\mathbb{R}_{\geq 0}$.

Suppose I is not closed in $\mathbb{R}_{\geq 0}$. Then $I = [0, s)$ for some $s \in \mathbb{R}_{>0}$. Take any sequence (x_n) in $B_s[a]$ such that $d(x_n, a) \to s$ as $n \to \infty$. For each $n \in \mathbb{N}_{\geq 1}$ and each $k \in \mathbb{N}_{\geq 1}$, $d(x_n, a) < s + \frac{1}{2k}$ so by (2.3) there exists $y_n(k) \in X$ such that $d(a, y_n(k)) < s - \frac{1}{2k}$ and $d(y_n(k), x_n) < \frac{1}{k}$. The sequence $(y_n(1))$ lies in $B_{s-\frac{1}{2}}[a]$ which is compact (since $s - \frac{1}{2}$ is in I) so it has a convergent subsequence $(y_{g_1(n)}(1))$ where $g_1 : \mathbb{N} \to \mathbb{N}$ is increasing. And the sequence $(y_{g_1(n)}(2))$ lies in $B_{s-\frac{1}{4}}[a]$ which is compact so it has a convergent subsequence $(y_{g_2(n)}(2))$ where the map $g_2 : \mathbb{N} \to \mathrm{im}(g_1)$ is increasing. Continuing in this way, we obtain a sequence of increasing maps (g_i) with domain $\mathbb{N}_{\geq 1}$ and $\mathbb{N}_{\geq 1} \supset \mathrm{im}(g_1) \supset \mathrm{im}(g_2) \supset \ldots$ such that for each $m \in \mathbb{N}_{\geq 1}$ and each $k \in \{1, 2, \ldots, m\}$, $(y_{g_m(n)}(k))$ converges as $n \to \infty$. Define $h : \mathbb{N}_{\geq 1} \to \mathbb{N}_{\geq 1}$ by $h(n) = g_n(n)$, also an increasing map. Then for each $k \in \mathbb{N}_{\geq 1}$, $h(n)$ is in $\mathrm{im}(g_k)$ for all $n \geq k$ so the sequence $(y_{h(n)}(k))$ converges as $n \to \infty$.

We claim the sequence $(x_{h(n)})$ is Cauchy: take any $\varepsilon > 0$. Then there exists $k \in \mathbb{N}_{\geq 1}$ with $\frac{1}{k} < \frac{\varepsilon}{3}$. The sequence $(y_{h(n)}(k))$ converges as $n \to \infty$ so it is Cauchy and there exists $N \in \mathbb{N}_{\geq 1}$ such that for all $m, n \geq N$,

$d(y_{h(m)}(k), y_{h(n)}(k)) < \frac{\varepsilon}{3}$. Also (by definition of y) $d(y_{h(n)}(k), x_{h(n)}) < \frac{1}{k}$ for all $n \in \mathbb{N}_{\geq 1}$, so for all $m, n \geq N$,

$$\begin{aligned} d(x_{h(m)}, x_{h(n)}) &\leq d(x_{h(m)}, y_{h(m)}(k)) + d(y_{h(m)}(k), y_{h(n)}(k)) \\ &\quad + d(y_{h(n)}(k), x_{h(n)}) \\ &< \frac{1}{k} + \frac{\varepsilon}{3} + \frac{1}{k} < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

proving the claim. As X is complete, $(x_{h(n)})$ converges, with limit in $B_s[a]$ as it is closed.

We have shown that for any sequence (x_n) in $B_s[a]$, if $d(x_n, a) \to s$ as $n \to \infty$ then (x_n) has a convergent subsequence with limit in $B_s[a]$; and otherwise it has a subsequence in $B_{s-\varepsilon}[a]$ for some $\varepsilon > 0$ which has a convergent subsequence since $s - \varepsilon \in I$. So $B_s[a]$ is compact and $s \in I$, which is a contradiction. Hence I is both open and closed in $\mathbb{R}_{\geq 0}$ (and non-empty) so $I = \mathbb{R}_{\geq 0}$ and it follows that X is proper.

(ii) Suppose X is proper, and take any $x, y \in X$. If $x = y$ then clearly there is a geodesic path in X between them. Suppose $x \neq y$ and let

$$L = d(x, y) > 0.$$

As X is a length space, for each $n \in \mathbb{N}_{\geq 1}$ there is a rectifiable path p_n in X from x to y with $l(p_n) < L + \frac{1}{n}$. By (2.2) and (2.1ii) we may assume that each p_n has domain $[0, 1]$ and is parametrized proportionally to arc length, i.e. for all $t, t' \in [0, 1]$ with $t' \geq t$, $\frac{l(p_n|_{[t,t']})}{l(p_n)} = t' - t$. It follows that $\{p_n : n \in \mathbb{N}_{\geq 1}\}$ is uniformly equicontinuous: for any $\varepsilon > 0$, let $\delta = \frac{\varepsilon}{L+1}$. Then for any $n \in \mathbb{N}_{\geq 1}$ and any $t, t' \in [0, 1]$ with $t' \geq t$, if $|t - t'| < \delta$ then

$$d(p_n(t), p_n(t')) \leq l(p_n|_{[t,t']}) \leq l(p_n|_{[t,t']}) \frac{L+1}{l(p_n)} = |t - t'| \cdot (L + 1) < \varepsilon.$$

And for all $n \in \mathbb{N}_{\geq 1}$, $\mathrm{im}(p_n) \subset B_{L+1}[x]$ which is compact as X is proper, so by the Arzelà-Ascoli theorem (2.5), there exists a subsequence $(p_{f(n)})$ of (p_n) which converges uniformly to a path $p : [0, 1] \to X$. As $p_n(0) = x$ for all $n \in \mathbb{N}_{\geq 1}$, we have $p(0) = x$ and similarly $p(1) = y$. The set L_p is bounded above by $L + 1$, so p is also rectifiable. By (2.6), for all $\varepsilon > 0$ there exists $N \in \mathbb{N}_{\geq 1}$ such that for all $n \geq N$, $l(p_{f(n)}) > l(p) - \varepsilon$, and for all $n \in \mathbb{N}_{\geq 1}$, $L \leq l(p_n) \leq L + \frac{1}{n}$ so $\liminf_{n\to\infty} l(p_{f(n)}) \geq l(p)$. But $l(p_n) \to L$ as $n \to \infty$ so $\liminf_{n\to\infty} l(p_{f(n)}) = L$ and thus $L \geq l(p)$. As $L = d(x, y)$, we also have $l(p) \geq L$, hence $l(p) = L$.

For any $t, t' \in [0, 1]$ with $t \leq t'$,

$$d(x, y) \leq d(x, p(t)) + d(p(t), p(t')) + d(p(t'), y) \leq l(p) = d(x, y)$$

so $d(x, y) = d(x, p(t)) + d(p(t), p(t')) + d(p(t'), y)$.
And $l(p|_{[0,t]}) \geq d(x, p(t))$, $l(p|_{[t,t']}) \geq d(p(t), p(t'))$, $l(p|_{[t',1]}) \geq d(p(t'), y)$.

But $l(p|_{[0,t]})+l(p|_{[t,t']})+l(p|_{[t',1]}) = l(p) = d(x,y)$ so $d(p(t),p(t')) = l(p|_{[t,t']})$. Hence, reparametrizing p by arc length if necessary (2.2), p is a geodesic path in X from x to y, and X is a geodesic space. □

We end this chapter by applying the above ideas to graphs. The vertex set of any connected graph can be made into a metric space by defining the distance between two vertices as the length of a shortest path between them. To make such a path (in the graph sense) correspond with a path in the metric space, however, we must extend the space to include points representing the edges.

Definition. Let $\Gamma = (V,E)$ be a connected graph and for each $v, v' \in V$ let $d_E(v,v') \in \mathbb{Z}_{\geq 0}$ be the length of a shortest path in Γ from v to v'. Let I be the open interval $(0,1) \subset \mathbb{R}$. A *geometric realization* of Γ is a complete geodesic space (X,d) together with an isometric embedding $\phi : (V,d_E) \to (X,d)$ and, for each $e = \{e_1,e_2\} \in E$, a geodesic path $p_e : [0,1] \to X$ between $\phi(e_1)$ and $\phi(e_2)$ such that $\{p_e(I) : e \in E\}$ forms a partition of $X \setminus \mathrm{im}(\phi)$ consisting of open subsets.

We first prove the existence of geometric realizations by constructing a model and then, separately, we prove their uniqueness purely in terms of the definition.

Proposition 2.8. *Let Γ be a connected graph. Then there exists a geometric realization of Γ, and if Γ is locally finite then it has a geometric realization which is a proper geodesic space.*

Proof. Let V be the vertex set of Γ and E its edge set. We first construct a metric space. Choose a map $\delta : E \times \{0,1\} \to V$ such that for each $e = \{v,w\} \in E$ we have $\{\delta(e,0),\delta(e,1)\} = \{v,w\}$ (imposing an arbitrarily chosen direction on each edge). Let $Y = E \times [0,1]$ and define a relation $\sim$ on Y by

$$(p,s) \sim (q,t) \Leftrightarrow (p,s) = (q,t) \text{ or } (s,t \in \{0,1\} \text{ and } \delta(p,s) = \delta(q,t))$$

(identifying endpoints of edges at the same vertex). Then $\sim$ is clearly an equivalence relation. Take any $x,y \in Y$, $n \in \mathbb{N}$, $p = (p_1,p_2,\ldots,p_n) \in E^n$ and any $s = (s_1,s_2,\ldots,s_n)$, $t = (t_1,t_2,\ldots,t_n) \in [0,1]^n$. For the purposes of this proof, we call (p,s,t,n) a *route* in Y from x to y if either $n = 0$ and $x \sim y$, or all of the following hold: $n > 0$, for all $i \in \{1,2,\ldots,n-1\}$ we have $t_i, s_{i+1} \in \{0,1\}$, for all $i \in \{1,2,\ldots,n\}$ the inequality $|t_i - s_i| > 0$ holds, $x \sim (p_1,s_1)$, for each $i \in \{1,2,\ldots,n-1\}$ we have $(p_i,t_i) \sim (p_{i+1},s_{i+1})$, and $(p_n,t_n) \sim y$. We define the *length* of a route (p,s,t,n) by

$$l(p,s,t,n) = \sum_{i=1}^{n} |t_i - s_i|$$

Figure 2: Visualisation of a route

$$x \sim (p_1, s_1) \overset{|t_1 - s_1|}{\text{——}} (p_1, t_1) \sim (p_2, s_2) \text{----} (p_{n-1}, t_{n-1}) \sim (p_n, s_n) \overset{|t_n - s_n|}{\text{——}} (p_n, t_n) \sim y$$

and call n the number of *steps*.

Let $X = Y/\sim$ and define $d : X \times X \to \mathbb{R}$ by

$$d(\bar{x}, \bar{y}) = \inf\{l(p, s, t, n) : (p, s, t, n) \text{ a route in } Y \text{ from } x \text{ to } y\}.$$

The infimum exists since Γ is connected, and each route is independent of the representatives x, y of $\bar{x}$ and $\bar{y}$ so d is well-defined.

We show that (X, d) is a metric space. For any $x, y \in Y$, $d(\bar{x}, \bar{y}) \geq 0$ and if $\bar{x} = \bar{y}$ then $d(\bar{x}, \bar{y}) = 0$. Conversely, take any $x = (q, r)$, $y = (q', r') \in Y$ and suppose $d(\bar{x}, \bar{y}) = 0$. Then for all $\varepsilon > 0$ there exists a route in Y from x to y with length less than ε. If $r \notin \{0, 1\}$ ($\bar{x}$ not at a vertex) then there exists a route with at most 1 step from x to y (take $\varepsilon < \min\{r, 1 - r\}$), and if it has 1 step then $x \sim (p_1, s_1)$ and $y \sim (p_1, t_1)$ for some $p_1 \in E$ and $s_1, t_1 \in [0, 1]$ with $|t_1 - s_1| > 0$, so the length of any route from x to y is at least $|t_1 - s_1|$, a contradiction. Thus a zero step route from x to y exists, giving $\bar{x} = \bar{y}$. The same argument holds if $r' \notin \{0, 1\}$, and if $r, r' \in \{0, 1\}$ then every route from x to y has integral length, so there exists a route of length 0 (put $\varepsilon = 1$) and thus $\bar{x} = \bar{y}$. Hence in all cases if $d(\bar{x}, \bar{y}) = 0$ then $\bar{x} = \bar{y}$. For any $x, y \in Y$, any route from x to y can be reversed with the same length (and vice versa) so $d(\bar{y}, \bar{x}) = d(\bar{x}, \bar{y})$. Take any $x, y, z \in Y$ and suppose $d(\bar{x}, \bar{z}) > d(\bar{x}, \bar{y}) + d(\bar{y}, \bar{z})$. Let $\varepsilon = d(\bar{x}, \bar{z}) - d(\bar{x}, \bar{y}) - d(\bar{y}, \bar{z}) > 0$. There exists a route from x to y with length $\lambda < d(\bar{x}, \bar{y}) + \frac{\varepsilon}{2}$ and a route from y to z with length $\mu < d(\bar{y}, \bar{z}) + \frac{\varepsilon}{2}$. So there exists a route from x to z with length at most $\lambda + \mu$ (concatenating and simplifying in the obvious way). Thus

$$d(\bar{x}, \bar{z}) \leq \lambda + \mu < d(\bar{x}, \bar{y}) + d(\bar{y}, \bar{z}) + \varepsilon = d(\bar{x}, \bar{z})$$

which is a contradiction. Hence $d(\bar{x}, \bar{z}) \leq d(\bar{x}, \bar{y}) + d(\bar{y}, \bar{z})$ and (X, d) is a metric space.

Take any $x, y \in Y$. If there exists a route from x to y with at most 1 step then $x \sim (e, r)$ and $y \sim (e, r')$ for some edge $e \in E$ and some $r, r' \in [0, 1]$, and by the triangle inequality $d(\bar{x}, \bar{y}) = |r - r'|$. If instead every route from x to y has at least 2 steps, then we claim there exists $z = (e, u) \in Y$ with $u \in \{0, 1\}$ such that $\bar{x} \neq \bar{z} \neq \bar{y}$ and $d(\bar{x}, \bar{y}) = d(\bar{x}, \bar{z}) + d(\bar{z}, \bar{y})$. Write $x = (p, s)$ and $y = (q, t)$ where $p, q \in E$ and $s, t \in [0, 1]$. Suppose $s \notin \{0, 1\}$. Let $z_0 = (p, 0)$, $z_1 = (p, 1)$ and choose $z \in \{z_0, z_1\}$ so that $d(\bar{x}, \bar{z}) + d(\bar{z}, \bar{y})$ is minimized. Then $\bar{x} \neq \bar{z} \neq \bar{y}$. For any route (p', s', t', n) from x to y, $(p'_1, s'_1) = (p, s)$ since $s \notin \{0, 1\}$ and $t'_1 \in \{0, 1\}$ since $n \geq 2$ so $(p'_1, t'_1) = z_j$

for some $j \in \{0,1\}$ and thus $d(\bar{z_j},\bar{y}) \le \sum_{i=2}^{n} |t'_i - s'_i|$. Also $d(\bar{x},\bar{z_j}) = |t'_1 - s'_1|$, so

$$d(\bar{x},\bar{z}) + d(\bar{z},\bar{y}) \le d(\bar{x},\bar{z_j}) + d(\bar{z_j},\bar{y}) \le l(p',s',t',n).$$

As the route was arbitrarily chosen, $d(\bar{x},\bar{y}) \ge d(\bar{x},\bar{z}) + d(\bar{z},\bar{y})$ by definition of the metric, and the triangle inequality gives $d(\bar{x},\bar{y}) = d(\bar{x},\bar{z}) + d(\bar{z},\bar{y})$. If $t \notin \{0,1\}$, the same argument applies. And if $s,t \in \{0,1\}$ then every route from x to y has integral length at least 2, so there exists a route (p',s',t',n) from x to y with length $d(\bar{x},\bar{y})$ and putting $z = (p'_1,t'_1)$ suffices. This proves the claim. Using the claim, it follows by a simple induction that (X,d) is a geodesic space.

Next we show that (X,d) is complete. Take any Cauchy sequence (x_n) in X. Suppose that there exists a subsequence of (x_n) which can be written in the form $\left(\overline{(q,t)}\right)_{n\in\mathbb{N}}$ for some fixed edge $q \in E$ and some sequence (t_n) in $[0,1]$. As $[0,1]$ is complete and the subsequence is also Cauchy, it converges to a limit $\overline{(q,t)}$ for some $t \in [0,1]$ and the sequence (x_n), being Cauchy, converges to $\overline{(q,t)}$ too. Suppose instead that no such subsequence exists. Then there exists a subsequence $\left(\overline{(q_n,t_n)}\right)_{n\in\mathbb{N}}$ of (x_n) with the following property: for all $m,n \in \mathbb{N}$ with $m \ne n$ there do not exist $e \in E$ and $r,r' \in [0,1]$ satisfying $(q_n,t_n) \sim (e,r)$ and $(q_m,t_m) \sim (e,r')$ (no distinct terms can be written as points on the same edge). So for any $m,n \in \mathbb{N}$ with $m \ne n$ there is no route from (q_m,t_m) to (q_n,t_n) with fewer than 2 steps. The subsequence $\left(\overline{(q_n,t_n)}\right)$ is also Cauchy, so there exists $N \in \mathbb{N}$ such that for all $m,n \ge N$ there is a route from (q_m,t_m) to (q_n,t_n) with length less than $\frac{1}{2}$, and for $m \ne n$ each such route therefore has exactly 2 steps (as any route with 3 or more steps has length greater than 1). It follows that there exist $r_N, r_{N+1}, r_{N+2}, \ldots$ in $\{0,1\}$ such that

$$(q_N,r_N) \sim (q_{N+1},r_{N+1}) \sim (q_{N+2},r_{N+2}) \sim \ldots.$$

So for any $m,n \ge N$ with $m \ne n$ we have the inequality

$$d\left(\overline{(q_m,t_m)},\overline{(q_N,r_N)}\right) \le d\left(\overline{(q_m,t_m)},\overline{(q_n,t_n)}\right).$$

Thus the subsequence converges to $\overline{(q_N,r_N)}$, as does the sequence (x_n) itself. Hence (X,d) is complete.

We can now define the isometric embedding and constituent paths. The graph Γ is connected so for each vertex $v \in V$ there exists an edge $e \in E$ with $v \in e$ and thus $\delta(e,r) = v$ for some $r \in \{0,1\}$. And for any $e' \in E$ and $r' \in \{0,1\}$, if $\delta(e',r') = v$ as well then $(e,r) \sim (e',r')$. So the map $\phi : V \to X$, $v \mapsto \overline{(e,r)}$ such that $\delta(e,r) = v$ is well-defined. Take any $v,w \in V$ and choose $x,y \in Y$ such that $\phi(v) = \bar{x}$ and $\phi(w) = \bar{y}$. Any path in Γ from v to w corresponds naturally with a route in Y from x to y having the same length and vice versa, so $d_E(v,w) = d(\bar{x},\bar{y})$ and ϕ is an isometric embedding. For any $e = \{e_1, e_2\} \in E$, define $p_e : [0,1] \to X$ by $t \mapsto \overline{(e,t)}$. Then p_e is a geodesic path between $\phi(e_1)$ and $\phi(e_2)$, and the subsets $p_e((0,1))$ for each $e \in E$ form a partition of X (from the definitions) and each is open (by the claim). Hence we have a geometric realization of Γ.

Suppose Γ is locally finite, and take any $x = (q,t) \in Y$. If $t \notin \{0,1\}$ then $\mathrm{im}(p_q)$ is a compact neighbourhood of $\bar{x}$ in X. Otherwise let S be the set $\{e \in E : \delta(e,0) = \delta(x) \text{ or } \delta(e,1) = \delta(x)\}$. Then S is finite so $\bigcup_{e\in S} \mathrm{im}(p_e)$ is a compact neighbourhood of $\bar{x}$. Thus X is locally compact (and complete) so by the Hopf-Rinow theorem (2.7i), X is proper. □

Proposition 2.9. *Let $\Gamma = (V,E)$ be a connected graph, and $(X, \phi, (p_e)_{e\in E})$, $(X', \phi', (p'_e)_{e\in E})$ geometric realizations of Γ. Then there exists a unique isometry ψ from X to X' such that $\psi \circ \phi = \phi'$.*

Proof. Let I be the open interval $(0,1) \subset \mathbb{R}$. We first show that for any $x \in X \setminus \mathrm{im}(\phi)$ there exists a unique pair $e \in E$ and $t \in I$ such that $p_e(t) = x$. Take any $x \in X \setminus \mathrm{im}(\phi)$. By definition of a geometric realization, the sets $p_e(I)$ for each $e \in E$ form a partition of $X \setminus \mathrm{im}(\phi)$ so there exists $e \in E$ such that x is an element of $p_e(I)$, and for any $f \in E$ if x is also in $p_f(I)$ then $p_f(I) = p_e(I)$. As $\mathrm{im}(p_f)$ and $\mathrm{im}(p_e)$ are compact and therefore closed, we have

$$\mathrm{im}(p_f) = \overline{p_f(I)} = \overline{p_e(I)} = \mathrm{im}(p_e)$$

and it follows that $f = e$, making e unique. Now x is an element of $p_e(I)$ so $x = p_e(t)$ for some $t \in I$, and p_e is a geodesic path so for any $u \in I$ if $x = p_e(u)$ as well then $|t-u| = d(p_e(t), p_e(u)) = 0$ and hence $u = t$, making t unique too.

Take any $e \in E$, $x \in \mathrm{im}(p_e)$ and $y \in X \setminus \mathrm{im}(p_e)$. We claim that any path in X from x to y runs via $p_e(0)$ or $p_e(1)$, (i.e. at least one of these points lies in the image of the path). By definition of a geometric realization, $p_e(I)$ is open in X, and $\mathrm{im}(p_e)$ is closed so $X \setminus \mathrm{im}(p_e)$ is open. Thus $p_e(I)$ and $X \setminus \mathrm{im}(p_e)$ form a partition of $X \setminus \{p_e(0), p_e(1)\}$ into open subsets with x in the first and y in the second. It follows that there is no path in $X \setminus \{p_e(0), p_e(1)\}$ from x to y, and this proves the claim.

For any $x, y \in X$, we have $x = p_e(t)$ and $y = p_f(u)$ for some $e, f \in E$ chosen equal if possible and some $t, u \in [0,1]$, since Γ is connected. If $e = f$ then $d(x,y) = |t-u|$ as p_e is a geodesic path. Otherwise there exists a geodesic path in X from x to y and by the claim it runs via $p_e(0)$ or $p_e(1)$ and similarly via $p_f(0)$ or $p_f(1)$. Hence $d(x,y) = d(p_e(t), p_f(u)) = \min(A)$ where

$$A = \{d(p_e(t), p_e(r)) + d(p_e(r), p_f(s)) + d(p_f(s), p_f(u)) : r, s \in \{0,1\}\}$$

which simplifies to

$$A = \{d(p_e(r), p_f(s)) + r(1-2t) + t + s(1-2u) + u : r, s \in \{0,1\}\} \quad (7)$$

In the same way, all the above holds for X' with ϕ' and p'. Without loss of generality, we may assume for each edge $e \in E$ that the parametrizations of p_e and p'_e are chosen so that for all $t \in \{0,1\}$ and all $v \in e$ we have

$$p_e(t) = \phi(v) \Leftrightarrow p'_e(t) = \phi'(v) \quad (8)$$

(replacing one of them with its reverse path if necessary).

We can now define $\psi : X \to X'$. For any $x \in X$, $x = p_e(t)$ for some $e \in E$ and $t \in [0,1]$. If $x \in \mathrm{im}(\phi)$ then $x = \phi(v)$ for some $v \in V$ and we must define $\psi(x) = \phi'(v)$ to get $\psi \circ \phi = \phi'$. Note that by (8) we can write this as $\psi(p_e(t)) = p'_e(t)$. Otherwise t is in I and by (7) the elements of $\mathrm{im}(\phi)$ at a distance less than 1 from x are precisely $p_e(0)$ and $p_e(1)$ (since distances between elements of $\mathrm{im}(\phi)$ are integral) so to make ψ an isometry we must define $\psi(x) = p'_e(t)$. Conclusion: the only candidate for ψ is the map $\psi : X \to X'$ given by $\psi(p_e(t)) = p'_e(t)$. This is well-defined and satisfies $\psi \circ \phi = \phi'$ by (8). The map ψ is also clearly surjective, so it remains to prove that ψ is an isometric embedding. Take any $x, y \in X$. If there exists $e \in E$ such that $x = p_e(t)$ and $y = p_e(u)$ for some $t, u \in [0,1]$ then

$$d(\psi(x), \psi(y)) = d(p'_e(t), p'_e(u)) = |t-u| = d(p_e(t), p_e(u)) = d(x,y)$$

as p_e and p'_e are isometric embeddings. Otherwise $x = p_e(t)$ and $y = p_f(u)$ for some distinct $e, f \in E$ and some $t, u \in [0,1]$. For any $r, s \in \{0,1\}$, $p_e(r) = \phi(v)$ and $p_f(s) = \phi(w)$ for some $v, w \in V$ so

$$\begin{aligned} d(p_e(r), p_f(s)) &= d(\phi(v), \phi(w)) = d_E(v,w) \\ &= d(\phi'(v), \phi'(w)) = d(p'_e(r), p'_f(s)) \end{aligned}$$

as ϕ and ϕ' are isometric embeddings and using (8). Hence by (7),

$$d(\psi(x), \psi(y)) = d(p'_e(t), p'_f(u)) = d(p_e(t), p_f(u)) = d(x,y)$$

and thus ψ is an isometry. □

3 Coarse geometry and group actions

In this chapter, we focus on the large-scale structure of metric spaces. We prove the theorem of Schwarz and Milnor, and apply it to finitely generated groups via the geometric realization of their Cayley graphs. Further results in this area can be found in [Roe] and [Bridson & Haefliger].

Definition. Let X be a set, (Y, d) a metric space and f, g maps from X to Y. We say f is *close* to g, denoted $f \sim g$, if there exists $\lambda \in \mathbb{R}_{>0}$ such that for all $x \in X$ we have $d(f(x), g(x)) < \lambda$.

This is clearly an equivalence relation.

Definition. Let (X, d_X) and (Y, d_Y) be metric spaces. A map $f : X \to Y$ is called *coarse* if it satisfies both of the following conditions:

1. for all $B \subset Y$, if B is bounded then $f^{-1}(B)$ is bounded;
2. $\forall \lambda > 0, \exists \mu > 0, \forall x, x' \in X, d_X(x, x') < \lambda \Rightarrow d_Y(f(x), f(x')) < \mu$.

We say X is *coarsely equivalent* to Y if there exist coarse maps $f : X \to Y$ and $g : Y \to X$ such that $g \circ f \sim \mathrm{id}_X$ and $f \circ g \sim \mathrm{id}_Y$.

It follows immediately that the composition of coarse maps yields a coarse map.

Proposition 3.1. *"is coarsely equivalent to" is an equivalence relation on any set of metric spaces.*

Proof. The relation is clearly reflexive and symmetric. Take any metric spaces X, Y and Z and suppose there exist coarse maps $X \overset{f_1}{\to} Y \overset{f_2}{\to} Z$ and $X \overset{g_1}{\leftarrow} Y \overset{g_2}{\leftarrow} Z$ such that $g_1 \circ f_1 \sim \mathrm{id}_X$, $f_1 \circ g_1 \sim \mathrm{id}_Y \sim g_2 \circ f_2$ and $f_2 \circ g_2 \sim \mathrm{id}_Z$. Then $f_2 \circ f_1 : X \to Z$ and $g_1 \circ g_2 : Z \to X$ are also coarse maps. As $g_1 \circ f_1 \sim \mathrm{id}_X$, there exists $\lambda_1 > 0$ such that for all $x \in X$ we have $d((g_1 \circ f_1)(x), x) < \lambda_1$. And $g_2 \circ f_2 \sim \mathrm{id}_Y$ so there exists $\lambda_2 > 0$ such that for all $y \in Y$, $d((g_2 \circ f_2)(y), y) < \lambda_2$. Also the map g_1 is coarse so there exists $\mu > 0$ such that for all $y, y' \in Y$,

$$d(y, y') < \lambda_2 \Rightarrow d(g_1(y), g_1(y')) < \mu.$$

Let $\lambda_3 = \lambda_1 + \mu$ and take any $x \in X$. Then $d((g_1 \circ f_1)(x), x) < \lambda_1$ and $d((g_2 \circ f_2)(f_1(x)), f_1(x)) < \lambda_2$ so

$$\begin{aligned} d((g_1 \circ g_2 \circ f_2 \circ f_1)(x), x) &\leq d((g_1 \circ g_2 \circ f_2 \circ f_1)(x), (g_1 \circ f_1)(x)) \\ &\quad + d((g_1 \circ f_1)(x), x) \\ &< \mu + \lambda_1 = \lambda_3 \end{aligned}$$

hence $(g_1 \circ g_2) \circ (f_2 \circ f_1) \sim \mathrm{id}_X$. In the same way, $(f_2 \circ f_1) \circ (g_1 \circ g_2) \sim \mathrm{id}_Z$. □

An alternative concept often used is that of quasi-isometry.

Definition. Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$ a map. We call f a *quasi-isometric embedding* if there exist $a \in \mathbb{R}_{\geq 1}$ and $b \in \mathbb{R}_{\geq 0}$ such that for all $x, x' \in X$,

$$\frac{1}{a} d_X(x, x') - b \leq d_Y(f(x), f(x')) \leq a d_X(x, x') + b. \tag{9}$$

We call f a *quasi-isometry* if it is a quasi-isometric embedding and there exists $c \in \mathbb{R}_{\geq 0}$ such that f has the following property: for all $y \in Y$ there exists $x \in X$ such that $d_Y(f(x), y) \leq c$. We say that X *is quasi-isometric with* Y, denoted $X \sim Y$, if there exists a quasi-isometry from X to Y.

Note that a quasi-isometric embedding need not be injective (in spite of the name). It is clear that the composition of quasi-isometric embeddings gives a quasi-isometric embedding.

Lemma 3.2. *Let X and Y be metric spaces, $f : X \to Y$ a quasi-isometric embedding and $g : Y \to X$ a map such that $f \circ g \sim \mathrm{id}_Y$. Then f and g are both quasi-isometries.*

Proof. The map f is a quasi-isometric embedding so there exist $a \geq 1$ and $b \geq 0$ such that for all $x, x' \in X$ the equation (9) holds. And $f \circ g \sim \mathrm{id}_Y$ so there exists $c > 0$ such that for all $y \in Y$ we have $d(f(g(y)), y) < c$. It follows immediately that f is a quasi-isometry. For all $y, y' \in Y$,

$$d(f(g(y)), f(g(y'))) \leq d(f(g(y)), y) + d(y, y') + d(y', f(g(y'))) < d(y, y') + 2c$$

and thus

$$d(g(y), g(y')) \leq a(d(f(g(y)), f(g(y'))) + b) < a(d(y, y') + 2c + b).$$

Similarly for all $y, y' \in Y$,

$$d(f(g(y)), f(g(y'))) \geq d(y, y') - d(f(g(y)), y) - d(f(g(y')), y') > d(y, y') - 2c$$

and hence

$$d(g(y), g(y')) \geq \frac{1}{a}\left(d(f(g(y)), f(g(y'))) - b\right) > \frac{1}{a}(d(y, y') - 2c - b)$$

making g a quasi-isometric embedding. Also for any $x \in X$, putting $y = f(x)$ gives

$$d(g(y), x) \leq a(d(f(g(y)), f(x)) + b) = a(d(f(g(y)), y) + b) < a(c + b)$$

and thus g is a quasi-isometry from Y to X. $\square$

Proposition 3.3. *The relation "is quasi-isometric with" is an equivalence relation on any set of metric spaces.*

Proof. The relation is clearly reflexive. Take any metrics spaces X and Y, and suppose there exists a quasi-isometry $f : X \to Y$. Then there exists $c \geq 0$ with the property that for every $y \in Y$ there is an element $x_y \in X$ such that $d(f(x_y), y) \leq c$. Let $g : Y \to X$ be the map $y \mapsto x_y$. Then $f \circ g \sim \mathrm{id}_Y$ so by (3.2) the map g is a quasi-isometry from Y to X.

For transitivity, take any metric spaces X, Y and Z, and suppose there exist quasi-isometries $X \xrightarrow{f} Y \xrightarrow{g} Z$. Define $h = g \circ f$. Then clearly h is a quasi-isometric embedding. The maps f and g are quasi-isometries, so there exist $a \geq 1$ and $b \geq 0$ such that for all $y, y' \in Y$,

$$\frac{1}{a}d(y, y') - b \leq d(g(y), g(y')) \leq ad(y, y') + b$$

and there exist $c, c' \geq 0$ such that the following both hold:

$$\forall y \in Y, \exists x \in X, d(f(x), y) \leq c \text{ and } \forall z \in Z, \exists y \in Y, d(g(y), z) \leq c'.$$

Thus for any $z \in Z$ there exist $y \in Y$ and $x \in X$ satisfying $d(f(x), y) \leq c$ and $d(g(y), z) \leq c'$ so

$$\begin{aligned} d(h(x), z) &\leq d(h(x), g(y)) + d(g(y), z) \\ &\leq ad(f(x), y) + b + d(g(y), z) \leq ac + b + c'. \end{aligned}$$

Hence h is a quasi-isometry from X to Z. □

Example. The inclusion map from $\mathbb{Z}$ to $\mathbb{R}$ is a quasi-isometry.

Lemma 3.4. *Let X be a length space, Y a metric space and $f : X \to Y$ a map. The the following are equivalent:*

(a) $\exists a, b > 0, \forall x, x' \in X, d(f(x), f(x')) < ad(x, x') + b$;

(b) $\forall a' > 0, \exists b' > 0, \forall x, x' \in X, d(x, x') < a' \Rightarrow d(f(x), f(x')) < b'$;

(c) $\exists a'', b'' > 0, \forall x, x' \in X, d(x, x') < a'' \Rightarrow d(f(x), f(x')) < b''$.

Proof. It is clear that (a)⇒(b) (take any $a' > 0$ and let $b' = aa' + b$), and (b)⇒(c). Suppose (c) holds and let $a = \frac{b''}{a''}$ and $b = b''$. Then for any $x, x' \in X$ there exists $n \in \mathbb{N}_{\geq 1}$ such that $(n-1)a'' \leq d(x, x') \leq na''$. And X is a length space so by (2.3) there exist $x_0, x_1, \ldots, x_n \in X$ such that $x_0 = x$, $x_n = x'$ and for each $i \in \{1, 2, \ldots, n\}$ we have $d(x_{i-1}, x_i) < a''$. Thus

$$\begin{aligned} d(f(x), f(x')) &\leq \sum_{i=1}^{n} d(f(x_{i-1}, f(x_i)) \\ &< nb'' \leq \frac{b''}{a''}d(x, x') + b = ad(x, x') + b. \end{aligned}$$

□

Proposition 3.5. *Let X and Y be metric spaces.*

(i) If X and Y are quasi-isometric then they are coarsely equivalent.

(ii) If X and Y are coarsely equivalent and they are length spaces, then they are quasi-isometric.

Proof. (i) Suppose there exists a quasi-isometry $f : X \to Y$. Then there exist $a \geq 1$ and $b \geq 0$ such that for all $x, x' \in X$ the equation (9) holds, and there exists $c \geq 0$ with the property that for every $y \in Y$ there is an element $x_y \in X$ satisfying $d(f(x_y), y) \leq c$. Define $g : Y \to X, y \mapsto x_y$. We show that f and g are coarse maps with $f \circ g \sim \mathrm{id}_Y$ and $g \circ f \sim \mathrm{id}_X$. For any bounded $B \subset Y$, there exists $r \geq 0$ such that for all $b, b' \in B$ we have $d(b, b') \leq r$. Thus for any $a, a' \in f^{-1}(B)$, $d(f(a), f(a')) \leq r$ so $d(a, a') \leq a(r + b)$ and $f^{-1}(B)$ is bounded. And for any $\lambda > 0$, putting $\mu = a\lambda + b$ gives for all $x, x' \in X$,

$$d(x, x') < \lambda \Rightarrow d(f(x), f(x')) < a\lambda + b = \mu.$$

Hence f is a coarse map. Moreover by choice of g we have $f \circ g \sim \mathrm{id}_Y$ so by (3.2) the map g is also a quasi-isometry, and therefore by the same reasoning g is a coarse map too. And for any $x \in X$, $d(f(g(f(x))), f(x)) \leq c$ so $d(g(f(x)), x) \leq a(c+b)$ giving $g \circ f \sim \mathrm{id}_X$. Hence X is coarsely equivalent to Y.

(ii) Suppose X and Y are coarsely equivalent length spaces. Then there exist coarse maps $f : X \to Y$ and $g : Y \to X$ such that $g \circ f \sim \mathrm{id}_X$ and $f \circ g \sim \mathrm{id}_Y$. We show that f is a quasi-isometry. By lemma (3.4), there exist $a, b > 0$ such that for all $y, y' \in Y$, $d(g(y), g(y')) < ad(y, y') + b$, and there exists $c > 0$ such that for all $x \in X$, $d(g(f(x)), x) < c$. So for all $x, x' \in X$,

$$\begin{aligned} d(x, x') &\leq d(x, g(f(x))) + d(g(f(x)), g(f(x'))) + d(g(f(x')), x') \\ &< ad(f(x), f(x')) + b + 2c. \end{aligned}$$

Also by lemma (3.4), there exist $a', b' > 0$ such that for all $x, x' \in X$, $d(f(x), f(x')) < a'd(x, x') + b'$, so f is a quasi-isometric embedding. And as $f \circ g \sim \mathrm{id}_Y$, f is a quasi-isometry as well (3.2). □

Using these concepts, we can examine the large-scale geometry of groups.

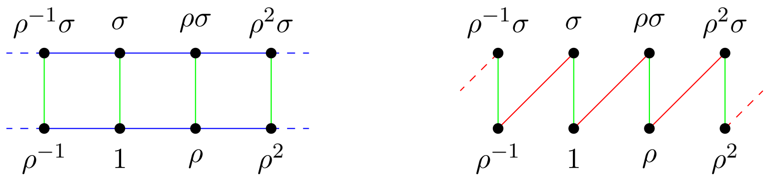
Definition. Let G be a group, $S \subset G$ and $\tilde{S} = (S \cup \{s^{-1} : s \in S\}) \setminus \{1\}$. The *Cayley graph* of G with respect to S is the graph with vertex set G and edge set $\left\{\{g, gs\} : g \in G, s \in \tilde{S}\right\}$.

Example. Let D_∞ be the infinite dihedral group

$$D_\infty = \langle \rho, \sigma | \sigma^2 = 1, \rho\sigma\rho\sigma = 1 \rangle.$$

The Cayley graph of D_∞ with respect to $S = \{\rho, \sigma\}$ and separately with respect to $S' = \{\sigma, \rho\sigma\}$ are pictured in figure 3.

Figure 3: Cayley graph of D_∞ with respect to S and S' respectively.



The Cayley graph of a group G with respect to a subset $S \subset G$ is connected if and only if S generates G.

Definition. Let G be a group generated by a subset $S \subset G$. The *word metric* on G with respect to S is the map $d_S : G \times G \to \mathbb{R}$ such that for each $g, h \in G$, $d_S(g, h)$ is the length of a shortest path from g to h in the Cayley graph of G with respect to S.

This is clearly a metric on G.

Definition. Let G be a group acting on a metric space (X, d). We say G acts on X *by isometries* if for each $g \in G$ the map from X to X given by $x \mapsto gx$ is an isometry.

Example. Let G be a group generated by a subset $S \subset G$. Then left multiplication is an action of G on the metric space (G, d_S) by isometries. In particular, for any $g, h \in G$ we have $d_S(g, h) = d_S(1, g^{-1}h)$.

Different generating sets for a group G can yield different metrics, but if the sets are finite then the resulting spaces are quasi-isometric, as we now show.

Proposition 3.6. *Let G be a group with $S, T \subset G$ finite subsets such that $\langle S\rangle = G = \langle T\rangle$. Then the identity map is a quasi-isometry from the metric space (G, d_S) to (G, d_T).*

Proof. If $G = \{1\}$ then all distances are 0 and the result follows immediately. Otherwise, the set $\{d_T(1, s) : s \in S\}$ is non-empty and finite, so it has a maximum element $a \in \mathbb{Z}_{\geq 1}$. For all $g, h \in G$, writing each element of S in terms of the generators from T gives

$$d_T(g, h) = d_T(1, g^{-1}h) \leq a d_S(1, g^{-1}h) = a d_S(g, h).$$

In the same way, if $b = \max\{d_S(1, t) : t \in T\}$, then for all $g, h \in G$ we have $d_S(g, h) \leq b d_T(g, h)$. Thus the identity map from (G, d_S) to (G, d_T) is a quasi-isometric embedding, and it is surjective so it is a quasi-isometry. $\square$

Definition. Let G be a group acting on a metric space (X, d). We say G acts *cocompactly* on X if there exists a compact subset $Y \subset X$ such that $\bigcup_{g\in G} gY = X$ (where $gY = \{gy : y \in Y\}$). We say G acts *properly* on X if for every compact subset $Y \subset X$ the set $\{g \in G : gY \cap Y \neq \varnothing\}$ is finite.

Examples. The group $I_2(\mathbb{R})$ of isometries of $\mathbb{R}^2$ acts on it cocompactly but not properly. Its subgroup $O_2(\mathbb{R})$ of orthogonal isometries acts on $\mathbb{R}^2$ neither cocompactly nor properly. Any finite subgroup of $I_2(\mathbb{R})$ acts on $\mathbb{R}^2$ properly but not cocompactly. The subgroup of $I_2(\mathbb{R})$ consisting of all translations by an integral offset in each co-ordinate acts both cocompactly and properly on $\mathbb{R}^2$.

Lemma 3.7. *Let G be a group generated by a finite set S and (X, d) a metric space such that G acts on (X, d) by isometries. Then for all $x \in X$ there exists $a > 0$ such that for all $g, g' \in G$, $d(gx, g'x) \leq a d_S(g, g')$ (where d_S is the word metric on G with respect to S).*

Proof. If $G = \{1\}$ then $a = 1$ suffices. Otherwise let $\tilde{S} = S \cup \{s^{-1} : s \in S\}$, take any $x \in X$ and let $a = \max\{d(x, sx) : s \in \tilde{S}\}$. Take any $g, g' \in G$ and write $g^{-1}g' = s_1 s_2 \ldots s_n$ where $n \in \mathbb{N}$ is chosen as small as possible and $s_1, s_2, \ldots, s_n \in \tilde{S}$. Then

$$\begin{aligned} d(gx, g'x) &= d(x, g^{-1}g'x) = d(x, s_1 s_2 \ldots s_n x) \\ &\leq d(x, s_1 x) + d(s_1 x, s_1 s_2 x) + \ldots + d(s_1 \ldots s_{n-1} x, s_1 \ldots s_n x) \\ &= d(x, s_1 x) + d(x, s_2 x) + \ldots + d(x, s_n x) \\ &\leq na = a d_S(1, g^{-1}g') = a d_S(g, g'). \end{aligned}$$

□

Theorem 3.8 (Schwarz, Milnor[4]). *Let (X, d) be a non-empty and proper geodesic space and G a group acting co-compactly, properly and by isometries on X. Then G is finitely generated and for any $a \in X$ and any finite subset $S \subset G$ with $\langle S \rangle = G$, the map $f : (G, d_S) \to (X, d)$ given by $g \mapsto ga$ is a quasi-isometry.*

Proof. The metric space X is non-empty so there exists $a \in X$, and the group G acts cocompactly on X so there exists a compact subset $Y \subset X$ such that $\bigcup_{g\in G} gY = X$. As Y is compact, there also exists $r > 0$ such that $Y \subset B_r[a]$, the closed ball of radius r with centre a. Let

$$S = \{g \in G : d(a, ga) \leq 2r+2\} \text{ and } S' = \{g \in G : gB_{r+1}[a] \cap B_{r+1}[a] \neq \varnothing\}.$$

We shall show that $S = S'$. For any $g \in S'$, there exists $x \in X$ such that $d(x, a) \leq r + 1$ and $d(x, ga) \leq r + 1$ (G acts on X by isometries so $gB_{r+1}[a] = B_{r+1}[ga]$) and thus $d(a, ga) \leq 2r + 2$ so $g \in S$. Hence $S' \subset S$.

[4]see [Bridson & Haefliger]

Conversely, take any $g \in S$. The metric space (X, d) is a geodesic space so there exists $x \in X$ such that

$$d(a, x) = d(x, ga) = \frac{1}{2} d(a, ga)$$

(midpoint on a geodesic path from a to ga). But $d(a, ga) \leq 2r + 2$ so $x \in B_{r+1}[a] \cap B_{r+1}[ga]$ and thus $g \in S'$. Hence $S \subset S'$ as well, so $S = S'$. As X is proper, $B_{r+1}[a]$ is compact, and G acts properly on X so the set $S = S'$ is finite.

We now show that S generates G: take any $g \in G$. There exists a geodesic path $p : [0, l] \to X$ from a to ga (where $l \geq 0$ is the length of p) and there exists a partition $0 = t_0 < t_1 < \ldots < t_n = l$ of $[0, l]$ such that for each $i \in \{1, 2, \ldots, n\}$ we have $d(p(t_{i-1}), p(t_i)) \leq 2$. Also $\bigcup_{g \in G} B_r[ga] = X$ so for each $i \in \{1, 2, \ldots, n-1\}$ there exists $g_i \in G$ such that $p(t_i) \in B_r[g_i a]$. Let $g_0 = 1$ and $g_n = g$. Then for each $i \in \{1, 2, \ldots, n\}$,

$$d(g_{i-1}a, g_i a) \leq d(g_{i-1}a, p(t_{i-1})) + d(p(t_{i-1}), p(t_i)) + d(p(t_i), g_i a) \leq 2r + 2$$

so $d(a, g_{i-1}^{-1} g_i a) \leq 2r + 2$ and thus $g_{i-1}^{-1} g_i \in S$. Hence

$$g = (g_0^{-1} g_1)(g_1^{-1} g_2) \ldots (g_{n-1}^{-1} g_n) \in \langle S \rangle$$

and G is therefore generated by the finite set S.

Define $f : (G, d_S) \to (X, d)$ by $g \mapsto ga$. We shall show that f is a quasi-isometry. Take any $g \in G$. Then $g = s_1 s_2 \ldots s_m$ for some $m \in \mathbb{N}$ chosen as small as possible and some $s_1, s_2, \ldots, s_m \in S$ (G acts on X by isometries, so it follows from the definition of S that for each $s \in S$, the element s^{-1} is also in S). If $\frac{1}{2} d(a, ga) \leq m - 1$ then using the above construction we can write g as a product of fewer than m elements of S, contradicting the minimality of m. So $d(a, ga) > 2m - 2 = 2d_S(1, g) - 2$. And G acts on both (G, d_S) and (X, d) by isometries, it follows that for all $g, g' \in G$, $2d_S(g, g') - 2 < d(ga, g'a)$. This together with lemma (3.7) makes f a quasi-isometric embedding, and $\bigcup_{g \in G} B_r[ga] = X$ so f is a quasi-isometry.

By (3.6), the same holds for any finite subset generating G. □

Example. Let the group $\mathbb{Z}$ act on the proper geodesic space $\mathbb{R}$ (with the Euclidean metric) by addition. Then $\mathbb{Z}$ acts by isometries, cocompactly and properly, and the theorem confirms that $\mathbb{Z}$ is finitely generated and that the inclusion map from $\mathbb{Z}$ to $\mathbb{R}$ is a quasi-isometry. Note that $\mathbb{Z}$ is the fundamental group of the circle S^1 and $\mathbb{R}$ its universal covering space (with the obvious projection). The Schwarz-Milnor theorem can be used to show that the fundamental group of any compact Riemannian manifold is finitely generated and quasi-isometric with its universal covering space (see [Roe]).

Corollary 3.9. *Let G be a finitely generated group and H a subgroup of finite index in G. Then H is also finitely generated, and H is quasi-isometric with G (using any word metrics).*

Proof. The group G is generated by a finite subset $S \subset G$ and the Cayley graph (G, E) of G with respect to S is connected and locally finite, so by (2.8) it has a proper geometric realization. Thus there exists a complete proper geodesic space (X, d) with an isometric embedding $\phi : (G, d_S) \to (X, d)$ and, for each edge $e = \{g, g'\} \in E$, a geodesic path $p_e : [0, 1] \to X$ between $\phi(g)$ and $\phi(g')$ such that $\{p_e((0, 1)) : e \in E\}$ is a partition of $X \setminus \operatorname{im}(\phi)$ into open subsets. The group G acts on (G, d_S) by left multiplication, and this action extends naturally to an action by isometries on X. By restriction, we get an action of H on X. The index of H in G is finite so

$$G = H \cup Ha_1 \cup Ha_2 \cup \ldots \cup Ha_n$$

for some $n \in \mathbb{N}$ and some $a_1, a_2, \ldots, a_n \in G$. Let $Y = \bigcup_{i=1}^{n} B_1[a_i]$. Then Y is compact and $\bigcup_{h \in H} hY = X$ so H acts on X cocompactly. Also, any compact subset $Z \subset X$ is totally bounded so it contains only finitely many elements of $\operatorname{im}(\phi)$ and it follows that $\{h \in H : hZ \cap Z \neq \varnothing\}$ is finite, so H acts properly on X, too. By the Schwarz-Milnor theorem (3.8), H is finitely generated and quasi-isometric with X. And X is quasi-isometric with G (take $H = G$), so H and G are quasi-isometric. $\square$

4 Polycyclic groups

In this chapter we look at the basic properties of polycyclic groups. We show that a group is polycyclic if and only if it is soluble and all its subgroups are finitely generated, we introduce the Hirsch length for virtually polycyclic groups, and we prove that any subgroup of such a group coincides with the intersection of the finite index subgroups containing it. For further results in this area, see [Wehrfritz], [Segal] and [Lennox & Robinson].

Definition. A *group property* is a class $\mathcal{P}$ of groups such that for any groups G_1 and G_2, if $G_1 \in \mathcal{P}$ and $G_1 \cong G_2$ then $G_2 \in \mathcal{P}$. We say a group G *has property* $\mathcal{P}$ if $G \in \mathcal{P}$.

For example, "abelian" and "soluble" are group properties.

Definition. Let G be a group, and $\mathcal{P}$, $\mathcal{Q}$ group properties.

1. We say G is *virtually* $\mathcal{P}$ if there exists a subgroup H with finite index in G such that H has property $\mathcal{P}$.
2. We say G is *residually* $\mathcal{P}$ if for every $g \in G \setminus \{1\}$ there exists $N \lhd G$ with $g \notin N$ such that G/N has property $\mathcal{P}$.
3. We say G is *poly-*$\mathcal{P}$ if there exist $r \in \mathbb{N}$ and subgroups
$$G = H_0 \supset H_1 \supset H_2 \supset \ldots \supset H_r = \{1\}$$
of G such that for each $i \in \{1, 2, \ldots, r\}$, $H_i \lhd H_{i-1}$ and H_{i-1}/H_i has property $\mathcal{P}$. In this case, we call the series a $\mathcal{P}$ series of G.
4. We say G is $\mathcal{P}$*-by-*$\mathcal{Q}$ if there exists a short exact sequence
$$1 \to K \to G \to H \to 1$$
such that K has property $\mathcal{P}$ and H has property $\mathcal{Q}$.
5. We call $\mathcal{P}$ *hereditary* if for every group G that has property $\mathcal{P}$, all subgroups of G also have property $\mathcal{P}$.

Lemma 4.1. *Let H be a subgroup of finite index in a group G. Then there exists a subgroup $N \subset H$ such that $N \lhd G$ and N has finite index in G.*

Proof. Take $N = \{g \in G : \text{for all } a \in G, gaH = aH\}$. □

Examples.

1. A group G is polyabelian if and only if it is soluble.
2. A group G is virtually trivial if and only if G is finite.

3. Let G be a group, $\mathcal{S}$ the set of all subgroups of finite index in G and $\mathcal{T}$ the set of all elements of $\mathcal{S}$ which are normal in G. Then

$$G \text{ is residually finite } \Leftrightarrow \bigcap_{N \in \mathcal{T}} N = \{1\} \Leftrightarrow \bigcap_{H \in \mathcal{S}} H = \{1\}$$

by lemma (4.1).

4. If a group property $\mathcal{P}$ is hereditary then a group G is virtually $\mathcal{P}$ if and only if G is $\mathcal{P}$-by-finite, again by (4.1).

Proposition 4.2. *Let $\mathcal{P}$ and $\mathcal{Q}$ be group properties.*

(i) If $\mathcal{P}$ is hereditary then so is poly-$\mathcal{P}$.

(ii) If $\mathcal{P}$ and $\mathcal{Q}$ are hereditary then so is $\mathcal{P}$-by-$\mathcal{Q}$.

Proof. (i) Suppose $\mathcal{P}$ is hereditary. Take any poly-$\mathcal{P}$ group G and any subgroup K of G. Then G has a $\mathcal{P}$ series $G = H_0 \supset H_1 \supset \ldots \supset H_r = \{1\}$. For each $i \in \{1, 2, \ldots, r\}$, $H_i \cap K$ is normal in $H_{i-1} \cap K$ and

$$(H_{i-1} \cap K) / (H_i \cap K) \cong H_i(H_{i-1} \cap K) / H_i$$

which is a subgroup of H_{i-1}/H_i and therefore has property $\mathcal{P}$. Thus

$$K = (H_0 \cap K) \supset (H_1 \cap K) \supset (H_2 \cap K) \supset \ldots \supset (H_r \cap K) = \{1\}$$

is a $\mathcal{P}$ series of K, and so K has property poly-$\mathcal{P}$.

(ii) Suppose $\mathcal{P}$ and $\mathcal{Q}$ are hereditary. Take any $\mathcal{P}$-by-$\mathcal{Q}$ group G and any subgroup G_0 of G. Then there exists a short exact sequence

$$1 \to K \xrightarrow{f} G \to H \to 1$$

such that K has property $\mathcal{P}$ and H has property $\mathcal{Q}$. The short sequence

$$1 \to G_0 \cap \operatorname{im}(f) \to G_0 \to G_0 / (G_0 \cap \operatorname{im}(f)) \to 1$$

is also exact and $\operatorname{im}(f)$ is isomorphic with K, so $G_0 \cap \operatorname{im}(f)$ is isomorphic with a subgroup of K and therefore has property $\mathcal{P}$. We also have

$$G_0 / (G_0 \cap \operatorname{im}(f)) \cong G_0 \operatorname{im}(f) / \operatorname{im}(f)$$

which is a subgroup of $(G/\operatorname{im}(f)) \cong H$ and therefore has property $\mathcal{Q}$. Hence G has property $\mathcal{P}$-by-$\mathcal{Q}$. □

Example. Any subgroup of a cyclic group is cyclic, so it follows by (4.2i) that any subgroup of a polycyclic group is polycyclic.

Lemma 4.3. *Let G be a soluble group such that every term in the derived series of G is finitely generated. Then G is polycyclic.*

Proof. Let $G \supset G' \supset G'' \supset G^{(3)} \supset \ldots$ be the derived series of G. The group G is soluble, so $G^{(r)} = \{1\}$ for some $r \in \mathbb{N}$. Induction on r.
Case $r \leq 1$: G is a finitely generated abelian group so $G = \langle\{g_1, g_2, \ldots, g_n\}\rangle$ for some $n \in \mathbb{N}$ and some $g_1, g_2, \ldots, g_n \in G$. For each $i \in \{0, 1, \ldots, n\}$, let $H_i = \langle\{g_1, g_2, \ldots, g_{n-i}\}\rangle$. Then $G = H_0 \supset H_1 \supset \ldots \supset H_n = \{1\}$ and for each $i \in \{1, 2, \ldots, n\}$ we have $H_i \lhd H_{i-1}$ (since G is abelian) and H_{i-1}/H_i is cyclic (generated by $g_{n-i+1}H_i$) so this is a cyclic series for G, making G polycyclic.
Case $r > 1$: the group G/G' is finitely generated and abelian, so by the first case it is polycyclic and has a cyclic series

$$G/G' = H_0/G' \supset H_1/G' \supset \ldots \supset H_k/G' = G'/G'.$$

Thus for each $i \in \{1, 2, \ldots, k\}$ we have $(H_i/G') \lhd (H_{i-1}/G')$ so $H_i \lhd H_{i-1}$ and $H_{i-1}/H_i \cong (H_{i-1}/G')/(H_i/G')$ is cyclic. And by the inductive hypothesis, G' is polycyclic so it has a cyclic series $G' = K_0 \supset K_1 \supset \ldots \supset K_m = \{1\}$ as well. It follows that

$$G = H_0 \supset H_1 \supset \ldots \supset H_k = G' = K_0 \supset K_1 \supset \ldots \supset K_m = \{1\}$$

is a cyclic series of G and hence G is polycyclic. The result follows by mathematical induction. □

Proposition 4.4. *Let G be a polycyclic group and*

$$G = H_0 \supset H_1 \supset \ldots \supset H_r = \{1\}$$

a cyclic series of G. For each $i \in \{1, 2, \ldots, r\}$, let a_iH_i generate the cyclic factor H_{i-1}/H_i. Then

$$G = \{a_1^{i_1} a_2^{i_2} \ldots a_r^{i_r} : i_1, i_2, \ldots, i_r \in \mathbb{Z}\}.$$

Proof. Take any $x \in G$. The left cosets of H_1 in $H_0 = G$ form a partition of G so $x \in a_1^{i_1}H_1$ for some $i_1 \in \mathbb{Z}$ and thus $x = a_1^{i_1}x_2$ for some $x_2 \in H_1$. And the left cosets of H_2 in H_1 form a partition of H_1 so $x_2 \in a_2^{i_2}H_2$ for some $i_2 \in \mathbb{Z}$ and thus $x_2 = a_2^{i_2}x_3$ for some $x_3 \in H_2$. Continuing in this way, we get the following: for each $j \in \{2, 3, \ldots, r\}$ there exist $i_j \in \mathbb{Z}$ and $x_{j+1} \in H_j$ such that $x_j = a_j^{i_j}x_{j+1}$. And $x_{r+1} \in H_r = \{1\}$ so $x = a_1^{i_1}a_2^{i_2}\ldots a_r^{i_r}$. □

Corollary 4.5. *Any polycyclic torsion group is finite.*

Proof. If G is a polycyclic group then by (4.4) there exist $r \in \mathbb{N}$ and $a_1, a_2, \ldots, a_r \in G$ such that every element of G can be written in the form $a_1^{i_1}a_2^{i_2}\ldots a_r^{i_r}$ where $i_1, i_2, \ldots, i_r \in \mathbb{Z}$. If G is also a torsion group then in particular $a_1, a_2, \ldots, a_r$ have finite order, so there are only finitely many elements of this form. □

Proposition 4.6. *Let G be a group. Then G is polycyclic if and only if G is soluble and every subgroup of G is finitely generated.*

Proof. Suppose G is polycyclic. Then G has a cyclic series, which is also an abelian series, so G is soluble. Also for any subgroup H of G, H is polycyclic as well (4.2i) so H is finitely generated (4.4).

Now suppose instead that G is soluble and every subgroup of G is finitely generated. Then in particular every term in the derived series of G is finitely generated, so G is polycyclic (4.3). □

Corollary 4.7. *Any quotient group of a polycyclic group is polycyclic.*

Proof. For any group G and any $N \lhd G$, if G is soluble then so is G/N, and if every subgroup of G is finitely generated then so is every subgroup of G/N. The result follows by (4.6). □

Corollary 4.8. *Any finitely generated nilpotent group is polycyclic.*

Proof. If G is a finitely generated nilpotent group then G is soluble and any subgroup of G is finitely generated[5], so G is polycyclic (4.6). □

Lemma 4.9. *Let G be a group and $N \lhd G$ such that N is finite and G/N is infinite cyclic. Then there exists $M \lhd G$ such that G/M is finite and M is infinite cyclic.*

Proof. The group G/N is infinite cyclic so $G/N = \langle aN \rangle$ for some $a \in G$ with infinite order, and $G = \langle a \rangle N$. Let $\sigma \in \mathrm{Aut}(N)$ be the automorphism $n \mapsto ana^{-1}$. Then $\sigma^k = \mathrm{id}_N$ for some $k \in \mathbb{Z}_{\geq 1}$ (as $\mathrm{Aut}(N)$ is finite). Let $M = \langle a^k \rangle \subset G$, which is infinite cyclic. Then $N \subset C_G(a^k)$ by choice of k, and $a \in C_G(a^k)$ so $G = \langle a \rangle N = C_G(a^k)$ and it follows that M is normal in G. We also have $[G : MN] = k$, $MN = NM$ and $[NM : M] \leq \#N$ so $[G : M] \leq k \cdot \#N$ and, in particular, G/M is finite. □

Note that the converse of the above lemma does not hold (example: D_∞).

Proposition 4.10. *Let G be a polycyclic group. Then G has a normal subgroup N with finite index in G such that N is poly-(infinite cyclic).*

Proof. The group G has a cyclic series $G = H_0 \supset H_1 \supset \ldots \supset H_r = \{1\}$ since G is polycyclic. Induction on r.
Case $r = 0$: $N = G$ suffices.
Case $r > 0$: by the induction hypothesis, there exists $K \lhd H_1$ with $[H_1 : K]$ finite such that K is poly-(infinite cyclic). Let $m = [H_1 : K] \in \mathbb{N}_{\geq 1}$ and let $M = \langle \{x^m : x \in H_1\} \rangle$. We claim that $M \subset K$ with $M \lhd G$ and that $[H_1 : M]$ is finite: the quotient group H_1/K has order m so for any $x \in H_1$ we have $x^m \in K$, and thus $M \subset K$. Also for all $x \in H_1$ and all $g \in G$, the element

[5] This is basic group theory, see for example [Macdonald].

$gx^m g^{-1} = (gxg^{-1})^m$ is in M since $H_1 \lhd G$, so M is normal in G. And H_1 is polycyclic so H_1/M is polycyclic too (4.7), but H_1/M is a torsion group, so it is finite (4.5), proving the claim. If $[G : H_1]$ is finite too, then putting $N = M$ suffices (the property poly-(infinite cyclic) is equivalent to poly-(infinite cyclic or trivial) which is hereditary by (4.2i)). Otherwise, G/H_1 is infinite cyclic, so $(G/M)/(H_1/M)$ is infinite cyclic with H_1/M finite, so by lemma (4.9) there exists $N/M \lhd G/M$ such that N/M is infinite cyclic and $(G/M)/(N/M)$ is finite. Thus $N \lhd G$ with $[G : N]$ finite, and N/M is infinite cyclic with M poly-(infinite cyclic), hence N is also poly-(infinite cyclic). The result follows by mathematical induction. □

Corollary 4.11. *Any poly-(infinite cyclic) group is torsion-free, and any polycyclic group is virtually torsion-free.*

Proof. If a group G is poly-(infinite cyclic) then it has an infinite cyclic series $G = H_0 \supset H_1 \supset \ldots \supset H_r = \{1\}$. Take any torsion element $x \in G$. Then xH_1 is a torsion element of G/H_1, which is infinite cyclic, so $xH_1 = 1H_1$ and $x \in H_1$. And xH_2 is a torsion element of H_1/H_2, which is infinite cyclic, so $x \in H_2$. Continuing in this way, we get $x \in H_r = \{1\}$. Hence G is torsion-free.

If G is polycyclic instead, then by (4.10) it has a poly-(infinite cyclic) subgroup N of finite index in G, and N is torsion-free by the same argument, so G is virtually torsion-free. □

Proposition 4.12. *Let G be an infinite group that is virtually polycyclic. Then there exists a non-trivial free abelian subgroup N of finite rank which is normal in G.*

Proof. The group G has a polycyclic subgroup H_1 with $[G : H_1]$ finite, the group H_1 has a poly-(infinite cyclic) subgroup H_2 with $[H_1 : H_2]$ finite (4.10) so $[G : H_2]$ is finite, and by (4.1) there exists a subgroup $M \subset H_2$ such that $M \lhd G$ and $[G : M]$ is also finite. As the property poly-(infinite cyclic) is hereditary, the group M is also poly-(infinite cyclic), so M is soluble (4.6) and $M^{(r)} = \{1\}$ for some $r \in \mathbb{N}$ which we choose as small as possible. The group M is infinite, so $r \geq 1$. Let $N = M^{(r-1)}$, the last non-trivial subgroup in the derived series of M. Then N is poly-(infinite cyclic) so it is finitely generated (4.4) and torsion-free (4.11). And $N' = M^{(r)} = \{1\}$ so N is also abelian. Thus, by the classification theorem for finitely generated modules over a principal ideal domain, N is free abelian of finite rank. Also, N is characteristic in M which is normal in G, hence $N \lhd G$. □

Proposition 4.13. *Let G be a virtually polycyclic group and*

$$G = H_0 \supset H_1 \supset \ldots \supset H_r = \{1\}, G = K_0 \supset K_1 \supset \ldots \supset K_s = \{1\}$$

both finite-or-(infinite cyclic) series of G. Then the number of infinite cyclic factors is the same in each case.

Proof. Take any $i \in \{1, 2, \ldots, r\}$ and any subgroup H of G with $H_i \subset H$ and $H \lhd H_{i-1}$. If $[H_{i-1} : H_i]$ is finite then so are $[H_{i-1} : H]$ and $[H : H_i]$. Otherwise H_{i-1}/H_i is an infinite cyclic group with H/H_i as a subgroup, so either H/H_i is trivial or it is infinite cyclic with finite index in H_{i-1}/H_i, making exactly one of the factors H_{i-1}/H and H/H_i infinite cyclic, and the other finite. The same holds if we insert any finite number of subgroups into the first series, and similarly for the second series, and the series therefore remain finite-or-(infinite cyclic) series with the number of infinite cyclic factors unaltered. But by the Schreier refinement theorem, there exist refinements of the two series that are equivalent (i.e. with a bijection between the sets of factors such that corresponding factors are isomorphic). It follows that both series have the same number of infinite cyclic factors. □

Definition. Let G be a virtually polycyclic group. The *Hirsch length* or *Hirsch number* of G, denoted $h(G)$, is the number of infinite cyclic factors in any finite-or-(infinite cyclic) series of G.

This is named after the German mathematician Kurt Hirsch[6].

Proposition 4.14. *Let G be a virtually polycyclic group and $N \lhd G$. Then N and G/N are also virtually polycyclic and $h(G) = h(G/N) + h(N)$.*

Proof. The property polycyclic is hereditary, so the property virtually polycyclic is equivalent to polycyclic-by-finite, which is hereditary by (4.2ii). Thus N is virtually polycyclic and has a finite-or-(infinite cyclic) series $N = K_0 \supset K_1 \supset \ldots \supset K_s = \{1\}$, in which $h(N)$ factors are infinite cyclic. The group G has a polycyclic subgroup G_1 with $[G : G_1]$ finite, so G_1N/N has finite index in G/N, and $G_1N/N \cong G_1/(G_1 \cap N)$ which is polycyclic (4.7), so G/N is also virtually polycyclic, with a finite-or-(infinite cyclic) series

$$G/N = H_0/N \supset H_1/N \supset \ldots \supset H_r/N = N/N$$

in which $h(G/N)$ factors are infinite cyclic. It follows that

$$G = H_0 \supset H_1 \supset \ldots \supset H_r = N = K_0 \supset K_1 \supset \ldots \supset K_s = \{1\}$$

is a finite-or-(infinite cyclic) series of G with $h(G/N) + h(N)$ infinite cyclic factors, and hence $h(G) = h(G/N) + h(N)$. □

Lemma 4.15. *Let A be a finitely generated abelian group (written additively). Then $\bigcap_{n \in \mathbb{Z}_{\geq 1}} nA = \{0\}$.*

[6] see [Wehrfritz]

Proof. Suppose first that A is cyclic: $A = \langle a \rangle$ for some $a \in A$. If a has finite order $k \in \mathbb{Z}_{\geq 1}$ then $\bigcap_{n \in \mathbb{Z}_{\geq 1}} nA \subset kA = \{0\}$. Otherwise, for any $m \in \mathbb{Z}$ and any $n \in \mathbb{Z}_{\geq 1}$, if $ma \in nA$ then n divides m and thus $m = 0$ or $n \leq |m|$, hence $ma \in \bigcap_{n \in \mathbb{Z}_{\geq 1}} nA \Rightarrow m = 0$ and the result follows.

In the general case, $A \cong B_1 \oplus B_2 \oplus \ldots \oplus B_m$ for some $m \in \mathbb{N}$ and some cyclic groups $B_1, B_2, \ldots, B_m$ (since A is a finitely generated abelian group). For each $j \in \{1, 2, \ldots, m\}$, $\bigcap_{n \in \mathbb{Z}_{\geq 1}} nB_j = \{0\}$ by the cyclic case, hence

$$\bigcap_{n \in \mathbb{Z}_{\geq 1}} nA = \bigcap_{n \in \mathbb{Z}_{\geq 1}} n \left(\bigoplus_{j=1}^{m} B_j \right) = \bigcap_{n \in \mathbb{Z}_{\geq 1}} \bigoplus_{j=1}^{m} nB_j = \bigoplus_{j=1}^{m} \bigcap_{n \in \mathbb{Z}_{\geq 1}} nB_j = \{0\}.$$

□

Theorem 4.16. *Let G be a virtually polycyclic group and K a subgroup of G. Let $\mathcal{S}$ be the set of all subgroups H of G such that $K \subset H$ and $[G : H]$ is finite. Then*

$$K = \bigcap_{H \in \mathcal{S}} H.$$

Proof. For any abelian group A and any $n \in \mathbb{Z}$, we write A^n in this proof for the subgroup $\{a^n : a \in A\}$ of A. Clearly $K \subset \bigcap_{H \in \mathcal{S}} H$. Suppose that G is abelian. Then $K \lhd G$ and G/K is a finitely generated abelian group. For each $n \in \mathbb{Z}_{\geq 1}$, $(G/K)^n$ is a finite index subgroup of G/K so it has the form H_n/K where H_n is a finite index subgroup of G containing K, i.e. each H_n is an element of $\mathcal{S}$. And by lemma (4.15) we have $\bigcap_{n \in \mathbb{Z}_{\geq 1}} (G/K)^n = \{1\}$, so $\bigcap_{H \in \mathcal{S}} H \subset \bigcap_{n \in \mathbb{Z}_{\geq 1}} H_n = K$.

We now prove the general case by induction on the Hirsch length $h(G)$.
Case $h(G) = 0$: G is finite so $K \in \mathcal{S}$ and $\bigcap_{H \in \mathcal{S}} H \subset K$.
Case $h(G) > 0$: the group G is infinite so by (4.12) there exists a non-trivial free abelian subgroup A of finite rank which is normal in G. Take any $g \in G \setminus K$. We shall show that $g \notin \bigcap_{H \in \mathcal{S}} H$. Suppose first that $g \notin KA$. Then g is not an element of the subgroup KA/A of G/A, and by (4.14) we have $h(G/A) < h(G)$ (since A is infinite) so by the induction hypothesis there exists a subgroup H/A of G/A such that $KA/A \subset H/A$, $[G/A : H/A]$ is finite but $gA \notin H/A$. As $A \subset H$, it follows that $K \subset H$ and $[G : H]$ is finite but $g \notin H$, so $g \notin \bigcap_{H \in \mathcal{S}} H$.

Suppose instead that $g \in KA$. Then $g = ka$ for some $k \in K$ and $a \in A$ but $g \notin K$ so $a \notin K \cap A$. By the abelian case, there exists a subgroup B of A such that $K \cap A \subset B$, $[A : B]$ is finite but $a \notin B$. As A/B is a finite group, we have $(A/B)^m = \{1\}$ for some $m \in \mathbb{Z}_{\geq 1}$ and thus $A^m \subset B$. Also A^m is characteristic in A, so $A^m \lhd G$. If $g \in KA^m$ then $ka = g = k_2 a_2$ for some $k_2 \in K$ and $a_2 \in A^m$ so the element $k_2^{-1} k = a_2 a^{-1}$ is in $K \cap A \subset B$, and $a_2 \in A^m \subset B$ giving $a \in B$, which is a contradiction. Thus g is not in KA^m, so gA^m is not an element of the subgroup KA^m/A^m of G/A^m. Moreover, A^m is infinite so $h(G/A^m) < h(G)$ and by the inductive hypothesis again there exists a subgroup H/A^m of G/A^m such that $KA^m/A^m \subset H/A^m$, $[G/A^m : H/A^m]$ is finite but $gA^m \notin H/A^m$. As $A^m \subset H$, it follows that $K \subset H$ and $[G : H]$ is finite but $g \notin H$, so $g \notin \bigcap_{H \in \mathcal{S}} H$ in this case too. And g was arbitrarily chosen, hence $\bigcap_{H \in \mathcal{S}} H \subset K$. The result follows by mathematical induction. □

Corollary 4.17. *Any virtually polycyclic group is residually finite.*

Proof. Put $K = \{1\}$ in theorem (4.16). □

5 Growth functions of finitely generated groups

In this chapter, we define the (element) growth function of a finitely generated group. We show that virtually nilpotent groups have polynomial growth and prove the converse in the case of soluble groups. The original papers for this material are [Wolf], [Milnor] and [Bass], and the extension of this result to all finitely generated groups came in [Gromov]. Other accounts can be found in [Mann] and [Drutu & Kapovich].

Definition. Let G be a group generated by a finite subset $S \subset G$, and $\tilde{S} = S \cup \{s^{-1} : s \in S\}$. The *growth function* of G with respect to S is the function

$$v : \mathbb{N} \to \mathbb{N}, n \mapsto \#\{s_1 s_2 \dots s_m : m \leq n, s_1, \dots, s_m \in \tilde{S}\}.$$

In terms of the word metric d_S on G, this is equivalent to saying

$$v(n) = \#\{g \in G : d_S(g, 1) \leq n\}$$

or more simply, in terms of closed balls, $v(n) = \#B_n[1]$.

Proposition 5.1. *Let G be a group generated by a finite subset $S \subset G$, and $v : \mathbb{N} \to \mathbb{N}$ the growth function of G with respect to S. Then there exists $a \in \mathbb{R}_{\geq 1}$ such that for all $n \in \mathbb{N}$ the inequality $v(n) \leq a^n$ holds.*

Proof. Let $k = \#S$ and $a = 2k + 1 \geq 1$. For each $n \in \mathbb{N}$ and each $g \in G$ with $d_S(g, 1) \leq n$, the element g can be written in the form $g = s_1 s_2 \dots s_n$ where for each $i \in \{1, 2, \dots, n\}$ there are at most $2k + 1$ possible values for s_i (since $s_i \in S$ or $s_i^{-1} \in S$ or $s_i = 1$) hence $v(n) \leq (2k + 1)^n = a^n$. □

Thus the growth function of any finitely generated group is bounded above by an exponential function.

Definition. Let G be a group generated by a finite subset $S \subset G$, and $v : \mathbb{N} \to \mathbb{N}$ the growth function of G with respect to S. We say G has *polynomial growth* with respect to S if there exist $a \in \mathbb{R}_{>0}$ and $d \in \mathbb{N}$ such that for all $n \in \mathbb{N}_{\geq 1}$ the inequality $v(n) \leq an^d$ holds. We say G has *exponential growth* with respect to S if there exists $b \in \mathbb{R}_{>1}$ such that for all $n \in \mathbb{N}$ the inequality $v(n) \geq b^n$ holds.

Lemma 5.2. *Let G be a group generated by a finite subset $S \subset G$, and v the growth function of G with respect to S. Then G has exponential growth with respect to S if and only if there exist $c \in \mathbb{R}_{>0}$, $d \in \mathbb{R}_{>1}$ and $N \in \mathbb{N}$ such that for all $n \geq N$ we have $v(n) \geq cd^n$.*

Proof. This is elementary analysis. □

Proposition 5.3. *Let G_1 and G_2 be groups generated by finite subsets $S_1 \subset G_1$ and $S_2 \subset G_2$ such that the metric spaces (G_1, d_{S_1}) and (G_2, d_{S_2}) are quasi-isometric. Then G_1 has polynomial growth with respect to S_1 if and only if G_2 has polynomial growth with respect to S_2, and similarly for exponential growth.*

Proof. Let $v_1 : \mathbb{N} \to \mathbb{N}$ be the growth function of G_1 w.r.t. S_1 and $v_2 : \mathbb{N} \to \mathbb{N}$ the growth function of G_2 w.r.t. S_2. For any $m, n \in \mathbb{N}$ it follows from the definition that $v_1(m+n) \leq v_1(m)v_1(n)$ and similarly for v_2. There exists a quasi-isometry $f : G_1 \to G_2$ so there exist $a \in \mathbb{R}_{\geq 1}$ and $b \in \mathbb{R}_{\geq 0}$ such that for all $g, h \in G_1$,

$$\frac{1}{a} d_{S_1}(g,h) - b \leq d_{S_2}(f(g), f(h)) \leq a d_{S_1}(g,h) + b.$$

Without loss of generality, we may assume $f(1) = 1$ (since G_2 acts on itself by isometries).

We first convert the above into an inequality between the growth functions. For any $g, h \in G_1$, if $d_{S_1}(g,h) > ab$ then $d_{S_2}(f(g), f(h)) > 0$ so $f(g) \neq f(h)$. And for any $h \in G_1$, putting $m_1 = \lfloor ab \rfloor$ gives

$$\#\{g \in G_1 : d_{S_1}(g,h) \leq ab\} = v_1(m_1).$$

Thus for all $g \in G_2$, we have $\#f^{-1}(\{g\}) \leq v_1(m_1)$. Also for any $n \in \mathbb{N}$ and any $g \in G_1$, if $d_{S_1}(g,1) \leq n$ then $d_{S_2}(f(g),1) \leq an + b \leq kn + m_2$ where $k = \lceil a \rceil$ and $m_2 = \lceil b \rceil$. It follows that for all $n \in \mathbb{N}$,

$$\frac{v_1(n)}{v_1(m_1)} \leq v_2(kn + m_2) \leq v_2(m_2)v_2(kn). \tag{10}$$

If G_2 has polynomial growth w.r.t. S_2 then there exist $c \in \mathbb{R}_{>0}$ and $d \in \mathbb{N}$ such that for all $n \in \mathbb{N}_{\geq 1}$ we have $v_2(n) \leq cn^d$. So for all $n \in \mathbb{N}_{\geq 1}$,

$$v_1(n) \leq v_1(m_1)v_2(m_2)v_2(kn) \leq (v_1(m_1)v_2(m_2)ck^d)n^d$$

and thus G_1 has polynomial growth w.r.t. S_1 too. The converse also holds by symmetry.

If G_1 has exponential growth w.r.t. S_1 then there exists $r \in \mathbb{R}_{>1}$ such that for all $n \in \mathbb{N}$ we have $v_1(n) \geq r^n$. Let $s = \frac{1}{rv_1(m_1)v_2(m_2)} > 0$ and $t = \sqrt[k]{r} > 1$. Then for all $n \in \mathbb{N}$,

$$\begin{aligned} v_2(n) &\geq v_2\left(k\left\lfloor\frac{n}{k}\right\rfloor\right) \geq \frac{v_1\left(\left\lfloor\frac{n}{k}\right\rfloor\right)}{v_1(m_1)v_2(m_2)} \text{ by (10)} \\ &\geq \frac{r^{\lfloor\frac{n}{k}\rfloor}}{v_1(m_1)v_2(m_2)} > \frac{r^{\frac{n}{k}-1}}{v_1(m_1)v_2(m_2)} = \frac{(\sqrt[k]{r})^n}{rv_1(m_1)v_2(m_2)} = st^n \end{aligned}$$

Hence by (5.2), G_2 has exponential growth w.r.t. S_2 as well. The converse also holds, again by symmetry. □

Thus polynomial growth is a quasi-isometric invariant of finitely generated groups, as is exponential growth.

Corollary 5.4. *Let G be a finitely generated group, and $S_1, S_2 \subset G$ finite subsets, each generating G. Then G has polynomial growth w.r.t. S_1 if and only if it has polynomial growth w.r.t. S_2, and similarly for exponential growth.*

Proof. By (3.6), the identity map is a quasi-isometry from the metric space (G, d_{S_1}) to (G, d_{S_2}), and the result follows by (5.3). □

Definition. Let G be a group generated by a finite subset $S \subset G$. We say that G has *polynomial growth* if it has polynomial growth with respect to S, and we say G has *exponential growth* if it has exponential growth with respect to S.

These are well-defined by (5.4).

Corollary 5.5. *Let G be a finitely generated group and H a subgroup of finite index in G. Then G has polynomial growth if and only if H has polynomial growth, and similarly for exponential growth.*

Proof. The subgroup H is finitely generated and quasi-isometric with G by the Schwarz-Milnor theorem (3.8). The result then follows immediately by (5.3). □

Lemma 5.6. *Let G be a finitely generated group that has polynomial growth. Then G does not have exponential growth.*

Proof. This is elementary analysis. □

It is worth noting that there exist finitely generated groups which have neither polynomial nor exponential growth: see [Grigorchuk & Pak].

Lemma 5.7. *Let G be a finitely generated nilpotent group and $N \lhd G$ with G/N infinite cyclic, generated by a coset aN. Let $T \subset N$ be a finite subset generating N (G is polycyclic (4.8) so N is finitely generated (4.6)), and let $S = T \cup \{a\}$ (which generates G). Let $r \in \mathbb{N}_{\geq 1}$ be the nilpotency class of G (note G is infinite so non-trivial). Then there exists $c > 0$ with the following property: for all $n \in \mathbb{N}_{\geq 1}$ and all $g \in G$ with $d_S(1, g) \leq n$ there exist $m \in \mathbb{Z}$ with $|m| \leq n$ and $h \in N$ with $d_T(1, h) \leq cn^r$ such that $g = ha^m$.*

Proof. Let $\tilde{S} = S \cup \{s^{-1} : s \in S\}$, $A_1 = \tilde{S}$ and for each $i \in \mathbb{N}_{\geq 1}$ let

$$A_{i+1} = \{[a^j, x] : j \in \{1, -1\}, x \in A_i\}.$$

Note for all $i \in \mathbb{N}_{\geq 1}$ that A_i is finite and for $i \geq 2$ that $A_i \subset N$ since G/N is abelian. Let

$$A = A_1 \sqcup A_2 \sqcup A_3 \sqcup \dots .$$

For any $t \in \mathbb{N}$ and any t-tuple $(x_1, x_2, \ldots, x_t)$ of elements of A, if there is no index i such that x_i is an element of $\{a, a^{-1}\} \subset A_1$ but x_{i+1} is not, then we call the tuple a *final string*. Otherwise, we define the *next string* to be the $(t+1)$-tuple

$$(x_1, x_2, \ldots, x_{i-1}, [x_i, x_{i+1}], x_{i+1}, x_i, x_{i+2}, x_{i+3}, \ldots, x_t)$$

where i is the greatest index such that $\{a, a^{-1}\}$ contains x_i but not x_{i+1}. Note that $x_{i+1} \in A_j \subset A$ for some $j \in \mathbb{N}_{\geq 1}$ and therefore that $[x_i, x_{i+1}]$ is an element of $A_{j+1} \subset A$.

Let $c = r \cdot \max\{d_T(1, x) : x \in \bigcup_{i=1}^{r} A_i\} > 0$. Take any $n \in \mathbb{N}_{\geq 1}$ and any $g \in G$ with $d_S(1, g) \leq n$. Then $g = x_1 x_2 \ldots x_t$ for some $t \in \mathbb{N}$ with $t \leq n$ and some $x_1, x_2, \ldots, x_t \in \tilde{S} = A_1$. Let s_1 be the tuple $(x_1, x_2, \ldots, x_t)$ of elements of $A_1 \subset A$. If s_1 is not a final string, let s_2 be its next string, and if s_2 is not a final string, let s_3 be its next string, etc. Continuing in this way, we eventually obtain a final string s_k for some $k \in \mathbb{N}_{\geq 1}$ since there are at most n occurrences of a or a^{-1} in s_1, and with each step the rightmost occurrence that is followed by some other element is moved past it, nearer to the end of the string. Moreover, for each $i \in \{1, 2, \ldots, k\}$ the product of entries in s_k is equal to g, by the insertion of the commutators. So s_k gives an expression for g of the form $g = ha^m$ where $m \in \mathbb{Z}$ with $|m| \leq n$ and $h \in N$. It remains to show that $d_T(1, h) \leq cn^r$. In moving the rightmost occurrence of a or a^{-1} (if any) to the end of the string, we inserted no more than n elements of $A_2 \subset A$. In moving the next rightmost occurrence to the penultimate position, we inserted at most n elements of $A_2 \subset A$ and at most n elements of $A_3 \subset A$. Continuing in this way, we see that for each $i \in \{0, 1, \ldots, n\}$, the expression for h contains at most $\binom{n}{i} n$ terms which are elements of $A_{i+1} \subset A$.

Let $G = \Gamma_1 \supset \Gamma_2 \supset \ldots$ be the lower central series of G. Then for each $i \in \mathbb{N}_{\geq 1}$ we have $A_i \subset \Gamma_i$. But $\Gamma_{r+1} = \{1\}$ so for all $i > r$, $A_i = \{1\}$. Thus the number of terms in the expression for h (not counting those equal to 1) is at most

$$\begin{aligned} & n + \binom{n}{1} n + \binom{n}{2} n + \ldots + \binom{n}{r-1} n \\ \leq \quad & n + n^2 + n^3 + \ldots + n^r \leq rn^r \end{aligned}$$

and each is an element of A_i for some $i \in \{1, 2, \ldots, r\}$, so by definition of c we have $d_T(1, h) \leq cn^r$. □

Theorem 5.8. *Let G be a finitely generated group. If G is virtually nilpotent then G has polynomial growth.*

Proof. Suppose G is virtually nilpotent. Then G is virtually polycyclic by (3.9) and (4.8). Induction on the Hirsch length $h(G)$.
Case $h(G) = 0$: the group G is finite so G has polynomial growth.

Case $h(G) > 0$: the group G has a nilpotent subgroup H with $[G : H]$ finite, and H is finitely generated too (3.9) so it is polycyclic (4.8). As H is infinite, it has a cyclic series of length at least 1

$$H \supset K \supset \ldots \supset \{1\}.$$

Moreover, we can choose H so that H/K is infinite cyclic, generated by aK for some $a \in H$. The subgroup K is finitely generated as well (4.6). Let T be a finite set generating K, and $S = T \cup \{a\}$ which generates H. We shall show that H has polynomial growth with respect to S. Let $r \in \mathbb{N}_{\geq 1}$ be the nilpotency class of H (the subgroup H is infinite and hence non-trivial). By lemma (5.7), there exists $c > 0$ with the property that for all $n \in \mathbb{N}_{\geq 1}$ and all $h \in H$ with $d_S(1, h) \leq n$ there exist $m \in \mathbb{Z}$ with $|m| \leq n$ and $k \in K$ with $d_T(1, k) \leq cn^r$ such that $h = ka^m$. Let $v_S : \mathbb{N} \to \mathbb{N}$ be the growth function of H w.r.t. S and $v_T : \mathbb{N} \to \mathbb{N}$ the growth function of K w.r.t. T. It follows that for all $n \in \mathbb{N}_{\geq 1}$, $v_S(n)$ is at most $(2n + 1)v_T(cn^r)$. And $h(K) < h(H)$ by (4.14) so by the induction hypothesis K has polynomial growth w.r.t. T: there exist $b \in \mathbb{R}_{>0}$ and $d \in \mathbb{N}$ such that for all $n \in \mathbb{N}_{\geq 1}$ we have $v_T(n) \leq bn^d$. Hence for all $n \in \mathbb{N}_{\geq 1}$,

$$v_S(n) \leq 3nb(cn^r)^d = (3bc^d)n^{rd+1}$$

and H has polynomial growth w.r.t. S. As $[G : H]$ is finite, G also has polynomial growth by (5.5). The result follows by mathematical induction.

□

We now turn to the question of extracting information about a group's structure from its growth type. The key is the following simple observation.

Lemma 5.9. *Let G be a finitely generated group, σ an inner automorphism of G and $a \in G$. If G does not have exponential growth then there exist $n \in \mathbb{N}_{\geq 1}$ and $i_1, \ldots, i_n, j_1, \ldots, j_n \in \{0, 1\}$ with $i_n \neq j_n$ such that*

$$\sigma(a)^{i_1}\sigma^2(a)^{i_2} \ldots \sigma^n(a)^{i_n} = \sigma(a)^{j_1}\sigma^2(a)^{j_2} \ldots \sigma^n(a)^{j_n}.$$

Proof. Choose $b \in G$ such that for all $x \in G$, $\sigma(x) = bxb^{-1}$. Suppose for all $n \in \mathbb{N}_{\geq 1}$ the 2^n elements $ba^{i_1}ba^{i_2} \ldots ba^{i_n}$ of G are distinct, where $i_1, i_2, \ldots, i_n \in \{0, 1\}$. Then, choosing a finite subset $S \subset G$ such that $b, ba \in S$ and $\langle S\rangle = G$, the growth function v of G with respect to S satisfies $v(n) \geq 2^n$ for all $n \in \mathbb{N}$ so G has exponential growth. Thus if G does not have exponential growth then there exist $n \in \mathbb{N}_{\geq 1}$ and $i = (i_1, \ldots, i_n)$, $j = (j_1, \ldots, j_n)$ both in $\{0, 1\}^n$ such that $i \neq j$ but $ba^{i_1}ba^{i_2} \ldots ba^{i_n}$ and $ba^{j_1}ba^{j_2} \ldots ba^{j_n}$ are equal. Taking n as small as possible, we have $i_n \neq j_n$. Now

$$\begin{aligned} ba^{i_1}ba^{i_2} \ldots ba^{i_n} &= ba^{i_1}b^{-1}b^2a^{i_2}b^{-2}b^3a^{i_3}b^{-3} \ldots b^na^{i_n}b^{-n}b^n \\ &= \sigma(a^{i_1})\sigma^2(a^{i_2})\sigma^3(a^{i_3}) \ldots \sigma^n(a^{i_n})b^n \\ &= \sigma(a)^{i_1}\sigma^2(a)^{i_2}\sigma^3(a)^{i_3} \ldots \sigma^n(a)^{i_n}b^n \end{aligned}$$

and similarly for $ba^{j_1}ba^{j_2}\dots ba^{j_n}$, hence

$$\sigma(a)^{i_1}\sigma^2(a)^{i_2}\dots\sigma^n(a)^{i_n} = \sigma(a)^{j_1}\sigma^2(a)^{j_2}\dots\sigma^n(a)^{j_n}.$$

□

Lemma 5.10. *Let G be a finitely generated group, and $N \lhd G$ with G/N cyclic. If G does not have exponential growth then N is also finitely generated.*

Proof. If $[G : N]$ is finite then N is finitely generated by (3.9). Suppose G/N is infinite cyclic: $G/N = \langle aN\rangle$ for some $a \in G$ with no non-zero power of a in N. Then $G = \langle a\rangle N$ and G is finitely generated so $G = \langle\{a, n_1, n_2, \dots, n_r\}\rangle$ for some $r \in \mathbb{N}$ and $n_1, n_2, \dots, n_r \in N$. Let $\mathcal{S} = \{H \lhd G : n_1, n_2, \dots, n_r \in H\}$ and define

$$M = \bigcap_{H\in\mathcal{S}} H \text{ and } K = \langle\{a^j n_i a^{-j} : j \in \mathbb{Z}, 1 \le i \le r\}\rangle.$$

We shall show that $N = M = K$. The subgroup M is normal in G and $M \subset N$ (since $N \in \mathcal{S}$). Also

$$G/M = \langle\{aM, n_1M, n_2M, \dots, n_rM\}\rangle = \langle aM\rangle$$

since $n_1, n_2 \dots, n_r \in M$. So G/M is infinite cyclic (no non-zero power of a lies in M as $M \subset N$). But $(G/M)/(N/M) \cong G/N$ is an infinite cyclic factor group of G/M, so $M = N$. Now for any $x \in K$, axa^{-1} and $a^{-1}xa$ are in K too by definition, and the other generators $n_1, n_2, \dots, n_r$ of G are in K, so $K \lhd G$ and thus $K \in \mathcal{S}$. Hence $N = M$ is a subset of K, but $K \subset N$ too (since $N \lhd G$) and so $K = N$.

Take any $i \in \{1, 2, \dots, r\}$ and let $K_i = \langle\{a^j n_i a^{-j} : j \in \mathbb{Z}\}\rangle$. Suppose G does not have exponential growth. We claim that K_i is finitely generated. Let $n = n_i$ and $\sigma : G \to G$ be the map $x \mapsto axa^{-1}$. By lemma (5.9), there exist $s \in \mathbb{N}_{\ge 1}$ and $i_1, \dots, i_s, j_1 \dots, j_s \in \{0, 1\}$ with $i_s \neq j_s$ such that

$$\sigma(n)^{i_1}\sigma^2(n)^{i_2}\dots\sigma^s(n)^{i_s} = \sigma(n)^{j_1}\sigma^2(n)^{j_2}\dots\sigma^s(n)^{j_s}.$$

Let $A = \{\sigma(n), \sigma^2(n), \dots, \sigma^{s-1}(n)\}$. As $\{i_s, j_s\} = \{0, 1\}$, $\sigma^s(n)$ is an element of $\langle A\rangle$ and, by a simple induction, it follows that $\sigma^j(n) \in \langle A\rangle$ for all $j \ge 1$. Similarly (with σ^{-1} instead of σ) there exists $t \ge 1$ such that for all $j \ge 1$ we have $\sigma^{-j}(n) \in \langle B\rangle$ where $B = \{\sigma^{-1}(n), \sigma^{-2}(n), \dots, \sigma^{-(t-1)}(n)\}$. Thus $K_i = \langle A \cup \{n\} \cup B\rangle$, proving the claim. It follows that

$$N = K = \langle K_1 \cup K_2 \cup \dots \cup K_r\rangle$$

which is finitely generated. □

Proposition 5.11. *Let G be a finitely generated group that does not have exponential growth, and $N \lhd G$ such that G/N is abelian. Then N is also finitely generated.*

Proof. If $N = G$ then the result is immediate. Otherwise the quotient group G/N is isomorphic with a direct product of 1 or more cyclic subgroups (since G/N is a finitely generated abelian group), each having the form H/N where H is a subgroup of G with $N \subset H$. By a simple induction, it follows that each such H is also finitely generated. Moreover, H does not have exponential growth and H/N is cyclic, hence N is finitely generated by (5.10). □

Corollary 5.12 (Milnor[7]). *Let G be a finitely generated soluble group. If G does not have exponential growth, then G is polycyclic.*

Proof. If G does not have exponential growth then, by a simple induction using lemma (5.11), all the terms of its derived series are finitely generated. And G is soluble, so G is polycyclic by (4.3). □

Lemma 5.13. *Let G be a finitely generated group and $n \in \mathbb{N}_{\geq 1}$. Then there exist only finitely many subgroups H of G with $[G : H] = n$.*

Proof. Take any subgroup H of G with index n in G. The group G acts on the set G/H of left cosets of H in G by left multiplication ($g : aH \mapsto gaH$), and $S(G/H) \cong S_n$ so there exists a homomorphism $\phi : G \to S_n$ with $\ker(\phi) \subset H$. The group G is generated by a finite subset $A \subset G$, and each homomorphism from G to S_n is completely determined by the images of the elements of A in S_n. As A and S_n are finite, it follows that there are only finitely many homomorphisms from G to S_n. Also for any homomorphism $\psi : G \to S_n$, there are only finitely many subgroups H of G with $\ker(\psi) \subset H$ since $G/\ker(\psi)$ is finite. Hence there are only finitely many candidates for H. □

The following proposition is a variant of (4.1).

Proposition 5.14. *Let G be a finitely generated group and H a subgroup of finite index in G. Then there exists a subgroup $N \subset H$ such that N is characteristic in G and N has finite index in G.*

Proof. Let $n = [G : H] \in \mathbb{N}_{\geq 1}$ and $\mathcal{S}$ be the set of all subgroups of G with index n in G. As G is finitely generated, the set $\mathcal{S}$ is finite by (5.13). Let $N = \bigcap_{K \in \mathcal{S}} K$. Then N is a characteristic subgroup of G and $N \subset H$ (since $H \in \mathcal{S}$). Write $\mathcal{S} = \{K_1, K_2, \dots, K_m\}$ and for each $j \in \{1, 2, \dots, m\}$ define $H_j = \bigcap_{i=1}^{j} K_i$. Then for each $j \in \{2, 3, \dots, m\}$, $[H_{j-1} : H_j] \leq [G : K_j]$ so

$$\begin{aligned}[G : N] &= [G : H_1][H_1 : H_2] \dots [H_{m-1} : H_m] \\ &\leq [G : K_1][G : K_2] \dots [G : K_m] = n^m\end{aligned}$$

[7] see [Milnor]

making $[G:N]$ finite. □

Lemma 5.15 (Kronecker[8]). *Let $\alpha \in \mathbb{C}$ be an non-zero algebraic integer such that α and all its conjugates over $\mathbb{Q}$ have modulus at most 1. Then α is a root of unity.*

Proof. There exists a monic irreducible polynomial $f \in \mathbb{Z}[X]$ with $f(\alpha) = 0$, and f is also irreducible over $\mathbb{Q}$ so it is the minimum polynomial of α over $\mathbb{Q}$. Also $f = (X-\alpha_1)(X-\alpha_2)\dots(X-\alpha_n)$ where $\alpha_1 = \alpha, \alpha_2, \dots, \alpha_n \in \mathbb{C}$ are the conjugates of α over $\mathbb{Q}$. Let $c = \alpha_1\alpha_2\dots\alpha_n$. By assumption, $|\alpha_j| \leq 1$ for each $j \in \{1, 2, \dots, n\}$, so $|c| \leq 1$. But $(-1)^n c$ is the constant term of f, which is a non-zero integer (since $\alpha \neq 0$) so $|c| \geq 1$ as well, hence $|c| = 1$ and it follows that $|\alpha_j| = 1$ for all $j \in \{1, 2, \dots, n\}$. For each $j \in \mathbb{N}_{\geq 1}$, define

$$f_j = (X - \alpha_1^j)(X - \alpha_2^j)\dots(X - \alpha_n^j).$$

Then for each $j \in \mathbb{N}_{\geq 1}$, each coefficient of f_j can be written as a symmetric expression in $\alpha_1, \dots, \alpha_n$ and thus is an element of $\mathbb{Z}$, making f_j an element of $\mathbb{Z}[X]$. For any $j \in \mathbb{N}_{\geq 1}$, $f = a_0 + a_1X + a_2X^2 + \dots + a_{n-1}X^{n-1} + X^n$ where $a_0, a_1, \dots, a_{n-1} \in \mathbb{Z}$. But for each $k \in \{1, 2, \dots, n\}$, a_{n-k} is the sum of the products taken in all possible ways of k of the values α_1^j, α_2^j,…,α_n^j, and each has modulus 1 so $|a_{n-k}| \leq \binom{n}{k}$. Thus the coefficients of the polynomials $f_1, f_2, f_3, \dots$ are bounded and integral, and hence the polynomials cannot all be different: $f_k = f_m$ for some $k > m$. Thus

$$(X - \alpha_1^k)(X - \alpha_2^k)\dots(X - \alpha_n^k) = (X - \alpha_1^m)(X - \alpha_2^m)\dots(X - \alpha_n^m)$$

and by unique factorization in $\mathbb{C}[X]$ there exists a permutation $\sigma \in S_n$ such that for all $j \in \{1, 2, \dots, n\}$ we have $\alpha_j^k = \alpha_{\sigma(j)}^m$. By a simple induction, for all $s \in \mathbb{N}$ and all $j \in \{1, 2, \dots, n\}$, we get $a_j^{k^s} = \alpha_{\sigma^s(j)}^{m^s}$. But $\sigma^{n!} = \mathrm{id}$ so for each $j \in \{1, 2, \dots, n\}$, $\alpha_j^{k^{n!}} = \alpha_j^{m^{n!}}$ hence $\alpha_j^{(k^{n!} - m^{n!})} = 1$ with $k > m$, making α_j a root of unity. In particular, this holds for $\alpha = \alpha_1$. □

The following theorem was first proved by Wolf (see [Wolf]), but we follow an argument by Bass (see [Bass]).

Theorem 5.16 (Wolf). *Let G be a polycyclic group. Then G has exponential growth or G is virtually nilpotent.*

Proof. Suppose G does not have exponential growth. Induction on the Hirsch length $h(G)$.
Case $h(G) = 0$: G is finite so virtually trivial and thus virtually nilpotent.
Case $h(G) > 0$: the group G has a cyclic series $G \supset H \supset \dots \supset \{1\}$ (with $H \lhd G$) and it suffices to prove that G is virtually nilpotent in the case that

[8] see [Murty & Esmonde]

G/H is infinite cyclic. So $G/H = \langle aH \rangle$ where $a \in G$ with no non-zero power of a in H. Also $h(H) < h(G)$ so by the inductive hypothesis H is virtually nilpotent. Moreover, as H is finitely generated by (4.6) and nilpotence is a hereditary property, there exists a nilpotent subgroup $N \subset H$ which is characteristic in H with $[H : N]$ finite (5.14). Also $H \lhd G$ so $N \lhd G$. Note that $[G : \langle a \rangle N]$ is finite ($G = \langle a \rangle H$ and $[H : N]$ is finite).

Let $\mathcal{S}$ be the set of all central series of N in which every term is normal in $\langle a \rangle N$. Then $\mathcal{S} \neq \varnothing$ (for example, each term of N's lower central series is characteristic in N so normal in $\langle a \rangle N$) and the number of infinite factors in any central series of N is bounded above by $h(N)$ since any such series can be refined to form a cyclic series of N and during this process the number of infinite factors cannot decrease. Moreover, for any series

$$N \supset \ldots \supset A \supset B \supset \ldots \supset \{1\}$$

in $\mathcal{S}$, if a factor A/B is infinite but not torsion-free then, putting T/B for its torsion subgroup, the series

$$N \supset \ldots \supset A \supset T \supset B \supset \ldots \supset \{1\}$$

is also an element of $\mathcal{S}$: any refinement of a central series remains central, and T/B is characteristic in A/B so T is characteristic in A and hence normal in $\langle a \rangle N$. This refinement also has the same number of infinite factors as before, A/B being infinite and torsion-free while T/B is finite (4.5). Conclusion: there exists a central series

$$N = N_0 \supset N_1 \supset \ldots \supset N_t = \{1\}$$

of N which is an element of $\mathcal{S}$, has the maximum number of infinite factors for an element of $\mathcal{S}$, and in which every factor is finite or torsion-free.

Take any $j \in \{1, 2, \ldots, t\}$ and define

$$\sigma : {}^{N_{j-1}}\!/\!_{N_j} \to {}^{N_{j-1}}\!/\!_{N_j}, xN_j \mapsto axa^{-1}N_j$$

which is well-defined since $N_j \lhd \langle a \rangle N$. We claim there exists $k_j \in \mathbb{Z}_{\geq 1}$ such that $\sigma^{k_j} = \mathrm{id}$. If N_{j-1}/N_j is finite then the claim holds immediately (σ permutes a finite number of cosets). Suppose N_{j-1}/N_j is infinite. Then, by choice of the series, N_{j-1}/N_j is torsion-free, and finitely generated so it is a free abelian group of finite rank at least 1. Let M be N_{j-1}/N_j but written additively, for clarity. The maps σ and σ^{-1} are elements of $\mathrm{End}(M)$ so M is a $\mathbb{Z}[\sigma, \sigma^{-1}]$-module. Taking the module of fractions with respect to $S = \mathbb{Z} \setminus \{0\}$, we get the $\mathbb{Q}[\sigma, \sigma^{-1}]$-module $S^{-1}M$, which we now show is simple. Clearly, $S^{-1}M \neq \{0\}$. Take any submodule $F \neq \{0\}$ of $S^{-1}M$. Then (embedding M in $S^{-1}M$ in the obvious way) $F \cap M$ is a submodule of the $\mathbb{Z}[\sigma, \sigma^{-1}]$-module M and $F \cap M \neq \{0\}$ so $F \cap M$ is a non-trivial subgroup of the abelian group M and it is σ-invariant and similarly for σ^{-1}. Thus (in the multiplicative notation) $F \cap M = A/N_j$ where the series

$$N \supset \ldots \supset N_{j-1} \supset A \supset N_j \supset \ldots \supset \{1\}$$

is also in $\mathcal{S}$. But the series without A already has the maximum number of infinite factors for an element of $\mathcal{S}$, and A/N_j is infinite (non-trivial subgroup of a torsion-free group) so N_{j-1}/A is finite. Thus (in the additive notation) $M/(F \cap M)$ is finite, so $S^{-1}M/(F \cap M)$ is a torsion group and therefore $S^{-1}M/F$ is also a torsion group, hence $F = S^{-1}M$ (both being modules over $\mathbb{Q}$). The module $S^{-1}M$ is therefore a simple module over $\mathbb{Q}[\sigma, \sigma^{-1}]$, so every endomorphism of $S^{-1}M$ except the zero map is an automorphism: its kernel is not $S^{-1}M$ so must be $\{0\}$, and its image is not $\{0\}$ so must be $S^{-1}M$. Conclusion: $\mathrm{End}(S^{-1}M)$ (over $\mathbb{Q}[\sigma, \sigma^{-1}]$) is a division ring.

Let $\alpha = S^{-1}\sigma \in \mathrm{End}(S^{-1}M)$, i.e. α is the endomorphism such that for each $\frac{m}{s} \in S^{-1}M$, $\alpha(\frac{m}{s}) = \frac{\sigma(m)}{s}$. The subring $\mathbb{Q}[\alpha] \subset \mathrm{End}(S^{-1}M)$ is commutative and $\mathrm{End}(S^{-1}M)$ is a division ring so $\mathbb{Q}[\alpha]$ is an integral domain. The free abelian group M has finite rank and therefore a finite basis B over $\mathbb{Z}$, and B is also a basis for $S^{-1}M$ over $\mathbb{Q}$. The matrix of α relative to B has all entries not just in $\mathbb{Q}$ but in $\mathbb{Z}$ (since for each $b \in B$, $\alpha(b) = \sigma(b) \in M$) so its characteristic polynomial P_α is a monic element of $\mathbb{Z}[X]$. By the Cayley-Hamilton theorem, $P_\alpha(\alpha) = 0 \in \mathrm{End}(S^{-1}M)$ and, in particular, $\mathbb{Q}[\alpha]$ is finite-dimensional as a vector space over $\mathbb{Q}$. It follows that $\mathbb{Q}[\alpha]$ is a field and hence an algebraic number field, and α is an algebraic integer in $\mathbb{Q}[\alpha]$.

Suppose that α is not a root of unity. We show that this leads to a contradiction. By Kronecker's proposition (5.15), there exists a field homomorphism $\phi : \mathbb{Q}[\alpha] \to \mathbb{C}$ such that $|\phi(\alpha)| > 1$ and hence $|\phi(\alpha^m)| > 2$ for some $m \in \mathbb{Z}_{\geq 1}$. Let $\beta = \alpha^m$ and let x be an element of M with $x \neq 0$. By assumption, G does not have exponential growth, so by lemma (5.9) (but written additively) there exist $r \in \mathbb{N}_{\geq 1}$ and $s_1, s_2 \ldots, s_r \in \{-1, 0, 1\}$ with $s_r \neq 0$ such that

$$s_1\beta(x) + s_2\beta^2(x) + \ldots + s_r\beta^r(x) = 0.$$

Clearly we can choose these so that $s_r = 1$. Let

$$\gamma = s_1\beta + s_2\beta^2 + \ldots + s_{r-1}\beta^{r-1} + \beta^r \in \mathbb{Q}[\alpha].$$

Then $\gamma(x) = 0 = \gamma(0)$ but $x \neq 0$ so γ has no inverse in the field $\mathbb{Q}[\alpha]$ hence $\gamma = 0$. Thus

$$\beta^r = -s_1\beta - s_2\beta^2 - \ldots - s_{r-1}\beta^{r-1}$$

and so

$$\begin{aligned}
|\phi(\beta)|^r &= |\phi(-s_1\beta - s_2\beta^2 - \ldots - s_{r-1}\beta^{r-1})| \\
&\leq |\phi(\beta)| + |\phi(\beta)|^2 + \ldots + |\phi(\beta)|^{r-1} \\
&= \frac{|\phi(\beta)|^r - |\phi(\beta)|}{|\phi(\beta)| - 1} < |\phi(\beta)|^r
\end{aligned}$$

since $|\phi(\beta)| > 2$, which is a contradiction. It follows that α is a root of unity: $\alpha^{k_j} = \mathrm{id}$ for some $k_j \in \mathbb{Z}_{\geq 1}$ so $\sigma^{k_j} = \mathrm{id}$ too, and this proves the claim.

Let $\tilde{N} = \langle a^k \rangle N$, where $k = k_1 k_2 \ldots k_t \in \mathbb{Z}_{\geq 1}$. Then $[\langle a \rangle N : \tilde{N}] = k$ and $[G : \langle a \rangle N]$ is finite so $[G : \tilde{N}]$ is finite. It remains to show that

$$\tilde{N} \supset N = N_0 \supset N_1 \supset \ldots \supset N_t = \{1\}$$

is a central series of $\tilde{N}$. The factor $\tilde{N}/N$ is isomorphic with $\langle a^k \rangle / (\langle a^k \rangle \cap N)$ which is isomorphic with $\langle a^k \rangle$ as $\langle a^k \rangle \cap N = \{1\}$, so $\tilde{N}/N$ is abelian and $[\tilde{N}, \tilde{N}] \subset N$. Take any $j \in \{1, 2, \ldots, t\}$. Then $[N_{j-1}, N] \subset N_j$ (central series of N) and $[N_{j-1}, \langle a^k \rangle] \subset N_j$ since $xN_j \mapsto a^k x a^{-k} N_j$ is the identity map on N_{j-1}/N_j. And N_j is normal in $\langle a \rangle N$, so for any $i \in \mathbb{Z}$, $x \in N_{j-1}$ and any $y \in N$ the elements $[x, a^{ik}]$, $[x, y]$ and $a^{ik}[x, y]a^{-ik}$ are all in N_j, hence

$$[x, a^{ik} y] = [x, a^{ik}] a^{ik} [x, y] a^{-ik} \in N_j$$

and thus $[N_{j-1}, \tilde{N}]$ is a subset of N_j. So $\tilde{N}$ is nilpotent with finite index in G, making G virtually nilpotent. The result follows by mathematical induction. □

Corollary 5.17. *Let G be a finitely generated soluble group. Then the following are equivalent:*

(a) G does not have exponential growth;

(b) G is polycyclic and does not have exponential growth;

(c) G is virtually nilpotent;

(d) G has polynomial growth.

Proof. (a)⇒(b) is Milnor's theorem (5.12), (b)⇒(c) is Wolf's theorem (5.16), (c)⇒(d) is (5.8) and (d)⇒(a) is (5.6). □

A Acknowledgements

I would like to thank my supervisor Lenny Taelman for all his help and encouragement with this thesis. I would also like to thank Hendrik Lenstra and Ronald van Luijk for their feedback and suggestions.

References

[Bass] H. Bass, *The degree of polynomial growth of finitely generated nilpotent groups*, Proceedings, London Mathematical Society (3) 25 (1972) 603-614.

[Bridson & Haefliger] Martin R. Bridson, André Haefliger, *Metric spaces of non-positive curvature*, volume 319 of Grundlehren der mathematischen Wissenschaften, Springer, 1999.

[Drutu & Kapovich] Cornelia Drutu, Michael Kapovich, *Lectures on geometric group theory*, in preparation.

[Grigorchuk & Pak] Rotislav Grigorchuk, Igor Pak, *Groups of intermediate growth: an introduction*, L'Enseignement Mathématique 54 (2008) 251-272.

[Gromov] Mikhael Gromov, *Groups of polynomial growth and expanding maps*, Publications Mathématiques, Institut des Hautes Études Scientifiques 53 (1981) 53-78.

[Lennox & Robinson] John C. Lennox, Derek J. S. Robinson, *The theory of infinite soluble groups*, Oxford University Press, 2004.

[Macdonald] Ian D. Macdonald, *The theory of groups*, Oxford University Press, 1968.

[Mann] Avinoam Mann, *How groups grow*, London Mathematical Society Lecture Note Series 395, Cambridge University Press, 2012.

[Milnor] John Milnor, *Growth of finitely generated solvable groups*, Journal of Differential Geometry 2 (1968) 447-449.

[Murty & Esmonde] M. Ram Murty, Jody Esmonde, *Problems in algebraic number theory*, volume 190 of Graduate Texts in Mathematics, second edition, Springer, 2005.

[Roe] John Roe, *Lectures on coarse geometry*, volume 31 of University lecture series, American Mathematical Society, 2003.

[Rudin] Walter Rudin, *Real and complex analysis*, third edition, McGraw-Hill, 1987.

[Segal] Daniel Segal, *Polycyclic groups*, Cambridge University Press, 1983.

[Wehrfritz] Bertram A.F. Wehrfritz, *Group and ring theoretic properties of polycyclic groups*, volume 10 of Algebra and applications, Springer, 2009.

[Wolf] Joseph A. Wolf, *Growth of finitely generated solvable groups and curvature of Riemannian manifolds*, Journal of Differential Geometry 2 (1968) 421-446.