# Notes on Coxeter groups

Christopher H. Cashen

Version: October 31, 2023

Faculty of Mathematics
University of Vienna
1090 Vienna, Austria
*Email address*: christopher.cashen@univie.ac.at
*URL*: http://www.mat.univie.ac.at/~cashen

# Contents

# Preface

These are notes on Coxeter groups from the viewpoint of Geometric Group Theory, used as the basis for a 1–semester "Topics in Algebra" class for early graduate students. We assume multivariable Calculus, Linear Algebra, and the basics of Group Theory and Topology. We do not assume familiarity with Differential Topology, Riemannian Geometry, or any prior experience with Geometric or Combinatorial Group Theory. We also do not assume much Algebraic Topology: the basics of covering spaces, the fundamental group, and the topology of cell complexes will suffice.

The plan is:

- Chapter 1: Basics of group presentations and how to recognize certain types of groups from their presentations. First examples of Coxeter groups.
- Chapter 2: Survey of geometric reflection groups.
- Chapter 3: Coxeter groups are linear, with some strong consequences. However, linearity does not give a nice geometric action when the group was not already a geometric reflection group.
- Chapter 4: Abstract group properties generalizing properties of geometric reflection groups. See what we can do with those in terms of Combinatorial Group Theory.
- Not being completely satisfied with the combinatorial approach, construct a nice geometric space, the "Davis complex" for abstract reflection groups to act on in Chapter 5. Profit; Get answers to questions such as: What are the solvable subgroups?, When is the group hyperbolic? When is it virtually free? When does it split? When does it contain free subgroups? Surface subgroups?
- Chapter 6: Specialize to right-angled Coxeter groups, focusing on how the interplay between geometry and combinatorics in CAT(0) cube complexes gives us stronger results in that case. Applications to the Quasiisometry Problem for right-angled Coxeter groups.

One might ask, "Why not use Davis's book [**11**]?" I did. You should. Chapter 6 focuses on things outside the scope of Davis's book and some more recent developments. Everything else represents a path that I picked through Davis's book that would cover the geometric reflection groups and the construction of the Davis complex and still leave time in the semester for Chapter 6. This is only a fraction of what is in [**11**]. The topics are presented in a different order than in Davis, and aimed at a less advanced audience. Some of the material has been chewed up and regurgitated in a slightly different form. I added examples and exercises. So [**11**] is an implicit citation for everything before Chapter 6. This will be made explicit for some notable results or those whose proofs we wish to skip. See [**11**] for original citations.

# Introduction

These notes are about Coxeter groups from the point of view of Geometric Group Theory.

What are Coxeter groups? There is a nice class of groups called 'geometric reflection groups' that we will survey in Chapter 2. These groups are naturally defined in terms of a concrete geometric action, but one could also write down an abstract presentation describing such a group. (Group presentations are reviewed in Chapter 1.) One could then deduce rules that such a presentation should obey, and consider the class of 'abstract reflection groups' defined by a presentation satisfying those rules. It turns out that that is the class of Coxeter groups, which we will see in Chapter 4. So Coxeter groups are groups defined by a certain kind of group presentation, generalizing the geometric reflection groups.

What kind of questions would we like to answer about groups?

(1) Isomorphism Problem: Given two finite presentations, do they define isomorphic groups?
   (a) Triviality Problem: Does a given finite presentation define the trivial group?
   (b) Commensurability Problem: Do two given finite presentations define groups with a common finite index subgroup?

(2) Word Problem: Given a finite presentation and a word in the generators, does the word represent the trivial element of the group?

(3) Conjugacy Problem: Given a finite presentation and two words in the generators, do they represent conjugate elements of the group?

(4) Membership Problem: Given a finite presentation, a word $w$ in the generators, and a finite set $\{v_1, \ldots, v_n\}$ of words in the generators, does the group element represented by $w$ belong to the subgroup generated by the elements represented by the words $v_i$?

(5) Subgroup Classification: Given a finite presentation, what kind of subgroups does the group have? What subgroups does it have that are free/Abelian/nilpotent/solvable? What kinds of finite index subgroups does it have?

One, minor, issue is that all of these problems are unsolvable: there is no algorithm that can answer them for arbitrary finite presentations. The first demonstration of the issue were constructions of Novikov and Boone in the 1950's showing that there exists a finitely presented group such that there is no algorithm that can decide, for arbitrary input word in the generators, whether or not the word represents the trivial element in the group. So there are finitely presented groups with unsolvable Word Problem. Not long after, Adian and Rabin showed how to use the Novikov-Boone result to prove unsolvability of the Triviality Problem.

The 'geometric' in Geometric Group Theory means that we will try to use some kind of geometry of the group to answer questions like the above, rather than working directly from a presentation. It turns out that all groups have some geometry to them. A first attempt to demonstrate this for finitely generated groups is the Cayley graph:

DEFINITION 0.0.1. The *Cayley graph* $\mathrm{Cay}(G, S)$ of the group $G$ generated by a finite set $S$ is the geometric realization of a directed labelled graph with one vertex for each element of $G$, and an edge $g \xrightarrow{s} h$ for $s \in S$ if $h = gs$.

By convention, if $s \in S$ has order 2 then we leave the $s$–edges undirected.

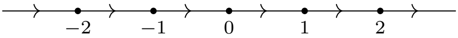
Two different Cayley graphs for $\mathbb{Z}$ are shown in Figure 1.

This means we abstractly take a collection of vertices, add some edges connecting them, and declare those edges to have length equal to 1. The distance between two points is defined to be the length of the shortest path connecting them, regardless of the orientation of edges. The orientation is only used to keep track of labels, with 'going the wrong way' across an edge corresponding to the inverse label. That is, following an edge $g \xrightarrow{s} h$ in the forward direction says that $gs = h$. Following it in the backward direction says $hs^{-1} = g$. The fact that the graph is connected is a consequence of $S$ being a generating set for the group.

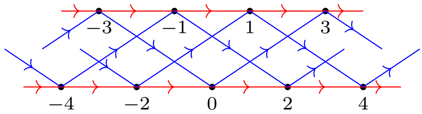
The group $G$ acts on $\mathrm{Cay}(G, S)$ by left multiplication: if $g \xrightarrow{s} h$ is an edge, so $gs = h$, then for any $f \in G$ we have $fhs = fg$, so $fg \xrightarrow{s} fh$ is also an edge. Thus the action is by graph isomorphisms, hence isometries. Notice furthermore that no nontrivial group element fixes a vertex.

We would like to define 'the geometry of $G$' as the geometry of such a Cayley graph, but the graph depends on the choice of generating set, not just the group. To make the geometry depend only the group, we introduce an equivalence relation:

DEFINITION 0.0.2. A map $\phi\colon X \to Y$ between metric spaces is a *quasi-isometry* if there are constants $L$ and $A$ such that:




(A) Cay($\mathbb{Z}$, {1})




(B) Cay($\mathbb{Z}$, {2, 3})

FIGURE 1. Two different Cayley graphs for $\mathbb{Z}$.

- $\frac{1}{L}d_X(x,x')-A \leqslant d_Y(\phi(x),\phi(x')) \leqslant Ld_X(x,x')+A$, for all $x,x' \in X$, and
- for all $y \in Y$ there exists $x \in X$ such that $d_Y(y,\phi(x)) \leqslant A$.

EXERCISE 0.0.3. Show the existence of a quasiisometry is an equivalence relation.

EXERCISE 0.0.4. Any two Cayley graphs of a fixed group are quasiisometric to one another. In fact, if you restrict to vertices the additive constant $A$ can be taken to be 0.

It turns out that the equivalence relation of quasiisometry captures more than just different choices of generating set. Here are two key examples:

- The inclusion map of a finite index subgroup into a finitely generated group is a quasiisometry.
- The fundamental group of a closed Riemannian manifold is quasiisometric to its universal cover.

More generally:

THEOREM 0.0.5 (Fundamental Theorem of Geometric Group Theory). *If $G$ is a finitely generated group acting properly discontinuously and cocompactly by isometries on a proper geodesic metric space $X$ then $G$ and $X$ are quasiisometric.*

Let's quickly unpack the terms:

- $X$ is *proper* means closed balls are compact.
- $X$ is *geodesic* means between any two points there exists a path whose length is equal to the distance between the points.
- $G \curvearrowright X$ *by isometries* means the action is defined by a homomorphism from $G$ into the isometry group of $X$.
- $G \curvearrowright X$ is *cocompact* means the quotient of $X$ by the $G$ action is compact.
- $G \curvearrowright X$ is *properly discontinuous* means for every compact set $K \subset X$ the set $\{g \in G \mid gK \cap K \neq \varnothing\}$ is finite.

The combination of hypotheses has a name:

Definition 0.0.6. $G \curvearrowright X$ is *geometric* if the action is properly discontinuous, cocompact, and by isometries.

The space $X$ on which $G$ acts geometrically is called a *geometric model* for $G$. The theorem says all geometric models for $G$ are quasiisometric to one another.

In practice, not all geometric models of a group are equally good for all purposes, and the theorem tells us we are free to use whichever one is most convenient. For instance, in the Riemannian manifold example it is often much better to work with the universal cover of the manifold then to work with a Cayley graph of the fundamental group, because the manifold has more structure, and there are a lot of tools available.

The geometric reflection groups are defined in terms of their (well understood!) geometric models. When we pass to the abstract reflection groups we lose those natural models. The goal of Chapter 5 is to build, in a systematic way, a nice geometric model space, the Davis complex, for an abstract Coxeter group.

The introduction of the equivalence relation of quasiisometry also adds a new question to our list:

(6) Quasiisometry Problem: Given two finite presentations, do they define quasiisometric groups?

In Chapter 6 we will talk about a special class of Coxeter group with particularly nice geometric models, and look at the beginnings of an approach to the Quasiisometry Problem within this class.

CHAPTER 1

# Groups from presentations

A group is the collection of symmetries of some object. What this means depends on context: if the object is just a set then symmetries could just be permutations, but if it is a topological space then we might ask for homeomorphisms, or if it is a metric space we might ask for isometries.

All symmetry groups obey certain rules, and we can define groups as algebraic structures that follow these rules. Thus, a *group* is a set of *elements* with an associative binary operation (product) such that there is an identity element and such that every element has an inverse. Defining groups in this way as abstract algebraic objects has some benefits: we can give concise descriptions of abstract groups in terms of group presentations, we can define and study families of groups with common properties, and we can focus on algebraic features of the groups, distinct from the complications of understanding the structure of a given object on which the group is acting by symmetries.

There is a serious drawback to a purely abstract understanding of groups: it is really hard. For most reasonable properties $\mathcal{P}$ the problem of taking as input a finite group presentation and returning an answer to whether or not the group defined by the presentation has property $\mathcal{P}$ is not algorithmically solvable. Even the property of being a nontrivial group is not algorithmically decidable.

In this section we review the basics of group presentations and actions. We find a few types of representations from which we can recognize the group on sight. Then we give definitions and first examples of Coxeter groups.

## 1. Free groups and free products

If $S$ is a subset of a group $G$, let $S^{-1} := \{s^{-1} \mid s \in S\}$ be the set of inverses of elements of $S$, and let $S^{\pm} := S \cup S^{-1}$. The subset $S$ is called a *generating set* if every element of the group can be obtained as a product of finitely many elements of $S^{\pm}$.

Definition 1.0.1. Let $S$ be a finite set. The *free group on $S$* is the group $F(S)$ satisfying the following universal property: There is an inclusion $\iota\colon S \hookrightarrow F(S)$ and for any group $G$ and any map $\phi\colon S \to G$ there exists a unique homomorphism $\psi\colon F(S) \to G$ such that $\phi = \psi \circ \iota$.

$$\begin{tikzcd} & F(S) \arrow[dr, "\psi"] & \\ S \arrow[ur, hook, "\iota"] \arrow[rr, "\phi"] & & G \end{tikzcd}$$

We should think of this as saying that the free group $F(S)$ is the simplest possible group containing the elements of $S$ as distinct elements.

We can make an explicit construction. Let $S = \{s_1, \dots, s_n\}$ be a set of $n$ symbols. Formally define $n$ new symbols $S^{-1} := \{s_1^{-1}, \dots, s_n^{-1}\}$. Call elements of $S^{\pm}$ *letters*, and call a finite list of letters a *word*. We also allow the *empty word* consisting of no letters.

A word is called *reduced* if it does not contain a pair of adjacent letters of the form $s_i s_i^{-1}$ or $s_i^{-1} s_i$. A word can be made into a reduced word by successively deleting pairs of consecutive inverse letters until no such pairs remain.

THEOREM 1.0.2. *Let $S = \{s_1, \dots, s_n\}$. The set of reduced words with letters $S^{\pm}$ forms a group with operation='concatenate and reduce', and this group is $F(S)$.*

PROOF. Observe that the reduced word obtained from an arbitrary word by successively deleting consecutive inverse letters does not depend on the choice of order in which the next pair to delete is chosen. This makes 'concatenate and reduce' a well-defined associative operation on the set of reduced words, with the empty word as identity element. The inverse to a word is obtained by reversing the order of the letters and then replacing each by its inverse.

To verify the universal property let us use quotation marks around words, so that $s_i \in S$ is an element of $S$ and '$s_i$' is the corresponding word of length 1. Words of length 1 are reduced, since a word with only one letter does not contain any consecutive pair of inverse letters, and the only words of length 1 are the $\iota$–images of $S^{\pm}$. Every word is a concatenation of its letters, so the length 1 words form a generating set.

Suppose $\phi\colon S \to G$ is a map from $S$ into a group $G$. Define $\iota$ by $s_i \mapsto$ '$s_i$'. We know how to define $\psi$ on the set of length 1 words: if we hope to have $\phi = \psi \circ \iota$ then the only choice is to send '$s_i$' to $\phi(s_i)$. But once $\psi$ is defined on a generating set there is a unique way to complete it to a group homomorphism:

$$\psi(\text{`}s_1^{\epsilon_1} \cdots s_k^{\epsilon_k}\text{'}) = \psi(\prod_{i=1}^{k}(\text{`}s_i\text{'})^{\epsilon_i}) = \prod_{i=1}^{k} \psi(\text{`}s_i\text{'})^{\epsilon_i} = \prod_{i=1}^{k} \phi(s_i)^{\epsilon_i}$$

Finally, observe that there is, up to isomorphism, only one free group on a given set of generators, because if $\iota\colon S \hookrightarrow F(S)$ and $\iota'\colon S \hookrightarrow F'(S)$ both satisfy the universal property, then there are unique maps $\psi\colon F(S) \to F'(S)$ and $\psi'\colon F'(S) \to F(S)$ such that $\iota' = \psi \circ \iota$ and $\iota = \psi' \circ \iota'$. But then for every $s \in S$ we have $\psi' \circ \psi \circ \iota(s) = \psi' \circ \iota'(s) = \iota(s)$, so $\psi' \circ \psi$ is the identity map on $F(S)$. By the same argument $\psi \circ \psi'$ is the identity on $F'(S)$. Thus, $\psi$ and $\psi'$ are inverse isomorphisms between $F(S)$ and $F'(S)$. $\square$

DEFINITION 1.0.3. The *free product* $G_1 * G_2$ of groups $G_1$ and $G_2$ is the group satisfying the following universal property: There are injections $\iota_i\colon G_i \hookrightarrow G_1 * G_2$ whose images generate $G_1 * G_2$ and for any homomorphisms $\phi_i\colon G_i \to Q$ there exists a unique homomorphism $\psi\colon G_1 * G_2 \to Q$ such that $\phi_i = \psi \circ \iota_i$.

$$\begin{tikzcd}
G_1 \arrow[r, hook, "\iota_1"] \arrow[dr, "\phi_i"'] & G_1 * G_2 \arrow[d, "\psi"] & G_2 \arrow[l, hook', "\iota_2"'] \arrow[dl, "\phi_2"] \\
& Q &
\end{tikzcd}$$

LEMMA 1.0.4. *The free product of free groups is a free group:* $F(S_1 \cup S_2) = F(S_1) * F(S_2)$.

PROOF. Consider the following commutative diagram:

$$\begin{tikzcd}
S_1 \arrow[r, hook] \arrow[d, hook] & S_1 \cup S_2 \arrow[d, hook] & S_2 \arrow[l, hook'] \arrow[d, hook] \\
F(S_1) \arrow[r] \arrow[dr, hook] & F(S_1 \cup S_2) & F(S_2) \arrow[l] \arrow[dl, hook'] \\
& F(S_1) * F(S_2) \arrow[u, "\psi"] &
\end{tikzcd}$$

In the diagram, the two upper squares are the universal property for $F(S_1)$ and $F(S_2)$, and the lower story is the universal property for $F(S_1) * F(S_2)$. Now observe that there is a map from $S_1 \cup S_2$ into $F(S_1) * F(S_2)$ by going around the outside of the diagram, to the left for elements of $S_1$ and to the right for elements of $S_2$. The universal property for $F(S_1 \cup S_2)$ gives $\psi'\colon F(S_1 \cup S_2) \to F(S_1) * F(S_2)$. But all of involved maps are the identity on $S_1$ and $S_2$, so $\psi$ and $\psi'$ are inverse isomorphisms. $\square$

## 2. Group presentations

If $S$ is a finite set and $R$ is a set of elements of $F(S)$, then a group presentation $\langle S \mid R \rangle$ describes a group $G = F(S)/\langle\langle R \rangle\rangle$, that is, the group that is the quotient of $F(S)$ by the normal subgroup $\langle\langle R \rangle\rangle$ of $F(S)$ generated by the elements of $R$. We say that $G$ is *generated by* $S$ and has *relators* $R$. Usually the notation will not distinguish between an element of $F(S)$ and the element of $G$ that it represents, but in this section we make the difference explicit with square brackets, so if $w \in F(S)$ then $[w]$ is the element of $G$ corresponding to the coset $w\langle\!\langle R\rangle\!\rangle$ of $F(S)/\langle\!\langle R\rangle\!\rangle$. We add a subscript if necessary to define the scope of the normal closure or equivalence relation, ie $\langle\!\langle R\rangle\!\rangle_{F(S)}$ is the smallest normal subgroup of $F(S)$ containing $R \subset F(S)$, and $[w]_G$ is the element of $G$ corresponding to $w\langle\!\langle R\rangle\!\rangle_{F(S)}$in $F(S)/\langle\!\langle R\rangle\!\rangle_{F(S)}$.

We make the following conventions on presentations. A word raised to an infinite power is to be ignored. This allows us some flexibility to use common notation when some parameter may be infinite. For example, we use $\mathcal{C}_m := \langle x \mid x^m\rangle$ to denote the cyclic group of order $m$. This covers both finite cyclic groups and the integers $\mathcal{C}_\infty = \langle x \mid\rangle = \langle x \mid x^\infty\rangle$.

In general it is hard to determine anything about a group from a finite presentation. One thing that we can do well from a presentation is construct homomorphisms: Let $P := \langle S \mid R\rangle$, let $G$ be a group, and let $\phi\colon S \to G$ be any map. Consider the diagram:

$$\begin{tikzcd}
1 \arrow[r] & \langle\!\langle R\rangle\!\rangle \arrow[r, hook] & F(S) \arrow[r, two heads, "q"] \arrow[dr, "\psi"] & P \arrow[r] \arrow[d, "\psi'"] & 1 \\
 & & S \arrow[u, hook] \arrow[r, "\phi"] & G &
\end{tikzcd}$$

The lower-left triangle is the universal property for $F(S)$. The map $\psi'$ making the upper-right triangle commute exists if and only if $\ker q \subset \ker\psi$, in which case $\psi'(x) := \psi(q^{-1}(x))$. But $\ker q = \langle\!\langle R\rangle\!\rangle$, and $\psi(\prod_{i=1}^k s_i^{\epsilon_i}) = \prod_{i=1}^k \psi(s_i)^{\epsilon_i} = \prod_{i=1}^k \phi(s_i)^{\epsilon_i}$. What this means is:

> *Having chosen, via $\phi$, images in $G$ for each of the generators of $P$, there exists a homomorphism $P \to G$ extending $\phi$ if and only if we can take each relator $r \in R$, evaluate its image letter-by-letter in $G$ using $\phi$, and see that it has trivial image.*

We give two lemmas and an exercise using this construction. The goal in all three is similar: we start with two groups presentations. We combine the resulting groups to make a new group by taking a direct product, free product, or semi-direct product. We would like to give a presentation for the resulting product in terms of the presentations we started with. In all three cases we start with a presentation whose generating set is the union of the generating sets of the two factors, so there is an obvious choice for the map $\phi$. We take as relators the union of the two given sets of relators, plus some more. We verify that $\phi$ has trivial image on each of the new relators, so it extends to a surjective homomorphism from the group $P$ defined by the presentation to the product group we are interested in. To show injectivity of this map the basic trick is the same in all three cases: we know a normal form for elements of the product. We show that the new relators added to define $P$ are precisely the ones necessary to derive a similar normal form for elements in $P$. Then injectivity follows easily.

The utility of these results is in the converse direction: given a group presentation of one of the three forms described below, we can immediately conclude that the group it defines splits as an appropriate combination of groups defined by subpresentations. If those subpresentations are simple enough that we can recognize the groups they define, then we can claim to understand the original group.

LEMMA 2.0.1. *For $i \in \{1,2\}$, let $G_i := \langle S_i \mid R_i \rangle$. Define:*

$$R_3 := \{sts^{-1}t^{-1} \text{ for all } s \in S_1,\ t \in S_2\}$$

*Then $\langle S_1, S_2 \mid R_1, R_2, R_3 \rangle$ is a presentation of $G_1 \times G_2$.*

PROOF. Let $P := \langle S_1, S_2 \mid R_1, R_2, R_3 \rangle$.

For $s_i \in S_1$ and $t_j \in S_2$, define $\phi([s_i]_P) := ([s_i]_{G_1}, 1_{G_2}) \in G_1 \times G_2$ and $\phi([t_j]_P) := (1_{G_1}, [t_j]_{G_2})$. We check that the relations of $P$ are all satisfied: If $r \in R_1$ then $r = \prod_{i=1}^k s_i^{\epsilon_i}$ for $s_1 \in S_1$, so:

$$\phi(r) := \prod_{i=1}^k \phi(s_i)^{\epsilon_i} = \prod_{i=1}^k ([s_i]_{G_1}, 1_{G_2})^{\epsilon_i} = ([\prod_{i=1}^k s_i^{\epsilon_i}]_{G_1}, 1_{G_2}) = ([r]_{G_1}, 1_{G_2})$$

But $r \in R_1$ means $[r]_{G_1} = 1_{G_1}$, so $\phi(r) = (1_{G_1}, 1_{G_2}) = 1_{G_1 \times G_2}$. A similar argument works for $R_2$.

For $sts^{-1}t^{-1} \in R_3$ we have:

$$\begin{aligned} \phi(sts^{-1}t^{-1}) &= ([s]_{G_1}, 1_{G_2})(1_{G_1}, [t]_{G_2})([s]_{G_1}^{-1}, 1_{G_2})(1_{G_1}, [t]_{G_2}^{-1}) \\ &= ([ss^{-1}]_{G_1}, [tt^{-1}]_{G_2}) \\ &= (1_{G_1}, 1_{G_2}) = 1_{G_1 \times G_2} \end{aligned}$$

Since all the relations of $P$ are satisfied, $\phi$ is a homomorphism, and it is surjective since its image contains generating sets of both factors of $G_1 \times G_2$.

Suppose $x \in \ker \phi$. The relations of $R_3$ imply that every element of $P$ can be represented by a word in $F(S_1, S_2)$ such that any letters from $S_1^{\pm}$ come first, and any letters from $S_2^{\pm}$ follow, so there are $w \in F(S_1)$ and $v \in F(S_2)$ such that $x = [wv]_P$. But $x \in \ker \phi$ means:

$$(1_{G_1}, 1_{G_2}) = 1_{G_1 \times G_2} = \phi([wv]_P) = ([w]_{G_1}, [v]_{G_2})$$

This shows $w \in \langle\!\langle R_1 \rangle\!\rangle_{F(S_1)}$ and $v \in \langle\!\langle R_2 \rangle\!\rangle_{F(S_2)}$, so $wv \in \langle\!\langle R_1, R_2, R_3 \rangle\!\rangle_{F(S_1,S_2)}$, and $x = [wv]_P = 1_P$. Thus, $\phi$ is injective. $\square$

EXAMPLE 2.0.2. $\langle a, b \mid aba^{-1}b^{-1} \rangle \cong \langle a \mid \rangle \times \langle b \mid \rangle = \mathcal{C}_\infty \times \mathcal{C}_\infty$.

EXERCISE 2.0.3. For $i \in \{1,2\}$, let $G_i := \langle S_i \mid R_i \rangle$. Prove $\langle S_1, S_2 \mid R_1, R_2 \rangle$ is a presentation of $G_1 * G_2$.

EXAMPLE 2.0.4. $\langle s, t \mid s^2, t^2\rangle \cong \langle s \mid s^2\rangle * \langle t \mid t^2\rangle = \mathcal{C}_2 * \mathcal{C}_2$

A *semi-direct product* $N \rtimes Q$ is a split extension of a group $Q$ by a group $N$, that is, it fits into a split short exact sequence $1 \to N \to N \rtimes Q \xrightarrow{q} Q \to 1$ that splits, in the sense that there is a homomorphism $\sigma\colon Q \to N \rtimes Q$ with $q \circ \sigma = \mathrm{Id}_Q$. Since $N$ is a normal subgroup of $N \rtimes Q$, $\sigma(Q)$ acts on it by conjugation, so there is a homomorphism $\rho\colon Q \to \mathrm{Aut}(N)$. This can be reflected explicitly in the notation as $N \rtimes_\rho Q$.

The easiest example of a semi-direct product is the case that $\rho$ is the trivial map, which results in a direct product: $N \rtimes_1 Q \cong N \times Q$.

When $Q = \mathcal{C}_m = \langle z \mid z^m\rangle$ is cyclic it is usual to replace $\rho$ with the $\rho$–image of a generator, so for $\alpha \in Aut(N)$ of order dividing $m$, there is a semi-direct product $N \rtimes_\alpha \mathcal{C}_m$ for which the map $\rho\colon \mathcal{C}_m \to \mathrm{Aut}(N)$ is $z \mapsto \alpha$.

LEMMA 2.0.5. *Let $G = \langle S \mid R\rangle$, $\mathcal{C}_m = \langle t \mid t^m\rangle$, and let $\alpha \in \mathrm{Aut}(G)$ be of order dividing $m$ (or of arbitrary order if $m = \infty$). Let $\tilde{\alpha}\colon S \to F(S)$ be any map satisfying $[\tilde{\alpha}(s)]_G = \alpha([s]_G)$ for all $s \in S$. Then*

$$\langle S, t \mid R,\, t^m,\, \tilde{\alpha}(s)ts^{-1}t^{-1} \text{ for } s \in S\rangle$$

*is a presentation for $G \rtimes_\alpha \mathcal{C}_m$.*

PROOF. Let $R' := \{R,\, t^m,\, \tilde{\alpha}(s)ts^{-1}t^{-1} \text{ for } s \in S\}$, and let $P := \langle S, t \mid R'\rangle$.

Elements of $G \rtimes_\alpha \mathcal{C}_m$ can be written as pairs $([w]_G, [t^a]_{\mathcal{C}_m})$ for $w \in F(S)$, with group operation:

$$([w]_G, [t^a]_{\mathcal{C}_m}) \cdot ([v]_G, [t^b]_{\mathcal{C}_m}) = ([w]_G \alpha^a([v]_G), [t^{a+b}]_{\mathcal{C}_m})$$

Define $\phi\colon S \cup \{t\} \to G \rtimes_\alpha \mathcal{C}_m$ by $\phi(s) := ([s]_G, 1_{\mathcal{C}m})$ for $s \in S$ and $\phi(t) := (1_G, [t]_{\mathcal{C}_m})$. To see that this extends to a homomorphism we check that the relations are satisfied. For $w \in F(S)$ we have $\phi(w) = ([w]_G, 1_{\mathcal{C}_m})$, which immediately shows that the relators of $R$ are satisfied, and similarly the relator $t^m$ is satisfied, so it only remains to show that $1_{G \rtimes_\alpha \mathcal{C}_m} = \phi(\tilde{\alpha}(s)ts^{-1}t^{-1})$ for all $s \in S$.

$$\begin{aligned}
\phi(\tilde{\alpha}(s)ts^{-1}t^{-1}) &= ([\tilde{\alpha}(s)]_G, 1_{\mathcal{C}_m}) \cdot (1_G, [t]_{\mathcal{C}_m}) \cdot ([s^{-1}]_G, 1_{\mathcal{C}_m}) \cdot (1_G, [t^{-1}]_{\mathcal{C}_m}) \\
&= ([\tilde{\alpha}(s)]_G, 1_{\mathcal{C}_m}) \cdot (\alpha([s]_G)^{-1}, [t]_{\mathcal{C}_m}) \cdot (1_G, [t^{-1}]_{\mathcal{C}_m}) \\
&= ([\tilde{\alpha}(s)]_G \cdot \alpha([s]_G)^{-1}, [t]_{\mathcal{C}_m} \cdot [t^{-1}]_{\mathcal{C}_m}) \\
&= (\alpha([s]_G) \cdot \alpha([s]_G)^{-1}, [t^{1-1}]_{\mathcal{C}_m}) = (1_G, 1_{\mathcal{C}_m}) = 1_{G \rtimes_\alpha \mathcal{C}_m}
\end{aligned}$$

Thus, $\phi\colon P \to G \rtimes_\alpha \mathcal{C}_m$ is a homomorphism. It is surjective, since the generators of each factor are in the image. It remains to show $\phi$ has trivial kernel. To do this we put the elements of $P$ into a normal form.

We claim that every element of $P$ can be written as the equivalence class of a word in $F(S,t)$ with all letters of $S^{\pm}$ coming first, followed by a power of $t$. Assuming this, it is easy to verify injectivity of $\phi$: Suppose $1_{G \rtimes_\alpha \mathcal{C}_m} = \phi([wt^a]_P) = ([w]_G, [t^a]_{\mathcal{C}_m})$ for some $w \in F(S)$. Then both coordinates are trivial, so $w \in \langle\!\langle R \rangle\!\rangle_{F(S)}$ and $t^a \in \langle t^m \rangle$. But then $wt^a \in \langle\!\langle R' \rangle\!\rangle_{F(S,t)}$, and $[wt^a]_P = 1_P$.

Let us prove the claim. The first step is to extend $\tilde{\alpha}$ to all of $F(S)$ in a way that remains compatible with $\alpha$. By the universal property for free groups, there is a unique extension of $\tilde{\alpha}$, and we have commutative diagram:

$$\begin{tikzcd}
F(S) \arrow[dd, "q"'] \arrow[drr, "\tilde{\alpha}"] & & \\
& S \arrow[ul, hook'] \arrow[r, "\tilde{\alpha}"] \arrow[dl] & F(S) \arrow[d, "q"] \\
G \arrow[rr, "\alpha"] & & G
\end{tikzcd}$$

The lower-right square commutes by construction. The square that is the outer boundary of the diagram commutes because $q \circ \tilde{\alpha}$ and $\alpha \circ q$ agree on the generating set $S$ of $F(S)$. Thus, for all $w \in F(S)$ we have $[\tilde{\alpha}(w)]_G = \alpha([w])_G$.

Next, we claim that the relation $\tilde{\alpha}(s) = tst^{-1}$ in $P$ extends to $S^{-1}$, hence to all of $F(S)$, as follows: $\tilde{\alpha}(s^{-1})tst^{-1} = ts^{-1}t^{-1}(\tilde{\alpha}(s)ts^{-1}t^{-1})^{-1}tst^{-1} \in \langle\!\langle R' \rangle\!\rangle_{F(S,t)}$, so $1_P = [\tilde{\alpha}(s^{-1})tst^{-1}]_P$. Then for $\epsilon \in \pm 1$ and $s \in S$ we have:

$$[ts^\epsilon]_P = [\tilde{\alpha}(s^\epsilon)ts^{-\epsilon}t^{-1}]_P \cdot [ts^\epsilon]_P = [\tilde{\alpha}(s^\epsilon)t]_P$$

So we can move positive powers of $t$ to the right letter-by-letter past any word in $F(S)$. If $m < \infty$ we are done; if $m = \infty$ then we also need to account for negative powers of $t$ on the left.

Given $s^\epsilon \in S^{\pm}$, there exists $v \in F(S)$ such that $\alpha([v]_G) = [s^\epsilon]_G$, since $\alpha$ is an automorphism. Equivalently, $s^\epsilon \in \tilde{\alpha}(v)\langle\!\langle R \rangle\!\rangle_{F(S)}$. But $\tilde{\alpha}(v)\langle\!\langle R \rangle\!\rangle_{F(S)} \subset \tilde{\alpha}(v)\langle\!\langle R' \rangle\!\rangle_{F(S,t)}$, so $[s^\epsilon]_P = [\tilde{\alpha}(v)]_P$. Finally:

$$[t^{-1}s^\epsilon]_P = [t^{-1}]_P \cdot [\tilde{\alpha}(v)tv^{-1}t^{-1}]_P^{-1} \cdot [\tilde{\alpha}(v)]_P = [vt^{-1}]_P \qquad \square$$

Definition 2.0.6. For $m \in \{1, \ldots, \infty\}$, the *dihedral group of order* $2m$ is the semi-direct product of the form $\mathcal{D}_m := \mathcal{C}_m \rtimes \mathcal{C}_2$, where the automorphism of $\mathcal{C}_m$ is inversion. By Lemma 2.0.5, $\mathcal{D}_m = \langle r, s \mid r^m, s^2, r^{-1}sr^{-1}s^{-1} \rangle$.

When $m = 1, 2$, inversion on $\mathcal{C}_m$ agrees with identity, so $\mathcal{D}_1 \cong \mathcal{C}_1 \times \mathcal{C}_2 = \mathcal{C}_2$ and $\mathcal{D}_2 \cong \mathcal{C}_2 \times \mathcal{C}_2$ are Abelian.

Definition 2.0.7. A *Tietze transform* is one of the following four operations on a group presentation $\langle S \mid R \rangle$ that results in an isomorphic group:

- Add a redundant relator: $\langle S \mid R\rangle$ becomes $\langle S \mid R \cup \{r\}\rangle$ where $r$ is an element of $\langle\langle R\rangle\rangle_{F(S)}$.
- Remove a redundant relator: $\langle S \mid R \cup \{r\}\rangle$ becomes $\langle S \mid R\rangle$ where $r$ is an element of $\langle\langle R\rangle\rangle_{F(S)}$.
- Add a redundant generator: $\langle S \mid R\rangle$ becomes $\langle S \cup \{s\} \mid R \cup \{sw^{-1}\}\rangle$, where $w \in F(S)$.
- Remove a redundant generator: $\langle S \cup \{s\} \mid R \cup \{sw^{-1}\}\rangle$, where $R \cup \{w\} \subset F(S)$, becomes $\langle S \mid R\rangle$.

EXAMPLE 2.0.8. Consider $\langle s \mid s^2, s^4\rangle$. The relator $s^4$ is redundant, since it can be written as a product of other relators, $s^4 = s^2 \cdot s^2$. Thus, $\langle s \mid s^2, s^4\rangle \cong \langle s \mid s^2\rangle$. The relator $s^2$ is not in $\langle\langle s^4\rangle\rangle$, so it is not a redundant relator, and $\mathcal{C}_2 \cong \langle s \mid s^2, s^4\rangle \not\cong \langle s \mid s^4\rangle \cong \mathcal{C}_4$.

EXERCISE 2.0.9. Given a presentation $\langle S \mid R\rangle$ with $r \in R$, there is a sequence of Tietze transformations that replaces $r$ with any conjugate of $r$ or of $r^{-1}$. If $R$ contains words $wv^{-1}$ and $xvy$ then there is a sequence of Tietze transformations that replaces $xvy$ by $xwy$.

EXAMPLE 2.0.10. Consider $\langle s, t \mid s^2,\ t^2,\ (st)^3\rangle$.

(1) Add a redundant generator $r = st$ to get $\langle s, t, r \mid s^2,\ t^2,\ (st)^3, r(st)^{-1}\rangle$.
(2) By Exercise 2.0.9 we can replace $(st)^3$ by $r^3$.
(3) Add a relator $r^{-1}sr^{-1}s^{-1}$, which is redundant, since:

$$r^{-1}sr^{-1}s^{-1} = t^{-1}s^{-1}(rt^{-1}s^{-1})^{-1}st \cdot t^{-2} \cdot s^{-1}(rt^{-1}s^{-1})^{-1}s$$

This gives: $\langle s, t, r \mid s^2,\ t^2,\ r^3, r(st)^{-1}, r^{-1}sr^{-1}s^{-1}\rangle$.
(4) The previous step makes $t^2$ redundant, so remove it.
(5) By Exercise 2.0.9 we can replace $rt^{-1}s^{-1}$ by $tr^{-1}s$, to get:

$$\langle r, s, t \mid s^2, r^3, r^{-1}sr^{-1}s^{-1}, tr^{-1}s\rangle$$

(6) Finally, we can remove the redundant generator $t$, since we have one relator $t(s^{-1}r)^{-1}$ and $t$ appears in no other relators. This leaves:

$$\langle r, s \mid s^2, r^3, r^{-1}sr^{-1}s^{-1}\rangle$$

By Lemma 2.0.5, we recognize this as a presentation of $\mathcal{D}_3 = \mathcal{C}_3 \rtimes \mathcal{C}_2$.

EXAMPLE 2.0.11. An argument similar to the previous example gives $\mathcal{D}_\infty = \langle r, s \mid s^2, r^{-1}sr^{-1}s^{-1}\rangle = \langle s, t \mid s^2, t^2\rangle$. By Exercise 2.0.3, this is a presentation of $\mathcal{C}_2 * \mathcal{C}_2$, so $\mathcal{D}_\infty \cong \mathcal{C}_2 * \mathcal{C}_2$.

## 3. Group actions and representations

A *representation* of a group is simply a homomorphism into some (hopefully!) better understood group. For instance, if $G = \langle S \mid R\rangle$ and we want to know whether the element of $G$ represented by the word $w \in F(S)$ is nontrivial, it suffices to find a representation $\rho\colon G \to H$ to some other group $H$ such that $\rho(w) \neq 1$.

Even better would be to find an injective representation. Such a representation is called *faithful*.

An *action* of a group $G$ on a structure $X$ is a representation of $G$ into the group of symmetries of $X$, where 'group of symmetries' has to be interpreted to preserve the structure of $X$. For example, if $X$ is just a set then its symmetries are permutations, but if is a group its symmetries are automorphisms, if it is a topological space its symmetries are homeomorphism, if it is a metric space its symmetries are isometries, etc.

If $G$ acts by homeomorphisms on a topological space $X$ we will take a *fundamental domain* for the action to mean a closed set $D$ such that every $G$–orbit meets $D$ and every $G$–orbit meets the interior of $D$ at most once.[1] A fundamental domain $D$ is a *strict fundamental domain* is every orbit meets $D$ exactly once.

If $D$ is a fundamental domain for $G \curvearrowright X$ then the translates of $D$ by the $G$–action cover $X$, and distinct translates of $D$ have disjoint interiors. In the case that $X$ is a surface such a covering is often called a *tessellation* or a *tiling*.

Example 3.0.1. $\mathbb{Z}$ acts on $\mathbb{R}$ by isometries by integer translations. The interval $[0,1]$ is a fundamental domain, but not a strict fundamental domain, since 0 and 1 are in the same $\mathbb{Z}$–orbit.

Example 3.0.2. $\mathcal{D}_\infty = \langle r, s \mid s^2, r^{-1}sr^{-1}s^{-1}\rangle$ admits a faithful isometric action on $\mathbb{R}$ with a strict fundamental domain by defining $\rho(r) := x \mapsto x+2$ and $\rho(s) := x \mapsto -x+1$.

First check that this definition extends to a homomorphism of $\mathcal{D}_\infty$ by checking that the relations are satisfied:

$$\rho(s^2) = \rho(s) \circ \rho(s) : x \mapsto -(-x+1)+1 = x$$
$$\rho(r^{-1}sr^{-1}s^{-1}) = \rho(r)^{-1}\circ\rho(s)\circ\rho(r)^{-1}\circ\rho(s)^{-1} : x \mapsto (-((-x+1)-2)+1)-2 = x$$

So $\rho$ takes each of the relators to the identity map on $\mathbb{R}$.

To see that the map is faithful, use the description of elements of $\mathcal{D}_\infty$ from Lemma 2.0.5: since $s$ has order 2, every element can be written either $r^a$ or $r^a s$ for some $a \in \mathbb{Z}$. But $\rho(r^a)(x) = x + 2a$, which is nontrivial if and

[1]Some authors take 'fundamental domain' to mean what we are calling a strict fundamental domain, and call what we are calling a fundamental domain a 'weak fundamental domain'.

only if $a \neq 0$, and $\rho(r^a s)(x) = -x + 1 + 2a$ fixes only the point $x = a + 1/2$, so it is always nontrivial.

Finally, $[-1/2, 1/2]$ is a strict fundamental domain. Powers of $r$ bring every element of $\mathbb{R}$ into the interval $[-1/2, 3/2]$, and then $s$ exchanges the two sides of that interval, fixing its midpoint. Furthermore, if $|x| \leqslant 1/2$ then a quick computation shows that $|\rho(r^a)(x)| > 1/2$ when $a \neq 0$ and $|\rho(r^a s)(x)| \geqslant 1/2$, with equality only in the cases $\rho(s)(1/2) = 1/2$ and $\rho(r^{-1}a)(-1/2) = -1/2$.

## 4. Definitions and first examples of Coxeter groups

### 4.1. Coxeter presentations.

DEFINITION 4.1.1. A *Coxeter matrix* $M = (m_{st})$ on a finite set $S$ is an $|S| \times |S|$ symmetric matrix with entries in $\bar{\mathbb{N}} := \{1, 2, \ldots\} \cup \{\infty\}$ such that $m_{ss} = 1$ for all $s \in S$ and $m_{st} \geqslant 2$ if $s \neq t$.

DEFINITION 4.1.2. A *Coxeter presentation* is a presentation of the form:

$$\langle S \mid (st)^{m_{st}} \text{ for all } s, t \in S \rangle$$

where $(m_{st})$ is a Coxeter matrix on $S$.

Recall the convention that a word raised to an infinite power in a presentation is to be ignored, so $(st)^\infty$ signifies no relation between $s$ and $t$.

DEFINITION 4.1.3. A *Coxeter group* is a group that admits a Coxeter presentation.

DEFINITION 4.1.4. A *Coxeter system* $(W, S)$ is a Coxeter group $W$ such that there is a Coxeter matrix $M$ on $S$ defining a Coxeter presentation of $W$. The set $S$ is called a *fundamental generating set* of $W$.

A priori we could have a Coxeter system $(W, S)$ such that for some $s, t \in S$ the corresponding relator is $(st)^{mn}$, but $st$ has order $m$ in $W$. In this case, we could perform a pair of Tietze transformations and replace the relator $(st)^{mn}$ by $(st)^m$ and get a new Coxeter presentation for $W$ with the same generating set. It will turn out, see Chapter 3 Proposition 2.0.3, that this does not actually happen: when a Coxeter presentation has relation $(st)^m$ the order of $st$ in $W$ is actually $m$, so choice of fundamental generating set determines the Coxeter presentation.

EXAMPLE 4.1.5. There is only one possible Coxeter matrix on a singleton $S = \{s\}$; it is $M = (1)$, and the corresponding Coxeter presentation is $\langle s \mid s^2 \rangle \cong \mathcal{C}_2$, the cyclic group of order 2.

EXAMPLE 4.1.6. Every Coxeter matrix on $S = \{s, t\}$ has the form $M = \begin{pmatrix} 1 & m \\ m & 1 \end{pmatrix}$ for some $m \geqslant 2$.

In the case that $m = 2$ the corresponding Coxeter presentation is: $\langle s, t \mid s^2,\, t^2,\, (st)^2 \rangle \xrightarrow{\cong} \langle s, t \mid s^2,\, t^2,\, sts^{-1}t^{-1} \rangle = \mathcal{C}_2 \times \mathcal{C}_2$, by Lemma 2.0.1.

In the case that $m = \infty$ the corresponding Coxeter presentation is: $\langle s, t \mid s^2,\, t^2 \rangle = \langle s \mid s^2 \rangle * \langle t \mid t^2 \rangle = \mathcal{C}_2 * \mathcal{C}_2$, by Exercise 2.0.3.

For $m > 2$ the argument of Example 2.0.10 generalizes to give:

$$\langle s, t \mid s^2,\, t^2,\, (st)^m \rangle \xrightarrow{r:=st} \langle r, s \mid s^2,\, r^m,\, r^{-1}sr^{-1}s^{-1} \rangle = \mathcal{D}_m$$

By Lemma 2.0.5 this is a presentation of the dihedral group $\mathcal{D}_m$.

The same transformation in the case $m = \infty$ gives:

$$\begin{aligned} \langle s, t \mid s^2,\, t^2 \rangle &\xrightarrow{r:=st} \langle r, s \mid s^2,\, (sr)^2 \rangle \\ &\longrightarrow \langle r, s \mid s^2,\, s^{-1}rs = r^{-1} \rangle \end{aligned}$$

**4.2. Coxeter graphs.** Notice that a Coxeter matrix is redundant; the $|S|(|S|-1)$–many entries above the diagonal already determine the matrix. There are two different conventions for representing this same information in terms of labelled graphs.

DEFINITION 4.2.1. Let $M$ be a Coxeter matrix on $S$. Define the *Coxeter graph* $\Gamma$ corresponding to $M$ to be the graph with vertex set $S$ and an edge labelled by $m_{st}$ between $s$ and $t$ if $m_{st} > 2$. For brevity, labels with value 3 are usually omitted.

DEFINITION 4.2.2. Let $M$ be a Coxeter matrix on $S$. Define the *presentation graph* $\Upsilon$ corresponding to $M$ to be the graph with vertex set $S$ and an edge between $s$ and $t$ if $m_{st} < \infty$, with the edge labelled by $m_{st}$ when $m_{st} > 3$.

Table 1.1 shows the graphs for the groups of Example 4.1.6.

We will most often use the Coxeter graph, but the presentation graph is more commonly used in the case of right-angled Coxeter groups:

DEFINITION 4.2.3. A *right-angled Coxeter group* (RACG) is a Coxeter group defined by a Coxeter matrix in which all off-diagonal entries are either 2 or $\infty$.

One reason that the graph definition is useful is that it makes it easy to describe a certain class of well-behaved subgroups:

DEFINITION 4.2.4. Let $(W, S)$ be a Coxeter system. The subgroup of $W$ generated by $T \subset S$ is called a *special subgroup*, and denoted $W_T$.

| $M$ | $\Gamma$ | $\Upsilon$ | $W$ |
|---|---|---|---|
| $(1)$ | $\bullet$ | $\bullet$ | $\mathcal{C}_2$ |
| $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ | $\bullet \qquad \bullet$ | $\bullet\!\!\text{———}\!\!\bullet$ | $\mathcal{C}_2 \times \mathcal{C}_2 = \mathcal{D}_2$ |
| $\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$ | $\bullet\!\!\text{———}\!\!\bullet$ | $\bullet\overset{3}{\text{———}}\bullet$ | $\mathcal{D}_3$ |
| $\begin{pmatrix} 1 & m \\ m & 1 \end{pmatrix}$ for $3 < m < \infty$ | $\bullet\overset{m}{\text{———}}\bullet$ | $\bullet\overset{m}{\text{———}}\bullet$ | $\mathcal{D}_m$ |
| $\begin{pmatrix} 1 & \infty \\ \infty & 1 \end{pmatrix}$ | $\bullet\overset{\infty}{\text{———}}\bullet$ | $\bullet \qquad \bullet$ | $\mathcal{C}_\infty \rtimes \mathcal{C}_2 = \mathcal{D}_\infty = \mathcal{C}_2 * \mathcal{C}_2$ |

TABLE 1.1. Coxeter and presentation graphs for 1 and 2–generator Coxeter systems

In the literature, special subgroups are sometimes called *standard subgroups* or *standard parabolic subgroups.* Their conjugates are called *parabolic subgroups.*

A *full subgraph* of a (labelled) graph is a subgraph that contains an (labelled) edge between two vertices if and only if the original graph does. If $(W, S)$ is a Coxeter system corresponding to Coxeter graph $\Gamma$, then $T \subset S$ corresponds to a subset of the vertices of $\Gamma$. This uniquely determines a full subgraph $\Gamma_T$ with vertex set $T$.

PROPOSITION 4.2.5. *Let $\Gamma$ be the Coxeter graph of a Coxeter system $(W, S)$. Let $T \subset S$. Let $\Gamma_T$ be the full subgraph of $\Gamma$ spanned by $T$, and let $(W_{\Gamma_T}, T)$ be the Coxeter system with Coxeter graph $\Gamma_T$. The natural map $W_{\Gamma_T} \to W_T$ defined by extending the identity map on $T$ is a surjection.*

PROOF. The identity map on $T$ hits all the generators of $W_T$, so we just need to check that the relations of $W_{\Gamma'}$ are satisfied in $W_T$. For every $t \in T$ we have $t^2 = 1$ in $W_{\Gamma'}$ and in $W_T$. If $t \neq t' \in T$ and there is an edge labelled $m_{tt'}$ in $\Gamma'$ then there is also an edge labelled $m_{tt'}$ in $\Gamma$, so the relation $(tt')^{m_{tt'}} = 1$ holds in both $W_{\Gamma'}$ and in $W_T$. If $t \neq t' \in T$ and there is no edge between $t$ and $t'$ in $\Gamma'$ then, since $\Gamma'$ is a full subgraph, there is no edge between $t$ and $t'$ in $\Gamma$, so the relation $(tt')^2 = 1$ holds in both $W_{\Gamma'}$ and in $W_T$. $\square$

It will turn out that the map of Proposition 4.2.5 is actually an *isomorphism*, so the special subgroups are themselves Coxeter groups. At the moment we can prove this only for a single vertex:

PROPOSITION 4.2.6. *If $(W, S)$ is a Coxeter system, every element of $S$ has order 2 in $W$.*

PROOF. For each $s \in S$, the Coxeter presentation has a relation $s^2 = 1$, so the order of $s$ is at most 2; we only have to show that $s \neq 1$. Define

$\rho\colon W \twoheadrightarrow \mathcal{C}_2$ by sending each generator to the nontrivial element of $\mathcal{C}_2$. This extends to a homomorphism of $W$ because all of the defining relations have even length. Since $s \in S$ has nontrivial image in $\mathcal{C}_2$, it was nontrivial in $W$. ☐

PROPOSITION 4.2.7. *Let $(W, S)$ be a Coxeter system and let $\Gamma$ and $\Upsilon$ be its Coxeter and presentation graphs, respectively. Then the decomposition of $\Gamma$ and $\Upsilon$ into connected components determine splittings of $W$ as a direct and free product, respectively:*

$$W = \prod_{\text{con. comp. } \Gamma' \subset \Gamma} W_{\Gamma'}$$

$$W = \mathop{\ast}_{\text{con. comp. } \Upsilon' \subset \Upsilon} W_{\Upsilon'}$$

PROOF. Since our groups are finitely generated it suffices to consider the case of two connected components and then induct.

Suppose that $\Gamma = \Gamma_1 \sqcup \Gamma_2$, with $\Gamma_i$ connected. For $i \in \{1, 2\}$, let $S_i$ be the vertices of $\Gamma_i$, and let $R_i$ be the set of relations $(st)^{m_{st}}$ for $s, t \in S_i$. For $s \in S_1$ and $t \in S_2$ there is no edge between $s$ and $t$ in $\Gamma$, so $m_{st} = 2$. Let $R_3$ be the set of relations $(st)^2$ for $s \in S_1$ and $t \in S_2$. The Coxeter presentation for $(W, S)$ is $\langle S_1, S_2 \mid R_1, R_2, R_3\rangle$ and the Coxeter presentation for $W_{\Gamma_i}$ is $\langle S_i \mid R_i\rangle$. Apply Lemma 2.0.1.

In the presentation graph case the argument is similar, but no edge between $s \in S_1$ and $t \in S_2$ means no relation between $s$ and $t$, ie $m_{st} = \infty$. Apply Exercise 2.0.3. ☐

DEFINITION 4.2.8. A Coxeter system $(W, S)$ is *irreducible* if its Coxeter graph is connected.

Conversely, a Coxeter system is reducible if its Coxeter graph is not connected, and Proposition 4.2.7 implies that reducibility implies that $W$ splits as a direct product. Unfortunately, the converse is **not true!**

EXAMPLE 4.2.9. Assume $m \geqslant 3$ is odd. $\mathcal{D}_{2m}$ is the Coxeter group $\bullet\overset{2m}{\text{———}}\bullet$. The graph is connected, so defines an irreducible Coxeter system. The dihedral presentation is $\langle r, s \mid r^{2m}, s^2, srs^{-1} = r^{-1}\rangle$. The subgroup $\langle r\rangle$ has exactly one non-trivial element, $r^m$, that is fixed by the conjugation $s$–action, so the center of $\mathcal{D}_{2m}$ is $Z(\mathcal{D}_{2m}) = \langle r^m\rangle \cong \mathcal{C}_2$. But $\mathcal{D}_{2m}/Z(\mathcal{D}_{2m}) = \langle r, s \mid r^{2m}, s^2, srs^{-1} = r^{-1}, r^m\rangle = \langle r, s \mid r^m, s^2, srs^{-1} = r^{-1}\rangle \cong \mathcal{D}_m$. This gives a short exact sequence:

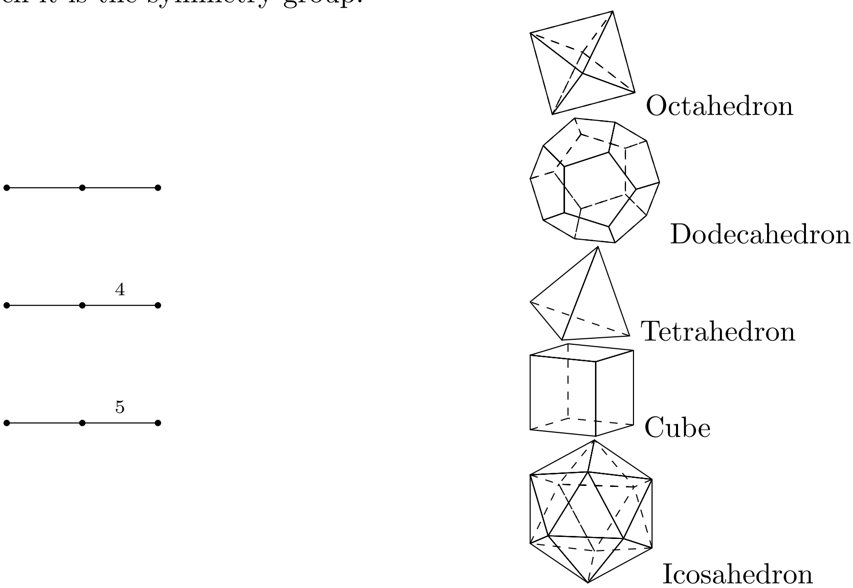
$$1 \to Z(\mathcal{D}_{2m}) \to \mathcal{D}_{2m} \to \mathcal{D}_m \to 1$$

This sequence splits, as follows. Define $\sigma(s) := s$ and $\sigma(r) := r^{m+1}$. Since $m$ is odd, $m(m+1)$ is a multiple of $2m$, so $(\sigma(r))^m = r^{m(m+1)} = 1_{\mathcal{D}_{2m}}$, and

$\sigma\colon \mathcal{D}_m \to \mathcal{D}_{2m}$ is a homomorphism such that $q \circ \sigma$ is the identity on $\mathcal{D}_m$. Thus, $\mathcal{D}_{2m}$ is a semi-direct product $\mathcal{C}_2 \rtimes \mathcal{D}_m$, but $\mathrm{Aut}(\mathcal{C}_2)$ is trivial, so this is actually a direct product: $\mathcal{D}_{2m} \cong \mathcal{C}_2 \times \mathcal{D}_m$. That is a Coxeter group defined by the Coxeter graph $\bullet \qquad \bullet\overset{m}{\rule[0.5ex]{3em}{0.4pt}}\bullet$.

This example also shows that if $(W, S)$ and $(W', S')$ are Coxeter systems with Coxeter graphs $\Gamma$ and $\Gamma'$, respectively, then $W \cong W'$ **does not imply** that $\Gamma$ and $\Gamma'$ are isomorphic as labelled graphs.

Exercise 4.2.10. Match the Coxeter group with the Platonic solid for which it is the symmetry group:




Exercise 4.2.11. Symmetric groups are Coxeter groups. Find a Coxeter presentation for the symmetric group on the set of $n$ things. Remark: Guessing a Coxeter presentation and an identification of the generators with permutations such that all of the relators of the presentation are satisfied is the easy part. You also have to show this map is *injective*. These are finite groups, so you can do this by counting.

CHAPTER 2

# Geometric reflection groups

Roughly, a geometric reflection group is going to mean a group generated by reflections in the sides of a convex polytope in $n$–dimensional spherical, Euclidean, or hyperbolic space, such that the original polytope is a strict fundamental domain for the action. We will make these terms precise, but the situation is simpler in dimensions 1 and 2, so we explore those first.

## 1. 1–dimensional geometric reflection groups

Hyperbolic and Euclidean space coincide in dimension 1, so there are only two spaces to consider, the unit circle $\mathbb{S}^1$ and the real line $\mathbb{R}^1$. In both cases a 'convex polytope' is just a closed interval, so the 1–dimensional geometric reflection groups are those that can be generated by reflections in the endpoints of a closed interval in either $\mathbb{S}^1$ or $\mathbb{R}^1$, such that the original closed interval is a strict fundamental domain for the resulting group action. Notice furthermore that $\mathbb{S}^1$ and $\mathbb{R}^1$ can be oriented, and reflection reverses orientation, so the images of the fundamental domain come in two flavors: those whose induced orientations match the ambient one, and those that do not.

First consider $\mathbb{R}^1$. All non-singleton compact intervals are equivalent under affine transformation, so all reflection groups defined on $\mathbb{R}^1$ are conjugate in the 1–dimensional affine group; that is, up to isomorphism there is only one reflection group on $\mathbb{R}^1$. We can define the interval to be $I = [-1/2, 1/2]$, and the generating reflections to be $x \mapsto -x - 1$ and $x \mapsto -x + 1$.

We claim the group generated by these two maps is $\mathcal{D}_\infty$. Define a map $\rho$ from $\mathcal{D}_\infty = \langle r, s \mid s^2,\, srs = r^{-1}\rangle \overset{r=st}{\cong} \langle s, t \mid s^2,\, t^2\rangle$ to this reflection group by $\rho(s) := x \mapsto -x + 1$ and $\rho(t) := x \mapsto -x - 1$. This is the faithful representation considered in Example 3.0.2.

Now consider $\mathbb{S}^1$. A reflection group action covers $\mathbb{S}^1$ by an even number of copies of a non-singleton closed interval $I$, alternating between those for which the induced orientation under the group action matches the ambient orientation, and those for which the opposite is true. Thus, the length of $I$ must be $\pi/m$ for some $m \in \mathbb{N}$. The case $m = 1$ is degenerate, since then the reflections through the two endpoints of the interval coincide, and the reflection group is $\mathcal{C}_2$, which is better illustrated as the reflection group of

$\mathbb{S}^0$, so we will assume $m \geqslant 2$. All closed intervals of the same length are equivalent under the isometry group of the circle, so the reflection group is determined by the choice of $m$.

Recall that the isometry group of the circle is the orthogonal group $\mathrm{O}(2)$, which we can realize as a matrix subgroup of $\mathrm{Isom}(\mathbb{E}^2)$ by fixing a basepoint and Cartesian coordinates, and thinking of $\mathbb{S}^1$ as the unit circle.

The orthogonal group splits as a semi-direct product of the rotations by the order two group generated by a reflection. In matrix form:

$$\text{(clockwise) rotation by angle } \theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

$$\text{reflection through the } x\text{–axis} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Reflection through the line through the origin with angle $\theta$ can then be expressed as a conjugate: rotate the line by $-\theta$ to move it to the $x$–axis, reflect through the $x$–axis, and then rotate by $\theta$ to get back to the original line:

(1)
$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \cos-\theta & -\sin-\theta \\ \sin-\theta & \cos-\theta \end{pmatrix} = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$$

Consider the Coxeter presentation $\langle s,\, t \mid s^2,\, t^2,\, (st)^m\rangle \cong \mathcal{D}_m$ for $2 \leqslant m < \infty$. Define $\rho(s) := \begin{pmatrix} \cos\pi/m & \sin\pi/m \\ \sin\pi/m & -\cos\pi/m \end{pmatrix}$, which is reflection through the line through the origin at angle $\frac{\pi}{2m}$. Define $\rho(t) := \begin{pmatrix} \cos\pi/m & -\sin\pi/m \\ -\sin\pi/m & -\cos\pi/m \end{pmatrix}$, which is reflection through the line through the origin at angle $-\frac{\pi}{2m}$.

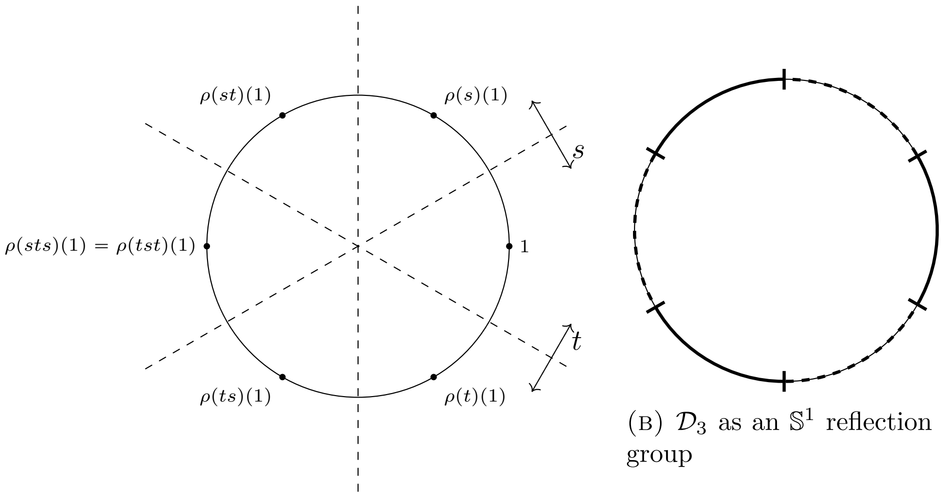
Both $\rho(s)$ and $\rho(t)$ have order 2, since they are reflections. The product $\rho(s)\rho(t) = \begin{pmatrix} \cos^2\frac{\pi}{m} - \sin^2\frac{\pi}{m} & -2\sin\frac{\pi}{m}\cos\frac{\pi}{m} \\ 2\sin\frac{\pi}{m}\cos\frac{\pi}{m} & \cos^2\frac{\pi}{m} - \sin^2\frac{\pi}{m} \end{pmatrix} = \begin{pmatrix} \cos\frac{2\pi}{m} & -\sin\frac{2\pi}{m} \\ \sin\frac{2\pi}{m} & \cos\frac{2\pi}{m} \end{pmatrix}$ is rotation through angle $2\pi/m$, which has order $m$. This shows that $\rho\colon \langle s,\, t \mid s^2,\, t^2,\, (st)^m\rangle \to \mathrm{O}(2)$ is a homomorphism, so we have defined an action of $\mathcal{D}_m$ on $\mathbb{S}^1$.

LEMMA 1.0.1. *The action $\mathcal{D}_m \curvearrowright \mathbb{S}^1$ described above is faithful.*

PROOF. $\mathcal{D}_m$ has order $2m$, so it is enough to demonstrate a point in $\mathbb{S}^1$ whose orbit under the action contains $2m$ distinct points. This is easier in complex coordinates. Translated to 1 complex coordinate, we have $\rho(s) = z \mapsto e^{i\pi/m}\cdot\bar{z}$ and $\rho(t) = z \mapsto e^{-i\pi/m}\cdot\bar{z}$. Extending inductively:

$$\begin{aligned}
\rho((st)^n) &= z \mapsto e^{i2n\pi/m} \cdot z \\
\rho((st)^n s) &= z \mapsto e^{i(2n+1)\pi/m} \cdot \bar{z} \\
\rho((ts)^n) &= z \mapsto e^{-i2n\pi/m} \cdot z \\
\rho((ts)^n t) &= z \mapsto e^{-i(2n+1)\pi/m} \cdot \bar{z}
\end{aligned}$$

The orbit of 1 contains the $2m$ many $2m$–th roots of 1. Figure 1a shows the case $m = 3$. □




(A) $\mathcal{D}_3$ acting on $\mathbb{S}^1 \subset \mathbb{E}^2$

(B) $\mathcal{D}_3$ as an $\mathbb{S}^1$ reflection group

FIGURE 1. $\mathcal{D}_3$ as a reflection group.

We have shown:

THEOREM 1.0.2. *One-dimensional geometric reflection groups are in bijection with rank 2 Coxeter systems, which are exactly the dihedral groups.*

## 2. 2–dimensional geometric reflection groups

In two dimensions we define a geometric reflection group to be a group acting on either the 2–sphere, the Euclidean plane, or the hyperbolic plane, $\mathbb{S}^2$, $\mathbb{E}^2$, or $\mathbb{H}^2$, generated by reflections in the sides of a convex polygon, such that the polygon is a strict fundamental domain for the action.

We will describe these 2–dimensional geometric reflection groups, and show that for each such group there is a Coxeter group that naturally surjects onto it, Theorem 2.6.1. The surjection is actually an isomorphism, but the proof of injectivity uses tools that will be developed in greater generality for higher dimensional cases in the following sections.

Many of these 2–dimensional geometric reflection groups are rank 3 Coxeter groups, but the analogue of Theorem 1.0.2 is not true: some 2–dimensional geometric reflection groups are Coxeter groups of rank greater than 3, and there are rank 3 Coxeter groups that are not geometric reflection groups, 2–dimensional or otherwise.

DEFINITION 2.0.1. A *Coxeter polygon* is a convex polygon in $\mathbb{S}^2$, $\mathbb{E}^2$, or $\mathbb{H}^2$ that is the fundamental domain of a geometric reflection group.

We will always assume we have described a polygon with the minimal possible number of vertices, so all dihedral angles are strictly less than $\pi$. Reflections in adjacent sides of a polygon fix the common vertex, so if the polygon is a fundamental domain for the action, the composition of adjacent reflections must tile out a neighborhood of a vertex with an integral number of copies of the corresponding corner of the polygon. Furthermore, reflections reverse the orientation of the polygon, so there actually have to be evenly many copies of the polygon incident to each vertex. This implies that the dihedral angle at each vertex of a Coxeter polygon must be an even integral fraction of $2\pi$; in other words, a proper integral submultiple of $\pi$.

The curvature $\kappa$ of the three model geometric spaces is:

$$\kappa(\mathbb{X}^2) = \begin{cases} +1 & \text{if } \mathbb{X}^2 = \mathbb{S}^2 \\ \phantom{+}0 & \text{if } \mathbb{X}^2 = \mathbb{E}^2 \\ -1 & \text{if } \mathbb{X}^2 = \mathbb{H}^2 \end{cases}$$

The Gauss-Bonnet theorem connects the curvature of a compact surface to its topology. If $P$ is a polygon with area $A(P)$, dihedral angles $\theta_i = \pi/m_i$, and Euler characteristic $\chi(P) = 1$, then:

$$A(P)\kappa(\mathbb{X}^2) + \sum_{i=1}^{n}(\pi - \pi/m_i) = 2\pi\chi(P) = 2\pi \tag{2}$$

If $P$ is a triangle then (2) implies $\sum_{i=1}^{3}\theta_i = \pi + A(P)\kappa(\mathbb{X}^2)$, so the sum of the dihedral angles of a geodesic triangle is strictly greater than $\pi$ in $\mathbb{S}^2$, equal to $\pi$ in $\mathbb{E}^2$, and strictly less than $\pi$ in $\mathbb{H}^2$.

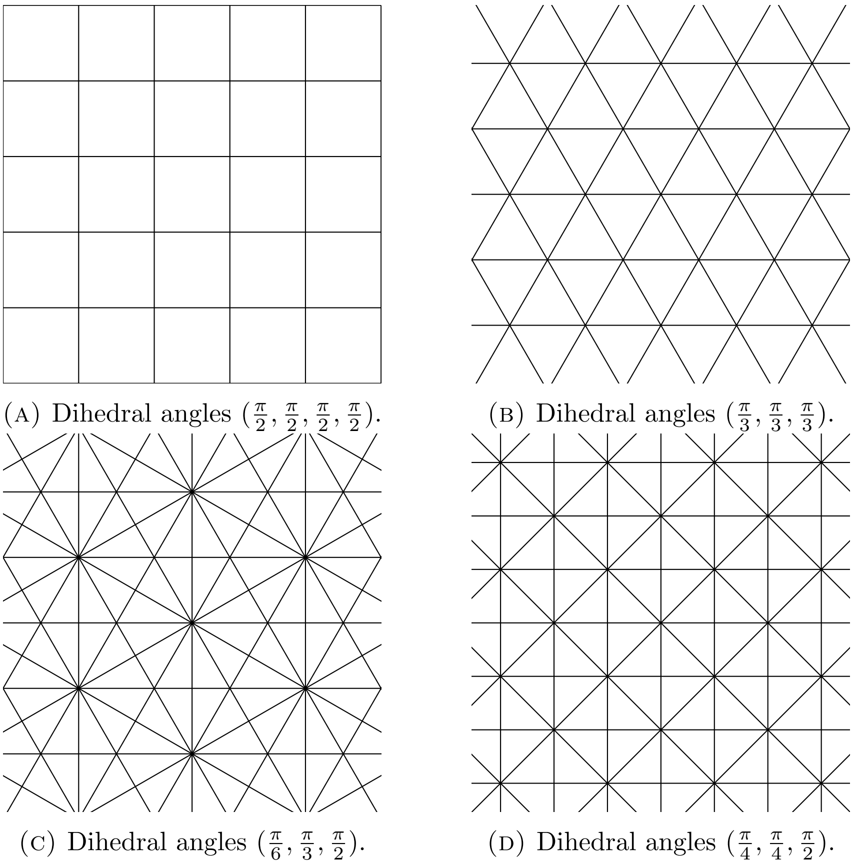
DEFINITION 2.0.2. A *triangle group* is a Coxeter group $\Delta(p, q, r)$ defined by Coxeter matrix:

$$M = \begin{pmatrix} 1 & p & r \\ p & 1 & q \\ r & q & 1 \end{pmatrix}$$

'Triangle group' is just another name for a rank 3 Coxeter system. We will see, in direct connection with (2), that triangle groups behave differently according to whether $S = \frac{1}{p} + \frac{1}{q} + \frac{1}{r}$ is less than 1, equal to 1, or greater than 1.

**2.1. Geometric reflection groups on the Euclidean plane.** In the case of $\mathbb{E}^2$ (2) gives $(n-2)\pi = \sum_{i=1}^n \pi/m_i$. Since $m_i \geqslant 2$, $(n/2)\pi \geqslant \sum_{i=1}^n \pi/m_i = (n-2)\pi$, so $n \leqslant 4$. Thus, the only possible Euclidean Coxeter polygons are that cases that $P$ is a 4–gon with all dihedral angles $\pi/2$, ie, $P$ is a rectangle, or $P$ is a triangle and the sum of the dihedral angles is $\pi$. There are only three combinations of three dihedral angles that are integer submultiples of $\pi$ and sum to $\pi$; they are $(\pi/3, \pi/3, \pi/3)$, $(\pi/6, \pi/3, \pi/2)$, and $(\pi/4, \pi/4, \pi/2)$.

All rectangles, and all triangles with three given angles, are equivalent in the affine group of $\mathbb{E}^2$, so up to isomorphism there are only four $\mathbb{E}^2$ reflection groups. They are shown in Figure 2.



(A) Dihedral angles $(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2})$.

(B) Dihedral angles $(\frac{\pi}{3}, \frac{\pi}{3}, \frac{\pi}{3})$.

(C) Dihedral angles $(\frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2})$.

(D) Dihedral angles $(\frac{\pi}{4}, \frac{\pi}{4}, \frac{\pi}{2})$.

FIGURE 2. $\mathbb{E}^2$ reflection groups

We claim that these are all Coxeter groups. The first is $\mathcal{D}_\infty \times \mathcal{D}_\infty$, and the others are the irreducible triangle groups $\Delta(p, q, r)$ such that $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$. It is easy to verify that the obvious map from the Coxeter group to the $\mathbb{E}^2$

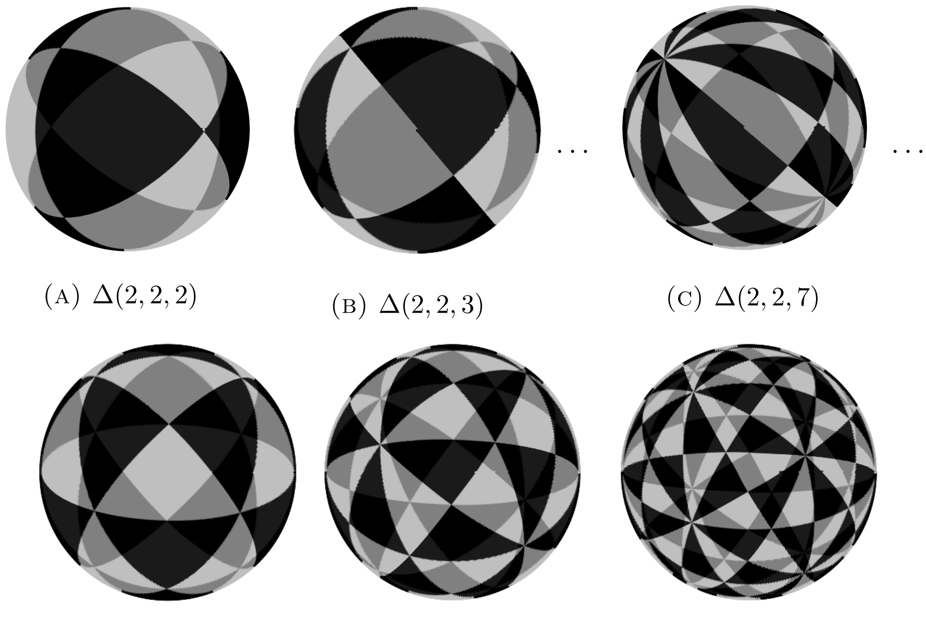
reflection group is a surjection. The fact that it is an injection will follow later from the general theory.

**2.2. Geometric reflection groups on the 2–sphere.** In the case of $\mathbb{S}^2$ (2) gives:

$$(n/2)\pi \geqslant \sum_{i=1}^{n} \pi/m_i = A(P) + (n-2)\pi > (n-2)\pi$$

The only solution is $n = 3$ with $\sum_{i=1}^{3} \pi/m_i = A(P) + \pi > \pi$, so the only possible spherical Coxeter polygons are triangles. Ignoring for a moment the area term, there are only a few possibilities to satisfy $\sum_{i=1}^{3} \pi/m_i > \pi$ with integers $m_i \geqslant 2$. There is an infinite family of solutions $(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{m})$ for $m \geqslant 2$, plus three more: $(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{3})$, $(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{4})$, and $(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{5})$. There are, in fact, spherical triangles with these triples of dihedral angles. For the infinite family, consider a globe with the equator marked out, and $m$–many evenly spaced great circles through the poles. Figure 3 shows the cases $m = 2, 3, 7$ of the infinite family as well as the three exceptional examples.

Recall from Exercise 4.2.10 that the three exceptional cases are also the symmetry groups of the Platonic solids. Can you see the relationship between the Platonic solid and the corresponding sphere in Figure 3?



(A) $\Delta(2,2,2)$ (B) $\Delta(2,2,3)$ (C) $\Delta(2,2,7)$

(D) $\Delta(2,3,3)$ (E) $\Delta(2,3,4)$ (F) $\Delta(2,3,5)$

FIGURE 3. $\mathbb{S}^2$ reflection groups.

Again, we claim that all of these spherical reflection groups are Coxeter groups: they are precisely the triangle groups $\Delta(p,q,r)$ with $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} > 1$.

As in the Euclidean case, it is trivial to check that there is a surjection from the triangle group to the corresponding spherical reflection group, and injectivity will follow later.

**2.3. Prelude to the hyperbolic case: The Poincaré Disc model of $\mathbb{H}^2$.** We are going to describe a way of metrizing a submanifold $M$ of $\mathbb{R}^n$. We will not give a formal definition of *submanifold*, as we will only encounter two simple cases:

(1) $M$ is an open subset of $\mathbb{R}^n$.
(2) $M = f^{-1}(r)$ is the level–$r$ set (or a connected component of a level set) of a smooth function $f\colon \mathbb{R}^n \to \mathbb{R}$ at a regular value $r$.

At every point $\mathbf{x} \in \mathbb{R}^n$ there is an $\mathbb{R}^n$ worth of directions that can be the initial velocity vector of a curve starting at $\mathbf{x}$. The space of such directions is called the *tangent space at* $\mathbf{x}$, and is denoted $T_{\mathbf{x}}\mathbb{R}^n$. We can naturally think of this space as having coordinates inherited from those of $\mathbb{R}^n$. If $M$ is some submanifold of $\mathbb{R}^n$ and $\mathbf{x} \in M$ then we can restrict to curves in $M$, and the space of initial velocity vectors to such curves forms a subspace $T_{\mathbf{x}}M < T_{\mathbf{x}}\mathbb{R}^n$. When $M$ is an open subset of $\mathbb{R}^n$, $T_{\mathbf{x}}M = T_{\mathbf{x}}\mathbb{R}^n$, and when $M$ is a level set $T_{\mathbf{x}}M$ is a subspace of dimension $n-1$.

The goal now is to metrize (a connected component of) $M$ as a length space. This should mean that the distance between $\mathbf{x}, \mathbf{y} \in M$ is the infinum of lengths of paths in $M$ from $\mathbf{x}$ to $\mathbf{y}$. The usual way of defining the length of a path $\gamma\colon I \to M$ in Calculus is the formula:

$$|\gamma| := \int_I |\gamma'(t)| dt$$

This implicitly assumes that we measure the length $|\gamma'(t)|$ of the tangent vector $\gamma'(t) \in T_{\gamma(t)}M$ to $\gamma$ at time $t$ using the standard Euclidean metric.

EXAMPLE 2.3.1. Euclidean space $\mathbb{E}^n$ is $M = \mathbb{R}^n$, which is an open subset of $\mathbb{R}^n$, where the tangent space at each point is endowed with the standard Euclidean inner product.

EXAMPLE 2.3.2. The $n$–sphere $\mathbb{S}^n$ is the level–1 set in $\mathbb{R}^{n+1}$ of the function $\mathbf{v} \mapsto |\mathbf{v}|$, where the tangent space at each point is endowed with the restriction of the standard Euclidean inner product. That is, $\mathbb{S}^n$ is the unit sphere in $\mathbb{R}^{n+1}$, where the distance between two points is the length of the shortest path between them that remains in the sphere.

**Taking the standard metric on every tangent space is not the only possible choice.**

By varying the choice of inner product on tangent spaces to points of $M$ we vary the way that lengths of, and angles between, tangent vectors are defined, resulting in different geometric structures that all have the same underlying topological manifold $M$.

Definition 2.3.3. The *Poincaré Disc model of the hyperbolic plane* takes $\mathbb{H}^2$ to be the open unit disk in $\mathbb{R}^2$ where the inner product on the tangent plane $T_{\mathbf{x}}\mathbb{H}^2$ at a point $\mathbf{x}$ is defined by[1]:

$$\langle \mathbf{v}, \mathbf{w} \rangle_{T_{\mathbf{x}}\mathbb{H}^2} := \frac{4\langle \mathbf{v}, \mathbf{w} \rangle_{\mathbb{E}^2}}{(1 - |\mathbf{x}|^2)^2}$$

This says that the usual Euclidean inner product on $T_{\mathbf{x}}\mathbb{H}^2$ is rescaled by a factor that depends on $\mathbf{x}$. Scaling the inner product does not change the angles between tangent vectors, but it changes their lengths:

$$|\mathbf{v}|_{T_{\mathbf{x}}\mathbb{H}^2} = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle_{T_{\mathbf{x}}\mathbb{H}^2}} = \sqrt{\frac{4\langle \mathbf{v}, \mathbf{v} \rangle_{\mathbb{E}^2}}{(1 - |\mathbf{x}|^2)^2}} = \sqrt{\frac{4|\mathbf{v}|^2_{\mathbb{E}^2}}{(1 - |\mathbf{x}|^2)^2}} = \frac{2|\mathbf{v}|_{\mathbb{E}^2}}{1 - |\mathbf{x}|^2} \tag{3}$$

Observe:

- In a small enough neighborhood of any point, the metric is very close to being a rescaled version of the Euclidean metric.
- Near the origin the metric is very close to being the Euclidean metric.
- As we move towards the boundary of the ball in $\mathbb{R}^2$, tangent vectors get scaled by arbitrarily large factors.

Example 2.3.4. Let for $a < 1$, let $\gamma\colon [0, a] \to \mathbb{H}^2 : t \mapsto (t, 0)$. We have $\gamma'(t) = (1, 0)$ for all $t$. In Euclidean terms $\gamma$ has constant unit speed.

$$\begin{aligned}
|\gamma|_{\mathbb{H}^2} &= \int_{[0,a]} |\gamma'(t)|_{T_{\gamma(t)}\mathbb{H}^2}\, dt \\
&= \int_{[0,a]} \frac{2|\gamma'(t)|_{\mathbb{E}^2}}{1 - |\gamma(t)|^2}\, dt \\
&= 2\int_{[0,a]} \frac{1}{1 - t^2}\, dt \\
&= \log \frac{1+t}{1-t}\Big|_0^a \\
&= \log \frac{1+a}{1-a}
\end{aligned}$$

Notice as $a$ gets close to 1, this curve becomes arbitrarily long.

[1] One might wonder why the factor of 4 is there. It is a normalization factor. Any positive number in that spot would yield a negatively curved plane, but there is a definition of 'curvature' for which the 4 gives this plane curvature equal to $-1$. This is analogous to the fact that any sphere gives a positively curved 2–dimensional space, but it is the unit sphere that gives a space of curvature equal to $+1$. Another explanation is that there are other natural models of hyperbolic space, and 4 is the correct factor to make this model equivalent to the others. See Theorem 3.1.7.

We cannot determine the distance from $\mathbf{0}$ to $(a, 0)$ from this example alone. We have only computed the length of one curve. Potentially there could be some shorter curve with the same endpoints. It turns out that that is not the case.

PROPOSITION 2.3.5. *Projection to any line through* $\mathbf{0}$ *in* $\mathbb{H}^2$ *is length non-increasing.*

PROOF. Let $\pi$ be orthogonal projection to a line $L$. Choose a unit vector in the direction of $L$ and extend it to an orthonormal basis of $\mathbb{R}^2$. Let $\gamma$ be any curve. The tangent vector to $\pi \circ \gamma$ at $t$ in the new coordinates is the same as $\gamma'(t)$, but with the second coordinate set to 0, so its Euclidean length does not increase. Furthermore, $\pi \circ \gamma(t)$ is no farther from the origin than $\gamma(t)$, so $|\pi \circ \gamma|_{\mathbb{H}^2} \leqslant |\gamma|_{\mathbb{H}^2}$. $\square$

EXAMPLE 2.3.6. We have seen a curve $\gamma$ from $\mathbf{0}$ to $(a, 0)$ of hyperbolic length $\log \frac{1+a}{1-a}$. For any curve $\delta$ from $\mathbf{0}$ to $(a, 0)$, the projection $\pi$ of $\delta$ to the first coordinate axis includes all the points of $\gamma([0, a])$, so $|\delta|_{\mathbb{H}^2} \geqslant |\pi \circ \delta|_{\mathbb{H}^2} \geqslant |\gamma|_{\mathbb{H}^2}$. Thus, $d_{\mathbb{H}^2}(\mathbf{0}, (a, 0)) = \inf_\delta |\delta|_{\mathbb{H}^2} = \log \frac{1+a}{1-a}$.

In particular, $d_{\mathbb{H}^2}(\mathbf{0}, (a, 0)) \xrightarrow{a \to 1} \infty$.

DEFINITION 2.3.7. A *geodesic* is a path whose length realizes the distance between its endpoints. A space if a *geodesic metric space* if every pair of points can be connected by a geodesic. A subspace is *convex* if every geodesic with endpoints in the subspace remains in the subspace.

LEMMA 2.3.8. *Up to reparameterization, the path* $t \mapsto (t, 0)$ *is the unique geodesic from* $\mathbf{0}$ *to* $(t, 0)$ *in* $\mathbb{H}^2$.

PROOF. Suppose $\gamma$ is a rectifiable path from $\mathbf{0}$ to $(t, 0)$. If $\gamma$ contains any points off the first coordinate axis then projection shortens $\gamma$, so $\gamma$ is not a geodesic. If $\gamma$ contains points on the first coordinate axis that are not between $\mathbf{0}$ and $(t, 0)$ then $\gamma$ has backtracking, so it is not a geodesic. $\square$

2.3.1. *The group of isometries of* $\mathbb{H}^2$. Next, we describe the group of isometries of $\mathbb{H}^2$, denoted $\mathrm{Isom}(\mathbb{H}^2)$. The structure we have defined, a space together with a vector space of tangent vectors at each point, and an inner product on each tangent space, is an example of a *Riemannian manifold*. The Myers–Steenrod theorem says that an isometry between connected Riemannian manifolds is continuously differentiable (in fact, smooth). Thus, if $\phi$ is an isometry of $M$ then at each point $\mathbf{x} \in M$ there is linear bijection $T_\mathbf{x}\phi \colon T_\mathbf{x}M \to T_{\phi(\mathbf{x})}M$ that is the derivative of $\phi$ at $\mathbf{x}$. Furthermore, the determinant of the derivative of $\phi$ is either positive at every point of $M$ or negative at every point of $M$. The isometries whose derivatives have positive determinant at every point are called *orientation preserving*, and they form a subgroup $\mathrm{Isom}^+(M)$ of index 2 in $\mathrm{Isom}(M)$.

LEMMA 2.3.9. *The stabilizer* $\mathrm{Stab}_{\mathrm{Isom}(\mathbb{H}^2)}(\mathbf{0})$ *of* $\mathbf{0}$ *in* $\mathrm{Isom}(\mathbb{H}^2)$ *is* $\mathrm{O}(2)$, *the 2–dimensional orthogonal group.*

PROOF. $\mathrm{O}(2)$ fixes the origin, preserves the Euclidean inner product, and preserves distance from the origin, so it preserves the hyperbolic distance. Thus, $\mathrm{O}(2) < \mathrm{Stab}_{\mathrm{Isom}(\mathbb{H}^2)}(\mathbf{0})$, and we need to show the opposite inclusion.

Since $\mathrm{O}(2)$ acts by isometries fixing $\mathbf{0}$ and we know one radial geodesic, every radial line starting at $\mathbf{0}$ is geodesic. Suppose that $\phi$ is a hyperbolic isometry fixing $\mathbf{0}$. Let $\mathbf{x} \neq \mathbf{0}$ be an arbitrary point of $\mathbb{H}^2$, and define:

$$\gamma_{\mathbf{x}} \colon [0, d_{\mathbb{E}^2}(\mathbf{0}, \mathbf{x})] \to \mathbb{H}^2 : t \mapsto (1-t)\mathbf{0} + t\mathbf{x}$$

This is the unique Euclidean-unit-speed radial hyperbolic geodesic from $\mathbf{0}$ to $\mathbf{x}$. $\phi(\gamma_{\mathbf{x}}) = \gamma_{\phi(\mathbf{x})}$ is therefore the Euclidean-unit-speed radial hyperbolic geodesic from $\mathbf{0}$ to $\phi(\mathbf{x})$, and $T_{\mathbf{0}}\phi\colon T_{\mathbf{0}}\mathbb{H}^2 \to T_{\mathbf{0}}\mathbb{H}^2$ takes the initial velocity vector of $\gamma_{\mathbf{x}}$ to the initial velocity vector of $\gamma_{\phi(\mathbf{x})}$, that is:

$$(T_{\mathbf{0}}\phi)(\gamma'_{\mathbf{x}}(0)) = \gamma'_{\phi(\mathbf{x})}(0)$$

Conversely, given a unit length initial velocity vector $\mathbf{v} \in T_{\mathbf{0}}\mathbb{H}^2$ and a distance $0 < D < 1$ there is a unique Euclidean-unit-speed radial path $[0, D] \to \mathbb{H}^2 : t \mapsto t\mathbf{v}$ with initial velocity $\mathbf{v}$ and Euclidean length $D$, so it is $\gamma_{D\mathbf{v}}$. Thus, $\phi(\mathbf{x}) = d_{\mathbb{E}^2}(\mathbf{0}, \mathbf{x}) \cdot (T_{\mathbf{0}}\phi)(\gamma'_{\mathbf{x}}(0))$. In other words, $\phi$ is determined by its derivative at $\mathbf{0}$.

Now, the inner product on $T_{\mathbf{0}}\mathbb{H}^2$ is the standard Euclidean inner product, so the group of linear maps preserving this inner product is $\mathrm{O}(2)$. Thus, if $\phi \in \mathrm{Stab}_{\mathrm{Isom}(\mathbb{H}^2)}(\mathbf{0})$, then $T_{\mathbf{0}}\phi : T_{\mathbf{0}}\mathbb{H}^2 \to T_{\mathbf{0}}\mathbb{H}^2$ is orthogonal. But if we just take $\psi := T_{\mathbf{0}}\phi \in \mathrm{O}(2)$ as a map of $\mathbb{R}^2$, then it preserves the open unit ball and, since it is linear, $T_{\mathbf{0}}\psi = \psi$. Since isometries stabilizing $\mathbf{0}$ are determined by their derivative at $\mathbf{0}$, we have $T_{\mathbf{0}}\phi = \psi = T_{\mathbf{0}}\psi \implies \phi = \psi \in \mathrm{O}(2)$. □

COROLLARY 2.3.10. *Every line through* $\mathbf{0}$ *in* $\mathbb{H}^2$ *is geodesic.*

PROOF. Isometries preserve geodesics. Our example shows that the first coordinate axis is geodesic. We can move this to any line through the origin with an appropriate element of $\mathrm{O}(2)$, and this is an isometry, by Lemma 2.3.9. □

It is convenient now to change notation and consider the plane as the complex numbers $\mathbb{C}$, with the hyperbolic plane consisting of the open unit disc.

Proposition 2.3.11. *Consider:*

$$\mathrm{PSU}(1,1) := \left\{ \left[\begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix}\right] \mid \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{C}),\ a\bar{a} - b\bar{b} = 1 \right\}$$

(1) $\mathrm{PSU}(1,1)$ *acts on* $\mathbb{H}^2$.
(2) $\mathrm{PSU}(1,1)$ *contains* $\mathrm{SO}(2) < \mathrm{Isom}^+(\mathbb{H}^2)$.
(3) $\mathrm{PSU}(1,1)$ *contains a 1-parameter family that acts by hyperbolic translation on the real axis.*
(4) $\mathrm{PSU}(1,1)$ *acts transitively on the set of unit tangent vectors to* $\mathbb{H}^2$.
(5) $\mathrm{Isom}^+(\mathbb{H}^2)$ *acts simply transitively on the set of unit tangent vectors to* $\mathbb{H}^2$.

*Consequently,* $\mathrm{Isom}^+(\mathbb{H}^2) \cong \mathrm{PSU}(1,1)$.

Proof. Let $\mathrm{SU}(1,1) := \left\{ \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{C}) \mid a\bar{a} - b\bar{b} = 1 \right\}$ act on $\mathbb{C} \cup \{\infty\}$ by Möbius transformations, so that:

$$\begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} : z \mapsto \frac{az+b}{\bar{b}z+\bar{a}}$$

Composition of Möbius transforms is compatible with matrix multiplication, so such a map is a bijection of $\mathbb{C} \cup \{\infty\}$. Observe that the center of $\mathrm{SU}(1,1)$ consists of plus and minus the identity matrix, and this is also the kernel of the action. $\mathrm{PSU}(1,1) := \mathrm{SU}(1,1)/\pm\mathrm{Id}$.

Take $\mathbb{H}^2$ to be the open unit disc in $\mathbb{C}$. First we check that the action of $\mathrm{SU}(1,1)$ preserves the open unit disc. Compute:

$$\left|\frac{az+b}{\bar{b}z+\bar{a}}\right|^2 = \frac{|z|^2 + |b|^2(|z|^2+1) + \bar{a}b\bar{z} + a\bar{b}z}{1 + |b|^2(|z|^2+1) + \bar{a}b\bar{z} + a\bar{b}z}$$

Thus, $|z| < 1 \iff \left|\frac{az+b}{\bar{b}z+\bar{a}}\right| < 1$. This proves Item (1).

Observe that for $\begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \in \mathrm{SU}(1,1)$ we have $1 = a\bar{a} - b\bar{b} = |a|^2 - |b|^2 \implies |a|^2 = 1 + |b|^2 > |b|^2 \geqslant 0$, so $|a| > |b|$ and $a \neq 0$. The point 0 is fixed if and only if $b = 0$, and in this case $|a| = 1$, so $a = e^{i\theta}$ for some $\theta \in [0, 2\pi)$, and the Möbius transform is $z \mapsto e^{i2\theta}z$, which is rotation about the origin by angle $2\theta$. This gives a bijection from the subgroup of $\mathrm{PSU}(1,1)$ whose representatives have $b = 0$ and the group of rotations $\mathrm{SO}(2)$. Rotations are hyperbolic isometries, by Lemma 2.3.9, so this proves Item (2).

Next, consider the case that $a, b \in \mathbb{R}$. Then a computation shows $\frac{az+b}{bz+a} \in \mathbb{R} \iff z \in \mathbb{R}$, so these maps preserve $(-1,1)$. Furthermore, they act simply transitively on $(-1,1)$, since for $r \in (-1,1)$ we can check that $0 \mapsto r$ if and only if $a = \pm\frac{1}{\sqrt{1-r^2}}$ and $b = ar$. Thus, for every $r \in (-1,1)$ there is a unique element of PSU$(1,1)$ represented by a real matrix and taking 0 to $r$. Let us check that an element of SU$(1,1)$ with real entries acts by a hyperbolic isometry. Let $a$, $b \in \mathbb{R}$ with $a^2 - b^2 = 1$, and let $\phi$ be the corresponding Möbius transform. For any $z \in \mathbb{C}$ with $|z| < 1$ and any vectors $\mathbf{v}$, $\mathbf{w} \in T_z\mathbb{H}^2$ we compute:

$$\begin{aligned}
\langle (T_z\phi)(\mathbf{v}), (T_z\phi)(\mathbf{w}) \rangle_{T_{\phi(z)}\mathbb{H}^2} &= \frac{4\langle (T_z\phi)(\mathbf{v}), (T_z\phi)(\mathbf{w}) \rangle_{\mathbb{E}^2}}{(1-|\phi(z)|^2)^2} \\
&= \frac{4|T_z\phi|^2 \langle \mathbf{v}, \mathbf{w} \rangle_{\mathbb{E}^2}}{(1-|\phi(z)|^2)^2} \\
&= \frac{|T_z\phi|^2(1-|z|^2)^2}{(1-|\phi(z)|^2)^2} \cdot \frac{4\langle \mathbf{v}, \mathbf{w} \rangle_{\mathbb{E}^2}}{(1-|z|^2)^2} \\
&= \frac{|T_z\phi|^2(1-|z|^2)^2}{(1-|\phi(z)|^2)^2} \cdot \langle \mathbf{v}, \mathbf{w} \rangle_{T_z\mathbb{H}^2}
\end{aligned}$$

This shows that $\phi$ is an isometry if and only if $\left(\frac{|T_z\phi|(1-|z|^2)}{1-|\phi(z)|^2}\right)^2 = 1$. When $a$ and $b$ are real with $a^2 - b^2 = 1$ this is a simple computation to verify.

We have already shown that the interval $(-1,1) \subset \mathbb{R} \subset \mathbb{C}$ is a hyperbolic geodesic. The hyperbolic isometry $\begin{pmatrix} 1/\sqrt{1-r^2} & r/\sqrt{1-r^2} \\ r/\sqrt{1-r^2} & 1/\sqrt{1-r^2} \end{pmatrix}$ preserves this geodesic and its orientation. Any such isometry acts on the geodesic by translation. Since the point 0 moves to the point $r$ at hyperbolic distance $\tau := \log \frac{1+r}{1-r}$ from 0, we conclude that every point on the geodesic is translated by distance $\tau$. This proves Item (3).

Combining these two types of isometries, rotations and translations, gets us a group of isometries that act transitively on the set of all unit tangent vectors to points in $\mathbb{H}^2$. First, to see that the group acts transitively on $\mathbb{H}^2$, consider a point $z = re^{i\theta} \in \mathbb{C}$ with $r \in [0,1)$. Then rotation about 0 by angle $-\theta$ is a hyperbolic isometry taking this point to $r$, and there is a unique, up to projectivization, element of $SU(1,1)$ with real entries taking $r$ to 0. Let $\phi$ be the composition of these two maps. Now let $\mathbf{v} \in T_z\mathbb{H}^2$ be a unit vector. The derivative $T_z\phi$ of $\phi$ at $z$ is a linear map $T_z\mathbb{H}^2 \to T_0\mathbb{H}^2$ that preserves the inner product, so it takes $\mathbf{v}$ to some unit vector in $T_0\mathbb{H}^2$. Linear maps are their own derivatives, so postcompose $\phi$ by a rotation that takes $(T_z\phi)(\mathbf{v})$ to $1 \in T_0\mathbb{H}^2$. This proves Item (4).

Finally, we claim that $\mathrm{Isom}^+(\mathbb{H}^2)$ acts simply transitively on unit vectors. From the previous item, we already know that it acts transitively, so it is enough to show that if an element of $\mathrm{Isom}^+(\mathbb{H}^2)$ fixes any unit tangent vector then it is the identity. Suppose $\phi \in \mathrm{Isom}^+(\mathbb{H}^2)$ such that $\phi(0) = 0$ and $(T_0\phi)(1) = 1 \in T_0\mathbb{H}^2$. By Lemma 2.3.9, $\phi \in \mathrm{SO}(2)$, but then $\phi = T_0\phi$

is a rotation, and a nontrivial rotations have no fixed points on the unit circle. □

Corollary 2.3.12. $\mathrm{Isom}(\mathbb{H}^2) \cong \mathrm{PSU}(1,1) \rtimes \mathcal{C}_2$

Proof. $\mathrm{Isom}(\mathbb{H}^2) \cong \mathrm{Isom}^+(\mathbb{H}^2) \rtimes \mathcal{C}_2$, where the $\mathcal{C}_2$ factor is generated by a reflection. We take this reflection to be $z \mapsto \bar{z}$. Conjugating a Möbius transformation by complex conjugation yields the Möbius transformation in which all entries have been replaced by their conjugate. □

Exercise 2.3.13. Show that $\mathrm{Isom}(\mathbb{H}^2)$ preserves the collection of Euclidean circles and straight lines through the origin (which should be thought of as circles through $\infty$).

Exercise 2.3.14. Show that $\mathrm{Isom}^+(\mathbb{H}^2)$ acts transitively on the set of distinct pairs of points in $\partial\mathbb{H}^2 := \mathbb{S}^1 \subset \mathbb{C}$.

Exercise 2.3.15. Show that the stabilizer of the (ordered) pair $(-1, 1)$ in $\mathrm{Isom}^+(\mathbb{H}^2)$ is the 1–parameter family described in Proposition 2.3.11 (3).

In the proof of Proposition 2.3.11 there were two one-real-parameter subgroups of isometries that played a distinguished role:

$$\text{rotations about 0:} \quad \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix}$$

$$\text{translations along } \mathbb{R}\text{:} \quad \begin{pmatrix} 1/\sqrt{1-r^2} & r/\sqrt{1-r^2} \\ r/\sqrt{1-r^2} & 1/\sqrt{1-r^2} \end{pmatrix}$$

We should mention one more family:

$$\text{parabolics fixing 1:} \quad \begin{pmatrix} 1+ir & -ir \\ ir & 1-ir \end{pmatrix}$$

These parabolic transformations fix the point 1 on the unit sphere, and fix no point in $\mathbb{H}^2$ when $r \neq 0$. Figures 4 and 5 illustrate the action of translations and parabolics.

Here are some properties of these subgroups of isometries:

| element | fixed points | trace of representative |
|---|---|---|
| nontrivial rotation about 0 | only 0 | in $(-2, 2)$ |
| nontrivial translation along $\mathbb{R}$ | $-1$ and $1$ | in $(-\infty, -2) \cup (2, \infty)$ |
| nontrivial parabolic fixing 1 | only 1 | $\pm 2$ |

Table 2.1. Properties of some isometries of $\mathbb{H}^2$

We can also put the translations into a more convenient form:

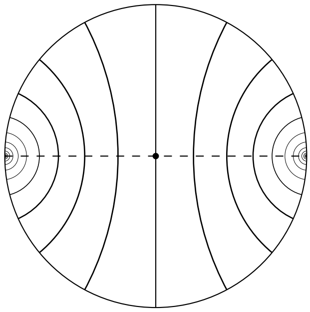


FIGURE 4. Images of the imaginary axis under some powers of the translation $\begin{pmatrix} 4/\sqrt{15} & 1/\sqrt{15} \\ 1/\sqrt{15} & 4/\sqrt{15} \end{pmatrix}$.

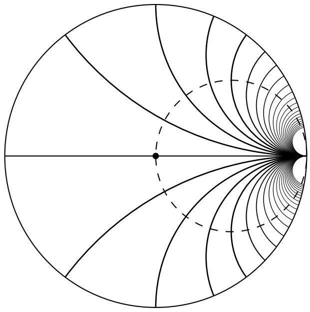


FIGURE 5. Images of the real axis under some powers of the parabolic transformation $\begin{pmatrix} 1+i/2 & -i/2 \\ i/2 & 1-i/2 \end{pmatrix}$. Orbits of points lie on circles tangent to the fixed point 1 in $\partial\mathbb{H}^2$.

EXERCISE 2.3.16. Show that hyperbolic translation along the real axis by hyperbolic distance $t$ is given by $\begin{pmatrix} \cosh(t/2) & \sinh(t/2) \\ \sinh(t/2) & \cosh(t/2) \end{pmatrix}$.

2.3.2. *Hyperbolic geodesics and polygons.*

PROPOSITION 2.3.17. *There is a unique unit speed geodesic between any two points of* $\mathbb{H}^2$*.*

PROOF. Let $z$ and $z'$ be distinct points of $\mathbb{H}^2$. As in the proof of Proposition 2.3.11, there is an isometry $\phi$ taking $z$ to 0. By Corollary 2.3.10, the radial line $\gamma$ from 0 to $\phi(z')$ is the unique geodesic between these points, up to reparameterization. Thus, $\phi^{-1}(\gamma)$ is the unique geodesic from $z$ to $z'$. $\square$

Thus far we have talked about finite geodesics, or geodesic segments, but there are also geodesic rays and bi-infinite geodesics. The prototypical bi-infinite geodesic is the interval $(-1,1) \subset \mathbb{R} \subset \mathbb{C}$.

LEMMA 2.3.18. *(All the uniqueness statements are up to reparameterization.)*

(1) *A point and a unit tangent vector at that point determine a unique bi-infinite geodesic.*
(2) *Every nontrivial geodesic segment in $\mathbb{H}^2$ is contained in a unique bi-infinite geodesic.*
(3) *A bi-infinite geodesic is uniquely determined by a pair of distinct points on the unit circle $\partial\mathbb{H}^2 := \mathbb{S}^1 \subset \mathbb{C}$.*
(4) *Every bi-infinite geodesic in $\mathbb{H}^2$ is an arc of a Euclidean circle orthogonal to the unit circle.*

PROOF. Let $z \in \mathbb{H}^2$ and $\mathbf{v}$ a unit tangent vector at $z$. By Proposition 2.3.11 (5) there is a unique $\phi \in \mathrm{Isom}^+(\mathbb{H}^2)$ taking $z$ to 0 and $\mathbf{v}$ to 1. By Lemma 2.3.8, $\gamma\colon\ t \mapsto t$ is the unique bi-infinite geodesic through 0 with initial velocity vector 1, so $\phi^{-1}(\gamma)$ is the unique bi-infinite unit speed geodesic through $z$ with initial velocity vector $\mathbf{v}$.

For Item (2), let $\delta\colon\ [0, L] \to \mathbb{H}^2$ be a geodesic segment, parameterized to have unit speed. Up to isometry, we may assume $\delta(0) = 0$ and $\delta(L)$ is on the positive real line. By Lemma 2.3.8, this implies every point of $\delta$ is on the real line, so $\delta$ is a subsegment of $\gamma$ from above.

For Item 3, let $\delta$ be a bi-infinite geodesic. Consider a compact subsegment. Up to $\phi \in \mathrm{Isom}^+(\mathbb{H}^2$, we may assume that the compact subsegment begins at 0 and ends on the positive real axis, so $\phi \circ \delta = \gamma$ has endpoints $(-1, 1)$ on $\partial\mathbb{H}^2$. But $\phi \in \mathrm{PSU}(1, 1)$ is well-defined homeomorphism of $\mathbb{C}$, not just of the open unit ball, so $\phi^{-1}(-1)$ and $\phi^{-1}(1)$ are the topological limits of $\delta$. So every bi-infinite geodesic has an associated pair of boundary points. To complete the proof of Item 3, we need to see that the boundary points determine the bi-infinite geodesic. This follows from Exercises 2.3.14 and 2.3.15.

Item 4 follows from Item 3, Exercise 2.3.13, and the fact that Möbius transformations are conformal, that is, they preserve angles. □

PROPOSITION 2.3.19. *Every hyperbolic geodesic triangle has angle sum strictly less than $\pi$. Conversely, given any angles $\theta_1$, $\theta_2$, $\theta_3 \in (0, \pi)$ such that $\theta_1 + \theta_2 + \theta_3 < \pi$, there exists a hyperbolic geodesic triangle, unique up to hyperbolic isometry, with angles $\theta_1$, $\theta_2$, and $\theta_3$.*

PROOF. The first part follows from the Gauss-Bonnet Theorem, recall (2). We prove the second part. Assume, by renumbering, if necessary, that $\theta_1 < \pi/2$.

Let $c_0$ be the point of intersection between the tangent line to the unit circle at $e^{i\theta_1}$ and the tangent to the unit circle at 1. Let $t \mapsto c_t := c_0 + te^{-i(\frac{\pi}{2}-\theta_1)}$ be a parameterization of the tangent line to the unit circle at $e^{i\theta_1}$, and for $t \geqslant 0$ define $R_t := d_{\mathbb{E}^2}(e^{i\theta_1}, c_t)$. Then the Euclidean circle $S_t$ of radius $R_t$ centered at $c_t$ is orthogonal to the unit circle and passes through $e^{i\theta_1}$. The parameterization of $c_t$ was chosen so that $S_t$ crosses the real line in the interval $(0, 1]$ when $t \geqslant 0$. Let $s_t$ be the point of intersection. The angle at which $S_t$ crosses $\mathbb{R}$ at $s_t$ is $\phi_t := \arccos(\frac{\mathfrak{Im}(c_t)}{R_t})$, which increases continuously from 0 for $t \geqslant 0$. As $t \to \infty$, $S_t$ limits to the line through 0 and $e^{i\theta_1}$, so $\phi_t \to \pi - \theta_1 > \theta_2$. Thus, there exists $t > 0$ such that $\phi_t = \theta_2$. See Figure 6.

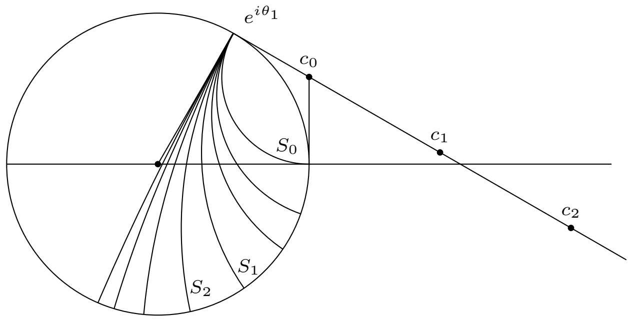



FIGURE 6. Construction of a geodesic with angle $\theta_2$ and end-point at $e^{i\theta_1}$.

Let $r_0$ be $s_t$ for $t$ such that $\phi_t = \theta_2$. For $t \in [0, r_0)$, define $r_t := r_t - t$. Let $\gamma_t$ be the hyperbolic geodesic ray starting at $r_t$ with initial velocity vector $e^{i(\pi-\theta_2)}$, so that $\gamma_t$ is a geodesic ray that meets $\mathbb{R}$ at $r_t$ with angle $\theta_2$. The geodesics $\mathbb{R}$, $[0, e^{i\theta_1}]$, and $\gamma_t$ form a hyperbolic geodesic triangle (ideal at $t = 0$) $\Delta_t$ whose angles are $\theta_1$, $\theta_2$, and $\psi_t$. See Figure 7.

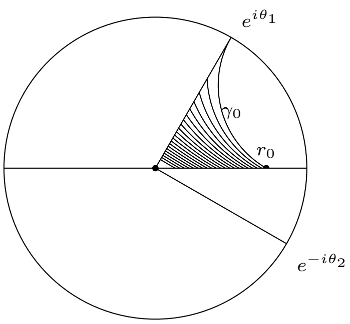



FIGURE 7. A sequence of geodesics forming hyperbolic geodesic triangles with base angles $\theta_1$ and $\theta_2$, and peak angle varying between 0 and $\pi - \theta_1 - \theta_2$.

At $t = 0$ we have by construction that $\psi_0 = 0$. Since $r_t \to 0$ as $t \to r_0$, the triangle $\Delta_t$ gets arbitrarily small, which means it gets closer and closer to being a Euclidean triangle. In particular $\psi_t \to \pi - \theta_1 - \theta_2$ as $t \to r_0$. Thus, as $t$ varies between 0 and $r_0$ the hyperbolic geodesic triangle $\Delta_t$ has angle $\psi_t$ strictly increasing from 0 to $\pi - \theta_1 - \theta_2 > \theta_3$. Thus, there is a unique $t$ as which $\Delta_t$ is a hyperbolic geodesic triangle with angles $\theta_1$, $\theta_2$, and $\theta_3$. □

EXERCISE 2.3.20. The *hyperbolic law of cosines* says that a hyperbolic triangle with angles $\alpha$, $\beta$ and $\gamma$ has side opposite $\alpha$ of length $a$, where:

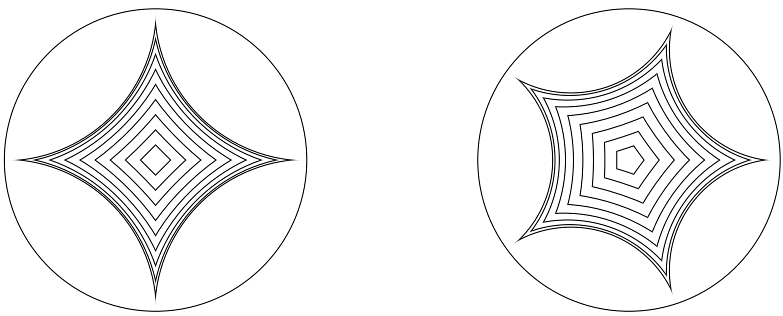
$$\cosh a = \frac{\cos\alpha - \cos\beta\cos\gamma}{\sin\beta\sin\gamma}$$

Choose angles $\theta_1$, $\theta_2$, and $\theta_3$ such that $\theta_1 + \theta_2 + \theta_3 < \pi$. Use the hyperbolic law of cosines to find $r$ and $s$ such that the hyperbolic triangle with vertices at the points 0, $r$, and $se^{i\theta_1}$, as in Figure 7, has angle $\theta_1$ at 0, angle $\theta_1$ at $r$, and angle $\theta_2$ at $se^{i\theta_1}$. The vertices of the triangle can be plotted, since we know from Example 2.3.4 how hyperbolic and Euclidean distances are related along radial segments. Two sides of the triangle are such radial segments. The third side is an arc of a Euclidean circle orthogonal to the unit circle, passing through $r$ and $se^{i\theta_1}$. Find the center and radius of this circle.

PROPOSITION 2.3.21. *For every $n \geqslant 3$ and every $0 < \theta < \frac{n-2}{n}\pi$ there is a unique, up to hyperbolic isometry, regular geodesic $n$–gon with all of its angles equal to $\theta$.*

*In particular, for all $n \geqslant 5$ there is a right-angled regular geodesic $n$–gon.*

PROOF. For $r \in (0,1)$ and $0 \leqslant k < n$ connect the points $re^{i\frac{2\pi}{n}k}$ and $re^{i\frac{2\pi}{n}(k+1)}$ by a hyperbolic geodesic. This makes a regular hyperbolic geodesic $n$–gon $P_r$ whose angles decrease continuously with $r$, from the corresponding Euclidean value $\frac{n-2}{n}\pi$ as $r \to 0$ to 0 as $r \to 1$. See Figure 8. □



(A) Hyperbolic geodesic squares. (B) Hyperbolic geodesic pentagons.

FIGURE 8. Some regular hyperbolic geodesic polygons.

Corollary 2.3.22. *For every $n \geqslant 3$ and every $m > 2n/(n-2)$ there is a tiling of the hyperbolic plane by regular $n$–gons, with $m$ many incident to each vertex.*

Proof. Choose a regular hyperbolic $n$–gon with angle $\theta = 2\pi/m$, which is possible by Proposition 2.3.21. □

Figure 9 shows some regular tilings of $\mathbb{H}^2$. Not all of these correspond to reflection groups: reflection groups should have an even number of faces at each vertex.

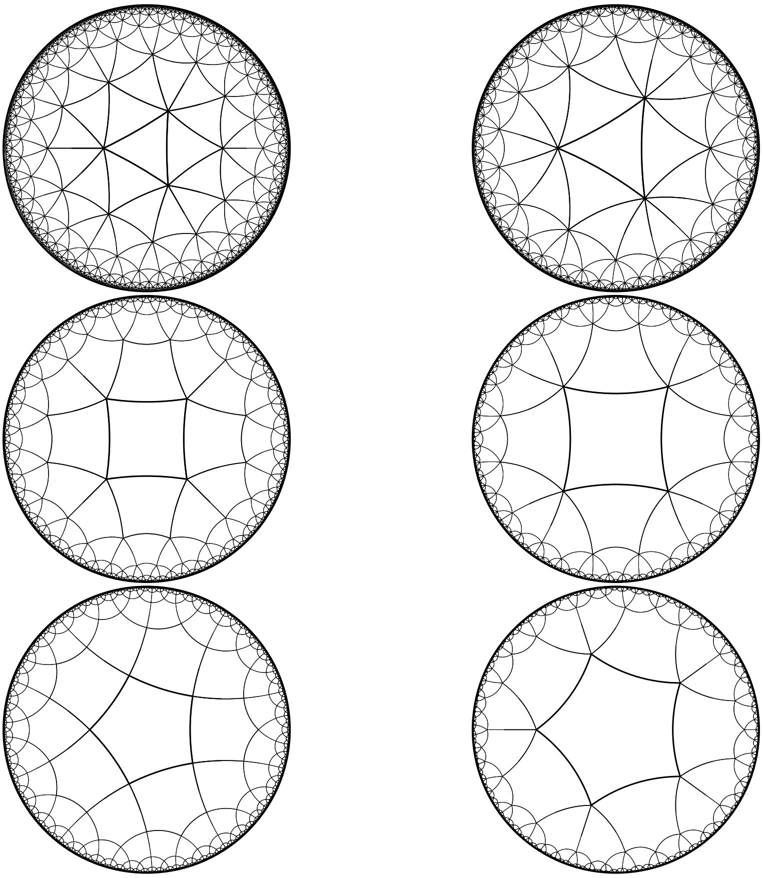


Figure 9. Some regular tilings of $\mathbb{H}^2$. Only the even valence ones correspond to geometric reflection groups.

Proposition 2.3.23. *Let $n \geqslant 3$ and $2m > 2n/(n-2)$. The Coxeter group with presentation graph an $n$–cycle and all edges labelled $m$ acts on $\mathbb{H}^2$ as a geometric reflection group with strict fundamental domain a regular hyperbolic $n$–gon with dihedral angles $\pi/m$.*

Proof. By Corollary 2.3.22 there is a tiling of $\mathbb{H}^2$ by regular hyperbolic $n$–gons with $2m$ faces incident to each vertex. Thus, the dihedral angles are

$2\pi/2m$. Define a map from generators of the Coxeter group to generators of the reflection group by sending consecutive vertices to reflections in consecutive faces of a fixed $n$–gon. The relators are satisfied by this map, so we get a surjection from the Coxeter group to the geometric reflection group. $\square$

EXERCISE 2.3.24. Show that the Coxeter group whose Coxeter graph is an unlabelled pentagon acts on $\mathbb{H}^2$ as a geometric reflection group.

EXERCISE 2.3.25. Given angles $0 < \theta_1, \theta_2, \theta_3, \theta_4 < \pi$ such that $\sum_i \theta_i < 2\pi$, show that for all sufficiently large $r < 1$ there exists a unique hyperbolic geodesic quadrilateral with one vertex at 0 with dihedral angle $\theta_1$, one vertex at $r$ with dihedral angle $\theta_2$, and the other two dihedral angles $\theta_3$ and $\theta_4$. Conclude there is a 1–parameter space of hyperbolic quadrilaterals with given dihedral angles.

EXERCISE 2.3.26. Show that for any $n \geqslant 3$ and $0 < \theta_i < \pi$ for $1 \leqslant i \leqslant n$ such that $\sum_{i=1}^n \theta_i < (n-2)\pi$ there exists a hyperbolic geodesic $n$–gon with dihedral angles $\theta_i$.

### 2.4. Geometric reflection groups on the hyperbolic plane.

PROPOSITION 2.4.1. *For all $n \geqslant 3$ and $2 \leqslant m_i \in \mathbb{N}$ for $1 \leqslant i \leqslant n$ such that $\sum_{i=1}^n \frac{1}{m_i} < n-2$, there is an $\mathbb{H}^2$ reflection group with fundamental domain a hyperbolic geodesic $n$–gon with dihedral angles $\theta_i = \pi/m_i$.*

*Every $\mathbb{H}^2$ reflection group is of this form.*

*The Coxeter graph with presentation graph an $n$–cycle with edges labelled by the $m_i$ surjects onto the geometric reflection group.*

Note that the condition $\sum_{i=1}^n \frac{1}{m_i} < n-2$ is vacuous for $n > 4$, and for $n = 4$ excludes only the case that all dihedral angles are $\pi/2$.

PROOF. The existence of the geodesic $n$–gon with given dihedral angles is Exercise 2.3.26.

If $P$ is the fundamental domain of an $\mathbb{H}^2$ reflection group then the dihedral angles of $P$ are integral submultiples of $\pi$, and Gauss-Bonnet (2), implies $\sum_{i=1}^n \frac{1}{m_i} < n-2$, so every $\mathbb{H}^2$ reflection group is of this form.

Finally, the usual check shows that sending generators of the Coxeter system to the corresponding reflection in faces of the fundamental domain satisfies all of the Coxeter relators, so we get the desired surjection. $\square$

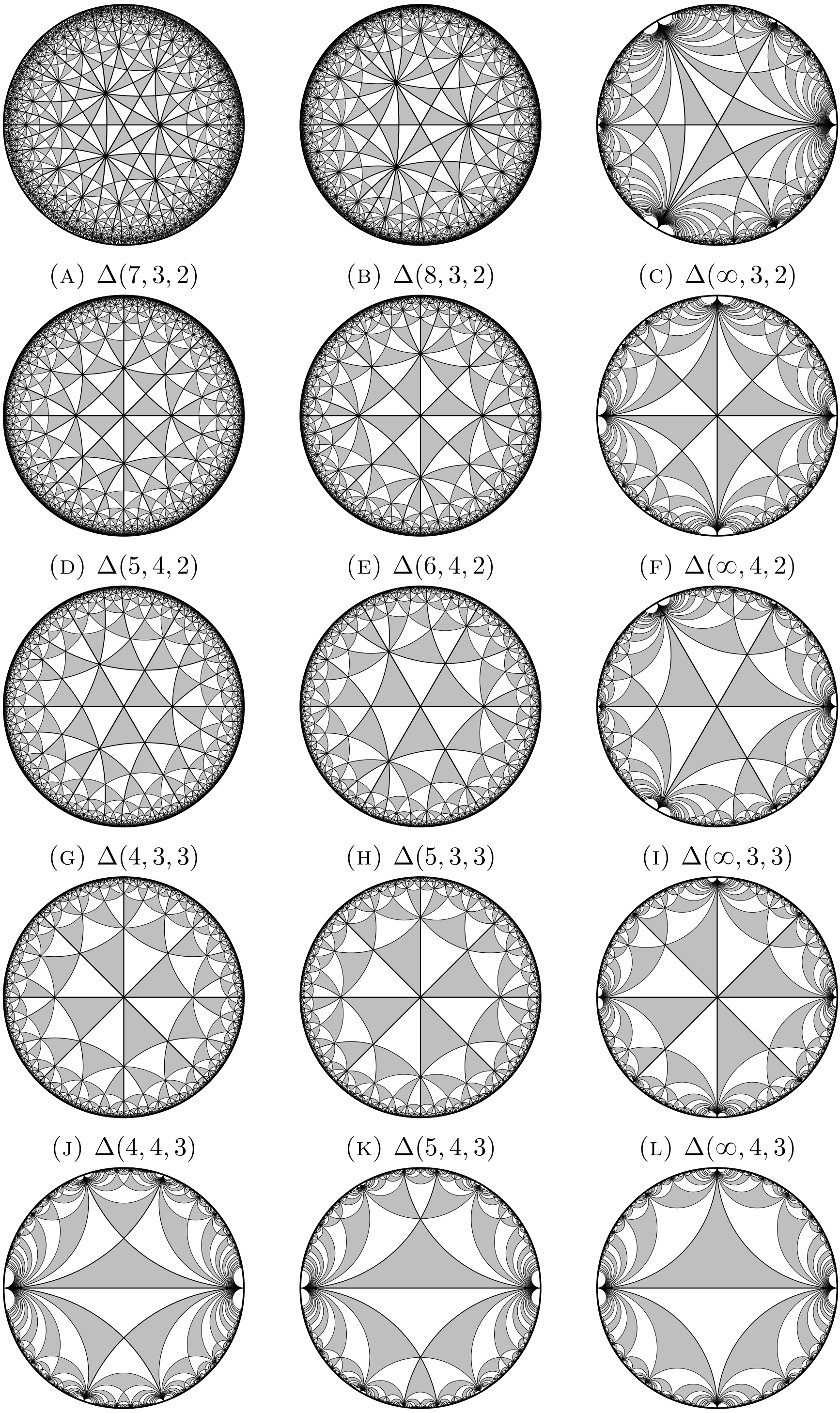
As in the Euclidean and spherical cases, the surjection in this homomorphism is actually an isomorphism.

A special case of Proposition 2.4.1 is the case $n = 3$, in which case the Coxeter groups are the triangle groups $\Delta(p,q,r)$ with $\frac{1}{p}+\frac{1}{q}+\frac{1}{r} < 1$ and $p,q,r < \infty$.

If we drop the condition that $p, q, r < \infty$ then $\Delta(p, q, r)$ does still act on $\mathbb{H}^2$ with fundamental domain a 'triangle', with each generator acting as reflection in one of the sides. The difference is that some of the vertices of the triangle are on the unit circle, and the adjacent edges are therefore infinite geodesics, not geodesic segments. Such an object is called an *ideal triangle*, and the vertices on the unit circle are called *ideal vertices*. An ideal triangle is not a compact set, so the group action in this case is not cocompact. There is, however, a way to extend the notion of $n$–dimensional volume to Riemannian manifolds of dimension $n$, 'area' when $n = 2$. It turns out that ideal hyperbolic triangles have finite area, so although the action is not cocompact is it 'finite co-volume'.

In fact, a similar thing works for the Euclidean triangle group $\Delta(2, 2, \infty) \cong \mathcal{D}_\infty \times \mathcal{C}_2$: This has an action on $\mathbb{E}^2$ where the $\mathcal{D}_\infty$ factor acts with its usual geometric action on the $x$–axis, and the $\mathcal{C}_2$ factor reflects through the $x$–axis. A fundamental domain is the region $\{(x, y) \mid -1/2 \leqslant x \leqslant 1/2,\ y \geqslant 0\}$, which can be thought of as an ideal Euclidean triangle. However, the Euclidean area of this region is infinite, so this action is neither cocompact nor of finite co-volume.

Figure 10 shows both actual and ideal triangular tilings of $\mathbb{H}^2$ for various triangle groups. The hyperbolic triangle groups with no infinite entries give us the hyperbolic reflection action pictured. For the ones with at least one infinite entry, the pictures only tell us that this particular action of the triangle group is not a geometric reflection action. It turns out that these groups are not geometric reflection groups at all. Consider the example $\Delta(\infty, \infty, \infty) = \langle s_1, s_2, s_3 \mid s_1^2, s_2^2, s_3^2\rangle$. The tiling pictured in Figure 10o has dual graph a tree (a graph with no loops). In fact, it is a bushy tree, in the sense that it is infinite and has branching at all its vertices, so it is not just a line. Correspondingly, $\Delta(\infty, \infty, \infty)$ has an index 2 subgroup $H := \langle s_1 s_2, s_1 s_3\rangle$ that acts on the tree freely and cocompactly without inverting edges. It is a fact that such a group is free. In this specific example it is also easy to verify by the Reidemeister-Schreier Algorithm that $H$ has presentation $\langle s_1 s_2, s_1 s_3\rangle$, so it is $F_2$. Non-cyclic free groups have a Cayley graph that is a bushy tree. Theorem 0.0.5 implies that a geometric reflection group is quasiisometric to either $\mathbb{S}^n$, $\mathbb{E}^n$, or $\mathbb{H}^n$, but it can be shown that none of these are quasiisometric to a bushy tree; for example there is a quasiisometry invariant called the 'number of ends' of a space, which is 0, 1, or 2 for $\mathbb{S}^n$, $\mathbb{E}^n$, and $\mathbb{H}^n$, but is infinite for a bushy tree. So it is not the case that $\Delta(\infty, \infty, \infty)$ might secretly be a geometric reflection group in some undiscovered way; there are Coxeter groups that just are not geometric reflection groups.



(A) $\Delta(7,3,2)$ (B) $\Delta(8,3,2)$ (C) $\Delta(\infty,3,2)$

(D) $\Delta(5,4,2)$ (E) $\Delta(6,4,2)$ (F) $\Delta(\infty,4,2)$

(G) $\Delta(4,3,3)$ (H) $\Delta(5,3,3)$ (I) $\Delta(\infty,3,3)$

(J) $\Delta(4,4,3)$ (K) $\Delta(5,4,3)$ (L) $\Delta(\infty,4,3)$

(M) $\Delta(\infty,\infty,2)$ (N) $\Delta(\infty,\infty,3)$ (O) $\Delta(\infty,\infty,\infty)$

FIGURE 10. Some tilings of $\mathbb{H}^2$ for triangle groups.

**2.5. Mirror Structures.** Let $(W, S)$ be a Coxeter system and let $X$ be a space. A *mirror structure* on $X$ over $S$ is a choice of closed subspace

$X_s$ of $X$ for each $s \in S$. The sets $X_s$ are the *mirrors*. For $x \in X$, let $S(x) := \{s \in S \mid x \in X_s\}$.

DEFINITION 2.5.1. Given a mirror structure on $X$ over $S$, define

$$\mathcal{U}(W, X) := (W \times X)/\sim$$

where $gx \sim hy$ if $x = y$ and $h^{-1}g \in W_{S(x)}$.

This definition means that $\mathcal{U}(W, X)$ has a copy of $X$ for each element of $W$, but some of the copies are glued together along mirrors. For instance, if $s \in S$ and $X_s \neq \varnothing$ then $X_s \subset X$ is identified with $sX_s \subset sX$.

EXAMPLE 2.5.2. If all of the mirrors are the empty set then $\mathcal{U}(W, X)$ is a disjoint union of copies of $X$, one for each element of $W$.

EXAMPLE 2.5.3. All of the pictures of tilings preserved by a geometric reflection group in the previous section are examples of $\mathcal{U}(W, X)$, where $X$ is an interval or polygon and the mirrors are the faces.

Subsets $(g, X)$ of $\mathcal{U}(W, X)$ are called *chambers*, and $(1, X)$ is called the *fundamental chamber*.

LEMMA 2.5.4. *Left multiplication of $W$ on itself extends to an action of $W$ on $\mathcal{U}(W, X)$ with strict fundamental domain $X$. If $X \neq \cup_{s \in S} X_s$ then the action is faithful.*

PROOF. The action of $W$ on $\mathcal{U}(W, X)$ is given by $w(gx) = (wg)x$. To see this is well defined, suppose $hy \sim gx$. This is true when $y = x$ and $h^{-1}g \in W_{S(x)}$. Consider $w(hy) = (wh)y$. Since $y = x$ and $(wh)^{-1}(wg) = h^{-1}g \in W_{S(x)}$, we have $w(hy) = (wh)y \sim (wg)x = w(gx)$, so the action is well defined.

Consider a point $gx \in \mathcal{U}(W, X)$. Its $W$ orbit contains $x \in X$. Now suppose $wx = y \in X$ for some $w \in W - \{1\}$. Then there are $w_0 = w, \ldots, w_n = 1$ and $x_0 = x, \ldots, x_n = y$ such that $w_i x_i \sim w_{i+1} x_{i+1}$ for all $i$. But then $x_i = x_{i+1}$ for all $i$, so $x = y$. Thus, $X$ is a strict fundamental domain. Notice also that $w = \prod_{i=0}^{n-1} w_{i+1}^{-1} w_i$ is a product of elements of $W_{S(x)}$, so $w \in W_{S(x)}$.

Suppose there exists $x_0 \in X - \cup_{s \in S} X_s$. We have $gx_0 \sim x_0$ if and only if $g \in W_{S(x_0)} = W_{\varnothing} = \{1\}$. $\square$

EXERCISE 2.5.5. Show that if $X$ is path connected and all of the mirrors are nonempty then $\mathcal{U}(W, X)$ is path connected.

EXAMPLE 2.5.6. Let $S = \{s_1, \ldots, s_{|S|}\}$, and let $X$ be a graph consisting of vertices $v_0, \ldots, v_{|S|}$ with an edge from $v_0$ to $v_i$ for all $1 \leqslant i \leqslant |S|$. Let

$X_{s_i} := v_i$. Then $\mathcal{U}(W, X)$ is a bipartite graph, where one part, the orbit of $v_0$, consists of valence $|S|$ vertices in bijection with $W$, and the other part, the orbits of the other $v_i$, consists of valence 2 vertices.

Now consider the *Cayley graph* $Cay(W, S)$ *of* $W$ *with respect to* $S$, which is the graph with a vertex for each element of $W$, and an edge labelled $s \in S$ between $g$ and $h$ when $gs = h$. The graph $\mathcal{U}(W, X)$ is $W$–equivariantly isomorphic to the barycentric subdivision of $Cay(W, S)$.

EXERCISE 2.5.7. Edge paths in $Cay(W, S)$ correspond to words in $F(S)$ by reading labels. That is, if $e_1$, $e_2$,…,$e_n$ is an edge path in $Cay(W, S)$ from $g$ to $h$, with $e_j$ labelled $s_{i_j} \in S$, then $f := s_{i_1} \cdots s_{i_n}$ is a word in $F(S)$ such that $gf = h$. Conversely, given a word $f = \prod_{j=1}^{n} s_{i_j}$, then starting from any vertex $g$ there is a unique edge path labelled $f$, and it leads from $g$ to $gf$.

Show that $f \in F(S)$ is trivial in $W$ if and only if $f$ labels a loop in $Cay(W, S)$.

Construct $Cay(W, S)$ for the Coxeter presentations of $\mathcal{D}_2$, $\mathcal{D}_3$, $\mathcal{D}_4$, and $\mathrm{Sym}_4$ (recall Exercise 4.2.11).

At this point $\mathcal{U}(W, X)$ is an abstract thing that we can define without understanding. Implicit in the construction is that we know how to distinguish elements of $W$, so we know when two copies of $X$ are distinct. This is something that we do not know how to do yet. However, $\mathcal{U}(W, X)$ will still be useful to us. The first reason is that it satisfies a universal property, explained in Lemma 2.5.8. The second reason is that we can understand $\mathcal{U}(W, X)$ *locally.* Points $x$ of $\mathcal{U}(W, X)$ that are not in any mirror have a neighborhood that is isometric to a neighborhood of the corresponding point in $X$. If we choose the mirrored space carefully we can also understand neighborhoods of points in the mirrors. In fact, in very special cases, see Theorem 2.5.9, we can choose the mirrored space geometrically in such a way that geometric local-to-global results tell us exactly what $\mathcal{U}(W, X)$ is.

LEMMA 2.5.8. *Let* $Y$ *be a space on which* $W$ *acts. Suppose* $\phi\colon X \to Y$ *is a continuous map such that for each* $s \in S$, $\phi(X_s)$ *is fixed by* $s$. *Then there is a unique continuous* $W$*–equivariant map* $\tilde{\phi}\colon \mathcal{U}(W, X) \to Y$ *extending* $\phi$.

PROOF. For all $x \in X$, $x \in \cap_{s \in S(x)} X_s$, so $\phi(x)$ is fixed by $W_{S(x)}$.

Define $\hat{\phi}\colon \coprod_{w \in W} X \to Y$ by taking the copy of $X$ corresponding to $w \in W$ to $w\phi(X)$. This map is continuous, since $\phi$ is. If $h^{-1}g \in W_{S(x)}$, so that $hx = gx$ in $\mathcal{U}(W, X)$, then $h^{-1}g\phi(x) = \phi(x)$, so:

$$\hat{\phi}(hx) = h\phi(x) = g\phi(x) = \hat{\phi}(gx)$$

Therefore, $\hat{\phi}$ factors through $\mathcal{U}(W, X)$ as $\tilde{\phi} \circ q$, where $q\colon \coprod_{w \in W} X \to \mathcal{U}(W, X)$ is the quotient map. $\square$

For the following theorem we will need the fact that if $(W, S)$ is a Coxeter system then for all $s, t \in S$, then the subgroup of $W$ that they generate is isomorphic to $\mathcal{D}_{m_{st}}$. This will be proven in Proposition 2.0.3.

THEOREM 2.5.9. *Let $P$ be a convex polygon in $\mathbb{X}^2$, where $\mathbb{X}^2$ is one of $\mathbb{S}^2$, $\mathbb{E}^2$, or $\mathbb{H}^2$. Suppose the dihedral angles of $P$ are integer submultiples of $\pi$. For faces $F_i$ and $F_j$ of $P$, if $F_i \cap F_j \neq \varnothing$ then let $m_{ij}$ be such that the dihedral angle between $F_i$ and $F_j$ is $\pi/m_{ij}$. If $F_i \cap F_j = \varnothing$ let $m_{ij} = \infty$. Let $W$ be the Coxeter group defined by Coxeter matrix $(m_{ij})$. Let $\bar{W}$ be the group generated by reflections in the faces of $P$. Then the map $\tilde{\phi}\colon \mathcal{U}(W, P) \to \mathbb{X}^2$ given by Lemma 2.5.8 is a homeomorphism, and the natural surjection $W \twoheadrightarrow \bar{W}$ is an isomorphism.*

PROOF SKETCH. Let $P$ be a convex polygon in $\mathbb{X}^2$ with faces $F_i$, and let $(W, S)$ be a Coxeter system with Coxeter matrix $(m_{ij})$ and $S = \{s_i\}$, where $F_i$ is the mirror for $s_i$.

The first step is to show that $\mathcal{U}(W, P)$ is locally isometric to $\mathbb{X}^2$. $\mathcal{U}(W, P)$ is made by gluing together copies of $P$. A small ball about an interior point in $P$ is isometric to a small ball about a point in $\mathbb{X}^2$. Every edge of a copy of $P$ in $\mathcal{U}(W, P)$ is glued to an edge of another copy of $P$. A point in the interior of such an edge has a small ball about it in $\mathcal{U}(W, P)$ that is isometric to two half balls about a point in $\mathbb{X}^2$, glued together along a geodesic boundary. This is isometric to a small ball about a point in $\mathbb{X}^2$. This leaves the vertices of $\mathcal{U}(W, P)$ to be checked.

Consider a vertex $v$ of $\mathcal{U}(W, P)$ that corresponds to the intersection of faces $F_i$ and $F_j$ in the copies of $P$ containing it. The dihedral angle of $v$ in $P$ is $0 < \pi/m_{ij} \leqslant \pi/2$, and by Proposition 2.0.3 the subgroup of $W$ generated by $s_i$ and $s_j$ is $\mathcal{D}_{m_{ij}}$, so there are $2m_{ij}$ copies of $P$ incident to $v$ in $\mathcal{U}(W, P)$. Furthermore, they are glued together cyclically as in Figure 1a, making a total angle of $2\pi$, as in Figure 1a. So a small ball about $v$ in $\mathcal{U}(W, P)$ looks like $2m_{ij}$ sectors of angle $\pi/m_{ij}$ in $\mathbb{X}^2$, glued together to make a full ball in $\mathbb{X}^2$. Thus, $\mathcal{U}(W, P)$ is locally isometric to $\mathbb{X}^2$.

The second step is to argue that the first step, that fact that $\mathcal{U}(W, P)$ is connected by Exercise 2.5.5, and compactness of $P$ imply that the universal cover of $\mathcal{U}(W, P)$ is metrically complete, so isometric to $\mathbb{X}^2$.

The third step is to argue that the continuous map $\tilde{\phi}$ provided by Lemma 2.5.8 is actually a covering map, but since $\mathbb{X}^2$ is simply connected, there are no nontrivial connected covers, so $\tilde{\phi}$ is a trivial covering, that is, a homeomorphism.

It follows that $W \twoheadrightarrow \bar{W}$ is injective because if not there are distinct copies of $P$ in $\mathcal{U}(W,P)$ that map to the same place in $\mathbb{X}^2$, contradicitng the fact that the map $\tilde{\phi}$ is a homeomorphism.

In this 2–dimensional case, we can replace the third step by a direct argument that $\mathcal{U}(W,P)$ is simply connected, as follows: Pick a point $x_0$ in the interior of $P$, a point $x_i$ in the interior of face $F_i$ for each $i$, and connect each $x_i$ to $x_0$ by a path in $P$. Call this space $X$, and consider it as a mirror structure, with the points $x_i$ for $i > 0$ being the mirrors. We saw in Example 2.5.6 that $\mathcal{U}(W,X)$ is the baryentric subdivision of the Cayley graph $Cay(W,S)$, and by construction we have $\mathcal{U}(W,X)$ sitting in $\mathcal{U}(W,P)$ as a dual graph to the 1–skeleton. Each vertex of $\mathcal{U}(W,P)$ corresponds to a coset of $W_{\{s_i,s_j\}}$, where $v = F_i \cap F_j$ is a vertex of $P$. By hypothesis, $W_{\{s_i,s_j\}} \cong \mathcal{D}_{m_ij}$, and a coset of this subgroup is a loop in $Cay(W,S)$ of length $2m_{ij}$ with edge labels alternating between $s_i$ and $s_j$, so it is a copy of one of the defining relators $(s_is_j)^{m_j}$. The corresponding loop in $\mathcal{U}(W,P)$ is filled by a disc, composed of $2m_{ij}$ copies of the corner of $P$ at $F_i \cap F_j$. The dual 2–complex to $\mathcal{U}(W,P)$ is isomorphic to the 2-complex obtained from $\mathcal{U}(W,X) \sim Cay(W,S)$ by adding a 2–cell for each distinct loop that is a translate of one of the defining relators of the form $(s_is_j)^{m_ij}$. This object is called the Cayley complex of the presentation (the given Coxeter presentation of $W$), and it is simply connected [**11**, Proposition 2.2.3]. $\square$

**2.6. Summary: the 2–dimensional geometric reflection groups.** We considered the three possible 2–dimensional spaces of constant curvature, $\mathbb{S}^2$, $\mathbb{E}^2$, and $\mathbb{H}^2$. A 2–dimensional geometric reflection group is a group that acts on one of these spaces by reflections in the sides of a convex polygon, such that the polygon is a strict fundamental domain for the action.

THEOREM 2.6.1. *The 2–dimensional geometric reflection groups are:*

- *spherical*
  - *spherical triangle groups*
- *Euclidean*
  - *Euclidean triangle groups with finite entries*
  - $\mathcal{D}_\infty \times \mathcal{D}_\infty$ *— The product of two 1–dimensional Euclidean reflection groups.*
- *hyperbolic*
  - *hyperbolic triangle groups with finite entries*
  - *Coxeter groups with presentation graph a square and at least one labelled edge*
  - *any Coxeter group whose presentation graph is a cycle of length at least 5*

*In particular, the 2–dimensional geometric reflection groups whose fundamental domain is a triangle are exactly the triangle groups with finite entries.*

PROOF. Using the Gauss-Bonnet Theorem, we showed that in the spherical case the polygon must be a triangle and in the Euclidean case the polygon can be a triangle or a rectangle. In the hyperbolic case Gauss-Bonnet does not restrict the number of sides, and we constructed examples to show that any number of sides is possible.

We have also shown, for each possibility, a Coxeter group $W$ with a natural surjection onto the corresponding reflection group. To show injectivity of this map, we apply Theorem 2.5.9. The homeomorphism $\tilde{\phi}$ of Theorem 2.5.9 conjugates the action of $W$ on $\mathcal{U}(W, P)$ to the action of the geometric reflection group on $\mathbb{X}^2$, so the geometric reflection group is isomorphic to the image of $W$ in $\mathrm{Isom}(\mathcal{U}(W, P))$. The action of $W$ on $\mathcal{U}(W, P)$ is faithful by Lemma 2.5.4, which means that $W$ is isomorphic to its image in $\mathrm{Isom}(\mathcal{U}(W, P))$. □

## 3. Higher dimensional geometric reflection groups

To make sense of higher dimensional geometric reflection groups, we need to define the higher dimensional analogues of convex polygons. In $\mathbb{R}^n$ with the Euclidean metric, any hyperplane, that is, any affine $(n-1)$–dimensional subspace, is a codimension 1 geodesic subspace that is isometric to $\mathbb{E}^{n-1}$. Call the closure of a complementary component of a hyperplane a *halfspace*.

DEFINITION 3.0.1. A *Euclidean convex polytope* is the compact intersection of finitely many halfspaces in $\mathbb{E}^n$.

We can make a similar definition for spheres.

FACT 3.0.2. *Our model for the $n$–sphere $\mathbb{S}^n$ as the unit sphere in $\mathbb{R}^{n+1}$ has the property that the intersection of $\mathbb{S}^n$ with a linear subspace of $\mathbb{R}^{n+1}$ of dimension $m+1$ is a geodesic subspace that is isometric to $\mathbb{S}^m$, and every such subspace arises in this way as the intersection of $\mathbb{S}^n$ with a linear subspace of $\mathbb{R}^{n+1}$.*

DEFINITION 3.0.3. A *spherical hyperplane* is the intersection of $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ with a linear hyperplane of $\mathbb{R}^{n+1}$. A *spherical halfspace* is the closure of a complementary component of a spherical hyperplane. A *spherical convex polytope* is the compact intersection of finitely many spherical halfspaces.

In Section 3.1 we will introduce the *hyperboloid model* of hyperbolic space, which has $\mathbb{H}^n \subset \mathbb{R}^{n+1}$ in such a way that the analogue of Fact 3.0.2

is true, see Proposition 3.1.19. This allows the following definition of hyperbolic polytopes:

DEFINITION 3.0.4. A *hyperbolic hyperplane* is the intersection of the hyperboloid model of $\mathbb{H}^n \subset \mathbb{R}^{n+1}$ with a linear hyperplane of $\mathbb{R}^{n+1}$. A *hyperbolic halfspace* is the closure of a complementary component of a spherical hyperplane. A *hyperbolic convex polytope* is the compact intersection of finitely many spherical halfspaces.

**3.1. Higher dimensional hyperbolic space.** The Poincaré Disc model of $\mathbb{H}^2$ was introduced in Section 2.3. The definition generalizes to all dimensions: The *Poincaré Ball model of* $\mathbb{H}^n$ is defined by taking $\mathbb{H}^n$ to be the open unit ball in $\mathbb{R}^n$ and taking the inner product on each tangent space to be:

$$\langle \mathbf{v}, \mathbf{w} \rangle_{T_\mathbf{x}\mathbb{H}^n} := \frac{4\langle \mathbf{v}, \mathbf{w} \rangle_{\mathbb{E}^n}}{(1-|\mathbf{x}|^2)^2}$$

Much of the theory from the Poincaré disc has direct analogues in higher dimensions. For example, the next theorem follows from the same proof strategy as in the $\mathbb{H}^2$ case:

THEOREM 3.1.1. *The Poincaré ball model of* $\mathbb{H}^n$ *is conformal, in the sense that angles measured with respect to the given inner product agree with angles measured with respect to the Euclidean inner product. Furthermore:*

- *The stabilizer of* $\mathbf{0}$ *in* $\operatorname{Isom}\mathbb{H}^n$ *contains (in fact, is equal to)* $\mathrm{O}(n)$.
- *Projection to a linear subspace is hyperbolic length nonincreasing, so linear subspaces of dimension* $m$ *are convex, isometrically embedded copies of* $\mathbb{H}^m$ *in* $\mathbb{H}^n$.
- *Every geodesic through* $\mathbf{0}$ *is contained in the intersection of the ball with a 1–dimensional linear subspace.*

However, in the Poincaré ball model we do not get the analogue of Fact 3.0.2, so do not get as convenient a characterization of hyperplanes. For this we introduce the hyperboloid model.

3.1.1. *The hyperboloid model of hyperbolic space.* For $n \geqslant 1$ let $C$ be an $(n+1) \times (n+1)$ diagonal matrix whose diagonal entries $\lambda_1, \dots, \lambda_{n+1}$ are nonzero. Let $B_C$ be the symmetric bilinear form on $\mathbb{R}^{n+1}$ defined by $C$; that is, $B_C(\mathbf{v}, \mathbf{w}) := \mathbf{v}^T C \mathbf{w} = \sum_{i=1}^{n+1} \lambda_i v_i w_i$. The function $f \colon \mathbb{R}^{n+1} \to \mathbb{R} : \mathbf{v} \mapsto B_C(\mathbf{v}, \mathbf{v})$ is smooth, with a single critical point at 0. For $r \neq 0$, the level set $f^{-1}(r)$ is an $n$–dimensional submanifold of $\mathbb{R}^{n+1}$. For $r < 0$ this submanifold has two components, while for $r > 0$ and $n > 1$ it is connected. See Figure 11.

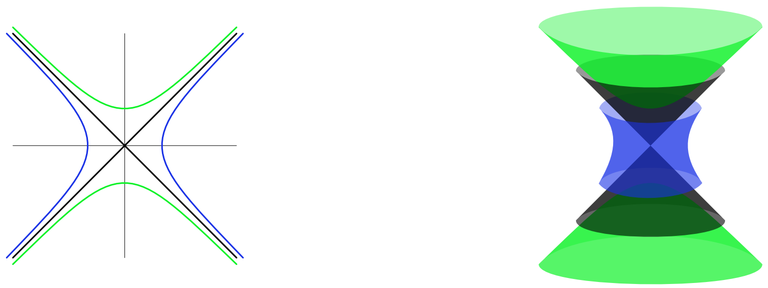


FIGURE 11. Parts of the 1, 0, and $-1$ level sets of the function $f$ for $n = 1, 2$.

EXERCISE 3.1.2. For $r \neq 0$ and $\mathbf{v} \in f^{-1}(r)$, the tangent space to $f^{-1}(r)$ at $\mathbf{v}$ is $T_{\mathbf{v}}(f^{-1}(r)) = \{\mathbf{w} \in \mathbb{R}^{n+1} \mid \mathbf{v}^T C\mathbf{w} = 0\}$. Hint: Compute the directional derivative to $f$ at $\mathbf{v}$ in the direction $\mathbf{w}$.

LEMMA 3.1.3. *Suppose that $r > 0$ and $\lambda_{n+1} < 0$ and $\lambda_i > 0$ for $1 \leqslant i \leqslant n$. For all $\mathbf{v} \in H := f^{-1}(-r)$, the restriction of $B_C$ to $T_{\mathbf{v}}H$ is positive definite.*

PROOF. Let $\mathbf{v} = (v_1, \ldots, v_{n+1}) \in H$ and $\mathbf{w} = (w_1, \ldots, w_{n+1}) \in T_{\mathbf{v}}H$. Since $-r = \mathbf{v}^T C\mathbf{v} = \sum_{i=1}^n \lambda_i v_i^2 + \lambda_{n+1} v_{n+1}^2$, we have:

$$v_{n+1} = \pm\sqrt{\frac{\sum_{i=1}^n \lambda_i v_i^2 + r}{|\lambda_{n+1}|}} \neq 0$$

By Exercise 3.1.2, $\mathbf{v}^T C\mathbf{w} = 0$, so:

$$w_{n+1} = \pm\frac{\sum_{i=1}^n \lambda_i v_i w_i}{\lambda_{n+1}\sqrt{\frac{\sum_{i=1}^n \lambda_i v_i^2 + r}{|\lambda_{n+1}|}}}$$

Thus:

$$-\lambda_{n+1} w_{n+1}^2 = \frac{(\sum_{i=1}^n \lambda_i v_i w_i)^2}{\sum_{i=1}^n \lambda_i v_i^2 + r}$$

Now compute:

$$
\begin{aligned}
0 \leqslant \mathbf{w}^T C \mathbf{w} = \sum_{i=1}^n \lambda_i w_i^2 - \frac{(\sum_{i=1}^n \lambda_i v_i w_i)^2}{\sum_{i=1}^n \lambda_i v_i^2 + r} \\
\iff 0 \leqslant (\sum_{i=1}^n \lambda_i w_i^2)(\sum_{i=1}^n \lambda_i v_i^2 + r) - (\sum_{i=1}^n \lambda_i v_i w_i)^2 \\
= r(\sum_{i=1}^n \lambda_i w_i^2) + (\sum_{i=1}^n \lambda_i w_i^2)(\sum_{i=1}^n \lambda_i v_i^2) - (\sum_{i=1}^n \lambda_i v_i w_i)^2 \\
= r(\sum_{i=1}^n \lambda_i w_i^2) + \sum_{1 \leqslant i < j \leqslant n} \lambda_i \lambda_j (v_i w_j - v_j w_i)^2
\end{aligned}
$$

Both sums have nonnegative terms, so $0 \leqslant \mathbf{w}^T C \mathbf{w}$, with equality only if all of the terms are 0, which happens only if all of the $w_i$ are 0. □

Definition 3.1.4. When $B_C$ is a positive definite form, let $|\mathbf{w}|_C := \sqrt{B_C(\mathbf{w}, \mathbf{w})}$.

Definition 3.1.5. Let $J$ be the $(n+1) \times (n+1)$ diagonal matrix with 1's on the diagonal, except for the last entry, which is $-1$. The corresponding symmetric bilinear form $B_J$ on $\mathbb{R}^{n+1}$ is called the *Minkowski form*. Let $f \colon \mathbb{R}^{n+1} \to \mathbb{R} : \mathbf{v} \mapsto B_J(\mathbf{v}, \mathbf{v})$, as above. Then $-1$ is a regular value of $f$, so $f^{-1}(-1)$ is an $n$–dimensional submanifold of $\mathbb{R}^{n+1}$. It is not hard to see that $f^{-1}(-1)$ has two connected components, distinguished by the sign of the last coordinate. Let $H$ be the component of $f^{-1}(-1)$ consisting of vectors with positive last coordinate. Consider the Riemannian manifold obtained from $H$ by defining the inner product on $T_{\mathbf{v}}H$, for each $\mathbf{v} \in H$, to be the restriction of $B_J$ to $T_{\mathbf{v}}H$. This Riemannian manifold is the *hyperboloid model of* $\mathbb{H}^n$.

Throughout this section, let $\mathbf{v}_0 := (0, \ldots, 0, 1) \in H$. Observe that the restriction of the Minkowski form to $T_{\mathbf{v}_0}H$ agrees with the standard Euclidean inner product. However, $\mathbf{v}_0$ is the only point of $H$ for which this is true, even up to rescaling. At every other point there exist tangent vectors whose Euclidean and Minkowski angles differ; for $n > 1$, the hyperboloid model of $\mathbb{H}^n$ is **not conformal**.

Exercise 3.1.6. For, say, $n = 2$, compute an explicit example of a vector $\mathbf{v} \in H$ and tangent vectors $\mathbf{w}_1, \mathbf{w}_2 \in T_{\mathbf{v}}H$ such that the Minkowski angle $\cos^{-1}(\frac{B_J(\mathbf{w}_1, \mathbf{w}_2)}{|\mathbf{w}_1|_J \cdot |\mathbf{w}_2|_J})$ between $\mathbf{w}_1$ and $\mathbf{w}_2$ is different than the Euclidean angle.

Theorem 3.1.7. *Define $\phi$ to be the projection map $\{(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_{n+1} > 0\} \to \{(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_{n+1} = 0\}$ that takes $\mathbf{x}$ to the point on the line between $\mathbf{x}$ and $(0, \ldots, 0, -1)$ with last coordinate equal* *to 0. The restriction of $\phi$ to the hyperboloid $H$ is an isometry between the hyperboloid model of $\mathbb{H}^n$ and the Poincaré ball model of $\mathbb{H}^n$.*

We will prove the theorem in the case $n = 2$.

PROOF. Let $\mathbf{x} = (x_1, x_2, x_3)$. We have $\phi(\mathbf{x}) = (\frac{x_1}{1+x_3}, \frac{x_2}{1+x_3}, 0)$, so:

$$T_{\mathbf{x}}\phi = \begin{pmatrix} \frac{1}{1+x_3} & 0 & -\frac{x_1}{(1+x_3)^2} \\ 0 & \frac{1}{1+x_3} & -\frac{x_2}{(1+x_3)^2} \\ 0 & 0 & 0 \end{pmatrix}$$

It is easy to see that $\phi$ is a continuous map that gives a bijection from $H$ to $P := \{(y_1, y_2, 0) \mid \sqrt{y_1^2 + y_2^2} < 1\}$. We must show additionally that the derivative of $\phi$ at each point of $\mathbf{x} \in H$ preserves the hyperbolic inner product on $T_{\mathbf{x}}H$. That is, the derivative at each point must take the inner product coming from the restriction of $B_J$ to $T_{\mathbf{x}}H$ to the inner product defined on $T_{\phi(\mathbf{x})}P$ for the Poincaré disc.

Assume that $\mathbf{x} \in H$, so that $x_3 = \sqrt{1 + x_1^2 + x_2^2}$. The vectors $\mathbf{v} := (\frac{x_3}{\sqrt{1+x_2^2}}, 0, \frac{x_1}{\sqrt{1+x_2^2}})$ and $\mathbf{w} := (\frac{x_1 x_2}{\sqrt{1+x_2^2}}, \sqrt{1 + x_2^2}, \frac{x_2 x_3}{\sqrt{1+x_2^2}})$ satisfy:

$$\begin{pmatrix} B_J(\mathbf{v},\mathbf{v}) & B_J(\mathbf{v},\mathbf{w}) & B_J(\mathbf{v},\mathbf{x}) \\ B_J(\mathbf{w},\mathbf{v}) & B_J(\mathbf{w},\mathbf{w}) & B_J(\mathbf{w},\mathbf{x}) \\ B_J(\mathbf{x},\mathbf{v}) & B_J(\mathbf{x},\mathbf{w}) & B_J(\mathbf{x},\mathbf{x}) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

This means that given $\mathbf{x} \in H$, the pair $(\mathbf{v}, \mathbf{w})$ is an $B_J$–orthonormal basis for $T_{\mathbf{x}}H$. Because $T_{\mathbf{x}}\phi$ is linear, it is enough to check that it sends $(\mathbf{v}, \mathbf{w})$ to an orthonormal basis of $T_{\phi(\mathbf{x})}P$.

Compute, using the fact that $x_3^2 = 1 + x_1^2 + x_2^2$:

$$(T_{\mathbf{x}}\phi)(\mathbf{v}\,\mathbf{w}) = \frac{1}{1+x_3} \begin{pmatrix} \frac{1+x_2^2+x_3}{(1+x_3)\sqrt{1+x_2^2}} & \frac{x_1 x_2}{(1+x_3)\sqrt{1+x_2^2}} \\ -\frac{x_1 x_2}{(1+x_3)\sqrt{1+x_2^2}} & \frac{1+x_2^2+x_3}{(1+x_3)\sqrt{1+x_2^2}} \\ 0 & 0 \end{pmatrix}$$

Let $\mathbf{v}'$ and $\mathbf{w}'$ be the column vectors of the matrix on the right-hand side. They are Euclidean orthogonal unit vectors. Recall that the Poincaré ball metric was defined by rescaling the Euclidean inner product, so $(T_{\mathbf{x}}\phi)(\mathbf{v})$ and $(T_{\mathbf{x}}\phi)(\mathbf{w})$ are orthogonal, and it remains only to show that they are hyperbolic unit vectors. The scaling factor for the inner product on $T_{\mathbf{y}}P$ is

$\frac{4}{(1-|\mathbf{y}|^2)^2}$, and:

$$\begin{aligned}\frac{(1-|\phi(\mathbf{x})|^2)}{2} &= \frac{1}{2}(1-\frac{x_1^2}{(1+x_3)^2}-\frac{x_2^2}{(1+x_3)^2}-0^2)\\ &= \frac{1+2x_3+x_3^2-x_1^2-x_2^2}{2(1+x_3)^2}\\ &= \frac{1+2x_3+1}{2(1+x_3)^2} = \frac{1}{1+x_3}\end{aligned}$$

Thus $|T_{\mathbf{x}}\phi(\mathbf{v})|_{T_{\phi(\mathbf{x})}P} = \frac{2|T_{\mathbf{x}}\phi(\mathbf{v})|}{1-|\phi(\mathbf{x})|^2} = \frac{2(\frac{1}{1+x_3})|\mathbf{v}'|}{1-|\phi(\mathbf{x})|^2} = 1$, and similarly for $\mathbf{w}$. □

3.1.2. *The isometry group of* $\mathbb{H}^n$. A nice feature of the hyperboloid model is that isometries are restrictions of linear maps. A consequence of this fact will be a convenient description of hyperbolic polytopes in Section 3.2.

DEFINITION 3.1.8. The *Lorentz group* $\mathrm{O}(n,1)$ is the group of linear maps of $\mathbb{R}^{n+1}$ that preserve $B_J$. The *orthochronous Lorentz group* $\mathrm{O}^+(n,1)$ is the index 2 subgroup of $\mathrm{O}(n,1)$ that preserves $H$.

We show in Lemma 3.1.9 that $\mathrm{O}^+(n,1)$ is a subgroup of $\operatorname{Isom}\mathbb{H}^n$, and in Theorem 3.1.16 that it is all of $\operatorname{Isom}\mathbb{H}^n$.

LEMMA 3.1.9. $\mathrm{O}^+(n,1) < \operatorname{Isom}\mathbb{H}^n$

PROOF. Let $\mathbf{x}$ and $\mathbf{y}$ be points of $H$. For any $\epsilon > 0$ there exists a path $\gamma \subset H$ from $\mathbf{x}$ to $\mathbf{y}$ of length less than $d_{\mathbb{H}^n}(\mathbf{x},\mathbf{y})+\epsilon$. If $\phi \in \mathrm{O}^+(n,1)$ then $\phi\circ\gamma$ is a path in $H$ from $\phi(\mathbf{x})$ to $\phi(\mathbf{y})$. Since $\phi$ is linear it is its own derivative, so $(\phi\circ\gamma)'(t) = \phi(\gamma'(t))$, but since $\phi$ preserves $B_J$, this means $|(\phi\circ\gamma)'(t)| = |\gamma'(t)|$ for all $t$. Thus, $d_{\mathbb{H}^n}(\phi(\mathbf{x}),\phi(\mathbf{y})) \leqslant |\phi\circ\gamma| = |\gamma| < d_{\mathbb{H}^n}(\mathbf{x},\mathbf{y})+\epsilon$. Since this is true for all $\epsilon > 0$, $d_{\mathbb{H}^n}(\phi(\mathbf{x}),\phi(\mathbf{y})) \leqslant d_{\mathbb{H}^n}(\mathbf{x},\mathbf{y})$. Conversely, $\phi^{-1} \in \mathrm{O}^+(n,1)$, so the same argument gives:

$$d_{\mathbb{H}^n}(\mathbf{x},\mathbf{y}) = d_{\mathbb{H}^n}(\phi^{-1}(\phi(\mathbf{x})),\phi^{-1}(\phi(\mathbf{y}))) \leqslant d_{\mathbb{H}^n}(\phi(\mathbf{x}),\phi(\mathbf{y}))$$

We have shown that there is a map $\mathrm{O}^+(n,1) \to \operatorname{Isom}\mathbb{H}^n$ obtained by restricting an element of $\mathrm{O}^+(n,1)$ to $H$. We must also say that this map is injective. This follows because $H$ spans $\mathbb{R}^{n+1}$. For example, the following set is a basis for $\mathbb{R}^{n+1}$ contained in $H$:

$$\{(1,0,\ldots,0,\sqrt{2}),(0,1,0,\ldots,0,\sqrt{2}),\ldots,(0,\ldots,0,1,\sqrt{2}),(0,\ldots,0,1)\}$$

Thus, any linear map that fixes $H$ pointwise is the identity. □

LEMMA 3.1.10. *The stabilizer* $\mathrm{Stab}_{\mathrm{O}^+(n,1)}(\mathbf{v}_0)$ *of* $\mathbf{v}_0$ *in* $\mathrm{O}^+(n,1)$ *is:*

$$\left\{ \begin{pmatrix} M & 0 \\ 0 & 1 \end{pmatrix} \mid M \in \mathrm{O}(n) \right\}$$

PROOF. Suppose $M \in \mathrm{Stab}_{\mathrm{O}^+(n,1)}(\mathbf{v}_0)$. Since $M\mathbf{v}_0 = \mathbf{v}_0$ is the last column of $M$, $M$ has the following form, where $M'$ is an $n \times n$ matrix and $\mathbf{w}$ is an $n$–dimensional vector:

$$\begin{pmatrix} M' & \mathbf{0} \\ \mathbf{w}^T & 1 \end{pmatrix}$$

For $\mathbf{u} = \begin{pmatrix} \mathbf{u}' \\ u_{n+1} \end{pmatrix}$, we have:

$$|\mathbf{u}'|^2 - u_{n+1}^2 = B_J(\mathbf{u}, \mathbf{u}) = B_J(M\mathbf{u}, M\mathbf{u}) = |M'\mathbf{u}'|^2 - (\mathbf{w}^T\mathbf{u}' + u_{n+1})^2$$

Thus:

$$|M'\mathbf{u}'|^2 - |\mathbf{u}'|^2 = \mathbf{w}^T\mathbf{u}'(\mathbf{w}^T\mathbf{u}' + 2u_{n+1})$$

If $\mathbf{w} \neq \mathbf{0}$ then for any choice of $\mathbf{u}'$ not orthogonal to $\mathbf{w}$ there is a unique solution for $u_{n+1}$ in terms of $\mathbf{w}$ and $\mathbf{u}'$ that makes the equation true. This would be a contradiction, since the equation should be true for every choice of $\mathbf{u}'$ and $u_{n+1}$, so we must have $\mathbf{w} = \mathbf{0}$.

When $\mathbf{w} = \mathbf{0}$ the above equation reduces to $|M'\mathbf{u}'| = |\mathbf{u}'|$ for all $\mathbf{u}'$, which is the condition that $M' \in \mathrm{O}(n)$. □

COROLLARY 3.1.11. *For every* $\phi \in \mathrm{Stab}_{\mathrm{Isom}\,\mathbb{H}^n}(\mathbf{v}_0)$ *there exists a choice of* $\psi \in \mathrm{Stab}_{\mathrm{O}^+(n,1)}(\mathbf{v}_0)$ *with* $T_{\mathbf{v}_0}\phi = T_{\mathbf{v}_0}\psi$.

PROOF. The tangent space $T_{\mathbf{v}_0}\mathbb{H}^n$ to $H$ at $\mathbf{v}_0$ is the first $n$–coordinate subspace of $\mathbb{R}^{n+1}$, and the restriction of $B_J$ to this subspace is the standard Euclidean inner product, so $T_{\mathbf{v}_0}\phi \in \mathrm{O}(n)$, and for $\psi := \begin{pmatrix} T_{\mathbf{v}_0}\phi & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \in \mathrm{Stab}_{\mathrm{O}^+(n,1)}(\mathbf{v}_0)$ we have $T_{\mathbf{v}_0}\psi = \psi|_{T_{\mathbf{v}_0}\mathbb{H}^n} = T_{\mathbf{v}_0}\phi$. □

LEMMA 3.1.12. *For every* $h \geqslant 1$, $\mathrm{O}^+(n,1)$ *contains* $M_h := \begin{pmatrix} \mathrm{I}_{n-1} & \mathbf{0} \\ \mathbf{0} & M_h' \end{pmatrix}$, *where* $M_h' := \begin{pmatrix} h & \sqrt{h^2-1} \\ \sqrt{h^2-1} & h \end{pmatrix}$.

PROOF. It is a simple computation to check that these matrices preserve $B_J$. They also take $\mathbf{v}_0$ to a vector with last coordinate $h \geqslant 1$, so they do not exchange the two sheets of $f^{-1}(-1)$. □

REMARK. Note that $M_h'$ can be rewritten as $M_\alpha'' := \begin{pmatrix} \cosh\alpha & \sinh\alpha \\ \sinh\alpha & \cosh\alpha \end{pmatrix}$ for $\alpha = \cosh^{-1} h$, and $M_\alpha'' M_\beta'' = M_{\alpha+\beta}''$. In fact, it will turn out that $M_\alpha''$

is a translation along a geodesic by hyperbolic distance $\alpha$. See Proposition 3.1.17.

LEMMA 3.1.13. $\mathrm{O}^+(n,1)$ *acts transitively on* $\mathbb{H}^n$*.*

PROOF. We will show that an arbitrary point in $H$ can be sent to $\mathbf{v}_0$. Let $\mathbf{w} = (\mathbf{w}', h) \in H$. There is a unique point in $H$ with last coordinate 1, so if $h = 1$ then $\mathbf{w}' = \mathbf{0}$ and there is nothing to show.

If $h > 1$ then consider the affine hyperplane $A$ with last coordinate equal to $h$. The intersection of $A$ with $H$ is an $n$–sphere of radius $\sqrt{h^2 - 1}$: $\{(w_1, \ldots, w_n, h) \mid \sum_{i=1}^n w_i^2 = h^2 - 1\}$. By Lemma 3.1.10, $\mathrm{Stab}_{\mathrm{O}^+(n,1)}(\mathbf{v}_0)$ preserves this $n$–sphere and acts transitively on it, so there exists $N \in \mathrm{Stab}_{\mathrm{O}^+(n,1)}(\mathbf{v}_0)$ such that $N\mathbf{w} = (0, \ldots, 0, \sqrt{h^2 - 1}, h)$. Now for $M_h$ as in Lemma 3.1.12, $M_h^{-1} N\mathbf{w} = \mathbf{v}_0$. $\square$

THEOREM 3.1.14. *The intersection of* $H$ *with a plane through the origin is a geodesic in the hyperboloid model of* $\mathbb{H}^n$*. Furthermore, every geodesic is a segment of such an intersection.*

PROOF. Let $F$ be a plane through $\mathbf{0}$ that intersects $H$. Let $\mathbf{w} \in H \cap F$. By Lemma 3.1.13, there exists $N_1 \in \mathrm{O}^+(n,1)$ taking $\mathbf{w}$ to $\mathbf{v}_0$. By Lemma 3.1.10, there exists $N_2 \in \mathrm{Stab}_{\mathrm{O}^+(n,1)}(\mathbf{v}_0)$ that takes $N_1 F$ to the plane $F_1$ spanned by $\mathbf{v}_0$ and $\mathbf{v}_1 := (1, 0, \ldots, 0)$. The intersection $H \cap N_2 N_1 F$ can be parameterized as $\gamma(t) := (\sinh t)\mathbf{v}_1 + (\cosh t)\mathbf{v}_0 = (\sinh t, 0, \ldots, 0, \cosh t)$. Let $\phi\colon H \to P$ be the projection map of Theorem 3.1.7, so that $\phi \circ \gamma(t) = (\frac{\sinh t}{1+\cosh t}, 0, \ldots, 0)$. By Theorem 3.1.1, that is a hyperbolic geodesic in the Poincaré ball, because it is the intersection of $P$ with a line through the origin. Since Theorem 3.1.7 says $\phi$ is an isometry, $\gamma$ is a geodesic in the hyperboloid model. But $N_2 N_1 \in \mathrm{Isom}\, \mathbb{H}^n$, so $H \cap F = N_1^{-1} N_2^{-1} \gamma$ is also a geodesic in the hyperboloid model.

Now we want to say that every geodesic has this form. Let $\gamma$ be a geodesic of the hyperboloid model. Let $F$ be the plane in $\mathbb{R}^{n+1}$ spanned by $\gamma(0)$ and $\gamma'(0)$. By the previous argument, $H \cap F$ is a hyperbolic geodesic that goes through $\gamma(0)$ and can be parameterized to have constant speed with velocity $\gamma'(0)$ at $\gamma(0)$, so it shares a position and velocity vector with $\gamma$. Uniqueness of geodesics in the Poincaré model pulls back to the hyperboloid model, so $\gamma$ is a subsegment of the geodesic $H \cap F$. $\square$

COROLLARY 3.1.15. *For every* $\mathbf{v} \in H$ *there is a bijection* $\exp_{\mathbf{v}}\colon T_{\mathbf{v}}H \to H$*, called the* exponential map at $\mathbf{v}$*, defined by sending* $\mathbf{w} \in T_{\mathbf{v}}H$ *to the endpoint of the unique constant speed geodesic* $[0,1] \to H$ *with initial point* $\mathbf{v}$ *and initial velocity* $\mathbf{w} \in T_{\mathbf{v}}H$*.*

Proof. Define $\exp_{\mathbf{v}}(\mathbf{0}) = \mathbf{v}$. If $\mathbf{w} \in T_{\mathbf{v}}H - \{\mathbf{0}\}$ then $\mathbf{v}$ and $\mathbf{w}$ span a plane $F$ in $\mathbb{R}^{n+1}$ containing $\mathbf{v}$, and $F \cap H$ is a geodesic. There is a unique constant speed parameterization of $F \cap H$ for which $\mathbf{v}$ is the initial position and $\mathbf{w}$ is the initial velocity, and $\exp_{\mathbf{v}}(\mathbf{w})$ is the point on this geodesic at $t = |\mathbf{w}|_J$. Surjectivity of this map is the existence of geodesics from $\mathbf{v}$ to any point in $H$, and injectivity is uniqueness of that geodesic. $\square$

Theorem 3.1.16. $\operatorname{Isom}\mathbb{H}^n = \mathrm{O}^+(n,1)$

Proof. By Lemma 3.1.9, $\mathrm{O}^+(n,1) < \operatorname{Isom}\mathbb{H}^n$, so we need to show the opposite. Suppose $\phi \in \operatorname{Isom}\mathbb{H}^n$. Exponential maps satisfy:

$$\phi \circ \exp_{\mathbf{v}_0} \circ (T_{\mathbf{v}_0}\phi)^{-1} = \exp_{\phi(\mathbf{v}_0)} \tag{4}$$

By Lemma 3.1.13 and Corollary 3.1.11, by composing with an element of $\mathrm{O}^+(n,1)$, we may assume $\phi$ fixes $\mathbf{v}_0$ and that $T_{\mathbf{v}_0}\phi$ is the identity map on $T_{\mathbf{v}_0}H$. But then (4) says $\phi \circ \exp_{\mathbf{v}_0} = \exp_{\mathbf{v}_0}$, so $\phi$ is the identity map on $H$. Thus, $\phi$ agrees with the restriction to $H$ of an element of $\mathrm{O}^+(n,1)$, namely, the identity. $\square$

We now justify the remark following Lemma 3.1.12:

Proposition 3.1.17. *For $\alpha > 0$ the matrix* $M_\alpha := \begin{pmatrix} \mathrm{I}_{n-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \cosh\alpha & \sinh\alpha \\ \mathbf{0} & \sinh\alpha & \cosh\alpha \end{pmatrix}$ *is hyperbolic translation by distance $\alpha$ in the hyperboloid model of $\mathbb{H}^n$.*

Proof. Consider the curve $\gamma(t) = (0,\dots,0,\sinh t,\cosh t)$, which is contained in the intersection of $H$ and the plane $F$ spanned by $\mathbf{u} = (0,\dots,0,1,0)$ and $\mathbf{v} = (0,\dots,0,1)$, so it is a geodesic, by Theorem 3.1.14. Furthermore, $\gamma'(t) = (0,\dots,0,\cosh t,\sinh t)$ and $|\gamma'(t)|_{T_{\gamma(t)}\mathbb{H}^n} = \cosh^2 t - \sinh^2 t = 1$, so $\gamma$ has constant unit speed. Now observe that $M_\alpha\gamma(t) = \gamma(t+\alpha)$. $\square$

A consequence is a simple distance formula in the hyperboloid model:

Proposition 3.1.18. *For any $\mathbf{x},\mathbf{y} \in H$, $d_{\mathbb{H}^n}(\mathbf{x},\mathbf{y}) = \cosh^{-1}(-B_J(\mathbf{x},\mathbf{y}))$.*

Proof. Let $M \in \mathrm{O}^+(n,1)$ be such that $M\mathbf{x} = \mathbf{v}_0$ and $M\mathbf{y} = \mathbf{w} := (0,\dots,0,\sinh t,\cosh t)$ for some $t$. Then $B_J(\mathbf{v}_0,\mathbf{w}) = -\cosh t$, and, by Proposition 3.1.17, $t = d_{\mathbb{H}^n}(\mathbf{v}_0,\mathbf{w})$. By definition of $\mathrm{O}^+(n,1)$, $B_J(\mathbf{x},\mathbf{y}) = B_J(M\mathbf{x},M\mathbf{y}) = B_J(\mathbf{v}_0,\mathbf{w})$. Since $\mathrm{O}^+(n,1) = \operatorname{Isom}\mathbb{H}^n$, $d_{\mathbb{H}^n}(\mathbf{v}_0,\mathbf{w}) = d_{\mathbb{H}^n}(\mathbf{x},\mathbf{y})$. Thus:

$$d_{\mathbb{H}^n}(\mathbf{x},\mathbf{y}) = \cosh^{-1}(-(-\cosh t)) = \cosh^{-1}(-B_J(\mathbf{v}_0,\mathbf{w})) = \cosh^{-1}(-B_J(\mathbf{x},\mathbf{y}))$$

$\square$

We now come to the hyperbolic analogue of the fact Fact 3.0.2 that $(m+1)$–dimensional linear subspaces of $\mathbb{R}^{n+1}$ intersected with $\mathbb{S}^n$ correspond to isometrically embedded copies of $\mathbb{S}^m$. This is what we will use to define hyperbolic polytopes.

PROPOSITION 3.1.19. *A nonempty intersection of $H \subset \mathbb{R}^{n+1}$ with a linear subspace of dimension $m+1$ is an isometrically embedded copy of $\mathbb{H}^m$ in $\mathbb{H}^n$. Furthermore, every isometrically embedded copy of $\mathbb{H}^m$ in $\mathbb{H}^n$ arises in this way.*

PROOF. Let $F$ be a linear subspace of dimension $m+1$ that intersects $H$. Any distinct $\mathbf{x}, \mathbf{y} \in F \cap H$ span a plane in $\mathbb{R}^{n+1}$ that intersects $H$, and by Theorem 3.1.14, the hyperbolic geodesic between $\mathbf{x}$ and $\mathbf{y}$ is contained in that plane, so contained in $F \cap H$. Thus, $F \cap H$ is convex.

Arguing as in Theorem 3.1.14, we may assume, up to changing $F$ by an element of $\mathrm{O}^+(n,1)$, that $\mathbf{v}_0 \in F$. Since $F$ is $(m+1)$–dimensional containing $\mathbf{v}$, and $T_{\mathbf{v}_0}H$ is the $J$–orthogonal complement of $\mathbf{v}_0$, the intersection of $F$ with $T_{\mathbf{v}_0}H$ is $m$–dimensional. Since the stabilizer of $\mathbf{v}_0$ is the full orthogonal group on $T_{\mathbf{v}_0}H$, we may further assume that this $m$–dimensional subspace consists of the final $m$ coordinates, so that $F$ is the linear subspace of $\mathbb{R}^{n+1}$ consisting of the last $m+1$ coordinates. Let $K = F \cap H$. Consider the inclusion $\iota$ of $\mathbb{R}^{m+1}$ into $\mathbb{R}^{n-m} \oplus \mathbb{R}^{m+1} = \mathbb{R}^{n+1}$ as the final $m+1$ coordinates. This carries the hyperboloid in $\mathbb{R}^{m+1}$ bijectively to $K$, and the push-forward of the Minkowski form from $\mathbb{R}^{m+1}$ is the restriction of the Minkowski form of $\mathbb{R}^{n+1}$, so paths in $\mathbb{H}^m \subset \mathbb{R}^{m+1}$ have the same hyperbolic lengths as their images in $\mathbb{H}^n$. Thus, $\iota$ does not increase distances. Conceivably there could be shortcuts in $\mathbb{H}^n$ between points of $K$, but we have already shown that $K$ is convex, so that does not happen. This shows that $K$ is an isometrically embedded copy of $\mathbb{H}^m$ in $\mathbb{H}^n$.

Conversely, suppose that $K \subset H$ is the image of an isometric embedding of $\mathbb{H}^m$ into $\mathbb{H}^n$. Let $\mathbf{v} \in K$. The exponential map in $H$ takes the subspace $T_{\mathbf{v}}K$ of $T_{\mathbf{v}}H$ bijectively to $K$, but the bi-infinite geodesic starting at $\mathbf{v}$ with initial velocity $\mathbf{w}$ is the intersection of $H$ with the plane spanned by $\mathbf{v}$ and $\mathbf{w}$, so $K$ is the intersection of $H$ with the subspace of $\mathbb{R}^{n+1}$ spanned by $\mathbf{v}$ and $T_{\mathbf{v}}K$, which has dimension $m+1$, since $T_{\mathbf{v}}K$ has dimension $m$ and is $J$–orthogonal to $\mathbf{v}$. $\square$

Recall that a hyperbolic hyperplane is defined to be the intersection of the hyperboloid $H \subset \mathbb{R}^{n+1}$ with a linear hyperplane of $\mathbb{R}^{n+1}$, a hyperbolic halfspace is the closure of a complementary component of a hyperplane, and a hyperbolic convex polytope is the compact intersection of finitely many halfspaces.

EXERCISE 3.1.20. For $\mathbf{v} \in \mathbb{R}^{n+1} - \mathbf{0}$, show the hyperplane $\mathbf{v}^\perp$, the Euclidean orthogonal complement of $\mathbf{v}$, intersects $H$ if and only if $B_J(\mathbf{v}, \mathbf{v}) > 0$.

EXERCISE 3.1.21. Suppose that $\mathbf{v}$ and $\mathbf{w}$ are vectors in $\mathbb{R}^{n+1}$ such that their Euclidean orthogonal complements $\mathbf{v}^\perp$ and $\mathbf{w}^\perp$ are hyperplanes intersecting $H$ and each other. Show the hyperbolic dihedral angle between $H \cap \mathbf{v}^\perp$ and $H \cap \mathbf{w}^\perp$ supplements the Minkowski angle between $\mathbf{v}$ and $\mathbf{w}$.

### 3.2. Coxeter polytopes.

DEFINITION 3.2.1. A *geometric reflection group* is a group generated by reflections in the faces of a convex polytope $P$ in $\mathbb{X}^n$, for $\mathbb{X}^n$ one of $\mathbb{S}^n$, $\mathbb{E}^n$, or $\mathbb{H}^n$, such that $P$ is a strict fundamental domain for the action.

DEFINITION 3.2.2. A *Coxeter polytope* is a convex polytope that is the fundamental domain for the action of a geometric reflection group.

We will describe in this section restrictions on the shapes of Coxeter polytopes. In the spherical and Euclidean cases these are strong enough that Coxeter was able to give complete enumerations of them. The hyperbolic case is more mysterious.

To motivate the results, recall that in dimension 2 there are the following possibilities:

- A spherical Coxeter 2–polytope is a triangle whose dihedral angle sum is greater than $\pi$.
- A Euclidean Coxeter 2–polytope is either a triangle whose dihedral angle sum is $\pi$, or a rectangle.
- A hyperbolic Coxeter 2–polytope has no restriction on its combinatorial type, but if it is a triangle its dihedral angle sum is strictly less than $\pi$, and if it is a quadrilateral then its dihedral angle sum is strictly less than $2\pi$.

If $P$ is a convex polytope, a *supporting hyperplane* $H$ is a hyperplane that intersects $P$, such that $P$ is contained in one of the halfspaces of $H$. The intersection of $P$ and a supporting hyperplane is a *face* of $P$. Each face is again a convex polytope, in some lower dimension. Faces of dimension 0 are called *vertices*.

The faces of a convex polytope form a finite poset with respect to inclusion. Two convex polytopes have the same *combinatorial type* if they have isomorphic face posets.

A convex polytope is an *n–simplex* if it has the combinatorial type of a simplex in $\mathbb{R}^{n+1}$; that is, of a convex polytope in $\mathbb{R}^{n+1}$ defined by $n + 1$ hyperplanes whose normal vectors are linearly independent.

Lemma 3.2.3. *A Coxeter polytope $P$ has dihedral angles that are proper integral submultiples of $\pi$.*

Proof. Consider the tangent space to a point $x$ in the relative interior of the intersection of distinct two codimension 1 faces $F_i$ and $F_j$ of $P$. Let $\theta_{ij}$ be the dihedral angle between $F_i$ and $F_j$. For $k \in \{i, j\}$, let $\sigma_k$ be the reflection in $F_k$. These two reflections fix $x$, so the subgroup they generate tiles out a neighborhood of $x$ with copies of $P$. Consider the tangent space $T_x\mathbb{X}^n$ of $x$. There is a hyperplane $H_k$ in $T_x\mathbb{X}^n$ corresponding to $F_k$, with $T_x\sigma_k$ acting on $T_x\mathbb{X}^n$ as a reflection though $H_k$. Consider the plane $V$ in $T_x\mathbb{X}^n$ orthogonal to $H_i \cap H_j$. Then $H_k \cap V$ is a line, $T_x\sigma_k$ is reflection through that line, and the angle between $T_x\sigma_i$ and $T_x\sigma_j$ is $\theta_{ij}$. The other copies of $P$ from $\langle\sigma_i, \sigma_j\rangle P$ contribute sectors of angle $\theta_{ij}$ in $T_x\mathbb{X}^n$, filling out the entire $2\pi$ worth of angle around $\mathbf{0}$. Furthermore, there are an even number $2m_{ij}$ of these sectors, because adjacent sectors have alternating orientations. Also, $m_{ij} \geqslant 2$, since if there were two sectors of angle $\pi$ that would mean that $F_i$ and $F_j$ coincide. Thus, $2\pi = 2m\theta_{ij}$, or $\theta_{ij} = \pi/m_{ij}$, for $m_{ij} \geqslant 2$. □

A convex polytope is *non-obtuse* if its dihedral angles are at most $\pi/2$, so Lemma 3.2.3 implies[2] that a Coxeter polytope is non-obtuse. It turns out that there are strong restrictions on the possible shapes of non-obtuse convex polytopes, particularly in the spherical and Euclidean cases.

Definition 3.2.4. The *link* of a vertex $v$ of a convex polytope $P$ in $\mathbb{X}^n$ is a convex polytope in $\mathbb{S}^{n-1}$ obtained by taking all unit vectors $\mathbf{u}$ in $T_v\mathbb{X}^n$ such that $\mathbf{u}$ is inward pointing; that is, the geodesic in $\mathbb{X}^n$ with initial point $v$ and initial velocity $\mathbf{u}$ has a nontrivial initial segment contained in $P$.

Definition 3.2.5. A convex polytope is *simple* if the link of every vertex is a spherical simplex.

The Platonic solids of Exercise 4.2.10 are convex polytopes in $\mathbb{E}^3$. The link of a vertex of each of them is illustrated in Figure 12. The tetrahedron,

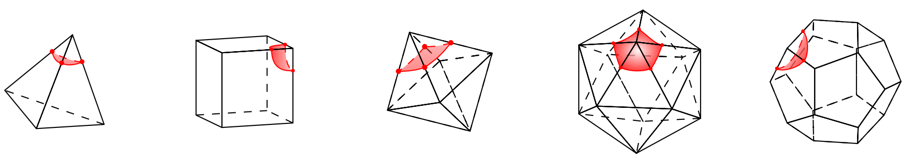


Figure 12. Links of vertices in the Platonic solids

cube, and dodecahedron are simple: their vertices all have links that are

[2]This isn't quite true, because we have only considered codimension 1 faces that intersect in a codimension 2 face. That turns out to be sufficient. See [**11**, Proposition 6.3.2].

2–simplices. The octahedron and icosahedron are not simple, since their vertices have links that are (filled) squares and pentagons, respectively.

Recall that in Section 2.2 we showed that a 2–dimensional spherical non-obtuse convex polytope must be a triangle. A similar result is true in higher dimensions:

LEMMA 3.2.6 ([**11**, Lemma 6.3.3]). *If $P$ is a non-obtuse convex polytope in $\mathbb{S}^n$ then $P$ is a simplex.*

COROLLARY 3.2.7 ([**11**, Proposition 6.3.9]). *If $P$ is a non-obtuse convex polytope in $\mathbb{S}^n$, $\mathbb{E}^n$, or $\mathbb{H}^n$ then $P$ is simple.*

PROOF. The link $L$ of a vertex in $P$ is a convex spherical polytope of one lower dimension, whose dihedral angles are a subset of the dihedral angles of $P$, so it is also non-obtuse. By Lemma 3.2.6, $L$ is a spherical simplex. □

A similar result to Lemma 3.2.6 was true in $\mathbb{E}^2$, except that there was an additional possibility that the polytope was a product of lower dimensional Euclidean polytopes. This result also holds in higher dimensions:

PROPOSITION 3.2.8 ([**11**, Corollary 6.3.11]). *If $P$ is a non-obtuse convex polytope in $\mathbb{E}^n$ then $P$ is a product of simplices.*

Look back at the simple Platonic solids from Figure 12 in light of Proposition 3.2.8: the tetrahedron is a Euclidean 3–simplex, the cube is a product of three 1–simplices, and the dodecahedron happens to be simple even though it is not non-obtuse. Another way to make a non-obtuse convex Euclidean polytope in dimension three would be to take a triangular prism, the product of a 1–simplex and a non-obtuse 2–simplex.

So far, a Coxeter polytope is a simple convex polytope whose dihedral angles are nontrivial integral submultiples of $\pi$. The next result says that these conditions are sufficient for the polytope $P$ to be a Coxeter polytope; that is, there really is a geometric reflection group on $\mathbb{X}^n$ generated by reflections through the faces of $P$ such that $P$ is a strict fundamental domain for the action.

THEOREM 3.2.9 ([**11**, Theorem 6.4.3]). *Let $P$ be a simple convex polytope in $\mathbb{X}^n$, where $\mathbb{X}^n$ is one of $\mathbb{S}^n$, $\mathbb{E}^n$, or $\mathbb{H}^n$ for $n \geqslant 2$. Suppose the dihedral angles of $P$ are integer submultiples of $\pi$. For codimension 1 faces $F_i$ and $F_j$ of $P$, if $F_i \cap F_j \neq \varnothing$ then let $m_{ij}$ be such that the dihedral angle between $F_i$ and $F_j$ is $\pi/m_i$. If $F_i \cap F_j = \varnothing$ let $m_{ij} = \infty$. Let $W$ be the Coxeter group defined by Coxeter matrix $(m_{ij})$. Let $\bar{W}$ be the group generated by reflections in the codimension 1 faces of $P$. Then the natural map $\mathcal{U}(W, P) \to \mathbb{X}^n$ is a homeomorphism and the natural surjection $W \twoheadrightarrow \bar{W}$ is an isomorphism.*

The idea of the proof is, like in Theorem 2.5.9, to argue that $\mathcal{U}(W, P)$ is locally isometric to $\mathbb{X}^n$, and then conclude that the natural map is a trivial covering. The locally isometric property is proved by induction on dimension. To get the idea, consider dimension 3. From the consideration of dihedral angles, every point of $\mathcal{U}(W, P)$ has a neighborhood isometric to the neighborhood of a point in $\mathbb{X}^3$ except possibly the vertices. Look at the neighborhood a vertex $v$. Since $P$ is simple, every copy of $P$ at $v$ has link at $v$ a spherical simplex. We would like to know that neighborhoods of $v$ in the copies of $P$ glue up to form a neighborhood of $v$ that is isometric to the neighborhood of a point in $\mathbb{X}^3$. The first claim is that this is true if all of the links glue together to make a 2–sphere. The second claim is that if we apply Theorem 2.5.9 to the group generated by reflection through faces containing $v$, then that gives an $\mathbb{S}^2$ reflection group whose fundamental domain is the link of $v$ in $P$, so the links do glue up to form a 2–sphere.

From Theorem 3.2.9, the same argument as in the 2–dimensional case gives:

Corollary 3.2.10. *Every geometric reflection group is a Coxeter group.*

There is a partial converse to this result called Lannér's Theorem.

Definition 3.2.11. A *simplicial Coxeter group* $W$ is one that acts properly on $\mathcal{U}(W, \Delta^n)$, where $\Delta^n$ is an $n$–simplex.

Theorem 3.2.12 (Lannér's Theorem [**11**, Theorem 6.9.1]). *If $W$ is a simplicial Coxeter group then it is a geometric reflection group with fundamental domain a simplex in either $\mathbb{S}^n$, $\mathbb{E}^n$, or $\mathbb{H}^n$.*

Lannér's Theorem is a *geometrization* theorem; it says that a topological hypothesis, properness of the $W$ action on $\mathcal{U}(W, \Delta^n)$, is enough to guarantee that it is possible to upgrade the topological simplex $\Delta^n$ to a geometric simplex without breaking the group action or the combinatorics of $\mathcal{U}(W, \Delta^n)$. Contrast with the examples of hyperbolic triangle groups with an infinite entry, as seen in the last column and row of Figure 10. The combinatorics of these figures are examples of a $\mathcal{U}(W, \Delta^2)$, but these complexes have infinite valence vertices, the $W$ action is not proper, and the group $W$ is not a geometric reflection group.

Recall that Theorem 1.0.2 says the 1–dimensional geometric reflection groups are exactly the dihedral groups of even order.

Theorem 2.6.1 says the irreducible 2–dimensional geometric reflection groups are either triangle groups with finite entries or they are hyperbolic with presentation graph a cycle. The triangle groups with finite entries are precisely the 2–dimensional simplicial Coxeter groups.

In higher dimensions we have, from Theorem 3.2.12, that simplicial Coxeter groups are geometric, and, from Lemma 3.2.6, Proposition 3.2.8, and Theorem 3.2.9, that irreducible spherical and Euclidean reflection groups are simplicial Coxeter groups. The spherical ones will be classified in Table 2.2 in Section 4.1. The Euclidean ones will be classified in Table 2.4 in Section 4.3.

Hyperbolic reflection groups need not be simplicial. In fact, there do not exist hyperbolic Coxeter simplices above dimension 4. Let us show that there are hyperbolic Coxeter simplices in dimension 3. A 3–dimensional simplex is combinatorially a tetrahedron. It has 4 faces, $f_i$ for $1 \leqslant i \leqslant 4$. For a moment, let us forget the fact that that dihedral angles should be integral submultiples of $\pi$ and just consider non-obtuse dihedral angles, so for all distinct $i$ and $j$, let $0 < \theta_{ji} = \theta_{ij} \leqslant \pi/2$. There are some easy constraints on the allowed angles. For distinct $i, j, k, \ell$, let $\phi_{ij,ik}$ be the angle in $f_i$ between the edges $f_i \cap f_j$ and $f_i \cap f_k$. The face $f_i$ is a geodesic triangle in an isometrically embedded copy of $\mathbb{H}^2$, so $\phi_{ij,ik} + \phi_{ik,i\ell} + \phi_{i\ell,ij} < \pi$. Also, the link of the vertex $f_i \cap f_j \cap f_k$ is a spherical simplex with angles $\theta_{ij}$, $\theta_{jk}$, and $\theta_{ki}$ and side lengths $\phi_{ij,ik}$, $\phi_{ij,jk}$, and $\phi_{jk,ij}$. Since it is a spherical simplex, $\theta_{ij} + \theta_{jk} + \theta_{ki} > \pi$. It turns out that these are the only constraints:

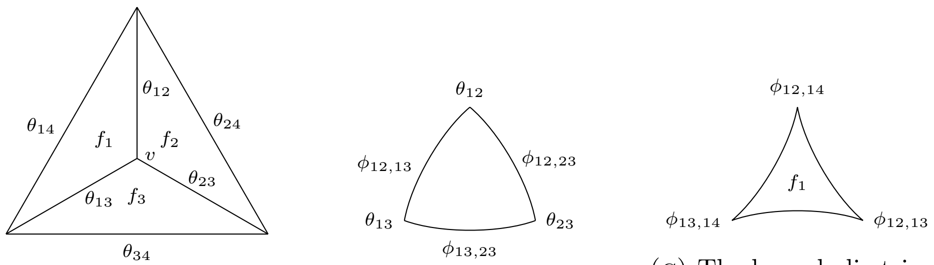
Theorem 3.2.13. *For distinct $i, j, k \in \{1, 2, 3, 4\}$, let $0 < \theta_{ji} = \theta_{ij} \leqslant \pi/2$, and define $\phi_{ij,ik} := \cos^{-1}\left(\frac{\cos\theta_{jk} + \cos\theta_{ij}\cos\theta_{ik}}{\sin\theta_{ij}\sin\theta_{ik}}\right)$.*

*The $\theta_{ij}$ are the dihedral angles of a hyperbolic tetrahedron if and only if for all distinct $i, j, k, \ell$:*

(1) $\theta_{ij} + \theta_{jk} + \theta_{ki} > \pi$
(2) $\phi_{ij,ik} + \phi_{ik,i\ell} + \phi_{i\ell,ij} < \pi$

*Furthermore, when the tetrahedron exists it is unique, up to hyperbolic isometry.*

Proof. Suppose a compact hyperbolic tetrahedron with the given dihedral angles exists. A schematic is shown in Figure 13. The tetrahedron with faces and dihedral angles in shown in Figure 13a. The link of the vertex $v = f_1 \cap f_2 \cap f_3$ is a spherical simplex with dihedral angles $\theta_{12}$, $\theta_{13}$, and $\theta_{14}$. This shows the necessity of condition (1). The side lengths of a spherical triangle can be computed in terms of the dihedral angles using the Spherical Law of Cosines, which is the relationship defining the $\phi_{ij,ik}$. This is depicted in Figure 13b. The $\phi_{ij,ik}$ are then the dihedral angles of the hyperbolic triangular faces of the tetrahedron, as shown in Figure 13c. Since these are hyperbolic triangles, this shows the necessity of condition (2). We also remark that condition (1) implies that the $\phi_{ij,ik}$ are all positive, since in the case $\phi_{ij,ik} = 0$ for an ideal tetrahedron, the corresponding ideal vertex link is a Euclidean triangle, not a spherical one.




(A) Tetrahedron with dihedral angles. $f_4$ is the rear face.

(B) The link of vertex $v$.

(C) The hyperbolic triangular face $f_1$.

FIGURE 13. Schematic of a hyperbolic tetrahedron.

Conversely, given $\theta$'s and $\phi$'s as above, there exist hyperbolic geodesic triangles $f_1$, $f_2$, $f_3$, and $f_4$ such that the angles of $f_i$ are $\phi_{ij,ik}$ $\phi_{ik,i\ell}$, $\phi_{i\ell,ij}$ and spherical triangles with angles $\theta_{ij}$, $\theta_{jk}$, and $\theta_{ki}$ and side lengths $\phi_{ij,ik}$, $\phi_{ij,jk}$, and $\phi_{jk,ij}$. It remains to show that there actually exists a hyperbolic tetrahedron with the given dihedral angles. Consider the following vectors in $\mathbb{R}^4$:

$$\mathbf{n}_1 := \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad \mathbf{n}_3 := \begin{pmatrix} \sin\theta_{13}\sin\phi_{12,13} \\ -\sin\theta_{13}\cos\phi_{12,13} \\ -\cos\theta_{13} \\ 0 \end{pmatrix}$$

$$\mathbf{n}_4 := \begin{pmatrix} -\sin\theta_{14}\sin\phi_{12,14}\cosh|f_1 \cap f_2| \\ -\sin\theta_{14}\cos\phi_{12,14} \\ -\cos\theta_{14} \\ \sin\theta_{14}\sin\phi_{12,14}\sinh|f_1 \cap f_2| \end{pmatrix} \qquad \mathbf{n}_2 := \begin{pmatrix} 0 \\ \sin\theta_{12} \\ -\cos\theta_{12} \\ 0 \end{pmatrix}$$

We claim that the halfspaces corresponding to the $\mathbf{n}_i$ side of $\mathbf{n}_i^\perp$ intersect to give a hyperbolic tetrahedron with the desired dihedral angles. To see why this is so, and where the formulae for the $\mathbf{n}_i$ came from, suppose that such a tetrahedron exists. Up to isometry, we may assume the three corners of $f_1$, which are $f_1 \cap f_2 \cap f_3$ and $f_1 \cap f_2 \cap f_4$ and $f_1 \cap f_3 \cap f_4$, are the points $\mathbf{v}_0 = (0,0,0,1)$, $\mathbf{v}_1 = (\sinh t, 0, 0, \cosh t)$ for some $t > 0$, and $\mathbf{v}_2 = (p, q, 0, \sqrt{1+p^2+q^2})$ for some $p$ and some $q > 0$, respectively. Finally, up to isometry not moving the previous three points, we may assume that the third coordinate of $\mathbf{v}_3 := f_2 \cap f_3 \cap f_4$ is positive. We will show that these choices uniquely determine the $\mathbf{n}_i$, up to scaling, which does not change the subspace $\mathbf{n}_i^\perp$. To get $\mathbf{n}_i \cap H \neq \varnothing$, we have $B_J(\mathbf{n}_i, \mathbf{n}_i) > 0$, so we may rescale and assume it is 1.

Note that $s \mapsto (\sinh s, 0, 0, \cosh s)$ is a geodesic, so $t = |f_1 \cap f_2|$ is the length of the edge $f_1 \cap f_2$. This is determined, in terms of the angles of $f_1$, by the Hyperbolic Law of Cosines:

$$\cosh t = \frac{\cos\phi_{13,14} + \cos\phi_{12,13}\cos\phi_{12,14}}{\sin\phi_{12,13}\sin\phi_{12,14}} \tag{5}$$

We could work out $p$ and $q$ as well, but we won't need these, except to know $q > 0$.

If $\mathbf{n}_1 = (a, b, c, d)$ then $\mathbf{v}_0 \in \mathbf{n}_1^\perp \implies d = 0$. Then it follows that $\mathbf{v}_1 \in \mathbf{n}_1^\perp \implies a = 0$ and then $\mathbf{v}_2 \in \mathbf{n}_1^\perp \implies b = 0$. Since $\mathbf{v}_3$ should be on the positive side of $\mathbf{n}_1^\perp$, $0 < \mathbf{n}_1 \cdot \mathbf{v}_3$, this is $c$ times the third coordinate of $\mathbf{v}_3$, which was assumed positive, so $c > 0$. Now $|\mathbf{n}_1| = 1 \implies c = 1$.

Suppose $\mathbf{n}_2 = (a, b, c, d)$. Since $\mathbf{v}_0, \mathbf{v}_1 \in \mathbf{n}_2^\perp$, $d = 0$ and $a \sinh t = 0 \implies a = 0$. To get the dihedral angle between $f_1$ and $f_2$ use Exercise 3.1.21:

$$\begin{aligned}
\theta_{12} &= \pi - \arccos\left(\frac{B_J(\mathbf{n}_1, \mathbf{n}_2)}{|\mathbf{n}_1|_J |\mathbf{n}_2|_J}\right) \\
&= \arccos(-B_J(\mathbf{n}_1, \mathbf{n}_2)) \\
&= \arccos(-c) \\
&\implies c = -\cos\theta_{12}
\end{aligned}$$

Since $\mathbf{n}_2$ is supposed to be inward pointing, $\mathbf{v}_2$ must be on the positive $\mathbf{n}_2$ side of $\mathbf{n}_2^\perp$, so $0 < \mathbf{n}_2 \cdot \mathbf{v}_2 = bq \overset{q>0}{\implies} b > 0$. Since $1 = |\mathbf{n}_2|^2 = b^2 + \cos^2\theta_{12}$ and $b > 0$, we take $b := \sin\theta_{12}$.

Suppose $\mathbf{n}_3 = (a, b, c, d)$. $\mathbf{v}_0 \in \mathbf{n}_3^\perp$, so $d = 0$. Arguments similar to $\mathbf{n}_2$ case show $c = -\cos\theta_{13}$ and

$$b = -\sin\theta_{13}\frac{\cos\theta_{23} + \cos\theta_{12}\cos\theta_{13}}{\sin\theta_{12}\sin\theta_{13}} = -\sin\theta_{13}\cos\phi_{12,13}$$

Now using $|\mathbf{n}_3| = 1$ and the fact that $\mathbf{v}_1$ is on the positive side of $\mathbf{n}_3^\perp$, so that $0 < \mathbf{v}_1 \cdot \mathbf{n}_3 = a \sinh t \implies a > 0$, gives $a = \sin\theta_{13}\sin\phi_{12,13}$.

Similar arguments give $\mathbf{n}_4$. $\square$

EXERCISE 3.2.14. Derive the equations expressions for $\mathbf{n}_4$ in the proof of Theorem 3.2.13.

EXAMPLE 3.2.15. Consider points $\mathbf{v}_0 := (0, 0, 0, 1)$, $\mathbf{v}_1 := (\sinh t, 0, 0, \cosh t)$, $\mathbf{v}_2 := (0, \sinh t, 0, \cosh t)$, and $\mathbf{v}_3 := (0, 0, \sinh t, \cosh t)$ for $0 < t < \infty$. These are the vertices of some hyperbolic tetrahedron.

Face $f_4$ is contained in $H \cap F$ for some linear subspace $F$ containing $v_1$, $v_2$, and $v_3$. So $\mathbf{n}_4 = (a, b, c, d)$ such that $B_J(\mathbf{n}_4, \mathbf{n}_4) = 1$ and $v_1$, $v_2$, and $v_3$ are Euclidean-orthogonal to $\mathbf{n}_4$. If $\mathbf{n}_4 = (a, b, c, d)$ then we find

$a = b = c = -d \coth t$. We should also have $d = \mathbf{v}_0 \cdot \mathbf{n}_4 > 0$ so that $\mathbf{n}_4$ is inward pointing. Thus, $B_J(\mathbf{n}_4, \mathbf{n}_4) = 1$ implies:

$$\mathbf{n}_4 = \begin{pmatrix} -\frac{\cosh t}{\sqrt{1+2\cosh^2 t}} \\ -\frac{\cosh t}{\sqrt{1+2\cosh^2 t}} \\ -\frac{\cosh t}{\sqrt{1+2\cosh^2 t}} \\ \frac{\sinh t}{\sqrt{1+2\cosh^t t}} \end{pmatrix}$$

Similar considerations give:

$$\mathbf{n}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{n}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{n}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Now we compute dihedral angles:

$$\theta_{12} = \theta_{13} = \theta_{23} = \pi/2$$

$$\theta := \theta_{14} = \theta_{24} = \theta_{34} = \pi - \arccos\left(-\frac{\cosh t}{\sqrt{1+2\cosh^2 t}}\right) = \arccos\left(\frac{\cosh t}{\sqrt{1+2\cosh^2 t}}\right)$$

We should expect that for $t$ close to 0 this tetrahedrahedron will be close to being Euclidean, and, indeed, $\lim_{t\to 0^+} \theta = \arccos(1/\sqrt{3})$, which is the value of the non-right dihedral angles of the Euclidean tetrahedron with vertices $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$.

In the other direction, $\lim_{t\to\infty} \theta = \arccos(1/\sqrt{2}) = \pi/4$. There is an *ideal* convex hyperbolic tetrahedron with dihedral angles $\pi/2$, $\pi/2$, $\pi/2$, $\pi/4$, $\pi/4$, $\pi/4$, but there is not a compact one.

For any $\theta$ such that $\pi/3 > \arccos(1/\sqrt{3}) > \theta > \pi/4$ there is a convex hyperbolic tetrahedron with dihedral angles $\pi/2$, $\pi/2$, $\pi/2$, $\theta$, $\theta$, $\theta$.

EXERCISE 3.2.16. Check that for $\pi/3 > \arccos(1/\sqrt{3}) > \theta > \pi/4$, using $\theta_{12} = \theta_{13} = \theta_{23} = \pi/2$ and $\theta_{14} = \theta_{24} = \theta_{34} = \theta$ satisfies the requirements of Theorem 3.2.13, and that these bounds are sharp, in the sense that as $\theta$ approaches $\arccos(1/\sqrt{3})$ or $\pi/4$ one of the hypotheses of Theorem 3.2.13 approaches failure.

Check that Theorem 3.2.13 gives the same four normal vectors defining the faces as computed in Example 3.2.15.

EXERCISE 3.2.17. Show there is no hyperbolic tetrahedron such that some face has all of its dihedral angles equal to $\pi/2$.

EXERCISE 3.2.18. Consider a hyperbolic Coxeter tetrahedron. Show:

- At every vertex there is at least one edge with dihedral angle $\pi/2$.

- There do not exist adjacent edges that both have dihedral angle strictly less than $\pi/3$.
- All dihedral angles are strictly greater than $\pi/6$.
- Some dihedral angle is strictly less than $\pi/3$.

EXERCISE 3.2.19. Translate the conditions of the previous two exercises to given conditions on the Coxeter graph of a 3–dimensional hyperbolic simplicial Coxeter group. (These conditions do not give a complete description of the 3–dimensional simplicial hyperbolic Coxeter groups. They leave 14 candidates. 5 more can be ruled out using Theorem 3.2.13, leaving the 9 examples shown in Table 2.3.)

Andreev's Theorem says that the combinatorial type of any simple 3–dimensional polytope with at least 5 faces is realizable as a hyperbolic polytope with non-obtuse dihedral angles. Moreover, there are linear conditions on non-obtuse dihedral angles, such that a polytope with these angles exists if and only if the conditions are satisfied, and in this case the polytope is unique up to isometry. See Davis [**11**, Theorem 6.10.2], and the surrounding discussion that explains how Andreev's conditions translate into hyperbolic geometry. In relation to Theorem 3.2.13, Andreev's conditions keep the first condition, drop the condition on face angles, and add some other conditions that are vacuous for the tetrahedron. The exclusion of the tetrahedron is important, even though the existence conclusion is the same. The point is that in the proof of Andreev's theorem one shows that for a fixed combinatorial type of simple 3–dimensional polytope, the space of allowed non-obtuse dihedral angles is itself a convex polytope whose sides are defined by the (linear) Andreev conditions. This is not true for the tetrahedron; the second condition of Theorem 3.2.13, when rewritten in terms of the dihedral angles, is not linear, and the space of possible hyperbolic tetrahedra with non-obtuse dihedral angles is not convex [**20**].

It turns out that there are also hyperbolic Coxeter simplices in dimension 4, but not in higher dimensions. The hyperbolic Coxeter simplices have been classified, see Table 2.3. In higher dimensions there are infinitely many isometry types of hyperbolic Coxeter polytopes up to dimension 6 [**1**]. In dimensions 7 and 8 there are examples of hyperbolic Coxeter polytopes, but it is not known if there are infinitely many isometry classes. In high enough dimension, there are no hyperbolic Coxeter polytopes at all:

THEOREM 3.2.20 (Vinberg's Theorem [**11**, Theorem 6.11.8]). *If $P$ is a convex polytope in $\mathbb{H}^n$ with all dihedral angles proper integral submultiples of $\pi$ then $n < 30$. If all dihedral angles are $\pi/2$ then $n \leqslant 4$.*

## 4. The classification of simplicial geometric reflection groups

We have seen, in Corollary 3.2.10, that every geometric reflection group is a Coxeter group. By analyzing the shapes of possible Coxeter polytopes, we have seen that every spherical reflection group is simplicial, Lemma 3.2.6, and every irreducible Euclidean reflection group is simplicial, Proposition 3.2.8. There are both simplicial and non-simplicial hyperbolic reflection groups, and, as our lack of knowledge about their existence in dimensions between 8 and 29 suggests, the non-simplicial ones are mysterious.

In this section we will realize the benefit of choosing geometric model spaces in which hyperplanes can be described by a single normal vector by describing geometric simplices in linear algebraic terms. This will lead to a classification of simplicial geometric reflection groups.

Recall that a convex polytope in $\mathbb{E}^n$, $\mathbb{S}^n$, or $\mathbb{H}^n$ is a *simplex* if it has the combinatorial type of a simplex in $\mathbb{R}^{n+1}$. A codimension 1 face $\sigma_i$ of a polytope $\sigma$ is contained in a hyperplane defined as the Euclidean-orthogonal space to an inward pointing unit normal vector $\mathbf{n}_i$.

Definition 4.0.1. The *Gram matrix* of a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is the matrix of their inner products $(\langle \mathbf{v}_i, \mathbf{v}_j \rangle)$. (In the hyperbolic case, this is the inner product with respect to the Minkowski form.) The *Gram matrix of a simplex* $\sigma$ is defined to be the Gram matrix of its set of inward pointing unit normal vectors.

Definition 4.0.2. If $M = (m_{ij})$ for $i, j \in I$ is a Coxeter matrix, define its *cosine matrix* to be $C := (c_{ij})$ with $c_{ij} := -\cos \frac{\pi}{m_{ij}}$, where $c_{ij} = -1 = -\cos 0$ if $m_{ij} = \infty$.

If $\theta_{ij}$ is the dihedral angle between codimension 1 faces of a simplex, then the inner product of the corresponding inward pointing unit normals satisfies $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \cos \pi - \theta_{ij} = -\cos \theta_{ij}$, so the product of reflections through those faces has order $m_{ij}$ for $\theta_{ij} = \pi/m_{ij}$. Thus, from what we have done so far, $(W, S)$ is a Coxeter system that acts as a geometric reflection group with elements of $S$ corresponding to reflections through the codimension 1 faces of a simplex $\sigma$ if and only if the cosine matrix of $(W, S)$ is the Gram matrix of $\sigma$. Our goal now is to identify properties of a matrix that would let us conclude that it is the Gram matrix of some geometric simplex.

We need some linear algebra:

Definition 4.0.3. An $n \times n$ real symmetric matrix $M$ is *positive semi-definite* if for all $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v}^T M \mathbf{v} \geqslant 0$. It is *positive definite* if it is positive semi-definite and equality in the condition is only achieved for $\mathbf{v} = \mathbf{0}$.

THEOREM 4.0.4. *If $M$ is a positive semi-definite matrix then there exists a unique positive semi-definite matrix $S$, called the* square root of $M$, *such that $S^2 = S^T S = M$. If $M$ is positive definite then so is its square root.*

PROOF SKETCH. Let $E$ be a matrix whose columns are an orthonormal basis consisting of eigenvectors $\mathbf{v}_i$ of $M$, let $\lambda_i$ be the eigenvalue corresponding to eigenvector $\mathbf{v}_i$, let $\Lambda$ be the diagonal matrix with entries $\lambda_i$, and let $\sqrt{\Lambda}$ be the diagonal matrix with entries $\sqrt{\lambda_i}$. Define $\sqrt{M} := E\sqrt{\Lambda}E^{-1}$. □

Here is what we will prove:

THEOREM 4.0.5. *Let $(W, S)$ be a Coxeter system with $S = \{s_i \mid i \in I\}$ and cosine matrix $C$. Then $W$ acts as a simplicial reflection group on $\mathbb{X}^n$, with each $s_i$ acting by reflection across a codimension 1 face and $n = |S| - 1$, in precisely the following cases:*

- *$\mathbb{X}^n = \mathbb{S}^n$ and $C$ is positive definite.*
- *$\mathbb{X}^n = \mathbb{H}^n$ and $C$ is of type $(n, 1)$ with every principal submatrix postive defininte.*
- *$\mathbb{X}^n = \mathbb{E}^n$ and $C$ is positive semi-definite of corank 1.*

The conditions on the cosine matrix are concrete (see Theorem 4.0.6), and have been used to give a complete enumeration of the possible Coxeter groups that can achieve them, so we have explicit lists of all possible simplicial geometric reflection groups.

THEOREM 4.0.6 (Sylvester's criteria). *An $n \times n$ matrix $C$ is positive definite if and only if for all $1 \leqslant i \leqslant n$ the upper-left $i \times i$ submatrix $C_i$ of $C$ has positive determinant.*

The matrices $C_i$ in the theorem are called the *principle submatrices*, and their determinants are called the *principle minors*.

EXERCISE 4.0.7. Show that the cosine matrix is positive definite for the Coxeter presentation of the symmetric group. (Recall Exercise 4.2.11.)

EXERCISE 4.0.8. Show that the cosine matrix of $\Delta(p, q, r)$ is positive definite if and only if $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} > 1$.

### 4.1. Spherical simplices.

LEMMA 4.1.1. *Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_{n+1}\}$ in $\mathbb{R}^{n+1}$ be a set of unit vectors. Let $H_i := \{\mathbf{v} \mid < \mathbf{v}, \mathbf{u}_i > \geqslant 0\}$ be the halfspace with bounding hyperplane $\mathbf{u}_i^{\perp}$ containing $\mathbf{u}_i$. Then $\sigma := \mathbb{S}^n \cap \bigcap_{i=1}^{n+1} H_i$ is a spherical $n$–simplex if and only if $\{\mathbf{u}_1, \ldots, \mathbf{u}_{n+1}\}$ is a basis of $\mathbb{R}^{n+1}$.*

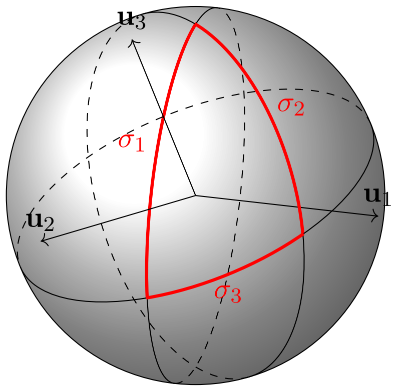



FIGURE 14. The codimension 1 faces and inward pointing unit normals for a spherical simplex in $\mathbb{S}^2$.

PROOF. Suppose $\sigma$ is a spherical $n$–simplex. Then $\bigcap_{i=1}^{n+1} H_i$ is $(n+1)$–dimensional simplicial cone, so $\{\mathbf{u}_1, \ldots, \mathbf{u}_{n+1}\}$ is linearly independent, hence a basis. Conversely, if $\{\mathbf{u}_1, \ldots, \mathbf{u}_{n+1}\}$ is a basis then the matrix $U$ whose columns are the $\mathbf{u}_i$ is an element of $\mathrm{GL}_{n+1}\mathbb{R}$ that takes the closed positive orthant to a simplicial cone such that the $\{\mathbf{u}_1, \ldots, \mathbf{u}_{n+1}\}$ are inward pointing normals to the codimension 1 faces. The intersection of $\mathbb{S}^n$ with this simplicial cone is a spherical $n$–simplex. $\square$

LEMMA 4.1.2. *A spherical simplex is determined up to isometry by its Gram matrix, or, equivalently, by its dihedral angles.*

PROOF. Suppose $\sigma$ and $\sigma'$ are two spherical simplices so that the corresponding sets of inward pointing unit normals are $\{\mathbf{n}_1, \ldots, \mathbf{n}_{n+1}\}$ and $\{\mathbf{n}'_1, \ldots, \mathbf{n}'_{n+1}\}$, and so that the Gram matrices agree. Since each set of inward pointing unit normals is a basis of $\mathbb{R}^{n+1}$ there is a unique element $\phi \in \mathrm{GL}(\mathbb{R}^{n+1})$ such that $\phi(\mathbf{n}_i) = \mathbf{n}'_i$ for all $i$. Since the Gram matrices agree, for all $i, j$ we have:

$$\langle \mathbf{n}_i, \mathbf{n}_j \rangle = \langle \mathbf{n}'_i, \mathbf{n}'_j \rangle = \langle \phi(\mathbf{n}_i), \phi(\mathbf{n}_j) \rangle$$

Thus, $\phi$ is a linear isometry of $\mathbb{R}^{n+1}$. $\square$

PROPOSITION 4.1.3. *Suppose $\theta_{ii} = \pi$ and $\theta_{ji} = \theta_{ij} \in (0, \pi)$ when $i \neq j$. Let $c_{ij} := -\cos\theta_{ij}$, and let $C = (c_{ij})$. There exists a spherical simplex with dihedral angles $\theta_{ij}$ if and only if $C$ is positive definite.*

PROOF. If $\sigma$ is a spherical simplex then the corresponding set of inward pointing unit normals $\{\mathbf{n}_1, \ldots, \mathbf{n}_{n+1}\}$ is an basis for $\mathbb{R}^{n+1}$, so there is $M \in \mathrm{GL}(\mathbb{R}^{n+1})$ such that $M\mathbf{e}_i = \mathbf{n}_i$. Thus, $\langle \mathbf{n}_i, \mathbf{n}_j \rangle = \mathbf{n}_i^T \mathbf{n}_j = \mathbf{e}_i^T M^T M \mathbf{e}_j$, which says that the Gram matrix $C$ is equal to $M^T M$. But $\mathbf{w}^T M^T M \mathbf{w} =$

$(M\mathbf{w})^T M\mathbf{w} \geqslant 0$ with equality if and only if $M\mathbf{w} = \mathbf{0}$, which, since $M$ is invertible, is true if and only if $\mathbf{w} = \mathbf{0}$. Thus, $C$ is positive definite.

Conversely, if $C$ is a positive definite symmetric matrix then let $\mathbf{n}_i$ be the column vectors of $\sqrt{C}$. Since $C$ is positive definite, so is $\sqrt{C}$, so the columns are linearly independent. Set $\sigma = \{\mathbf{v} \in \mathbb{S}^n \mid \langle \mathbf{v}, \mathbf{n}_i \rangle \geqslant 0 \text{ for all } i\}$. □

PROPOSITION 4.1.4. *If $(W, S)$ is a Coxeter system with $S = \{s_i \mid i \in I\}$ the its cosine matrix $C$ is positive definite if and only if there is a spherical simplex $\sigma \subset \mathbb{S}^n \subset \mathbb{R}^{n+1}$ for $n+1 = |I|$ whose Gram matrix is $C$ such that $W$ is isomorphic to the spherical simplicial reflection group with fundamental domain $\sigma$.*

PROOF. If $(W, S)$ is a spherical simplicial reflection group then its cosine matrix is equal to the Gram matrix of the simplex, so is positive definite.

If $C$ is positive definite then by Proposition 4.1.3 there is a spherical simplex $\sigma$ whose Gram matrix is $C$. By Theorem 3.2.9, $W$ is isomorphic to the the spherical reflection group with fundamental domain $\sigma$ generated by reflections in the codimension 1 faces. □

Coxeter gave a complete classification of irreducible spherical Coxeter groups: there are four infinite families and six additional exceptional examples.

THEOREM 4.1.5 (Classification of irreducible spherical Coxeter groups [**10**] [**11**, Table 6.1]). *The irreducible spherical Coxeter groups are those appearing in Table 2.2.*

In Table 2.2, $\bullet\!\!-\!\!\bullet\cdots\bullet\!\!-\!\!\bullet$ indicates a segment containing at least one vertex, and $n$ in the symbol indicates the number of vertices. So, $\mathrm{D}_3 = A_3 \cong \mathrm{Sym}_4$, and $\mathrm{B}_2 = \mathrm{I}_2(4) \cong \mathcal{D}_4$.

EXERCISE 4.1.6. Without appealing to Table 2.2, prove that if the Coxeter graph of a irreducible spherical reflection group has more than one edge then it does not have an edge labelled $m$ for any $m > 5$.

**4.2. Hyperbolic simplices.** The hyperbolic case is very similar to the spherical case, so we will be brief. We use the hyperboloid model of hyperbolic space, so, as in the spherical case, a simplex is determined by a linearly independent collection of inward pointing unit normal vectors in the enclosing real vector space. Not every set of linearly independent vectors will do, as the orthogonal hyperplane needs to meet the hyperboloid. It turns out that there is a nice characterization of the Gram matrix of a hyperbolic simplex. The $i$–th *principle submatrix* of a square matrix $C$ is the matrix obtained by deleting the $i$–th row and column.

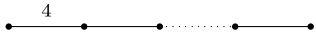
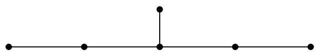
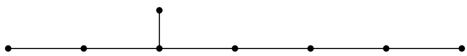
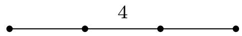
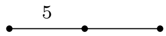
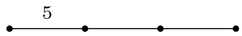
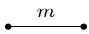
| Symbol | Coxeter diagram |
|---|---|
| $\mathrm{A}_n$ |  |
| $\mathrm{B}_n$ |  |
| $\mathrm{D}_n$ |  |
| $\mathrm{E}_6$ |  |
| $\mathrm{E}_7$ |  |
| $\mathrm{E}_8$ |  |
| $\mathrm{F}_4$ |  |
| $\mathrm{H}_3$ |  |
| $\mathrm{H}_4$ |  |
| $\mathrm{I}_2(m)$ |  |

TABLE 2.2. Irreducible spherical Coxeter groups

PROPOSITION 4.2.1. *Let $(W, S)$ be a Coxeter system with $S = \{s_i \mid i \in I\}$ such that no $m_{ij} = \infty$. Let $C$ be its cosine matrix. The following are equivalent:*

- *$C$ is of type $(|I| - 1, 1)$ and every principal submatrix is positive definite.*
- *$W$ is a simplicial hyperbolic reflection group on $\mathbb{H}^{|I|-1}$ such that each $s_i$ acts by reflection across a codimension 1 face of hyperbolic simplex.*

PROOF. The key fact is [**11**, Lemma 6.8.4], which says that the cosine matrix of a collection of nonobtuse dihedral angles is the Gram matrix of a hyperbolic simplex if and only if it is of type $(n, 1)$ and every principal submatrix is positive definite. One direction of the proposition follows immediately.

Conversely, if $C$ is of type $(|I|-1, 1)$ with every principal submatrix positive definite, then there is a hyperbolic simplex $\sigma$ whose dihedral angles are $\theta_{ij}$ with $c_{ij} = -\cos\theta_{ij}$. By Theorem 3.2.9, $W$ is isomorphic to the simplicial hyperbolic reflection group generated by reflections in the codimension 1 faces of $\sigma$. $\square$

As in the spherical case, one can compute explicitly which combinations of proper integral submultiples of $\pi$ can occur as dihedral angles satisfying the necessary conditions on $C$. These are shown in Table 2.3. As mentioned previously, there are no hyperbolic Coxeter simplices above dimension 4, and it is only in dimension 2 that there are infinitely many different isometry types.

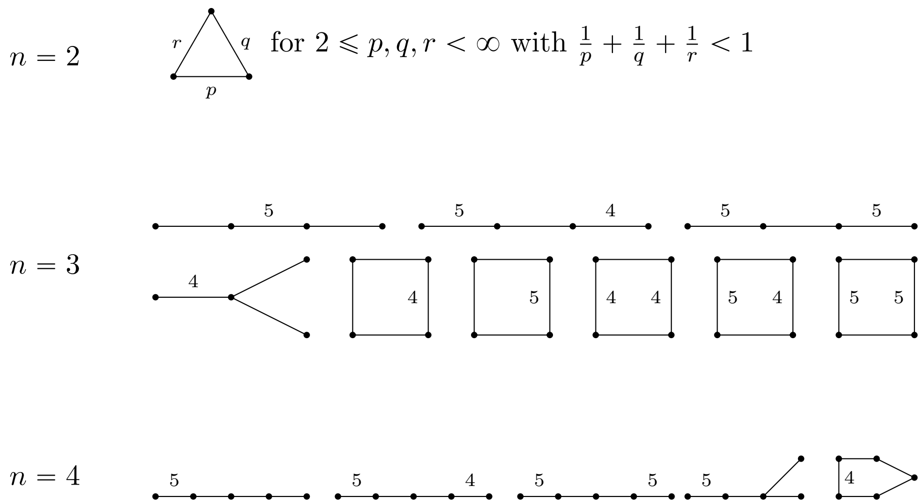



TABLE 2.3. Coxeter graphs of the hyperbolic simplicial Coxeter groups

**4.3. Euclidean simplices.** A *Euclidean $n$–simplex* $\sigma$ is the convex hull of $n+1$ affinely independent points $\{\mathbf{v}_0, \dots, \mathbf{v}_n\}$ in $\mathbb{E}^n$. Let $\sigma_i$ be the codimension 1 face containing $\{\mathbf{v}_0, \dots, \mathbf{v}_n\} - \{\mathbf{v}_i\}$, and let $H_i$ be the affine hyperplane containing $\sigma_i$. This is equivalent to our earlier definition of a Euclidean $n$–simplex as a convex Euclidean polytope that has the combinatorial type of a simplex.

Let $\mathbf{n}_i$ be the inward pointing unit normal vector to $H_i$.

LEMMA 4.3.1. *The $\mathbf{n}_i$ determine $\sigma$ up to homothety.*

PROOF. Impose Cartesian coordinates with $\mathbf{v}_0 = \mathbf{0}$ and the standard inner product. Up to rescaling all of the $\mathbf{v}_i$ by $1/d(H_0, \mathbf{0})$, we may assume $d(H_0, \mathbf{0}) = 1$. Then the simplex is $\sigma = \{\mathbf{v} \in \mathbb{R}^n \mid \langle \mathbf{v}, \mathbf{n}_i \rangle \geqslant 0 \text{ for } 1 \leqslant i \leqslant n \text{ and } \langle \mathbf{v}, \mathbf{n}_0 \rangle \geqslant -1\}$, as in Figure 15.

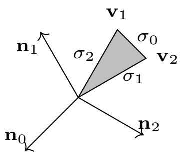



FIGURE 15. Euclidean 2–simplex with its unit normal vectors.

□

LEMMA 4.3.2. *A set of unit vectors $\{\mathbf{n}_0, \ldots, \mathbf{n}_n\}$ spanning $\mathbb{R}^n$ is the set of inward point unit normal vectors to a Euclidean simplex if and only if there exists a set of positive numbers $\{c_0, \ldots, c_n\}$ such that $\sum_i c_i\mathbf{n}_i = \mathbf{0}$.*

PROOF. Suppose $\{\mathbf{n}_0, \ldots, \mathbf{n}_n\}$ is the set of normals to a simplex. By translation, we may assume $\mathbf{v}_0 = \mathbf{0}$. There are $n+1$ vectors, so there is an equation $\sum_i c_i\mathbf{n}_i = \mathbf{0}$ with all $c_i \neq 0$. By multiplying by $-1$ if necessary, assume $c_0 > 0$. Let $D = d(H_0, \mathbf{0})$, so that $\langle \mathbf{n}_0, \mathbf{v}_i\rangle = -D$ for all $i \neq 0$. Then for $1 \leqslant j \leqslant n$:

$$0 = \langle \sum_i c_i\mathbf{n}_i, \mathbf{v_j}\rangle = \sum_i c_i\langle \mathbf{n}_i, \mathbf{v}_j\rangle = -Dc_0 + c_j\langle \mathbf{n}_j, \mathbf{v}_j\rangle \tag{6}$$

So $c_j = \frac{Dc_0}{\langle \mathbf{n}_j, \mathbf{v}_j\rangle} > 0$.

Conversely, suppose we have a set of unit vectors $\{\mathbf{n}_0, \ldots, \mathbf{n}_n\}$ spanning $\mathbb{R}^n$ and positive constants $\{c_0, \ldots, c_n\}$ such that $\sum_i c_i\mathbf{n}_i = \mathbf{0}$. We may assume $c_0 = 1$. Set $\mathbf{v}_0 = \mathbf{0}$. Since the vectors span and $\mathbf{n}_0$ is a linear combinations of the others, $\{\mathbf{n}_1, \ldots, \mathbf{n}_n\}$ is linearly independent, so the hyperplanes $H_i := \mathbf{n}_i^\perp$ cut out a simplicial cone. Let $L_i$ be the line containing the extremal ray of this cone opposite to $H_i$. Let $v_i \in L_i$ be the point such that $\langle \mathbf{v}_i, \mathbf{n}_0\rangle = -1$. Compute as in (6) that $\langle \mathbf{n}_i, \mathbf{v}_i\rangle = \frac{1}{c_i} > 0$ for all $i > 0$. Thus, $\{\mathbf{v}_0, \ldots, \mathbf{v}_n\}$ are, as in Lemma 4.3.1, the vertices of the simplex:

$$\{\mathbf{v} \in \mathbb{R}^n \mid \langle \mathbf{v}, \mathbf{n}_i\rangle \geqslant 0 \text{ for } 1 \leqslant i \leqslant n \text{ and } \langle \mathbf{v}, \mathbf{n}_0\rangle \geqslant -1\}$$

□

PROPOSITION 4.3.3. *For $1 \leqslant i, j \leqslant n+1$, let $\theta_{ij} = \theta_{ji} \in (0, \pi)$ when $i \neq j$, and $\theta_{ii} = \pi$. There is a Euclidean simplex $\sigma$ such that the dihedral angle between $\sigma_{i-1}$ and $\sigma_{j-1}$ is $\theta_{ij}$ if and only if the matrix $C = (-\cos\theta_{ij})$ is positive semi-definite of corank 1 and the nullspace is spanned by a vector with positive coordinates.*

PROOF. We saw in Lemma 4.3.2 that the nullspace of $C$ is spanned by a vector with positive coordinates.

Conversely, suppose $C$ is positive semi-definite of corank 1 and the nullspace is spanned by a vector $\mathbf{v}$ with positive coordinates $c_0, \ldots, c_n$. Let $U$ be the square root of $C$. Let $\mathbf{n}_i$ be the columns of $U$, which are unit vectors since $C$ has 1's on the diagonal. The vector $\mathbf{v}$ is killed by $U$, so $\sum_i c_i \mathbf{n}_i = 0$. By Lemma 4.3.2, the $\mathbf{n}_i$ are in the inward pointing unit normals to a Euclidean simplex. $\square$

PROPOSITION 4.3.4. *Let $(W, S)$ be an irreducible Coxeter system with $S = \{s_i \mid i \in I\}$ such that no $m_{ij} = \infty$. Let $C$ be its cosine matrix. The following are equivalent:*

- *$C$ is positive semi-definite with corank 1.*
- *$W$ is a Euclidean simplicial reflection group on $\mathbb{E}^{|I|-1}$ such that $s_i$ acts by reflection across a codimension 1 face of Euclidean simplex.*

PROOF. If $W$ is a Euclidean simplicial reflection group then $C$ is positive semi-definite and corank 1 by Proposition 4.3.3.

Suppose $C$ is positive semi-definite with corank 1. Since $C$ is a cosine matrix of a Coxeter system, the off-diagonal entries are all negative. By [**11**, Lemma 6.3.7], its nullspace is spanned by a vector with positive coordinates, so we can apply Proposition 4.3.3 to get a Euclidean simplex $\sigma$ whose dihedral angles are $\theta_{ij}$ with $c_{ij} = -\cos\theta_{ij}$. By Theorem 3.2.9, $W$ is isomorphic to the simplicial Euclidean reflection group with fundamental domain $\sigma$ generated by reflections in the codimension 1 faces. $\square$

As in the spherical case, Coxeter gave a complete enumeration of irreducible Coxeter graphs that define Euclidean simplicial reflection groups. It turns out that every such graph can be obtained from a Coxeter graph defining an irreducible spherical Coxeter group by adding only one additional edge. Conversely:

EXERCISE 4.3.5. If $\Gamma$ is in Table 2.4 then removing any single edge from $\Gamma$ results in either one or a disjoint union of two graphs from Table 2.2.

THEOREM 4.3.6 (Classification of irreducible Euclidean Coxeter groups [**10**] [**11**, Table 6.1]). *The irreducible Coxeter groups that are Euclidean simplicial reflection groups are those appearing in Table 2.4.*

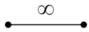
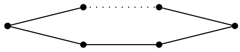
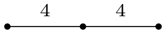
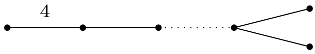
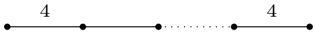
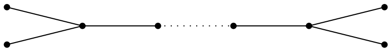
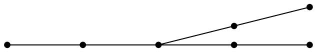
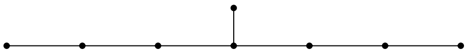
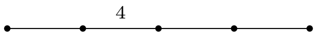
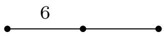
| Symbol | Coxeter diagram |
|---|---|
| $\tilde{\mathrm{A}}_1$ |  |
| $\tilde{\mathrm{A}}_n$ |  |
| $\tilde{\mathrm{B}}_2$ |  |
| $\tilde{\mathrm{B}}_n$ |  |
| $\tilde{\mathrm{C}}_n$ |  |
| $\tilde{\mathrm{D}}_n$ |  |
| $\tilde{\mathrm{E}}_6$ |  |
| $\tilde{\mathrm{E}}_7$ |  |
| $\tilde{\mathrm{E}}_8$ |  |
| $\tilde{\mathrm{F}}_4$ |  |
| $\tilde{\mathrm{G}}_2$ |  |

TABLE 2.4. Irreducible Euclidean Coxeter groups

CHAPTER 3

# Linear representations

Let $H_1$ and $H_2$ be two lines through the origin in $\mathbb{E}^2$ that differ by angle $\pi/m$. Let $\mathbf{e}_i$ be a unit normal to $H_i$ as in Figure 1.

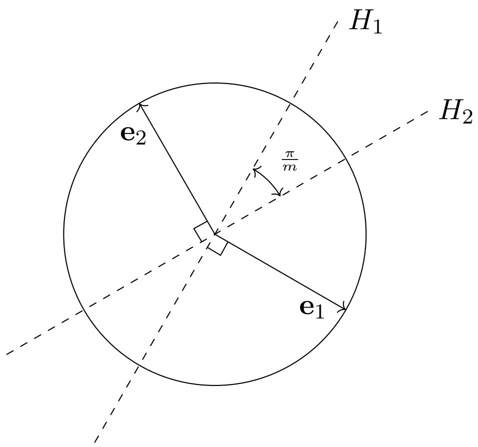



FIGURE 1

Observe that the angle between $\mathbf{e}_1$ and $\mathbf{e}_2$ is $\pi - \pi/m$. We can rewrite the plane as the 2–dimensional vector space spanned by $\mathbf{e}_1$ and $\mathbf{e}_2$. To retain the conditions that $\mathbf{e}_1$ and $\mathbf{e}_2$ are unit vectors with angle $\pi - \pi/m$, we define an inner product $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = c_{ij}$ where $C = (c_{ij}) = \begin{pmatrix} 1 & -\cos\frac{\pi}{m} \\ -\cos\frac{\pi}{m} & 1 \end{pmatrix}$. Notice this is the Gram matrix of a 1–dimensional spherical reflection group, and thus coincides with the cosine matrix of the Coxeter system. Now we will try to go in the other direction: start with the cosine matrix and use it to define a symmetric bilinear form on a vector space spanned by vectors corresponding to the Coxeter generators. This gives a linear representation of the Coxeter system. In the simple example above, the representation is conjugate to representation from the realization as a geometric reflection group, so we did not gain any new information. However, it turns out that the new linear representation is faithful even in the non-geometric case, and this has strong consequences.

## 1. Consequences of linearity

A group is called *linear* if it admits a faithful representation into $\mathrm{GL}(V)$ for some finite dimensional vector space $V$. Being a linear group implies several strong properties:

THEOREM 1.0.1. *Let $G$ be a finitely generated linear group.*

- *$G$ has solvable word problem.*
- *$G$ is residually finite.*
- *$G$ has a finite index torsion-free subgroup.*
- *Every infinite, finitely generated subgroup of $G$ contains an infinite order element.*
- *$G$ satisfies the* Tits alternative*: every finitely generated subgroup of $G$ either contains a rank 2 free subgroup or has a finite index subgroup that is solvable.*

We will show (Corollary 4.0.5) that finitely generated Coxeter groups have all of these properties.

## 2. The canonical representation

Let $(W, S)$, $S = \{s_i \mid i \in I\}$, be a Coxeter system with Coxeter matrix $M = (m_{ij})$ and cosine matrix $C = (c_{ij} := -\cos \pi/m_{ij})$. Since $M$ is symmetric with 1's on the diagonal and non-diagonal entries at least 2, $C$ is symmetric with 1's on the diagonal and non-diagonal entries in $[-1, 0]$. Define a symmetric bilinear form $B_C$ on the vector space $\mathbb{R}^I$ spanned by vectors $\mathbf{e}_i$ for $i \in I$ by extending $B_C(\mathbf{e}_i, \mathbf{e}_j) := c_{ij}$ bilinearly. For each $i \in I$, define $H_i := \{\mathbf{v} \mid B_C(\mathbf{v}, \mathbf{e}_i) = 0\}$, which we think of as the hyperplane that is $B_C$–orthogonal to $\mathbf{e}_i$. Define a 'reflection' through this hyperplane by:

$$\rho_i(\mathbf{v}) := \mathbf{v} - 2B_C(\mathbf{v}, \mathbf{e}_i)\mathbf{e}_i \tag{7}$$

We need to check that $\rho(s_i) := \rho_i$ extends to a homomorphism $\rho\colon W \to \mathrm{GL}(\mathbb{R}^I)$. To do this, we should check that the defining relations of $W$ are satisfied in the image; that is, we need to check that for each $i, j \in I$ the map $\rho_i\rho_j$ has order (dividing) $m_{ij}$.

EXERCISE 2.0.1. Verify the following properties of $\rho_i$:

(1) $\rho_i$ is linear.
(2) $H_i = \{\mathbf{v} \mid \rho_i(\mathbf{v}) = \mathbf{v}\}$
(3) $\forall \mathbf{v} \in \mathbb{R}^I$, $\rho_i^2(\mathbf{v}) = \mathbf{v}$
(4) $B_C$ is $\rho_i$–invariant, that is: $\forall \mathbf{v}, \mathbf{w} \in \mathbb{R}^I$, $B_C(\rho_i(\mathbf{v}), \rho_i(\mathbf{w})) = B_C(\mathbf{v}, \mathbf{w})$
(5) $\rho_i$ preserves any vector subspace containing $\mathbf{e}_i$.

Let $P_{ij}$ be the plane spanned by $\mathbf{e}_i$ and $\mathbf{e}_j$ for each $i \neq j$. The restriction of $B_C$ to $P_{ij}$ is $\begin{pmatrix} 1 & -c_{ij} \\ -c_{ij} & 1 \end{pmatrix}$. If $m_{ij} < \infty$ then this is positive definite, so $B_C$ restricts to an inner product on $P_{ij}$ with respect to which:

(1) $\mathbf{e}_i$ and $\mathbf{e}_j$ are unit vectors that differ by angle $\pi - \pi/m_{ij}$.
(2) $H_i \cap P_{ij}$ is a line orthogonal to $\mathbf{e}_i$, so the angle between $H_i \cap P_{ij}$ and $H_j \cap P_{ij}$ is $\pi/m_{ij}$.
(3) $\rho_i|_{P_{ij}}$ is reflection through $H_i \cap P_{ij}$.

This is exactly our model situation of reflections in a plane through lines the differ by angle $\pi/m_{ij}$. In this situation we know that *generic points*, that is, points of $P_{ij} \backslash (H_i \cup H_j)$, have orbits of size $m_{ij}$ under the action by the group generated by $\rho_i|_{P_{ij}}$ and $\rho_j|_{P_{ij}}$. This shows that the order of $\rho_i\rho_j$ is a multiple of $m_{ij}$, but to see it is really $m_{ij}$ we consider the rest of the space.

Consider $\mathbf{e}_k$ for $k \neq i, j$. Since $\rho_i$ and $\rho_j$ act by adding multiples of $\mathbf{e}_i$ and $\mathbf{e}_j$, we have $(\rho_i\rho_j)(\mathbf{e}_k) = \mathbf{e}_k + \mathbf{p}_k$ for some $\mathbf{p}_k \in P_{ij}$. Iterating, we see for $n \geqslant 0$:

$$(\rho_i\rho_j)^n(\mathbf{e}_k) = \mathbf{e}_k + \sum_{l=0}^{n-1} (\rho_i\rho_j)^l(\mathbf{p}_k) \tag{8}$$

For $d_k$ equal to the size of the $\rho_i\rho_j$ orbit of $\mathbf{p}_k$, (8) says $(\rho_i\rho_j)^{d_k}(\mathbf{e}_k)$ is $\mathbf{e}_k$ plus the sum of all of the vectors in the $\rho_i\rho_j$–orbit of $\mathbf{p}_k$. The sum of all the vectors in a $\rho_i\rho_j$ orbit is fixed by $\rho_i\rho_j$, but $\rho_i\rho_j$ is a rotation on $P_{ij}$, so the only fixed point is $\mathbf{0}$, so $(\rho_i\rho_j)^{d_k}(\mathbf{e}_k) = \mathbf{e}_k$. Now, every point of $P_{ij}$ has orbit of size dividing $m_{ij}$, so for all $k \neq i, j$, $(\rho_i\rho_j)^{m_{ij}}(\mathbf{e}_k) = \mathbf{e}_\mathbf{k}$, so the order of $\rho_i\rho_j$ on $\mathbb{R}^I$ is $m_{ij}$.

Now suppose $m_{ij} = \infty$, so that the restriction of $B_C$ to $P_{ij}$ is $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$. This matrix is only positive semi-definite, so the restriction of $B_C$ to $P_{ij}$ is not an inner product, so our intuition of how ‘reflections’ behave may not be accurate.

We have $H_i \cap P_{ij} = H_j \cap P_{ij} = \mathbb{R}(\mathbf{e}_i + \mathbf{e}_j)$.

Since $\rho_i$ is linear and fixes $H_i$ pointwise, we compute:

$$\rho_i(a\mathbf{e}_i + b\mathbf{e}_j) = \rho_i(b(\mathbf{e}_1 + \mathbf{e}_j) + (a-b)\mathbf{e}_i) = b(\mathbf{e}_i + \mathbf{e}_j) - (a-b)\mathbf{e}_i$$

If we think of $\mathbf{e}_i$ and $\mathbf{e}_j$ as the standard basis for $\mathbb{R}^2$, this says the map $\rho_i|_{P_{ij}}$ is that map that flips horizontally through the main diagonal. The matrix representation of $\rho_i$ in $\mathbf{e}_i$, $\mathbf{e}_j$ coordinates is $\begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix}$. A similar computation for $\rho_j$ gives matrix $\begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}$. Then the product $\rho_i \circ \rho_j$ has matrix $\begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix}$. This is a shear matrix; it is conjugate via $\pi/4$ rotation to the basic shear matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. In particular, points of $H_i \cap P_{ij} = H_j \cap P_{ij}$ are fixed pointwise and points of $P_{ij} \backslash (H_i \cup H_j)$ have infinite orbits, so $\rho_i \rho_j$ has infinite order.

We have verified that all of the relations of the Coxeter presentation of $(W, S)$ are satisfied in the image of $\rho$, so $\rho$ is a homomorphism.

DEFINITION 2.0.2. The *canonical representation* (also known as the *standard representation* or *Tits representation*) of a Coxeter system $(W, S)$ with $S = \{s_i \mid i \in I\}$ is the homomorphism $\rho \colon W \to \mathrm{GL}(\mathbb{R}^I)$ defined on the generators of $W$ by $\rho(s_i) := \rho_i$, where the $\rho_i$ are as defined in (7).

PROPOSITION 2.0.3. *Let $(W, S)$ be a Coxeter system with $S = \{s_i \mid i \in I\}$ and Coxeter matrix $M = (m_{ij})$. For all $i$ and $j$ the subgroup of $W$ generated by $s_i$ and $s_j$ is isomorphic to the dihedral group $\mathcal{D}_{m_{ij}}$.*

This improves Proposition 4.2.6 to rank 2 special subgroups. Recall that we needed this fact in Theorem 2.5.9.

PROOF. For any $i$ and $j$, let $\mathcal{D}_{m_{ij}} = \langle s, t \mid s^2, t^2, (st)^{m_{ij}} \rangle$. The map defined by $s \mapsto s_i \mapsto \rho_i$ and $t \mapsto s_j \mapsto \rho_j$ is an isomorphism that factors through the subgroup of $W$ generated by $s_i$ and $s_j$, so the surjection $\mathcal{D}_{m_{ij}} \twoheadrightarrow \langle s_i, s_j \rangle$ is an isomorphism. □

## 3. Finiteness criterion

We use some representation theory to derive a criterion for a Coxeter group to be finite.

Consider a representation $\rho \colon G \to \mathrm{GL}(V)$, where $V$ is a vector space. A *stable subspace* is a vector subspace $U < V$ such that $\rho(g)(U) < U$ for every $g \in G$. There are always the stable subspaces $V$ and $\mathbf{0}$; the representation is called *irreducible* if there are no others. It is *semi-simple* if $V$ splits into a direct sum $V = \oplus V_i$ of stable subspaces such that the representation of $G$ obtained by restricting $\rho(G)$ to $V_i$ is irreducible for each summand. For finite dimensional $V$ an equivalent definition of semi-simple is that every stable subspace of $V$ is a direct summand of $V$.

LEMMA 3.0.1. *Any representation of a finite group $G$ into a finite dimensional real vector space admits a $G$–invariant positive definite symmetric bilinear form.*

Proof. Average the standard inner product $\langle\cdot,\cdot\rangle$ over the $G$–action:

$$B(\mathbf{v},\mathbf{w}) := \frac{1}{|G|}\sum_{g\in G}\langle\rho(g)(\mathbf{v}),\rho(g)(\mathbf{w}))\rangle \qquad \square$$

Corollary 3.0.2. *Any representation of a finite group $G$ into a finite dimensional real vector space $V$ is semi-simple.*

Proof. Let $B$ be the bilinear form of Lemma 3.0.1. Suppose $U$ is a stable subspace of $V$. Then the subspace of $V$ that is $B$–orthogonal to $U$ is also stable, so $V = U \oplus U^{\perp_B}$. $\square$

Proposition 3.0.3 ([**11**, Proposition 6.12.7]). *Suppose $(W, S)$ is an irreducible Coxeter system and $C$ is its cosine matrix. Consider the canonical representation $\rho\colon W \to \mathrm{GL}(V)$, where $V = \mathbb{R}^I$. Define the 'kernel of $B_C$' to be $V_0 := \{\mathbf{v} \in V \mid \forall \mathbf{w} \in V,\ B_C(\mathbf{v},\mathbf{w}) = 0\}$. Then $W$ acts trivially on $V_0$ and any stable subspace of $V$ is contained in $V_0$.*

Proof. If $\mathbf{v} \in V_0$ then $\rho_i(\mathbf{v}) = \mathbf{v} - 2B_C(\mathbf{v},\mathbf{e}_i)\mathbf{e}_i = \mathbf{v} - 0\mathbf{e}_i = \mathbf{v}$, so $W$ acts trivially on $V_0$.

Suppose that $V'$ is a proper stable subspace. First suppose $\mathbf{e}_i \in V'$ for some $i \in I$. Suppose $m_{ij} > 2$, so $i$ and $j$ correspond to adjacent vertices in the Coxeter graph for $(W, S)$, and $c_{ij} = -\cos\pi/m + ij \neq 0$. Then $\rho_j(\mathbf{e}_i) = \mathbf{e}_i - 2B_C(\mathbf{e}_i,\mathbf{e}_j)\mathbf{e}_j = \mathbf{e}_i + 2c_{ij}\mathbf{e}_j$. So $\rho_j$ sends $\mathbf{e}_i$ into the span$(\mathbf{e}_i,\mathbf{e}_j)$ but not into span$(\mathbf{e}_i)$. Since $\mathbf{e}_i$ is contained in the stable subspace $V'$, we must also have span$(\mathbf{e}_i,\mathbf{e}_j) \subset V'$. But since a similar argument applies for any edge of the Coxeter graph, and the Coxeter graph is connected by hypothesis, this means that if one $\mathbf{e}_i$ is in $V'$ then they all are. But we cannot have all of the $\mathbf{e}_i$ contained in $V'$, because that would mean that $V'$ is all of $V$, and it was assumed to be a proper subspace. Thus, $V'$ contains none of the $\mathbf{e}_i$.

Now, $V = \mathbb{R}\mathbf{e}_i \oplus H_i$, so suppose $a\mathbf{e}_i + \mathbf{w} \in V'$ for $w \in H_i$. Since $V'$ is stable, $\rho_i(a\mathbf{e}_i + \mathbf{w})$ is also in $V'$, as is:

$$a\mathbf{e}_i + \mathbf{w} - \rho_i(a\mathbf{e}_i + \mathbf{w}) = a\mathbf{e}_i + \mathbf{w} - (-a\mathbf{e}_i + \mathbf{w}) = 2a\mathbf{e}_i$$

But since $\mathbf{e}_i \notin V'$, this means $a = 0$, so $V' < H_i$. Since this is true for every $i$, we have $V' \subset \bigcap_{i\in I} H_i = V_0$. $\square$

Corollary 3.0.4. *Suppose $(W, S)$ is an irreducible Coxeter system with cosine matrix $C$.*

1. *If $B_C$ is nondegenerate then $\rho$ is irreducible.*
2. *If $B_C$ is degenerate then $\rho$ is not semi-simple.*

Proof. $B_C$ is degenerate when its kernel is nontrivial. Proposition 3.0.3 says any proper stable subspace is contained in the kernel, so when the kernel is trivial then $\rho$ is irreducible.

For the second item, suppose that $B_C$ has nontrivial kernel $V_0$. Suppose $V = V_0 \oplus V'$. Then $V'$ is a proper stable subspace that is not contained in $V_0$, which contradicts Proposition 3.0.3. Thus, $V_0$ is an stable subspace that is not a summand, which cannot exist for a semi-simple representation. $\square$

Lemma 3.0.5 ([**11**, Lemma 6.12.2]). *Let $\rho\colon G \to \mathrm{GL}(V)$ be an irreducible finite dimensional representation. If there is a $g \in G$ such that $\mathrm{Id} - \rho(g)$ has 1–dimensional image then any two nonzero $G$–invariant bilinear forms on $V$ are proportional.*

Proposition 3.0.6. *If $(W, S)$ is a Coxeter system with cosine matrix $C$ and $W$ is finite then $C$ is positive definite.*

Proof. For a reducible Coxeter system, if we order generators by component then the cosine matrix splits diagonally as the cosine matrices of the factors, so it is positive definite if and only if every factor has positive definite cosine matrix. So we may pass to factors, and assume $(W, S)$ is irreducible.

Consider the canonical representation $\rho\colon W \to \mathrm{GL}(V)$. By Lemma 3.0.1 there exists a $W$–invariant positive definite symmetric bilinear form $B'$. By Corollary 3.0.2, $\rho$ is semi-simple, so by Corollary 3.0.4, $B_C$ is nondegenerate and $\rho$ is irreducible. For each $i \in I$, $Id - \rho(s_i) = Id - \rho_i$ has 1–dimensional image $\mathbb{R}\mathbf{e}_i$, so by Lemma 3.0.5, $B_C$ is proportional to $B'$. Since $B'$ is positive definite, $B_C$ is either positive definite or negative definite. But $B_C$ is certainly not negative definite, because $B_C(\mathbf{e}_i, \mathbf{e}_i) = 1$. Finally, $B_C$ is positive definite if and only if $C$ is. $\square$

We combine this result with our knowledge of geometric reflection groups to get an "if and only if" criterion:

Theorem 3.0.7 (Finiteness criterion). *Let $(W, S)$ be a Coxeter system with $S = \{s_i \mid i \in I\}$ and cosine matrix $C$. The following are equivalent:*

(1) *$W$ is finite.*
(2) *$C$ is positive definite.*
(3) *$W$ is a spherical simplicial reflection group on $\mathbb{S}^{|I|-1}$ such that $s_i$ acts by reflection across a codimension 1 face of a spherical simplex.*

Proof. (1) $\implies$ (2) is Proposition 3.0.6. (2) $\implies$ (3) is Proposition 4.1.4. Finally, the reflection group described in (3) is finite because it is a discrete subgroup of $\mathrm{Isom}(\mathbb{S}^{|I|-1}) = \mathrm{O}(|I| - 1)$, which is compact. $\square$

## 4. The geometric representation

It turns out that there is a different linear representation obtained by dualizing the canonical one that is in some senses ‘better’. This is called the ‘geometric representation’, and in the case of geometric reflection groups it is closely related to the representation as a subgroup of $\operatorname{Isom}\mathbb{X}^n$. It is also easier to prove faithfulness of the geometric representation than to prove it directly for the canonical representation, but we will not provide the details of either proof.

The dual of a real vector space $V$ is the vector space of real valued linear functions on $V$. If $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is a basis for $V$ then a basis for the dual $V^*$ is $\{\mathbf{e}_1^*, \dots, \mathbf{e}_n^*\}$ where $\mathbf{e}_i^*$ is the linear function that is determined by its values on basis elements by:

$$\mathbf{e}_i^*(\mathbf{e_j}) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

A linear transformation $\phi \colon V \to V$ dualizes to a linear transformation $\phi^* \colon V^* \to V^*$ by $\phi^*(\nu)(\mathbf{v}) := \nu(\phi(\mathbf{v}))$.

A representation $\rho \colon W \to \mathrm{GL}(V)$ dualizes to a representation $\rho^* \colon W \to \mathrm{GL}(V^*)$ by $\rho^*(w)(\nu)(\mathbf{v}) := \nu(\rho(w)^{-1}(\mathbf{v}))$.

Definition 4.0.1. The *geometric representation* $\rho^* \colon W \to \mathrm{GL}(V^*)$ of a Coxeter system $(W, S)$ with $S = \{s_i \mid i \in I\}$ is the dual to the canonical representation $\rho \colon W \to \mathrm{GL}(V)$.

Define $\rho_i^* := \rho^*(s_i)$. Since $\rho_i$ has order 2, we have :

$$\rho_k^*(\mathbf{e}_j^*)(\mathbf{e}_i) = \mathbf{e}_j^*(\rho_k^{-1}(\mathbf{e}_i)) = \mathbf{e}_j^*(\rho_k(\mathbf{e}_i)) \tag{9}$$

$$= \mathbf{e}_j^*(\mathbf{e}_i - 2c_{ik}\mathbf{e}_k) = \delta_{ij} - 2\delta_{jk}c_{ik} \tag{10}$$

(9) says the matrix expression for $\rho_i^*$ in terms of the basis $\{\mathbf{e}_i^*\}$ is the transpose of the matrix expression for $\rho_i$ in terms of the $\{\mathbf{e}_i\}$ basis.

Define $\xi_i \in V^*$ to be the linear function that is determined by its values on basis elements by $\xi_i(\mathbf{e}_j) := c_{ij}$. Observe:

$$\begin{aligned}(\mathbf{e}_j^* - 2\mathbf{e}_j^*(\mathbf{e}_k)\xi_k)(\mathbf{e}_i) &= \mathbf{e}_j^*(\mathbf{e}_i) - 2\mathbf{e}_j^*(\mathbf{e}_k)\xi_k(\mathbf{e}_i) \\ &= \delta_{ij} - 2\delta_{jk}c_{ik}\end{aligned}$$

This is the same result as (10), so $\rho_k^*(\mathbf{e}_j^*)$ and $\mathbf{e}_j^* - 2\mathbf{e}_j^*(\mathbf{e}_k)\xi_k$ are linear functions on $V$ that take the same value on every basis element, hence are the same function. Since the $\mathbf{e}_j^*$ form a basis for $V^*$, extend linearly to get:

$$\rho_k^*(\nu) = \nu - 2\nu(\mathbf{e}_k)\xi_k \tag{11}$$

LEMMA 4.0.2. *For all $i \in I$, $\rho_i^*$ fixes the $i$–th coordinate hyperplane and exchanges the corresponding half-spaces.*

PROOF. The $i$–th coordinate hyperplane is $\{\sum_{j\in I} a_j \mathbf{e}_j^* \mid a_i = 0\} = \{\nu \in V^* \mid \nu(\mathbf{e}_i) = 0\}$, and the positive and negative half-spaces are covectors $\nu$ that take positive and negative values on $\mathbf{e}_i$, respectively.

Using (11) and the fact that $\xi_i$ is nontrivial, since, eg, $\xi_i(\mathbf{e}_i) = 1$, we have $\nu(\mathbf{e}_i) = 0 \iff \rho_i^*(\nu) = \nu - 2\nu(\mathbf{e}_i)\xi_i = \nu$.

To see that half-spaces are exchanged, observe:

$$\rho_i^*(\nu)(\mathbf{e}_i) = (\nu - 2\nu(\mathbf{e}_i)\xi_i)(\mathbf{e}_i) = \nu(\mathbf{e}_i) - 2\nu(\mathbf{e}_i) \cdot 1 = -\nu(\mathbf{e}_i) \qquad \square$$

THEOREM 4.0.3 (Tits, see [**11**, Theorem D.1.1]). *Let $D := \{\sum_{i\in I} a_i \mathbf{e}_i^* \mid a_i \geqslant 0, \forall i \in I\} = \{\nu \in V^* \mid \nu(\mathbf{e}_i) \geqslant 0, \forall i \in I\}$ be the closed positive orthant in $V^*$, and let $\mathring{D}$ be its interior, defined by strict inequalities. Then:*

$$\rho^*(w)\mathring{D} \cap \mathring{D} \neq \varnothing \iff w \text{ is trivial in } W$$

COROLLARY 4.0.4. *$\rho^*$ is faithful.*

COROLLARY 4.0.5. *$\rho$ is faithful.*

$D$ is called the *fundamental chamber*, and the set $\bigcup_{w\in W} \rho^*(w)D$ of $W$–translates of $D$ is called the *Tits cone*. The theorem implies that the fundamental chamber is a fundamental domain for the action of $W$ on the Tits cone. Furthermore:

THEOREM 4.0.6 ([**11**, Theorem D.2.7]). *The Tits cone is convex.*

However, the Tits cone might not be all of $V^*$.

COROLLARY 4.0.7 (Coxeter complex). *There is a connected simplicial complex of dimension $|I| - 1$ on which $W$ acts, with fundamental domain a single top dimensional simplex.*

PROOF. Let the *fundamental simplex* be the standard $|I| - 1$ dimensional simplex $\sigma := \{\sum_{i\in I} a_i \mathbf{e}_i^* \mid \sum_{i\in I} a_i = 1\}$, which is contained in the fundamental chamber and contains exactly one point for every ray through $\mathbf{0}$ contained in the fundamental chamber. Consider $\Sigma := \bigcup_{w\in W} \rho^*(w)\sigma$. By Theorem 4.0.3, $\sigma$ is a fundamental domain for the $W$–action on $\Sigma$, and by Theorem 4.0.6, $\Sigma$ is connected. By Lemma 2.5.8 there is a continuous $W$–equivariant bijection $\mathcal{U}(W, \sigma) \to \Sigma$. $\square$

Note that in general $\mathcal{U}(W, \sigma)$ and $\Sigma$ are not homeomorphic and the action of $W$ on $\mathcal{U}(W, \sigma)$ is not proper. See Example 5.0.4.

PROPOSITION 4.0.8. *When $C$ is invertible, $\rho$ and $\rho^*$ are conjugate representations into* $\mathrm{GL}(R^{|I|})$*. Furthermore, they are conjugate to a representation that preserves the signature of $C$. In particular, if $C$ is positive definite then $\rho$ and $\rho^*$ are conjugate to an orthogonal representation, and if $C$ has signature $(|I|-1,1)$ then $\rho$ and $\rho^*$ are conjugate to a representation preserving the Minkowski form on $\mathbb{R}^{|I|}$.*

PROOF. If $C$ is invertible then $C\colon V \to V^* : \mathbf{e_i} \mapsto \xi_i$ is an isomorphism. It conjugates $\rho$ to $\rho^*$, since for all $i$ and $j$:

$$C\rho_i C^{-1}(\xi_j) = C\rho_i(\mathbf{e}_j) = C(\mathbf{e}_j - 2c_{ij}\mathbf{e}_i) = \xi_j - 2c_{ij}\xi_i \overset{(11)}{=} \rho_i^*(\xi_j)$$

We proceed similarly to Theorem 4.0.4. Since $C$ is symmetric, there is an orthonormal basis consisting of eigenvectors. Let $E$ be a matrix whose columns are such a set of eigenvectors, and let $\Lambda$ be the corresponding diagonal matrix of eigenvalues. Let $\sqrt{|\Lambda|}$ be the diagonal matrix whose entries are positive square roots of the absolute values of the entries of $\Lambda$.

If $C$ is invertible then none of the eigenvalues are 0, so $\sqrt{|\Lambda|}$ is invertible and the signature matrix of $C$ is the diagonal matrix $J$ where the entries of $\Lambda$ are replaced by their sign. This setup gives us $E^{-1}CE = \Lambda = \sqrt{|\Lambda|}J\sqrt{|\Lambda|}$. Putting these together with $\rho_i^T C\rho_i = C$ gives:

$$\begin{aligned} E\sqrt{|\Lambda|}J\sqrt{|\Lambda|}E^{-1} &= C \\ &= \rho_i^T C\rho_i \\ &= \rho_i^T E\sqrt{|\Lambda|}J\sqrt{|\Lambda|}E^{-1}\rho_i \end{aligned}$$

So:

$$\begin{aligned} EJE^{-1} &= E\sqrt{|\Lambda|}^{-1}E^{-1}\rho_i^T E\sqrt{|\Lambda|}J\sqrt{|\Lambda|}E^{-1}\rho_i E\sqrt{|\Lambda|}^{-1}E^{-1} \\ &= E\sqrt{|\Lambda|}^{-1}E^{-1}\rho_i^T E\sqrt{|\Lambda|}E^{-1}EJE^{-1}E\sqrt{|\Lambda|}E^{-1}\rho_i E\sqrt{|\Lambda|}^{-1}E^{-1} \\ &= \left(E\sqrt{|\Lambda|}E^{-1}\rho_i E\sqrt{|\Lambda|}^{-1}E^{-1}\right)^T EJE^{-1}\left(E\sqrt{|\Lambda|}E^{-1}\rho_i E\sqrt{|\Lambda|}^{-1}E^{-1}\right) \end{aligned}$$

Thus, $\check{\rho}_i := E\sqrt{|\Lambda|}E^{-1}\rho_i E\sqrt{|\Lambda|}^{-1}E^{-1}$ preserves $EJE^{-1}$. If $J = I$ then $EJE^{-1} = I = J$ and $\check{\rho}_i = \sqrt{C}\rho_i\sqrt{C}$ preserves $I$.

If $J$ is the Minkowski signature then $\check{\rho}_i$ preserves the Minkowski form conjugated by the orthogonal matrix $E$, so $\hat{\rho}_i := \sqrt{|\Lambda|}E^{-1}\rho_i E\sqrt{|\Lambda|}^{-1}$ preserves the Minkowski form. □

When $C$ is not invertible the two representations $\rho$ and $\rho^*$ are different. We will see examples in the next subsection, but the Euclidean case deserves special consideration.

Recall from Proposition 4.0.8, that when $C$ is invertible we can think of it as defining an isomorphism $V \to V^*$ conjugating $\rho$ to $\rho^*$. In the case of a Euclidean reflection group, $C$ is not invertible, so the proof of Proposition 4.0.8 does not work: $C$ as a map $V \to V^*$ has a 1–dimensional kernel, and maps $V$ onto a codimension 1 linear subspace $L$ spanned by the covectors $\xi_i$. Equation (11) shows that for all $\nu \in V^*$ and all $k$, $\rho_k^*(\nu) - \nu$ is contained in $L$. So $\rho^*$ preserves $L$, but even more, it preserves any affine subspace $\nu + L$ parallel to $L$. The benefit is that while $\rho^*(W)$ fixes a point of $L$, the origin, if we instead choose $\nu$ in the fundamental chamber $D$, then Theorem 4.0.3 implies that $\rho^*(W)$ acts cocompactly by affine transformations on the $|I| - 1$ dimensional Euclidean space $\nu + L$, with fundamental domain the simplex $D \cap (\nu + L)$.

EXAMPLE 4.0.9. Consider $\Delta(3,3,3)$. Then $C = \begin{pmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{pmatrix}$ has nullspace spanned by $\mathbf{v} = (1,1,1)$. A choice of orthonormal eigenvector matrix is $E = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & -2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$, with corresponding $\Lambda = \begin{pmatrix} 3/2 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. We would like to try the same thing as in Proposition 4.0.8, but $\Lambda$ is not invertible. Define $\Sigma := \begin{pmatrix} \sqrt{3/2} & 0 & 0 \\ 0 & \sqrt{3/2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and let $\hat{\rho}_i := \Sigma E^{-1} \rho_i^* E \Sigma^{-1}$. For example:

$$\rho_1 = \begin{pmatrix} -1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \rho_1^* = \begin{pmatrix} -1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \hat{\rho}_1 = \begin{pmatrix} -1/2 & -\sqrt{3}/2 & -3/2 \\ -\sqrt{3}/2 & 1/2 & -\sqrt{3}/2 \\ 0 & 0 & 1 \end{pmatrix}$$

Observe that $\hat{\rho}_1$ restricts to the plane $z = 1$ in $\mathbb{R}^3$ to be an affine isometry: it is orthogonal reflection through the line $(x, -\sqrt{3}(x+1), 1)$. Similarly, $\hat{\rho}_2$ and $\hat{\rho}_3$ restricted to the affine plane $z = 1$ are affine reflections through lines that differ from this by $\pi/3$, so we see $\Delta(3,3,3)$ as a simplicial Euclidean reflection group, not on all of $\mathbb{R}^3$, but on a 2–dimensional (affine) subspace.

This example is particularly nice because with $\mathbf{v} = (1,1,1)$, the fundamental simplex is contained in a single affine plane preserved by $\rho^*(W)$, and $C$ has repeated eigenvalues, so there is not even any relative distortion between the non-zero eigenvectors. In this case we could have already seen

$W$ acting as a Euclidean reflection group just by looking at the Coxeter complex, shown in Figure 2. Compare with Figure 2b.

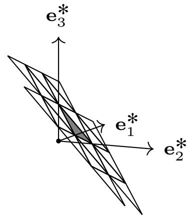



FIGURE 2. Part of the Coxeter complex for $\Delta(3,3,3)$. (Fundamental simplex shaded.)

## 5. Examples of canonical vs geometric representations

EXAMPLE 5.0.1 (Canonical vs geometric representation; spherical case). Consider $W = \langle s_1, s_2 \mid s_1^2, s_2^2, (s_1s_2)^3\rangle \cong \mathcal{D}_3$. Its cosine matrix is $C = \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix}$.

In terms of the basis $(\mathbf{e}_1, \mathbf{e}_2)$ of $V = \mathbb{R}^2$ we have:

$$\rho_1 = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \rho_2 = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$$

In terms of the basis $(\mathbf{e}_1^*, \mathbf{e}_2^*)$ of $V^*$ we have:

$$\rho_1^* = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} \qquad \rho_2^* = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$$

Figures 3a and 3b show these two actions, which are conjugate via $C$. In both cases we can choose a sector between two lines, $H_1$ and $H_2$ in the canonical representation or the $x$ and $y$ axes in the geometric representation, to be a fundamental domain for the action.

In the geometric representation, the Tits cone is all of $\mathbb{R}^2$, and we can see the Coxeter complex sitting inside $\mathbb{R}^2$ by taking translates of the fundamental simplex to get a hexagon. Since the canonical representation is conjugate to the geometric representation, we also see the $W$–action on a hexagon there.

Finally, as in Proposition 4.0.8 we consider the conjugate orthogonal representation, obtained from:

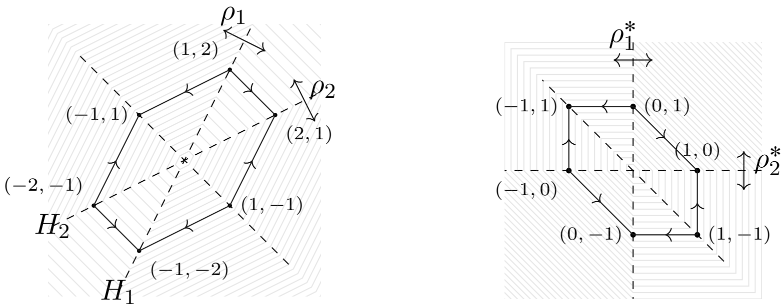
$$\sqrt{C} = \begin{pmatrix} \frac{1+\sqrt{3}}{2\sqrt{2}} & \frac{1-\sqrt{3}}{2\sqrt{2}} \\ \frac{1-\sqrt{3}}{2\sqrt{2}} & \frac{1+\sqrt{3}}{2\sqrt{2}} \end{pmatrix}$$

By:

$$\check{\rho}_1 := \sqrt{C}\rho_1\sqrt{C}^{-1} = \begin{pmatrix} \cos\frac{5\pi}{6} & \sin\frac{5\pi}{6} \\ \sin\frac{5\pi}{6} & -\cos\frac{5\pi}{6} \end{pmatrix}$$

$$\check{\rho}_2 := \sqrt{C}\rho_2\sqrt{C}^{-1} = \begin{pmatrix} \cos\frac{\pi}{6} & \sin\frac{\pi}{6} \\ \sin\frac{\pi}{6} & -\cos\frac{\pi}{6} \end{pmatrix}$$

From (1), recognize that $\check{\rho}_1$ is reflection through a line at angle $\frac{5\pi}{12}$, and $\check{\rho}_2$ is reflection through a line at angle $\frac{\pi}{12}$, and these differ by $\frac{\pi}{3}$. Thus, $\rho$ and $\rho^*$ are conjugate to the standard action of $\mathcal{D}_3$ on the plane described at the beginning of Section 1.




(A) Canonical rep.

(B) Geometric rep.

FIGURE 3. Canonical and geometric representations of $\mathcal{D}_3$. Note that the maps $\rho_i$ and $\rho_j^*$ each fix a line and exchange its sides, but are not orthogonal reflections.

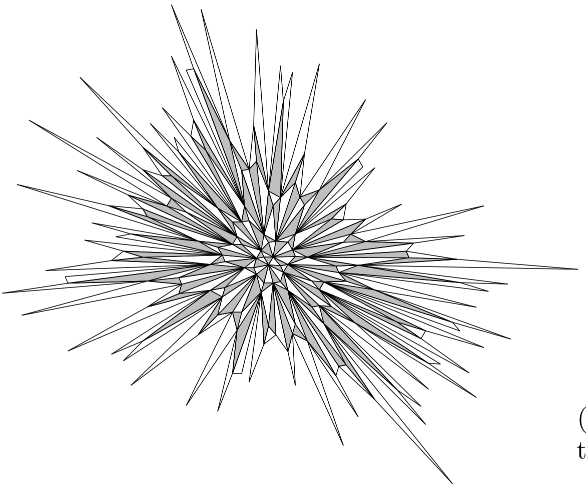
EXAMPLE 5.0.2 (Canonical vs geometric representation; hyperbolic simplex case). Consider the hyperbolic reflection group $W := \Delta(5,5,2)$. Its cosine matrix is $C = \begin{pmatrix} 1 & -\frac{1+\sqrt{5}}{4} & 0 \\ -\frac{1+\sqrt{5}}{4} & 1 & -\frac{1+\sqrt{5}}{4} \\ 0 & -\frac{1+\sqrt{5}}{4} & 1 \end{pmatrix}$; the canonical representation takes the generators to $\rho_1 = \begin{pmatrix} -1 & \frac{1+\sqrt{5}}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $\rho_2 = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1+\sqrt{5}}{2} & -1 & \frac{1+\sqrt{5}}{2} \\ 0 & 0 & 1 \end{pmatrix}$, and $\rho_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1+\sqrt{5}}{2} & -1 \end{pmatrix}$, and the geometric representation takes the generators to the linear maps whose matrices are the transposes of those. A part of the Coxeter complex is shown in Figure 4a.

The matrix $C$ has eigenvalues $\lambda_1 = 1$, $\lambda_2 = 1+\frac{1+\sqrt{5}}{2\sqrt{2}}$, and $\lambda_3 = 1-\frac{1+\sqrt{5}}{2\sqrt{2}}$. A matrix of corresponding unit eigenvectors is given by:

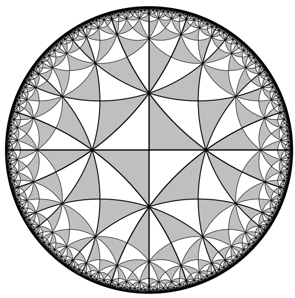
$$E := \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

The first two eigenvalues are positive, and the third is negative, so $C$ is of type (2,1). As in Proposition 4.0.8, there is a conjugate representation preserving the Minkowski form, where $\hat{\rho}_i := \left(E\sqrt{|\Lambda|}^{-1}\right)^{-1} \rho_i E\sqrt{|\Lambda|}^{-1}$. Further, $\sqrt{|\Lambda|}E^{-1}C^{-1}$ defines a map $V^* \to \mathbb{R}^3$ conjugating $\rho^*$ to $\hat{\rho}$ from $W$ to a group of maps preserving the Minkowski form. Restricting to the hyperboloid gives a representation of $W$ as a group of isometries of the hyperbolic plane generated by hyperbolic reflections through the three sides of a triangle.

Figure 4 shows the Coxeter complex for this group in $V^*$, and a conjugate representation into the Poincaré disc. Notice that combinatorially the two complexes of Figure 4 are identical. The action on $\mathbb{H}^2$ is by hyperbolic isometries, while the action on the Coxeter complex is only by $B_C$–isometries, not isometries of $\mathbb{R}^3$ with the Euclidean metric.



(A) Part of the Coxeter complex of $\Delta(5,5,2)$ in $V^*$, viewed from $(1,1,1)$.



(B) A hyperbolic simplicial tiling for $\Delta(5,5,2)$.

FIGURE 4. Two complexes for $\Delta(5,5,2)$.

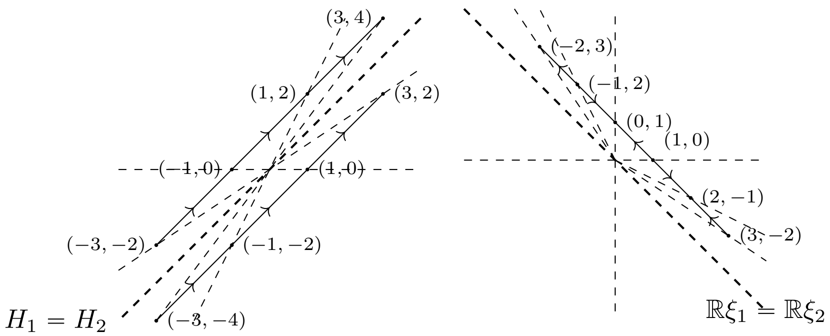
EXAMPLE 5.0.3 (Canonical vs geometric representation; Euclidean case). Consider $W = \langle s_1, s_2 \mid s_1^2,\, s_2^2 \rangle \cong \mathcal{D}_\infty$. Its cosine matrix is $C = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$.

In terms of the basis $(\mathbf{e}_1, \mathbf{e}_2)$ of $V = \mathbb{R}^2$ we have:

$$\rho_1 = \begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix} \qquad \rho_2 = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}$$

For this representation we have $H_1 = H_2 = \mathbb{R}(\mathbf{e}_1 + \mathbf{e}_2)$, the main diagonal. This line is fixed by the action. This is bad news; it does not leave a natural candidate for a fundamental domain for the action. The action is shown in Figure 5a. In this figure we see that it is possible to choose a sector to be a fundamental domain for the action on $\mathbb{R}^2 - H_1$. (Such a sector is relatively closed in $\mathbb{R}^2 - H_1$.) Pairs of lines equidistant from the diagonal are preserved by the action of $\mathcal{D}_\infty$, such that $s$ and $t$ exchange the pair and $st$ translates along both of them.

Compare this to the geometric representation in Figure 5b. In the previous cases we had an isomorphism $C\colon V \to V^*$, but in this case $C$ is not invertible, it has 1–dimensional image, sending $V$ to the subspace $\mathbb{R}\xi_1 = \mathbb{R}\xi_2$, which is preserved by the $\rho^*$ action. But the action does not *fix* $\mathbb{R}\xi_1$, it reverses it, and does not exchange complementary halfspaces. The fundamental chamber is the first quadrant, and the Tits cone is $\{a\mathbf{e}_1^* + b\mathbf{e}_2^* \mid b > -a\}$, so everything above and right of the fixed line $\mathbb{R}\xi_1$. The translates of the fundamental simplex $\{a\mathbf{e}_1^* + b\mathbf{e}_2^* \mid a + b = 1\}$ form a line parallel to the anti-diagonal on which $\mathcal{D}_\infty$ acts. This is a copy of the Coxeter complex, and the action restricted to this line is conjugate to the action of $\mathcal{D}_\infty$ on $\mathbb{R}$ described in Section 1.




(A) Canonical (B) Geometric

FIGURE 5. Linear representations of $\mathcal{D}_\infty$.

Not all Coxeter group are geometric reflection groups. In the next example we see that the copy of the Coxeter complex in $V^*$ can be distorted, and some points have infinite stabilizers. These problems will motivate us to construct a geometric complex with a more well behaved $W$–action in Chapter 5.

Example 5.0.4 (Canonical vs geometric representation; a non-geometric case). Consider $\langle r, s, t \mid r^2, s^2, t^2, (rt)^2 \rangle \cong \mathcal{D}_2 * \mathcal{C}_2 \cong \Delta(2, \infty, \infty)$.

$$\rho_1^* = \begin{pmatrix} -1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \rho_2^* = \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \qquad \rho_3^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}$$

The fundamental simplex is $\sigma := \{a\mathbf{e}_1^* + b\mathbf{e}_2^* + c\mathbf{e}_3^* \mid a + b + c = 1\}$. Figure 6 shows a portion of the Coxeter complex, viewed from along the diagonal.

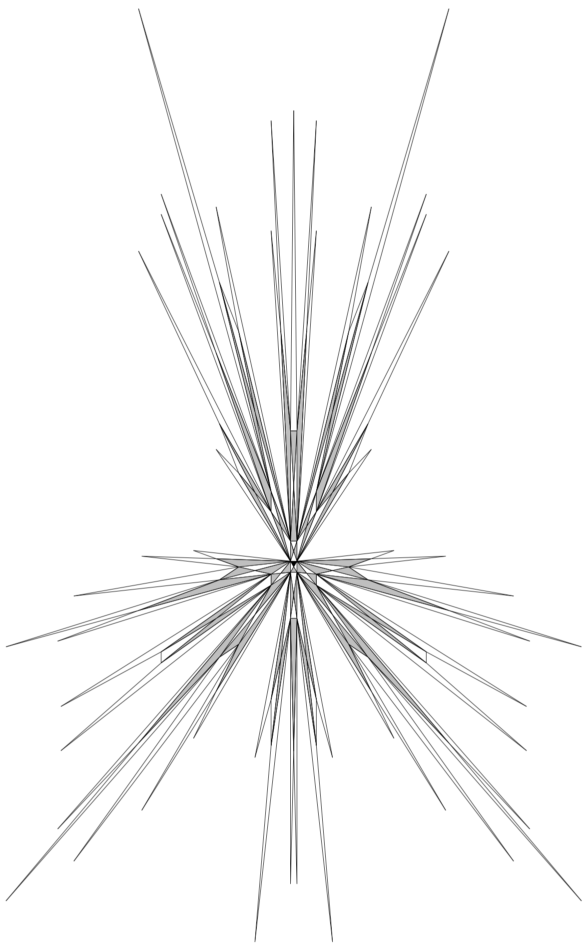


Figure 6. Part of the Coxeter complex of $\mathcal{D}_2 * \mathcal{C}_2$; radius = 7 reflections of fundamental simplex (black).

Observe that $\rho_1^*$ and $\rho_3^*$ both fix the vector $\mathbf{e}_2^*$ and commute with each other. For the $W$–action we have that $\mathcal{D}_2 \cong \mathcal{C}_2 \times \mathcal{C}_2 \cong \langle r, t \rangle$ fixes a vertex of the fundamental simplex, and this vertex has valence 4 in the Coxeter complex.

The other two rank 2 special subgroups are $\langle r, s \rangle \cong \mathcal{D}_\infty \cong \langle s, t \rangle$. Observe that for their images, $\rho_1^*$ and $\rho_2^*$ both fix the vector $\mathbf{e}_3^*$, and $\rho_2^*$ and $\rho_3^*$ both fix the vector $\mathbf{e}_1^*$. These two vertices of the fundamental simplex have infinite valence in the Coxeter complex.

Note that the topology of $\bigcup_{w\in W}\rho^*(w)\sigma$ induced from $\mathbb{R}^3$ is not the same as the topology of the corresponding simplicial complex near the infinite valence vertices. We can see this explicitly. The matrices $\rho_1^*\rho_2^*$ and $\rho_2^*\rho_1^*$ both have Jordan normal form $J := \begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}$. Observe $J^n\begin{pmatrix}0\\1\\2\end{pmatrix} = \begin{pmatrix}n^2\\ 1+2n\\ 2\end{pmatrix}$.

Let $C_{12} := \begin{pmatrix}0 & -1/2 & 1/4\\ 0 & 1/2 & 0\\ 1 & 0 & 0\end{pmatrix}$ and $C_{21} := \begin{pmatrix}0 & 1/2 & -1/4\\ 0 & -1/2 & 1/2\\ 1 & 0 & 0\end{pmatrix}$, so that $\rho_i^*\rho_j^* = S_{ij}JS_{ij}^{-1}$.

Consider the sequence of points $\mathbf{u}_n := (1-\frac{1}{2n^2})\mathbf{e}_3^* + \frac{1}{2n^2}\mathbf{e}_2^*$, which are all on one edge of $\sigma$. Compute $C_{12}J^nC_{12}^{-1}(\frac{1}{2n^2}\mathbf{e}_2^*) = \begin{pmatrix}-1/n\\ 1/n+1/(2n^2)\\ 1\end{pmatrix}$ and $C_{21}J^nC_{21}^{-1}(\frac{1}{2n^2}\mathbf{e}_2^*) = \begin{pmatrix}1/n\\ -1/n+1/(2n^2)\\ 1\end{pmatrix}$.

Conclude that $\rho_1^*\rho_2^*(\mathbf{u}_n)$ and $\rho_2^*\rho_1^*(\mathbf{u}_n)$ are sequences of points in the Coxeter complex, on distinct edges incident to $\mathbf{e}_3^*$, that in $\mathbb{R}^3$ limit to the same point $2\mathbf{e}_3^* \neq \mathbf{e}_3^*$. In the simplicial complex topology a sequence of points on distinct edges incident to a common vertex $v$ converges only if it converges to $v$.

Recall that we saw a different picture for $\Delta(2,\infty,\infty)$ in Figure 10m. Combinatorially, this is isomorphic to the complex in Figure 6 with the infinite valence vertices removed (pushed out to infinity). **This trades one problem for another: the action of $W$ on the Coxeter complex is cocompact but not proper, it has vertices with infinite stabilizer; the action on the hyperbolic plane is proper but not cocompact.**

CHAPTER 4

# Abstract reflection groups

## 1. Three definitions of abstract reflection group

In each of the next three subsections we will make a definition for a group to be an "abstract reflection group". It will turn out that all three will be equivalent to the group being a Coxeter group. In each case we make a first guess at a definition and then strengthen it.

This chapter is independent of Chapter 3. In particular, we will give an independent proof of the fact that a special subgroup of a Coxeter group is the Coxeter group defined by the corresponding subpresentation.

**1.1. Algebraic ARGs.** To start, the most basic property we could ask for in an "abstract reflection group" is that it be generated by elements of order 2.

Definition 1.1.1. A *preCoxeter system* is a pair $(W, S)$ where $W$ is a group generated by a set of involutions $S$.

Definition 1.1.2. Given a preCoxeter system $(W, S)$, the *associated Coxeter system* is $(\tilde{W}, \tilde{S})$ where $\tilde{S} = \{\tilde{s} \mid s \in S\}$ is a formally defined set of symbols and $\tilde{W}$ is the Coxeter group generated by $\tilde{S}$ with defining relations $(\tilde{s}\tilde{t})^{m_{st}}$ where $s, t \in S$ and $m_{st}$ is the order of $st$ in $W$.

By construction, the map $\tilde{s} \to s$ for $s \in S$ extends to a surjection $\tilde{W} \twoheadrightarrow W$, so every group generable by involutions is a quotient of some Coxeter group.

Definition 1.1.3 (Algebraic definition of ARG). Say that a group is *algebraically an abstract reflection group* if it admits a choice generating set of involutions such that the resulting surjection from the associated Coxeter group is an isomorphism.

Proposition 1.1.4. *Every Coxeter group is an algebraic abstract reflection group.*

Proof. Recall, as in the remark following the definition of Coxeter system, Chapter 1 Definition 4.1.4, that a Coxeter group has a Coxeter presentation in which a relation $(st)^m$ implies $st$ has order $m$ in the group. (In fact, by Chapter 3 Proposition 2.0.3, there is a unique Coxeter presentation for the choice of fundamental generating set, but this fact is unnecessary here.) For this Coxeter presentation, the Coxeter presentation of the associated Coxeter system is identical, so the associated Coxeter group is isomorphic to the one we started with. $\square$

PROPOSITION 1.1.5. *Let $(W, S)$ be a preCoxeter system, and let $s$ and $t$ be distinct elements of $S$ such that $st$ has order $m$ in $W$. Then $\langle s, t\rangle \cong \mathcal{D}_m$.*

PROOF. The relations $s^2 = 1$, $t^2 = 1$, and $(st)^m = 1$ are all satisfied in $W$ by hypothesis, so the dihedral group $\mathcal{D}_m$ surjects onto $\langle s, t\rangle < W$. But $\langle s, t\rangle$ cannot be a proper quotient of $\mathcal{D}_m$, because no proper quotient of $\mathcal{D}_m$ contains an element of order $m$ that is a product of two order 2 elements. $\square$

**1.2. Geometric ARGs.** In this section we try to be a little more geometric about the definition of abstract reflection group, by considering the geometry of a group action on a graph. In Section 1.2.1 we recall a construction of group actions on graphs. In Section 1.2.2 we add some geometric constraints on the action.

1.2.1. *Simple coset graphs.*

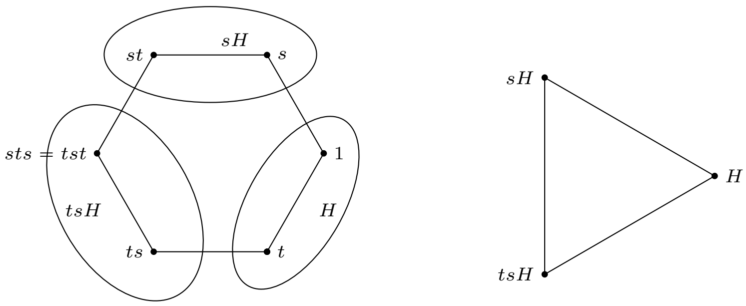
DEFINITION 1.2.1. Given a group $G$, subgroup $H < G$, and set $S \subset G$ such that $H \sqcup S$ generates $G$, define the *simple coset graph* $\mathcal{S}(G, H, S)$ to be the graph whose vertices are cosets $gH \in G/H$, with an edge between distinct vertices $g_1H$ and $g_2H$ if $g_1H$ and $g_2H$ are adjacent in $Cay(G, H \sqcup S)$, which is true if and only if there exist $h_1, h_2 \in H$ and $s \in S^{\pm}$ with $g_1h_1s = g_2h_2$.

EXAMPLE 1.2.2. If $H$ is trivial then $S$ is a generating set and $\mathcal{S}(G, H, S)$ is just the Cayley graph $Cay(G, S)$.

EXAMPLE 1.2.3. Figure 1 shows a Cayley graph and nontrivial coset graph for $\mathcal{D}_3$.

PROPOSITION 1.2.4. *Let $G$ be a group that acts on a connected simple graph $\Omega$, transitively on vertices. Let $v_0$ be a base vertex of $\Omega$, and let $H$ be the stabilizer of $v_0$. Let $S \subset G$ such that $\{sv_0 \mid s \in S\}$ contains exactly one point in each $H$–orbit of neighbors of $v_0$. Then $H \sqcup S$ generates $G$ and $\Omega$ is $G$–equivariantly isomorphic to $\mathcal{S}(G, H, S)$.*

PROOF. Transitivity of the action of $G$ on vertices of $\Omega$ implies that $\phi(gH) := gv_0$ defines a $G$–equivariant bijection between vertices.




(A) $Cay(\mathcal{D}_3, \{s,t\})$ with cosets of $H = \langle t \rangle = \{1,t\}$ circled.

(B) $\mathcal{S}(\mathcal{D}_3, H, \{s\})$

FIGURE 1. A Cayley graph and coset graph for $\mathcal{D}_3$.

To see that $\phi$ preserves the existence of edges between a pair of vertices, observe:

$$\begin{aligned}
&[g_1H, g_2H] \text{ is an edge of } \mathcal{S}(G,H,S)\\
&\qquad \iff \exists h_1, h_2 \in H,\ s \in S^{\pm}, \quad g_1h_1 = g_2h_2s\\
&\qquad \iff \text{up to exchanging subscripts, } \exists s \in S, \quad g_2^{-1}g_1 \in HsH\\
&\qquad \iff \exists s \in S, \quad g_2^{-1}g_1v_0 \in Hsv_0\\
&\qquad \iff [g_2^{-1}g_1v_0, v_0] \text{ is an edge of } \Omega\\
&\qquad \iff [g_1v_0, g_2v_0] \text{ is an edge of } \Omega
\end{aligned}$$

Since both graphs are simple, this implies $\phi$ is an isomorphism.

Finally, for any $g \in G$, since $\Omega$ is connected it is possible to choose an edge path $e_1, \ldots, e_k$ starting at $v_0$ and ending at $gv_0$. Label the vertices so that $e_i = [v_{i-1}, v_i]$. Since $v_1$ is adjacent to $v_0$, there exist $h_1 \in H$ and $s_1 \in S$ such that $v_1 = h_1s_1v_0$. Translate the path by $(h_1s_1)^{-1}$. Now $(h_1s_1)^{-1}v_2$ is adjacent to $(h_1s_1)^{-1}v_1 = v_0$, so there exist $h_2 \in H$ and $s_2 \in S$ such that $(h_1s_1)^{-1}v_2 = h_2s_2v_0$, or, $v_2 = h_1s_1h_2s_2v_0$. Continuing in this way, find that $gv_0 = v_k = h_1s_1 \cdots h_ks_kv_0$. Thus, $g \in h_1s_1 \cdots h_ks_kH \subset \langle H \sqcup S \rangle$. $\square$

Now consider a Coxeter system $(W,S)$, a subgroup $H < W$, and a set $S' \subset S$ of representatives of nontrivial elements of $H \backslash S / H$. Let us examine the action of $W$ on $\mathcal{S}(W,H,S')$.

LEMMA 1.2.5. *For $s \in S'$, the stabilizer of the edge $[H, sH]$ is $\langle s \rangle (H \cap sHs^{-1})$.*

PROOF. The pointwise stabilizer is $H \cap sHs^{-1}$, and this is a subgroup of index at most 2 in the stabilizer. But $s$ is an element of the stabilizer that is not in the pointwise stabilizer. $\square$

COROLLARY 1.2.6. *If $H \cap sHs^{-1} = 1$ for all $s \in S'$ then for $R$ the set of conjugates of elements of $S'$, every edge of $\mathcal{S}(W, H, S')$ is flipped by a unique element of $R$ and every element of $R$ flips some edge.*

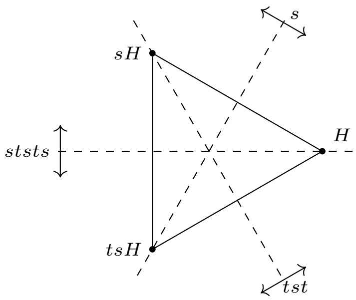
1.2.2. *Prereflection and reflection systems.*

DEFINITION 1.2.7. Let $W$ be a group. A *prereflection system* $(W, R, \Omega, v_0)$ for $W$ consists of a subset $R \subset W$ and an action $W \curvearrowright \Omega$ on a connected simple graph with base vertex $v_0$, satisfying the following conditions:

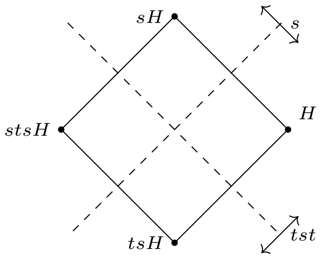
(1) Every element of $R$ is an involution.
(2) $R$ is closed under conjugation.
(3) For every edge of $\Omega$ there is a unique element of $R$ that "flips" it — acts by preserving the edge and exchanging its endpoints — and every element of $R$ flips at least one edge of $\Omega$.
(4) $R$ generates $W$.

By Proposition 1.2.4, every graph with transitive group action is a simple coset graph. Consider the following example and non-example of prereflection systems.

EXAMPLE 1.2.8. Figure 2 show two examples of dihedral groups $\mathcal{D}_m = \langle s, t \mid s^2, t^2, (st)^m \rangle$ acting on simple coset graphs. Figure 2a is a prereflection system, but Figure 2b is not, because in that case $R$ does not generate $\mathcal{D}_4$.




(A) $\mathcal{D}_3 \curvearrowright \mathcal{S}(\mathcal{D}_3, \langle t \rangle, \{s\})$,
$R = \{s, tst, ststs\}$




(B) $\mathcal{D}_4 \curvearrowright \mathcal{S}(\mathcal{D}_4, \langle t \rangle, \{s\})$,
$R = \{s, tst\}$

FIGURE 2. Two examples of $\mathcal{D}_m$ acting on one of its simple coset graphs. The $m = 3$ case is a prereflection system, but the $m = 4$ case is not, because in that case $R$ does not generate $\mathcal{D}_4$.

EXERCISE 1.2.9. For $2 \leqslant m \leqslant \infty$, can you classify subgroups $H$ of $\mathcal{D}_m$ for which $\mathcal{D}_m \curvearrowright \mathcal{S}(\mathcal{D}_m, H, S')$ is a prereflection system?

LEMMA 1.2.10. *Let* $(W, R, \Omega, v_0)$ *be a prereflection system, and let* $S := \{r \in R \mid r \text{ flips an edge incident to } v_0\}$. *For all* $k \in \mathbb{N}$ *there is a bijection between edge paths of length* $k$ *in* $\Omega$ *starting from* $v_0$ *and tuples* $\bar{s} = (s_1, \ldots, s_k) \in S^k$.

*Moreover, the edge path corresponding to* $\bar{s}$ *ends at* $s_1 \cdots s_k v_0$.

PROOF. Given $\bar{s} = (s_1, \ldots, s_k)$, define $w_0 = 1$ and $w_i = s_1 \cdots s_i$ for each $1 \leqslant i \leqslant k$. By definition, for each $s \in S$ there exists an edge incident to $v_0$ that is flipped by $s$, so it has vertices $v_0$ and $sv_0$. Since $\Omega$ is simple, an edge is determined by its vertices, so $[v_0, sv_0]$ is the unique edge incident to $v_0$ flipped by $s$. For any $w \in W$, there is a corresponding edge $[wv_0, wsv_0]$ flipped by $wsw^{-1}$. If we define elements $r_i := w_{i-1} s_i w_{i-1}^{-1}$ for $1 \leqslant i \leqslant k$, then $r_i$ flips the edge $e_i := [w_{i-1}v_0, w_i v_0]$. These edges fit together to make an edge path from $v_0$ to $w_k v_0$.

Conversely, if $e_1,\ldots,e_k$ is an edge path starting at $v_0$, then for each $i$ there is a unique $r_i \in R$ that flips $e_i$. Since the edge path starts at $v_0$, we can number the vertices so that $e_i = [v_{i-1}, v_i]$. Define $s_1 := r_1$, which is in $S$ since $e_1$ is incident to $v_0$. Then $e_1 = [v_0, s_1 v_0]$. Apply $s_1$ to the last $k-1$ edges to get an edge path $s_1 e_2, \ldots, s_1 e_k$ starting from $s_1 v_1 = v_0$ whose edges are (uniquely) flipped by $s_1 r_2 s_1, \ldots, s_1 r_k s_1$, respectively. Define $s_2 := s_1 r_2 s_1 \in S$ flipping the first edge, so the first edge is $[v_0, s_2 v_0] = s_1 e_2 = [s_1 v_1, s_1 v_2]$. This shows that $e_2 = [s_1 v_0, s_1 s_2 v_0]$. Now apply $s_2$ to the last $k-2$ edges of the new path to get another path $s_2 s_1 e_3,\ldots,s_2 s_1 e_k$ starting at $s_2 s_1 v_2 = v_0$. The edge $s_2 s_1 e_3$ is flipped by $s_3 := s_2 s_1 r_3 s_1 s_2$. Continuing in this way, define $\bar{s}$ with $s_i := w_{i-1}^{-1} r_i w_{i-1} \in S$ and see that the edge $e_i$ is $[w_{i-1}v_0, w_i v_0]$, which is expected $i$–th edge from the first part of the argument. □

DEFINITION 1.2.11. For $\bar{s} = (s_1, \ldots, s_k) \in S^k$, let $w_i(\bar{s}) := s_1 \cdots s_i$ and $w(\bar{s}) := w_k(\bar{s})$.

COROLLARY 1.2.12. *The subgroup of* $W$ *generated by* $S$ *acts transitively on the vertices of* $\Omega$.

PROOF. $\Omega$ is connected, so for any vertex $v$ there is an edge path starting from $v_0$ that leads to $v$. Take $\bar{s}$ associated with this edge path. Then $v = w(\bar{s})v_0$. □

COROLLARY 1.2.13. $S$ *generates* $W$*, and* $R$ *is the set of conjugates of* $S$.

PROOF. Let $r \in R$. It flips some edge $e = [v, v']$. There is an element $w$ of $\langle S \rangle$ taking $v$ to $v_0$, so $we$ is flipped by some $s \in S$. Conversely, this means

$e = w^{-1}we$ is flipped by $w^{-1}sw \in R$. But there is a unique element of $R$ flipping $e$, so $r = w^{-1}sw$. Not only does this express $r$ as a conjugate of an element of $S$, the conjugator was in $\langle S\rangle$, so $R \subset \langle S\rangle$. Since $R$ generates $W$ by hypothesis, $S$ generates $W$. □

Recall that the *word length* $|w|_S$ of an element $w \in W$ with respect to the generating set $S$ is the minimal length of an expression of $w$ as a product of elements of $S$ (In general we should say products of elements of $S$ or their inverses, but in this case elements of $S$ are involutions.).

Definition 1.2.14. Given a prereflection system as above, for each $r \in R$ define the *wall corresponding to* $r$ to be the set:

$$\Omega^r := \{\text{midpoint of } e \mid r \text{ flips } e\}$$

Let $x$ and $y$ be vertices of $\Omega$, and define the *set of walls separating $x$ and $y$* to be:

$$\mathcal{W}(x, y) := \{r \in R \mid x \text{ and } y \text{ are in different components of } \Omega^r\}$$

Define the *wall distance* between $x$ and $y$ to be $d_{\mathcal{W}}(x, y) := \#\mathcal{W}(x, y)$.

The graph distance, that is, the minimum length of an edge path between two vertices, will be denoted $d_\Omega$.

Lemma 1.2.15. *For all $w \in W$, $d_{\mathcal{W}}(v_0, wv_0) \leqslant d_\Omega(v_0, wv_0) \leqslant |w|_S$.*

Proof. If a wall separates two vertices then every edge path between them must cross the wall. By definition of a prereflection system, every edge crosses a unique wall. Thus, $d_{\mathcal{W}} \leqslant d_\Omega$.

By Lemma 1.2.10, a minimal expression of $w \in W$ in terms of the generating set $S$ gives an edge path from $v_0$ to $wv_0$ of the same length, so $d_\Omega(v_0, wv_0) \leqslant |w|_S$. □

Definition 1.2.16. A prereflection system $(W, R, \Omega, v_0)$ is a *reflection system* if the following conditions, which will shown to be equivalent in Lemma 1.2.23, are satisfied:

- Every wall has exactly two complementary components.
- For all $w \in W$, $d_{\mathcal{W}}(v_0, wv_0) = d_\Omega(v_0, wv_0) = |w|_S$.

In this case, the closure of a complementary component of $\Omega - \Omega^r$ is a *halfspace*, and the component containing $v_0$ is the *positive halfspace for* $r$.

Exercise 1.2.17. Show that $\mathcal{D}_m \curvearrowright Cay(\mathcal{D}_m, \{s, t\})$ is a reflection system. The case $m = 3$ is shown in Figure 3.

Definition 1.2.18. A tuple $\bar{s} \in S^k$ is *minimal* if $|w(\bar{s})|_S = k$.

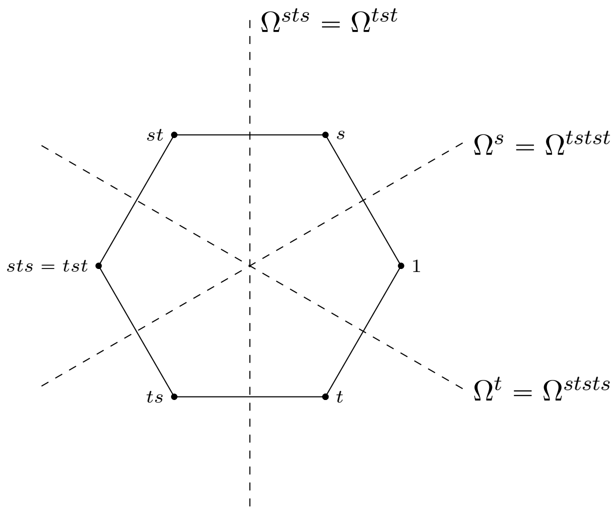



FIGURE 3. $Cay(\mathcal{D}_3, \{s,t\})$ as a reflection system

DEFINITION 1.2.19. With notation as in Lemma 1.2.10, define $\Phi\colon S^k \to R^k$ by $\bar{s} \mapsto (r_1, \dots, r_k)$, so that $r_i = w_{i-1}s_iw_{i-1}^{-1}$ is the unique element of $R$ that flips the $i$–th edge in the edge path corresponding to $\bar{s}$.

LEMMA 1.2.20. *If $\Phi(\bar{s}) = (r_1, \dots, r_k)$ then the sequence of walls crossed by the edge path corresponding to $\bar{s}$ is $\Omega^{r_1}, \dots, \Omega^{r_k}$.*

LEMMA 1.2.21. *In a prereflection system, if an edge path starting at $v_0$ crosses some wall more than once then the corresponding $\bar{s}$ is not minimal.*

*More specifically, let $\bar{s} = (s_1, \dots, s_k) \in S^k$ and $\Phi(\bar{s}) = (r_1, \dots, r_k)$, and suppose for some $i < j$ that $r_i = r_j$, and that $j$ is the smallest index for which this is true. Let $\bar{s}' := (s_1, \dots, \hat{s}_i, \dots, \hat{s}_j, \dots, s_k) \in S^{k-2}$, where $\hat{s}_i$ means 'omit this entry'. Then the edge paths corresponding to $\bar{s}$ and $\bar{s}'$ have the same endpoints, and $w(\bar{s}) = w(\bar{s}')$.*

PROOF. Let $r := r_i = r_j$. Let $\gamma$ be the edge path corresponding to $\bar{s}$, and let $\delta := e_{i+1} + \cdots + e_{j-1}$ be the subpath of $\gamma$ consisting of edges between $e_i$ and $e_j$. Minimality of $j$ implies that $\delta$ does not cross $\Omega^r$. Consider $r\delta$. Whereas $\delta$ starts from the terminal vertex of $e_i$ and ends at the initial vertex of $e_j$, since $r$ flips both of these edges, $r\delta$ starts at the initial vertex of $e_i$ and ends at the terminal vertex of $e_j$. See Figure 4. Since $\delta$ and $r\delta$ have the same length, the path $\gamma' := e_1 + \cdots + e_{i-1} + r\gamma + e_{j+1} + \cdots + e_k$ has the same endpoints as $\gamma$, but is shorter by two edges.

Let $w_\ell := s_1 \cdots s_\ell$, and let $w'_\ell$ be the product of the first $\ell$ elements of $\bar{s}'$. Let $s'_\ell$ be the $\ell$–th element of $\bar{s}'$.

To see that $\gamma'$ is the edge path corresponding to $\bar{s}'$, we must show that $r'_\ell := w'_{\ell-1}s'_\ell(w'_{\ell-1})^{-1}$ flips the $\ell$–th edge of $\gamma'$. This is clear for the first $i-1$

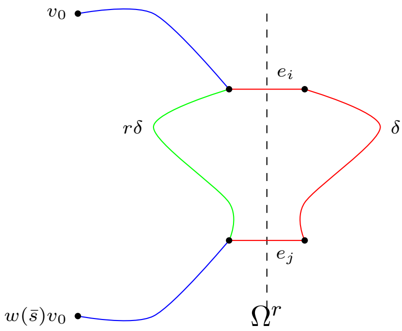



FIGURE 4. The subpath between consecutive wall crossings can be shortened.

entries, since $\bar{s}$ and $\bar{s}'$ agree on those entries, and $\gamma$ and $\gamma'$ share an initial subpath of length $i-1$. Moreover, $w_\ell = w'_\ell$ for $\ell < i$.

If $i \leqslant \ell < j-1$, then $\bar{s}'_\ell = s_{\ell+1}$.

$$\begin{aligned} r_{\ell+1} &= s_1 \cdots s_\ell s_{\ell+1} s_\ell \cdots s_1 \\ &= w_{i-1} s_i w_{i-1}^{-1} w_{i-1} s_{i+1} \cdots s_\ell s_{\ell+1} s_\ell \cdots w_{i-1}^{-1} w_{i-1} s_i w_{i-1}^{-1} \\ &= r_i s_1 \cdots s_{i-1} s_{i+1} \cdots s_\ell s_{\ell+1} s_\ell \cdots s_{i+1} s_{i-1} \cdots s_1 r_i \\ &= r w'_\ell s_{\ell+1} (w'_\ell)^{-1} r \\ &= r w'_\ell s'_\ell (w'_\ell)^{-1} r \\ &= r r'_\ell r \end{aligned}$$

Now $r_{\ell+1}$ flips $e_{\ell+1}$, which is the $(\ell+1-i)$–th edge in $\delta$, so its conjugate $r r_{\ell+1} r = r'_\ell$ flips $r e_{\ell+1}$, which is the $(\ell+1-i)$–th edge of $\delta' = r\delta$, so is the $\ell$–th edge of $\gamma'$.

Now suppose $\ell \geqslant j-1$. Observe that $w_{i-1} s_i (w_{i-1})^{-1} = r_i = r_j = w_{j-1} s_j (w_{j-1})^{-1}$, so:

$$s_i = (w_{i-1})^{-1} w_{j-1} s_j (w_{j-1})^{-1} w_{i-1} = s_i s_{i+1} \cdots s_{j-1} s_j s_{j-1} \cdots s_{i+1} s_i$$

Thus:

$$s_i s_{i+1} \cdots s_{j-1} s_j = s_{i+1} \cdots s_{j-1} \tag{12}$$

Since $\ell \geqslant j-1$, $s'_\ell = s_{\ell+2}$, and:

$$\begin{aligned} w'_\ell &= s_1 \cdots s_{i-1} \hat{s}_i s_{i+1} \cdots s_{j-1} \hat{s}_j s_{j+1} \cdots s_{\ell+2} \\ &\overset{(12)}{=} s_1 \cdots s_{i-1} s_i s_{i+1} \cdots s_{j-1} s_j s_{j+1} \cdots s_{\ell+2} \\ &= w_{\ell+2} \end{aligned}$$

This gives:

$$\begin{aligned} r_{\ell+2} &= w_{\ell+1} s_{\ell+2} (w_{\ell+1})^{-1} \\ &= w'_{\ell-1} s_{\ell+2} (w'_{\ell-1})^{-1} \\ &= w'_{\ell-1} s'_{\ell} (w'_{\ell-1})^{-1} \\ &= r'_{\ell} \end{aligned}$$

Since $r_{\ell+2}$ flips the $(\ell+2)$–nd edge of $\gamma$, which is the $\ell$–th edge of $\gamma'$, we are done. □

LEMMA 1.2.22. *Let $(W, R, \Omega, v_0)$ be a prereflection system. For all $r \in R$, $\Omega - \Omega^r$ has either 1 or 2 connected components and if there are 2 then they are exchanged by the $r$–action.*

PROOF. Since $R$ consists of conjugates of $S$, every wall is a translate of some $\Omega^s$ for $s \in S$, so it suffices to prove the lemma for $r = s \in S$.

For any vertex $v \in \Omega$, choose a minimal length edge path from $v_0$ to $v$. By Lemma 1.2.21, the edge path crosses the wall $\Omega^s$ at most once, since if it crossed at least twice we could replace it by a shorter path with the same endpoints.

If the path does not cross $\Omega^s$ then $v$ and $v_0$ are in the same component of $\Omega - \Omega^s$.

Suppose the path crosses $\Omega^s$ once. Then it can be written $\gamma + e + \delta$, where $\gamma$ and $\delta$ are edge paths that do not cross $\Omega^s$, and $e$ is an edge that does. Let $e = [x, y]$. Consider the path $[v_0, sv_0] + s\gamma + \delta$. This is a path, since $s\gamma$ begins at $sv_0$ and ends at $sx = y$. See Figure 5. Its first edge

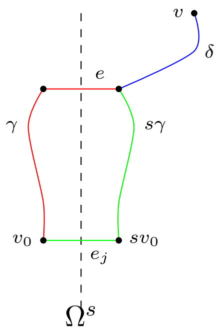



FIGURE 5. When a path crosses a wall once it can be replaced by a path with the same length and endpoints that crosses the wall on its first step.

$[v_0, sv_0]$ crosses $\Omega^s$. The segment $\delta$ does not cross $\Omega^s$ by definition. The segment $s\gamma$ also does not cross $\Omega^s$, because if it did then an edge crossing $\Omega^s$ would be fixed by the $s$–action, so $\gamma$ also would have had an edge crossing

$\Omega^s$. Therefore, we can replace the path from $v_0$ to $v$ by a path with the same endpoints and the same length, such that the crossing of $\Omega^s$ occurs on its first edge $[v_0, sv_0]$. Thus, every vertex of $\Omega$ is in the same component of $\Omega - \Omega^s$ with either $v_0$ or $sv_0$. These are not necessarily exclusive, but if $v_0$ and $sv_0$ are in different components of $\Omega^s$ then, since the $s$–action exchanges $v_0$ and $sv_0$, it exchanges the two components. □

LEMMA 1.2.23. *Let $(W, R, \Omega, v_0)$ be a prereflection system. The following are equivalent:*

- *Every wall has exactly two complementary components.*
- *For all $w \in W$, $d_{\mathcal{W}}(v_0, wv_0) = d_\Omega(v_0, wv_0) = |w|_S$.*

PROOF. By Lemma 1.2.22, if some wall does not have two complementary components then it has only one. Since the action is transitive on vertices, we may assume that wall is adjacent to $v_0$, so that there is $s \in S$ such that $\Omega^s$ has only one complementary component. There is an edge path from $v_0$ to $sv_0$ consisting of a single edge $[v_0, sv_0]$ that crosses only $\Omega^s$, so no other wall separates $v_0$ and $sv_0$. But $\Omega^s$ also does not separate $v_0$ and $sv_0$, since it only has one complementary component, so no walls separate $v_0$ and $sv_0$. Thus $d_{\mathcal{W}}(v_0, sv_0) = 0 < 1 = d_\Omega(v_0, sv_0) = |s|_S$.

Now suppose that every wall has two complementary components. Suppose that $\bar{s} = (s_1, \ldots, s_k)$ is minimal, so that $|w(\bar{s})|_S = k$. Lemma 1.2.21 says the edge path for a minimal tuple crosses no wall more than once, so the walls $\Omega^{r_i}$ for $\Phi(\bar{s}) = (r_1, \ldots, r_k)$ are distinct. Since each of these has two complementary components, the edge path crosses from the $v_0$ side of $\Omega^{r_i}$ to the other side, and does not return. Thus, each wall $\Omega^{r_i}$ separates $v_0$ from $w(\bar{s})v_0$. This shows $k \leqslant d_{\mathcal{W}}(v_0, w(\bar{s})v_0)$, but we already knew $d_{\mathcal{W}}(v_0, w(\bar{s})v_0) \leqslant d_\Omega(v_0, w(\bar{s})v_0) \leqslant |w|_S = k$, so these are equalities. □

COROLLARY 1.2.24. *If $(W, R, \Omega, v_0)$ is a reflection system then $\bar{s}$ is minimal if and only if the corresponding edge path crosses no wall more than once.*

PROOF. Lemma 1.2.21 says the edge path for a minimal tuple crosses no wall more than once. Conversely, $\bar{s}$ has length $k$ and the edge path for $\bar{s}$ crosses no wall more than once and each of them has two components then $k = d_{\mathcal{W}}(v_0, w(\bar{s})v_0)$, but, by Lemma 1.2.23, $d_{\mathcal{W}}(v_0, w(\bar{s})v_0) = |w(\bar{s})|_S$. □

COROLLARY 1.2.25. *If $(W, R, \Omega, v_0)$ is a reflection system then halfspaces are convex.*

PROOF. Consider a wall $\Omega^r$ and one of its halfspaces $H$. A path between two points of $H$ crosses $\Omega^r$ an even number of times, but a geodesic crosses at most once, so a geodesic between points of $H$ crosses $\Omega^r$ zero times; that is a geodesic between points of $H$ stays in $H$. $\square$

Corollary 1.2.26. *If $(W, R, \Omega, v_0)$ is a reflection system then the action of $W$ on vertices of $\Omega$ is free.*

Proof. Since the action on vertices is transitive, it suffices to show $v_0$ has trivial stabilizer. Suppose $w$ fixes $v_0$. Then $d_\Omega(v_0, wv_0) = 0$, but, by Lemma 1.2.23, $d_\Omega(v_0, wv_0) = |w|_S$, so $|w|_S = 0$, which means $w = 1$. $\square$

Corollary 1.2.27. *If $(W, R, \Omega, v_0)$ is a reflection system then $\Omega$ is isomorphic to the Cayley graph $Cay(W, S)$ of $W$ with respect to $S$.*

Exercise 1.2.28. Consider $2 \leqslant m < \infty$, $S = \{s, t\}$, $W = \mathcal{D}_m \cong \langle s, t \mid s^2, t^2, (st)^m \rangle$. Show that $(W, R, Cay(W, S), 1)$ for $R$ the set of conjugates of $S$ is a reflection system. Show that $Cay(W, S)$ is a $2m$–gon such that every wall consists of midpoints of edges on opposite sides.

Exercise 1.2.29. Consider the Coxeter system $(W, S)$ defined by Coxeter graph $\overset{a}{\bullet}\!\!\longrightarrow\!\!\!\!-\!\!\overset{b}{\bullet}\!\!\longrightarrow\!\!\!\!-\!\!\overset{c}{\bullet}$. (Recall Figure 3d of Chapter 2.) Draw $Cay(W, S)$, and show that it is a reflection system. Show that there is a unique vertex at maximal distance from 1.

Definition 1.2.30 (Geometric definition of ARG). Say a group $W$ is *geometrically an abstract reflection group* if it admits a choice of generating set of involutions $S$ such that for $R$ the set of all conjugates of generators, $(W, R, Cay(W, S), 1)$ is a reflection system.

Theorem 1.2.31. *[Algebraic ARG $\implies$ geometric ARG] If $(W, S)$ is a Coxeter system and $R$ is the set of conjugates of elements of $S$ then $(W, R, Cay(W, S), 1)$ is a reflection system.*

The idea of the proof is the following: If the Cayley graph is a reflection system, then each wall defines a positive and negative halfspace. $W$ acts on the collection of halfspaces, which we can identify with $R \times \{\pm 1\}$. This action is a homomorphism from $W$ into the permutation group of $R \times \{\pm 1\}$. In the proof we will define such a homomorphism from scratch, and then conclude that it must have arisen from an action on a collection of halfspaces.

Proof. For $s \in S$ define $\phi(s)$ on $(r, \epsilon) \in R \times \{\pm 1\}$ as follows:

$$\phi(s)(r, \epsilon) = \begin{cases} (srs, \epsilon) \text{ if } r \neq s \\ (srs, -\epsilon) \text{ if } r = s \end{cases}$$

This is an order 2 permutation of $R \times \{\pm 1\}$, which we reinterpret as follows: The first coordinate is given by the conjugation action in $W$. The second coordinate multiplies the existing $\epsilon$ by a power of $-1$, where the power is the number of times the edge path corresponding to $(s)$ crosses the wall $\Omega^r$.

Extend this to tuples by $\phi(s_1, \ldots, s_k) := \phi(s_k) \circ \phi(s_{k-1}) \circ \cdots \circ \phi(s_1)$. We will prove by induction that $\phi(\bar{s})$ takes $(r, \epsilon)$ the pair $(w(\bar{s})^{-1} r w(\bar{s}), (-1)^p \epsilon)$, where $p$ is the number of times the path corresponding to $\bar{s}$ crosses $\Omega^r$. We have already established the case $|\bar{s}| = 1$, so let $\bar{s}' = (s_1, \ldots, s_{k-1})$, and assume the claim is true for $\phi(\bar{s}')$, so that $\phi(\bar{s}) = \phi(s_k) \circ \phi(\bar{s}')$. The first coordinate of $\phi(s_k) \circ \phi(\bar{s}')(r, \epsilon)$ is $s_k w(\bar{s}')^{-1} r w(\bar{s}') s_k = w(\bar{s})^{-1} r w(\bar{s})$, as desired. For the second coordinate, the sign changes when applying $\phi(s_k)$ if and only if $s_k$ is equal to the first coordinate of the output of $\phi(\bar{s}')(r, \epsilon)$, which is $w(\bar{s}')^{-1} r w(\bar{s})$. So the sign changes if and only if $r = w(\bar{s}') s_k w(\bar{s}')^{-1}$. Now observe that the edge path corresponding to $\bar{s}$ is the edge path corresponding to $\bar{s}'$ plus one additional edge that is flipped by $w(\bar{s}') s_k w(\bar{s}')^{-1}$. So we conclude the sign on the second coordinate changes when applying $\phi(s_k)$ if and only if the edge path corresponding to $\bar{s}$ crosses $\Omega^r$ on its last edge, so has exactly one more $\Omega^r$ crossing than does its initial subpath corresponding to $\bar{s}'$.

We can think of the map defined so far as a homomorphism from the free product of $|S|$ copies of $\mathcal{C}_2$ to permutations of $R \times \{\pm 1\}$. We want to see that it factors through $W$, so we should check that the relations of $W$ are satisfied. Let $s, t \in S$ and suppose that $st$ has order $m < \infty$ in $W$, so that $(st)^m$ a relation in the Coxeter presentation. Let $\bar{s}$ be the alternating tuple $(s, t, \ldots)$ of length $2m$, and consider $\phi(\bar{s})(r, \epsilon)$. The first coordinate of the result is given by conjugation by $w(\bar{s}) = 1$, so yields $r$. The second coordinate is multiplied by $-1$ to the power the number of times the path for $\bar{s}$ crosses $\Omega^r$. We must show this number is even. The element the flips the $i$–th edge in this path is $w_{i-1}(\bar{s}) s_i w_{i-1}(\bar{s})^{-1}$, so this number is certainly 0 if $r \notin \langle s, t \rangle$. Suppose instead that $r \in \langle s, t \rangle$. Since $\langle s, t \rangle \cong \mathcal{D}_m$, by Exercise 1.2.28, the edge path corresponding to $w(\bar{s})$ goes once around the Cayley graph for $\mathcal{D}_m$, crossing every wall twice.

It remains to show that every wall has two complementary components. By Lemma 1.2.22 and using the group action, this is true unless there is $s \in S$ such that $\Omega^s$ has only one complementary component. Suppose this is so. Then there is some edge path from 1 to $s$ that does not cross $\Omega^s$. Let $\bar{t}$ be the corresponding tuple, and let $\bar{s} = (s)$. Compute $\phi(\bar{s})(s, 1) = (s, -1) \neq (s, 1) = \phi(\bar{t})(s, 1)$. This is a contradiction, because we have shown that $\phi$ factors though $W$, so since $w(\bar{t}) = w(\bar{s})$, $\phi(\bar{t})$ and $\phi(\bar{s})$ must agree on all inputs. $\square$

**1.3. Combinatorial ARGs.** In this section we consider a preCoxeter system as a collection of words with combinatorial rewriting rules.

DEFINITION 1.3.1. For a preCoxeter system $(W, S)$, define the following rewriting conditions for tuples in $S^*$:

(D) *The deletion condition*: If $\bar{s}$ is not minimal then there are indices $1 \leqslant i < j \leqslant k = |\bar{s}|$ such that for $\bar{s}' = (s_1, \dots, \hat{s}_i, \dots, \hat{s}_j, \dots, s_k)$ we have $w(\bar{s}) = w(\bar{s}')$.

(E) *The exchange condition*: If $\bar{s} = (s_1, \dots, s_k)$ is minimal and there is an element $s_0 \in S$ such that $(s_0, s_1, \dots, s_k)$ is not minimal then there is an index $1 \leqslant i \leqslant k$ such that for $\bar{s}' = (s_0, s_1, \dots, \hat{s}_i, \dots, s_k)$ we have $w(\bar{s}) = w(\bar{s}')$.

(F) *The folding condition* Let $w \in W$, $s, t \in S$ such that $|sw|_S = |w|_S + 1$, $|wt|_S = |w|_S + 1$, and $|swt|_S < |w|_S + 2$. Then $swt = w$.

The first two conditions mirror the properties of prereflection systems of Figure 4 in Lemma 1.2.21 and of Figure 5 in Lemma 1.2.22.

THEOREM 1.3.2 ([**11**, Theorem 3.2.16]). *Conditions (D), (E), and (F) are equivalent.*

We will prove the equivalence of (D) and (E).

PROOF. Suppose condition (D) holds. Suppose $\bar{s}$ is minimal and there is an element $s_0 \in S$ such that $\bar{s}'' := (s_0, s_1, \dots, s_k)$ is not minimal. Apply (D) to get indices $0 \leqslant i < j \leqslant k$ whose entries can be deleted without changing $w(\bar{s}'')$. Since $\bar{s}$ is minimal $i = 0$, so $w(\bar{s}'') = s_0 s_1 \cdots s_k = s_1 \cdots \hat{s}_j \cdots s_k$. Thus, $s_0 s_1 \cdots \hat{s}_j \cdots s_k = s_0 w(\bar{s}'') = w(\bar{s})$, which establishes (E).

Suppose condition (E) holds and $\bar{s}$ is not minimal. Let $\bar{s}'$ be the shortest nonminimal suffix of $\bar{s}$. Length 1 tuples are minimal, so $|\bar{s}'| > 1$, so $\bar{s}' = (s_i, \dots, s_k)$ for some $i < k = |\bar{s}|$ is nonminimal, but $\bar{s}'' := (s_{i+1}, \dots, s_k)$ is minimal. Apply (E) to $\bar{s}''$ and conclude there is an index $i + 1 \leqslant j \leqslant k$ such that $w(\bar{s}'') = s_i s_{i+1} \cdots \hat{s}_j \cdots s_k$. Conclude (D), since:

$$w(\bar{s}) = s_1 \cdots s_i w(\bar{s}'') = s_1 \cdots s_i s_i s_{i+1} \cdots \hat{s}_j \cdots s_k = s_1 \cdots \hat{s}_i \cdots \hat{s}_j \cdots s_k \quad \square$$

DEFINITION 1.3.3 (Combinatorial definition of ARG). Say that a group $W$ is *combinatorially an abstract reflection group* if it admits a choice of generating set of involutions $S$ such that the preCoxeter system $(W, S)$ satisfies the equivalent conditions of Theorem 1.3.2.

DEFINITION 1.3.4. Let $(W, S)$ be a preCoxeter system such that for $s, t \in S$ the order of $st$ is $m_{st}$. An *elementary $M$–operations* on a tuple $\bar{s} = s_1, \dots, s_k \in S^k$ is one of the two operations:

(I) Delete a subword $(s, s)$.

(II) For $s, t \in S$, replace an alternating subword $(s, t, \dots)$ of length $m_{st}$ by the alternating word $(t, s, \dots)$ of length $m_{ts}$.

A tuple $\bar{s}$ is *$M$–reduced* if it cannot be shortened by a sequence of elementary $M$–operations.

Tuples that are related by a sequence of elementary $M$–operations define the same element of $W$, since these two elementary operations correspond to relations in the group. Tits showed the converse was true:

THEOREM 1.3.5 (Word Problem for combinatorial ARG's). *If $(W, S)$ is a preCoxeter system satisfying condition $(E)$ then:*

(1) *A tuple $\bar{s}$ is minimal if and only if it is $M$–reduced.*
(2) *Two minimal tuples $\bar{s}$ and $\bar{t}$ represent the same element of $W$ if and only if they are related by a sequence of Type II moves.*

PROOF. First prove Item (2). Suppose $\bar{s} = (s_1, \ldots, s_k)$ and $\bar{t} = (t_1, \ldots, t_k)$ are two minimal expressions of the same element $w$ of $W$. If $k = 1$ then $w = s_1 = t_1$ and we are done: no moves are necessary. Proceed by induction on $k$, so suppose Item (2) is true for all tuples of length less than $k$. First suppose that $s_1 = t_1$. Then $(s_2, \ldots, s_k)$ and $(t_2, \ldots, t_k)$ are minimal expressions of $s_1 w = t_1 w$, so by the induction hypothesis, they differ by a sequence of Type II moves. But then $\bar{s}$ and $\bar{t}$ differ by the same sequence of Type II moves, as desired.

Thus, we may suppose that $s_1 \neq t_1$. Let $m$ be the order of $s_1 t_1$ in $W$. The strategy is to show that $\bar{s}$ may be replaced by a tuple $\bar{u}$ that begins with an alternating word $(s_1, t_1, \ldots)$ of length $m$. To this word we apply a Type II move swapping the length $m$ prefix with the alternating word $(t_1, s_1, \ldots)$. Call the result $\bar{u}'$. Now, regardless of how we discovered $\bar{u}$, the first case, when the leading letters are the same, says there are sequence of Type II moves changing $\bar{s}$ to $\bar{u}$ and changing $\bar{u}'$ to $\bar{t}$, so $\bar{s}$ and $\bar{t}$ differ by a sequence of Type II moves.

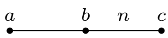
We justify the strategy by producing $\bar{u}$. Observe that $|t_1 w|_S < |w|_S$, since $w$ has a minimal expression staring with $t_1$. Apply condition (E): there exists $i$ such that $s_i = t_1$ and we get a new expression for $w$ by moving that $t_1$ to the front, $\bar{s}' = (t_1, s_1, \ldots, \hat{s}_i, \ldots, s_k)$. Now we want to show that this process can be repeated. Let $\bar{a}_q$ be the alternating tuple $(s_1, t_1, \ldots)$ of length $q$ if $q$ is odd, or the alternating tuple $(t_1, s_1, \ldots)$ of length $q$ if $q$ is even. So $\bar{s}$ starts with $\bar{a}_1$ and $\bar{s}'$ starts with $\bar{a}_2$. Suppose that we have built a minimal tuple $\bar{s}_{q-1}$ representing $w$ that starts with $\bar{a}_{q-1}$. Let $r \in \{s_1, t_1\}$ be the element that is not the first letter of $\bar{s}_{q-1}$. We have $|rw|_S < |w|_S$ because, whichever element $r$ happens to be, we know a minimal tuple for $w$ that starts with $r$, so condition $(E)$ says there is an index $i$ with entry $r$ whose term can be moved to the front without changing the element $w$. If

$i > q-1$ then we have succeeded in constructing a new representative $\bar{s}_q$ of $w$ that starts with $\bar{a}_q$.

We claim that $i$ must be greater than $q-1$ when $q \leqslant m$. By Proposition 1.1.5, $\langle s_1, t_1 \rangle \cong \mathcal{D}_m$. If $i < q \leqslant m$. This would mean we have a relation $s_1 t_1 \cdots = t_1 s_1 \ldots$ in $\mathcal{D}_m$, where both sides have length $i$. But $\mathcal{D}_m$ has no such relations: the Cayley graph of $\mathcal{D}_m$ with respect to $s_1$ and $t_1$ is a circle of length $2m$, so there are no reduced relations of total length less than $2m$.

We have shown that whenever $q \leqslant m$ we can find a minimal tuple $\bar{s}_q$ representing $w$ that starts with $\bar{a}_q$. This implies $m \leqslant k$ is finite. If $m$ is odd then $\bar{a}_m$ begins and ends with $s_1$, so take $\bar{u} := \bar{s}_m$. If $m$ is even then $\bar{a}_m$ begins with $t_1$ and ends with $s_1$, so take $\bar{u}' := \bar{s}_m$ and let $\bar{u}$ be the tuple obtained from $\bar{u}'$ by a type II move exchanging $\bar{a}_m$ for the opposite alternating word.

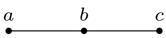
Now prove Item (1). A minimal tuple cannot be shortened without changing the element $w$, so a minimal tuple is $M$–reduced. Conversely, suppose $\bar{s} = (s_1, \ldots, s_k)$ is $M$–reduced. If $k = 1$ then $\bar{s}$ is minimal, so assume by induction that the claim is true for tuples shorter than $k$. The suffix $\bar{s}' = (s_2, \ldots, s_k)$ is shorter, and it certainly $M$–reduced, since any reductions applied to a $\bar{s}'$ could have been applied directly to $\bar{s}$. By induction, $\bar{s}'$ is minimal. Then we have $w = s_1 \cdots s_k$, $w' = s_2 \cdots s_k$, and $k-1 = |w'|_S$. If $\bar{s}$ is not minimal then $|w'|_S = k-1 \geqslant |w|_S = |s_1 w'|_S$, so (E) implies $\bar{s}'$ can be replaced by another minimal tuple $\bar{s}''$ that starts with $s_1$ and still represents $w'$. By Item (2), $\bar{s}'$ and $\bar{s}''$ are related by a sequence of Type II moves. Apply these moves to $\bar{s}'$ as a suffix of $\bar{s}$. This converts $\bar{s}$ to a tuple representing the same element $w \in W$, but the new tuple starts with $(s_1, s_1)$, so it can be reduced by a Type I move, contradicting the hypothesis that $\bar{s}$ is $M$–reduced. $\square$

EXERCISE 1.3.6. Suppose that the word *bcababcabacacacbabac* is trivial in the group defined by Coxeter graph 

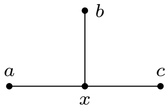
. What is $n$?

EXERCISE 1.3.7. Show that a Coxeter diagram containing a loop defines an infinite Coxeter group.

EXERCISE 1.3.8. Show that the only nondegenerate squares in $Cay(W, S)$ are those that arise from commuting generators. In particular, if $rsrt = 1$ for $r, s, t \in S$ with $r \neq s$ then $s = t$ and $m_{rs} = 2$.

EXERCISE 1.3.9. What is the order of the element $abc$ in the group defined by Coxeter graph 

?

EXERCISE 1.3.10. Show that the order of $abcx$ is 6 in the group defined by Coxeter graph: 


DEFINITION 1.3.11. If $(W, S)$ is a Coxeter system, let $S = (s_1, \ldots, s_n)$ be any ordering of the fundamental generators. Their product $s_1 \cdots s_n$ is a *Coxeter element.*

The two previous exercises are examples of Coxeter elements. It is immediate from Theorem 1.3.5 that Coxeter elements are minimal.

THEOREM 1.3.12 (Speyer [**24**, Theorem 1]). *If $(W, S)$ is irreducible and nonspherical then every nontrivial power of a Coxeter element is minimal.*

**1.4. Equivalence of the three definitions.** We can finally show that all the definitions of abstract reflection group are equivalent:

THEOREM 1.4.1. *If $(W, S)$ is a preCoxeter system then the following are equivalent:*

- *$(W, S)$ is a Coxeter system, that is, $W$ is an algebraic ARG.*
- *$(W, S)$ is a reflection system, that is, $W$ is a geometric ARG.*
- *$(W, S)$ has condition (E), that is, $W$ is a combinatorial ARG.*

PROOF. That a Coxeter system is a reflection system was Theorem 1.2.31.

Suppose $(W, S)$ is a reflection system. Suppose $\bar{s} = (s_1, \ldots, s_k)$ with $w = s_1 \cdots s_k$ is minimal and there is an $s \in S$ such that $|sw|_S \leqslant |w|_S = k$. Then $\bar{s}' = (s, s_1, \ldots, s_k)$ represents $sw$, but is nonminimal, since it has length $k + 1 > k$. Since $(W, S)$ is a reflection system, minimality of a tuple is equivalent to the condition that the associated edge path does not cross any wall more than once. That means the edge path corresponding to $\bar{s}'$ crosses some wall more than once, but the edge path corresponding to $\bar{s}$ does not. The edge path corresponding to $\bar{s}'$ consists of the edge $[1, s]$ crossing $\Omega^s$, followed by the $s$–translate of the edge path corresponding to $\bar{s}$. The latter part has no repeated wall crossings, so the repeated wall must be $\Omega^s$. Suppose the $(i+1)$–st edge crosses $\Omega^s$. Then Lemma 1.2.21 says that $w'$ is represented by $(\hat{s}, s_1, \cdots, \hat{s}_i, \ldots s_k)$. But then $w = sw'$ is represented by $(s, \hat{s}, s_1, \ldots, \hat{s}_i, \ldots s_k) = (s, s_1, \ldots, \hat{s}_i, \ldots s_k)$. This establishes condition (E).

Now suppose that $(W, S)$ has condition (E), so that we have the solution to the word problem in $W$ from Theorem 1.3.5. Let $q\colon \tilde{W} \to W$ be the canonical surjection defined by $q(\tilde{s}) = s$ for each $s \in S$. Note that $(\tilde{W}, \tilde{S})$ is a Coxeter system, so by what we have already shown it is also a combinatorial ARG. Moreover, $(W, S)$ and $(\tilde{W}, \tilde{S})$ have the same Coxeter matrix $M$, so we have exactly the same $M$–operations on $S^*$ and $\tilde{S}^*$. Suppose $\tilde{w} \in \ker q$, and let $\bar{\tilde{s}} = (\tilde{s}_1, \ldots, \tilde{s}_k)$ be a minimal tuple for $\tilde{w}$. Assume $\tilde{w}$ is nontrivial, so $k > 0$. Since $\tilde{w} \in \ker q$, $\bar{s} = (s_1, \ldots, s_k)$ represents the trivial word in $W$, so is not minimal, so it is not $M$–reduced. But a sequence of elementary $M$–operations that reduces $\bar{s}$ also reduces $\bar{\tilde{w}}$, contradicting the fact that $\bar{\tilde{s}}$ is minimal. □

## 2. Special subgroups and convexity

Recall, from Definition 4.2.4, that if $(W, S)$ is a Coxeter system and $T \subset S$ then the subgroup $W_T$ generated by $t \in T$ is called a *special subgroup*. If $\Gamma$ is the Coxeter graph of $(W, S)$ then the Coxeter group defined by the full subgraph spanned by $T$ surjects onto $W_T$. In Proposition 2.0.3 we proved that when $|T| = 2$ this map is an isomorphism. Then general case follows from Corollary 4.0.5, but now we can give a different proof.

THEOREM 2.0.1. *Let $(W, S)$ be a Coxeter system, and let $T \subset S$. Then $(W_T, T)$ is a Coxeter system and the inclusion $Cay(W_T, T) \hookrightarrow Cay(W, S)$ is an isometric embedding with convex image.*

PROOF. $(W_T, T)$ is a preCoxeter system. Let $M_T$ be the Coxeter matrix. Let $M_S$ be the Coxeter matrix for $(W, S)$. Notice that in the solution to the Word Problem for combinatorial ARGs, the elementary $M_T$–operations are a subset of the elementary $M_S$–operations. Consider the surjection $(\tilde{W}_T, \tilde{T}) \twoheadrightarrow (W_T, T)$ from the associated Coxeter system. Take an element of the kernel. This is a nontrivial tuple $\bar{t}$ in $T$ that represents the trivial element in $W_T$, hence in $W$. Theorem 1.3.5 gives a sequence of elementary $M_S$–operations that reduces $\bar{t}$ to the empty tuple. But notice that the elementary $M_S$–operations that can be applied to a tuple in $T$ are only the elementary $M_T$–operations, so $\bar{t}$ can be reduced to the empty word by a sequence of elementary $M_T$–operations. This shows that $\bar{t}$ represents the trivial element in $\tilde{W}_T$, so $\tilde{W}_T \cong W_T$.

For the geometric claim, suppose that $\gamma$ is a geodesic between two elements of $W_T$. Up to the group action, we may assume $\gamma$ is a geodesic from 1 to $w \in W_T$. Since $w \in W_T$, there is an edge path $\delta$ from 1 to $w$ that stays in $W_T$. Let $\bar{t} \in T^*$ be the tuple associated to the edge path $\delta$, and let $\bar{s} \in S^*$ be the tuple associated to the edge path $\gamma$. Since $\bar{s}$ is minimal, $\bar{t}$ can be transformed to $\bar{s}$ by a sequence of elementary $M_S$–operations, but, as above, no $M_S$–operation introduces new letters, so the set of letters appearing in $\bar{s}$ is a subset of $T$, which means $\gamma$ was contained in $W_T$ all along. □

PROPOSITION 2.0.2. *Let $(W, S)$ be a Coxeter system. A subset of $Cay(W, S)$ is convex if and only if it is an intersection of halfspaces.*

PROOF. Halfspaces are convex by Corollary 1.2.25, and intersections of convex sets are convex, so intersections of halfspaces are convex.

Suppose $U$ is a convex set. Suppose $v \notin U$ is vertex in the intersection of halfspaces containing $U$. Let $u \in U$ be a closest point to $v$, and let $\alpha$ be a geodesic from $v$ to $u$. The first edge of $\alpha$ crosses some wall $\Omega^r$, which does not separate $v$ from $U$, so there is a point $w \in U$ on the same side of $\Omega^r$ as $v$.

Since $u$ and $w$ are on different sides of $\Omega^r$, a geodesic from $u$ to $w$ must cross $\Omega^r$. By the exchange condition $(E)$, we may assume it is the first edge, so there is $s \in S$ conjugate to $r$ such that the edge from $u$ to $us$ is the first edge of a geodesic from $u$ to $w$, and it is an edge of $\Omega^r$. By convexity of $U$, $us \in U$. However, this gives a contradiction to the choice of $u$ as closest to $v$, because now $us$ is a point of $U$ with $\mathcal{W}(u,v) = \mathcal{W}(us,v) \sqcup \{\Omega^r\}$, so $d(v,us) = \#\mathcal{W}(v,us) < \#\mathcal{W}(u,v) = d(u,v)$. $\square$

DEFINITION 2.0.3. If $Y < X$, the *convex hull of* $Y$, $\mathcal{H}(Y)$, is the smallest convex subset of $X$ containing $Y$.

By Proposition 2.0.2 the convex hull of $Y \subset Cay(W,S)$ is the intersection of halfspaces containing $Y$.

LEMMA 2.0.4. *Let $(W,S)$ be a Coxeter system. Let $u, w \in W$. Then:*

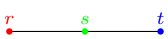
$$\mathcal{H}(\{u,w\}) = \{v \in W \mid v \textit{ is on some geodesic between } u \textit{ and } w\}$$

The $\supset$ direction is clear, but in a general geodesic metric space the right-hand side can fail to be convex. If $\alpha$ and $\beta$ are two different geodesics between $u$ and $w$, then they are contained in $\mathcal{H}(\{u,w\})$, but so are geodesics between interior points of $\alpha$ and $\beta$, which are not necessarily included in the right-hand side.

PROOF. Suppose $v \in \mathcal{H}(\{u,w\})$, which, by Proposition 2.0.2, is the intersection of halfspaces containing $\{u,w\}$. Then there are no walls separating $v$ from both $u$ and $w$, which implies $\mathcal{W}(u,v) \sqcup \mathcal{W}(v,w) = \mathcal{W}(u,w)$. Thus, $d(u,v) + d(v,w) = \#\mathcal{W}(u,v) + \#\mathcal{W}(v,w) = \#\mathcal{W}(u,w) = d(u,w)$, so the concatenation of a geodesic from $u$ to $v$ with one from $v$ to $w$ is a geodesic from $u$ to $w$ containing $v$. $\square$

## 3. Longest elements

Let $(W,S)$ be a Coxeter system, and let $R$ be the conjugates of $S$, so that $(W,R,Cay(W,S),1)$ is a reflection system.

EXAMPLE 3.0.1. Consider $Cay($

$)$, shown in Figure 6. If we imagine this as the surface of a polytope, the codimension 1 faces are cosets of the rank 2 special subgroups, which are $\langle r, s\rangle \cong \langle s, t\rangle \cong \mathcal{D}_3$ or $\langle r, t\rangle \cong \mathcal{C}_2 \times \mathcal{C}_2$. The edges are cosets of the rank 1 special subgroups $\langle r\rangle \cong \langle s\rangle \cong \langle t\rangle \cong \mathcal{C}_2$.

Call a coset of a special subgroup a 'special coset'. The figure is organized

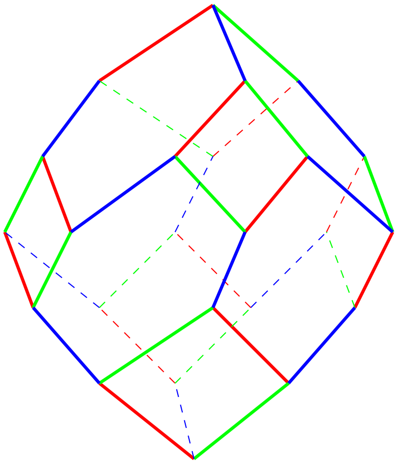


FIGURE 6

so that if the bottom vertex is the identity, then the height of each vertex is the edge distance from that vertex to the identity. Notice several things:

- There is a unique highest vertex.
- There are no horizontal edges, every edge moves strictly higher or lower.
- Special cosets are convex.
- Special cosets have a unique local maximum and a unique local minimum.
- Every path starting from 1 and moving up on each edge is a geodesic.
- Geodesics from 1 to a vertex $w$ are not in general unique, but the ambiguity can be described as follows: The edges coming in to $w$ from below are the topmost edges of some special coset. Let $v$ be the lowest vertex of this coset. Then any geodesic from 1 to $v$ concatenated with a geodesic from $v$ to $w$ is a geodesic from 1 to $v$.

We will show in this section and the next that all of these are general phenomena.

EXERCISE 3.0.2. Show that the walls of the reflection system consisting of multiplication of the group of Figure 6 on itself can be described as follows: if a wall intersects a special coset then it intersects it precisely in a pair of opposite edges. For instance, most vertices do not have a unique minimal labelling in terms of $r, s, t$, but you can show that the wall $\Omega^r$ separates vertices that have some minimal labelling starting with $r$ from vertices that do not have any such minimal labelling.

Show that there are exactly 6 walls, and that all of them separate 1 from the unique vertex farthest from 1.

LEMMA 3.0.3. *For all $w \in W$ and $r \in R$, $|rw| \neq |w| \neq |wr|$.*

PROOF. If $r \in \mathcal{W}(1, w)$, consider a shortest edge path from 1 to $w$. It crosses $\Omega^r$ at some edge $e$, so that the path is $\alpha + e + \beta$. Then $\alpha + r\beta$ is a path from 1 to $rw$ whose length is 1 shorter, so $|rw| \leqslant |w| - 1$.

If $r \notin \mathcal{W}(1, w)$ then 1 and $w$ are on one side of $\Omega^r$ and $r$ and $rw$ are on the other. So 1 and $rw$ are on different sides of $\Omega^r$. The previous case says $|w| = |r \cdot rw| \leqslant |rw| - 1$, so $|rw| \geqslant |w| + 1$.

We have shown $|rw| \neq |w|$. Given $r \in R$ and $w \in W$, let $r' := wrw^{-1} \in R$. Then $wr = r'w$, and by the previous argument $|wr| = |r'w| \neq |w|$. ☐

COROLLARY 3.0.4.

$$|w| > |rw| \iff r \in \mathcal{W}(1, w)$$

COROLLARY 3.0.5. *For an element $\Delta \in W$, the following are equivalent:*

(1) $\forall w \in W$, $|\Delta| = |w| + |w^{-1}\Delta|$
(2) $\forall r \in R$, $|r\Delta| < |\Delta|$

PROOF. Item (1) says that every vertex $w$ lies on some geodesic from 1 to $\Delta$. By Corollary 3.0.4, the Item (2) says that every wall separates 1 from $\Delta$.

Suppose a wall $\Omega^r$ does not separate 1 and $\Delta$. Then any path from 1 to $\Delta$ that goes through vertex $r$ must cross $\Omega^r$ from the 1–side to the $r$–side to get to $r$, and then cross back to the 1–side to get to $\Delta$. Geodesics cross each wall at most once, so no geodesic from 1 to $\Delta$ goes through $r$.

Conversely, suppose there is an element $w$ that does not lie on any geodesic from 1 to $\Delta$. Let $\alpha$ be a geodesic from 1 to $w$, and let $\beta$ be a geodesic from $w$ to $\Delta$. By hypothesis, the path $\alpha + \beta$ from 1 to $\Delta$ is not a geodesic, so it crosses some wall $\Omega^r$ more than once. But $\alpha$ and $\beta$ are geodesic, so each crosses $\Omega^r$ at most once, so they each cross it exactly once, and the path $\alpha + \beta$ starts and ends on the same side of $\Omega^r$. Thus, there is a wall, $\Omega^r$, that does not separate 1 from $\Delta$. ☐

PROPOSITION 3.0.6. *An element $\Delta$ as in Corollary 3.0.5 exists if and only if $W$ is finite. It is the* longest element *of $W$.*

*When such a $\Delta$ exists it has the following properties:*

(1) *$\Delta$ is unique.*
(2) $|\Delta| = |R|$

(3) $\Delta$ *is an involution.*
(4) $\Delta S \Delta = S$

PROOF. Item (1) of Corollary 3.0.5 says $\Delta$ is strictly longer than every other element of $W$, so if $\Delta$ exists $W$ is finite and there is a unique $\Delta$ that can have this property.

Conversely, if $W$ is finite let $\Delta$ be an element of maximal length. By Lemma 3.0.3, for all $r \in R$, $|\Delta| \neq |r\Delta|$, so, since nothing is longer that $\Delta$, $r\Delta$ must be shorter, so $r \in W(1, \Delta)$, by Corollary 3.0.4. Thus, $\Delta$ satisfies Item (2) of Corollary 3.0.5, and since the graph distance is the same as the wall distance, $|\Delta| = |R|$.

$|\Delta^{-1}| = |\Delta|$, but $\Delta$ is the unique longest element, so $\Delta = \Delta^{-1}$.

For $s \in S$, $|s\Delta| < |\Delta|$, so $|s\Delta| = |\Delta| - 1$. Since $w := s\Delta$ lies on a geodesic from 1 to $\Delta$, we have $|\Delta| = |w| + |w^{-1}\Delta|$, so $|w^{-1}\Delta| = |\Delta| - |w| = 1$. But $w^{-1} = \Delta^{-1}s^{-1} = \Delta s$ since $\Delta$ and $s$ are both involutions. Therefore, $|\Delta s \Delta| = 1$, which means $\Delta s \Delta \in S$. $\square$

In fact, Item (2) of Corollary 3.0.5 can be relaxed to only checking $S$:

LEMMA 3.0.7 ([**11**, Lemma 4.6.2]). *Suppose $\Delta \in W$ has the property that for all $s \in S$, $|\Delta| > |s\Delta|$. Then $W$ is finite and $\Delta$ is the longest element.*

COROLLARY 3.0.8. *If $\Delta \in W$ is a local maximum for $w \mapsto |w|$ then $W$ is finite and $\Delta$ is the longest element.*

PROOF. The neighbors of $w$ are $wS$. Suppose for all $s \in S$ that $|\Delta| \geqslant |\Delta s|$. By Lemma 3.0.3, these inequalities are strict, so $|\Delta^{-1}| = |\Delta| > |\Delta s| = |s\Delta^{-1}|$. The previous result says $W$ is finite and $\Delta^{-1}$ is the longest element, but Proposition 3.0.6 says the longest element is an involution, so $\Delta = \Delta^{-1}$. $\square$

## 4. How special cosets fit together

The following is a corollary to Tits' solution to the word problem, Theorem 1.3.5.

COROLLARY 4.0.1. *Let $(W, S)$ be a Coxeter system. For $w \in W$:*

$$S(w) := \{s \in S \mid s \text{ appears in some minimal length expression of } w\}$$

*Then for every minimal length expression $\bar{s}$ of $w$, the set of elements of $S$ that appear in $\bar{s}$ is $S(w)$.*

COROLLARY 4.0.2. *For all $T \subset S$ we have $W_T = \{w \in W \mid S(w) \subset T\}$.*

The following corollary will be key in the next chapter.

COROLLARY 4.0.3. *Let $(W, S)$ be a Coxeter system. For all $T, T' \subset S$ and $w \in W$, $W_T \subset wW_{T'}$ if and only if $T \subset T'$ and $w \in W_{T'}$.*

PROOF. One direction is obvious. For the other, assume $W_T \subset wW_{T'}$. Since $1 \in W_T \subset wW_{T'}$, $w^{-1} \in W_{T'}$, so $w \in W_{T'}$ and $W_T \subset W_{T'}$. Now apply the previous corollary. □

EXERCISE 4.0.4. Show that if $\Gamma$ is a Coxeter graph that is connected with at least two edges, and some edge has label at least 6, then $W_\Gamma$ is infinite.

LEMMA 4.0.5 (Bridge Lemma). *Let $(W, S)$ be a Coxeter system and $A, B \subset S$. Suppose $w_0 \in W$ is of minimal length in $W_A w_0 W_B$. Then every element $z \in W_A w_0 W_B$ can be written $z = xw_0 y$ for some $x \in W_A$ and $y \in W_B$ such that $|z| = |x| + |w_0| + |y|$. In particular, for every $w \in W$ the set $W_A w W_B$ has a unique element of minimal length.*

PROOF. Suppose $z = xw_0 y \in W_A w_0 W_B$. Let $\bar{a}$ be a minimal tuple in $A$ representing $x$, let $\bar{s}$ be a minimal tuple in $S$ representing $w_0$, and let $\bar{b}$ be a minimal tuple in $B$ representing $y$.

Suppose $\bar{a} + \bar{s} + \bar{b}$ is not minimal. By the deletion condition (D), there are a pair of entries of $\bar{a} + \bar{s} + \bar{b}$ that can be deleted without changing the element of $W$ represented by the tuple. But $\bar{a}$, $\bar{s}$, and $\bar{b}$ were minimal, so it is not the case that both deleted entries come from a single one of these. Furthermore, if one of the deleted entries is in $\bar{s}$ and $\bar{s}'$ is the resulting tuple of length one less, then $\bar{s}'$ represents a word $w$ in $W$ that is shorter than $w_0$, but still in $W_A w_0 W_B$, which is a contradiction. Therefore, one deleted entry comes from $\bar{a}$ and one comes from $\bar{b}$. Let $\bar{a}'$ and $\bar{b}'$ be the tuples resulting from deleting this pair of entries from $\bar{a}$ and $\bar{b}$. We have that $\bar{a}' \in A^*$ and $\bar{b}' \in B^*$ such that $z$ is represented by $\bar{a}' + \bar{s} + \bar{b}'$. If $\bar{a}' + \bar{s} + \bar{b}'$ is not minimal repeat the argument and shorten $\bar{a}'$ and $\bar{b}'$ by one letter each. Continue in this way until we arrive at $\bar{a}'' \in A^*$ and $\bar{b}'' \in B^*$ such that $z$ is represented by $\bar{a}'' + \bar{s} + \bar{b}''$ and $\bar{a}'' + \bar{s} + \bar{b}''$ is minimal. Let $x'$ be the element of $W_A$ represented by $\bar{a}''$, and let $y'$ be the element of $W_B$ represented by $\bar{b}''$. We have $|z| = |\bar{a}'' + \bar{s} + \bar{b}''| = |\bar{a}''| + |\bar{s}| + |\bar{b}''| = |x'| + |w_0| + |y'|$. □

In $Cay(W, S)$, for all $w \in W$ and $s \in S$, $w$ and $ws$ are adjacent, so their lengths differ by at most one. By Lemma 3.0.3, their lengths are not the same, so we can partition $S$ into two sets:

DEFINITION 4.0.6. For $w \in W$ define:

$$\mathrm{In}(w) := \{s \in S \mid |ws| = |w| - 1\} = \{s \in S \mid |ws| < |w|\}$$

$$\mathrm{Out}(w) := \{s \in S \mid |ws| = |w| + 1\} = \{s \in S \mid |ws| > |w|\}$$

In other words, if we call an edge of $Cay(W, S)$ "incoming" if it the last edge of a geodesic from 1 to $w$ and "outgoing" if it the first edge after $w$ of a geodesic starting at 1, then every edge incident to $w$ is either incoming or outgoing, $\mathrm{In}(w)$ is the set of labels of the incoming edges, and $\mathrm{Out}(w)$ is the set of labels of the outgoing edges.

LEMMA 4.0.7. *For $T \subset S$ and $w \in W_T$, $\mathrm{In}(w) \subset T$.*

PROOF. This follows immediately from Corollary 4.0.1. □

LEMMA 4.0.8. *For all $w \in W$, the subgroup $W_{\mathrm{In}(w)}$ is finite.*

PROOF. By Lemma 4.0.5, for every $w \in W$ there exists a unique $x$ of minimal length in $W_\varnothing w W_{\mathrm{In}(w)}$ such that for all $z \in W_{\mathrm{In}(w)}$, $|xz| = |x| + |z|$.

Let $y := x^{-1}w \in W_{\mathrm{In}(w)}$, so $|w| = |x| + |y|$. For all $s \in \mathrm{In}(w)$, we have $|ws| = |xys| = |x| + |ys|$, but, by definition of $\mathrm{In}(w)$, $|ws| = |w| - 1 = |x| + |y| - 1$, so $|ys| = |y| - 1$.

This says that $y$ satisfies Corollary 3.0.5, so $y$ is the longest element of $W_{\mathrm{In}(w)}$. By Proposition 3.0.6, the existence of a longest element implies $W_{\mathrm{In}(w)}$ is finite. □

DEFINITION 4.0.9. For $T \subset S$, define $W^T := \{w \in W \mid \mathrm{In}(w) = T\}$.

LEMMA 4.0.10. *If $W_T$ is finite then its longest element $\Delta_T$ is in $W^T$. If $W_T$ is infinite then $W^T = \varnothing$.*

PROOF. If $w \in W^T$ then $\mathrm{In}(w) = T$, so $W_T = W_{\mathrm{In}(w)}$ is finite, by Lemma 4.0.8. Conversely, if $W_T$ is finite then there exists a longest element $\Delta_T \in W_T$, by Proposition 3.0.6. Thus, for every $t \in T$, $|\Delta_T| > |\Delta_T t| \geqslant |\Delta_T| - 1$, so $|\Delta_T| = |\Delta_T t| + |t|$. This means a geodesic from 1 to $\Delta_T t$ concatenated with a single edge labelled $t$ is a geodesic from 1 to $\Delta_T$, so $t \in \mathrm{In}(\Delta_T)$. This shows $T \subset \mathrm{In}(\Delta_T)$. By Lemma 4.0.7, $\mathrm{In}(\Delta_T) \subset T$. □

LEMMA 4.0.11. *Let $\Delta_T$ be the longest element of $W_T$, take $s \in S - T$, and let $T' := T \cup \{s\}$. Then $\mathrm{In}(s\Delta_T) \in \{T, T'\}$, and the following are equivalent:*

- *$\mathrm{In}(s\Delta_T) = T'$*
- *$s\Delta_T = \Delta_{T'}$ is the longest element of $W_{T'}$.*
- *$s$ and $\Delta_T$ commute.*
- *$s$ commutes with $t$ for all $t \in T$.*

PROOF. By Lemma 4.0.7, $\mathrm{In}(s\Delta_T) \subset T'$. By Lemma 4.0.5, $|s\Delta_T| = |\Delta_T| + 1$, since $s$ is the shortest element of $W_\varnothing s W_T$. By Lemma 4.0.10, $\mathrm{In}(\Delta_T) = T$, so for every $t \in T$ there exists a geodesic $\alpha_t$ from 1 to $\Delta_T$ in $W_T$ ending with $t$. Consider $s + s\alpha_t$. It is a path of length $1 + |\alpha_t| = |s\Delta_T|$

from 1 to $s\Delta_T$, so it is a geodesic from 1 to $s\Delta_T$ ending with $t$. Thus, $T \subset \mathrm{In}(s\Delta_T)$.

If $\mathrm{In}(s\Delta_T) = T'$ then $s\Delta_T \in W_{T'}$ is the longest element of $W_{T'}$ by Corollary 3.0.8, since it is a local maximum for $w \mapsto |w|$ in $W_{T'}$. But $s \notin \mathrm{In}(\Delta_T) = T$ implies $s$ is an outgoing edge from $\Delta_T$, so $|\Delta_T s| = |\Delta_T| + 1$. So $\Delta_T s \in W_{T'}$ has the same length at the longest element of $W_{T'}$, so it is the longest element, so $\Delta_T s = s\Delta_T$.

If $s$ and $\Delta_T$ commute then, since $s \in \mathrm{In}(\Delta_T s)$, we have $s \in \mathrm{In}(s\Delta_T)$, so $\mathrm{In}(s\Delta_T) = T'$.

If $s$ commutes with all $t \in T$ then it certainly commutes with $\Delta_T \in W_T$.

Now suppose that $s$ commutes with $\Delta_T$ and consider $t \in T$. Consider the coset $\Delta_{T'} W_{\{s,t\}}$. It is contained in $W_{T'}$, and $\Delta_{T'}$ is the longest element, so for all $w \in W_{\{s,t\}}$, $|\Delta_{T'} w| = |\Delta_{T'}| - |w|$. In particular, look at $w = sts$. For this element $|\Delta_{T'}| - |w| = |\Delta_{T'} w| = |\Delta_T ts| = |\Delta_T|$, since $\Delta_T t \in W_T$ is one step closer to 1 from $\Delta_T$ and $s \notin \mathrm{In}(\Delta_T t) \subset W_T$. But then $|w| = |\Delta_{T'}| - |\Delta_T| = 1$. By Exercise 1.3.8 this is only possible when $m_{st} = 2$. Since $t \in T$ was arbitrary, $s$ commutes with every element of $T$. □

Corollary 4.0.12 ([**11**, Lemma 4.7.5]). *Let $(W, S)$ be a Coxeter system with Coxeter graph $\Gamma$. If $W^T = \{\Delta_T\}$ then $\Gamma = \Gamma_T \sqcup \Gamma_{S-T}$, so $W = W_T \times W_{S-T}$.*

Proof. Suppose $s \in S - T$, and let $T' := T \cup \{s\}$. Since $s\Delta_T \notin W^T = \{\Delta_T\}$, by Lemma 4.0.11, $\mathrm{In}(s\Delta_T) = T'$ and $s$ commutes with every $t \in T$. Since this was true for every $s \in S - T$, we have $m_{st} = 2$ for all $s \in S - T$ and $t \in T$, so there are no edges between $\Gamma_T$ and $\Gamma_{T-S}$. □

Definition 4.0.13. For $(W, S)$ a Coxeter system and $T \subset S$ define:

$$A_T := \{w \in W \mid w \text{ is of minimal length in } W_T w = W_T w W_\varnothing\}$$

$$B_T := \{w \in W \mid w \text{ is of minimal length in } w W_T = W_\varnothing w W_T\}$$

Lemma 4.0.14. *Let $(W, S)$ be a Coxeter system and $T \subset S$.*

(1) *Every $w \in W$ admits a unique factorization $w = xy$ with $y \in W_T$ and $x$ of minimal length in $xW_T$. For this factorization, $|w| = |x| + |y|$.*
(2) *Every $w \in W$ admits a unique factorization $w = xy$ with $x \in W_T$ and $y$ of minimal length in $W_T y$. For this factorization, $|w| = |x| + |y|$.*

(3)

$$\begin{aligned} B_T &= \{w \in W \mid \forall t \in T,\ |wt| = |w| + 1\} \\ &= \{w \in W \mid \textit{no minimal expression of } w \textit{ ends with } t \in T\} \\ &= \{w \in W \mid \mathrm{In}(w) \subset S - T\} \end{aligned}$$

(4)

$$\begin{aligned} A_T &= \{w \in W \mid \forall t \in T,\ |tw| = |w| + 1\} \\ &= \{w \in W \mid \textit{no minimal expression of } w \textit{ begins with } t \in T\} \\ &= B_T^{-1} \end{aligned}$$

Items (2) and (1) say $W$ admits a sort of 'orthogonal splitting' as $W = W_T A_T$, or as $B_T W_T$. In particular the index of $W_T$ in $W$ is $[W : W_T] = |A_T| = |B_T|$.

Item (3) says that for any $w \in W$ and $T \subset S$, $w$ has minimal length in $wW_T$ if and only if it is shorter than every element of $wT$, which are precisely the neighbors of $w$ in $wW_T$. So an element is the unique global minimum of the function $wW_T \to \mathbb{N} : z \mapsto |z|$ if and only if it is a local minimum.

PROOF. We'll prove (2) and (4).

For (2), take $y$ to be the unique minimal length element of $W_T w W_\varnothing$, and take $x = wy^{-1} \in W_T$. If $w = x'y'$ with $x' \in W_T$ is a different factorization then $y' = (x')^{-1}xy \in W_T y$, so $y'$ does not have minimal length in $W_T y' = W_T y$. Thus, the factorization is unique. The second claim follows from the Bridge Lemma.

The 'only if' direction of (4) is clear, so assume $|tw| = |w| + 1$ for all $t \in T$. Let $w = xy$ be the factorization given by part (2).

For any $t \in T$, consider $tw = txy \in W_T y$. By part (2), it admits a unique factorization $tw = x'y$, with the same $y$, since $y$ is the unique minimal length element of $W_T y$.

(13)

$$|x'| + |y| = |tw| = |w| + 1 = |x| + |y| + 1 \implies |x'| = |x| + 1 \overset{x' = tx}{\Longrightarrow} |tx| = |x| + 1$$

Since $x \in W_T$, there exists a minimal $\bar{t} \in T^*$ representing $x$. Suppose $\bar{t}$ is nonempty, so it has a leading entry $t_1 \in T$. Then $|t_1 x| = |x| - 1$. Since (13) is true for all $t \in T$, this is a contradiction, so $\bar{t}$ is empty and $x = 1$. Thus, $w = y$ is of minimal length in $W_T w = W_T y$. $\square$

Exercise 4.0.15. For $t \in T$, show:

$$A_{\{t\}} = \{w \in W \mid w \text{ and } 1 \text{ are on the same side of } \Omega^t\}$$

Show $A_T = \bigcap_{t \in T} A_{\{t\}}$.

Definition 4.0.16. Let $(W, S)$ be a Coxeter system, let $\mathcal{S}$ be the collection of spherical subsets of $S$. The *nerve* $L = L(W, S)$ of $(W, S)$ is the simplicial complex whose vertices are $S$, and such that $\varnothing \neq T \subset S$ spans a simplex $\sigma_T$ when $T \in \mathcal{S}$.

We mention here another result that says that when $W$ is infinite then Corollary 4.0.12 is the only way that special subgroups have finite index:

Theorem 4.0.17 (Hosaka [**17**]). *If $(W, S)$ is a Coxeter system with Coxeter graph $\Gamma$ and $W$ infinite then a special subgroup $W_T$ has finite index in $W$ if and only if $\Gamma = \Gamma_T \sqcup \Gamma_{S-T}$ with $S - T$ spherical.*

The proof can be reduced to the case that $(W, S)$ is irreducible and $S = \{s\} \cup T$. Hosaka supposes that $A_T s \cap W_T$ is finite, and shows that the nerve of $W_T$ can be built inductively as joins of simplices based on spheres of decreasing radius about $s$ in $\Gamma$. Joins of simplices are simplices, so the nerve of $W_T$ is simplex, which means that $W_T$ is finite. This is a contradiction, since if $W$ is finite and $[W : W_T]$ is finite then $W_T$ is infinite. Thus, $A_T s \cap W_T$ is infinite, which implies $[W : W_T] = |A_T| = \infty$.

CHAPTER 5

# The Davis complex

The goal of this chapter is, given a Coxeter system $(W, S)$, to construct a geometric action $W \curvearrowright \Sigma$, where $\Sigma$ is a CAT(0) space. First we will define CAT(0) spaces, and, more generally, CAT(k) spaces, and talk about why we care about actions on such spaces. Then we will build the space $\Sigma$ for a Coxeter group.

## 1. CAT(k) spaces

This section is a short introduction to CAT($k$) spaces. See [**6**] for more.

### 1.1. Model spaces.

DEFINITION 1.1.1. For each $k \in \mathbb{R}$ and $n \in \mathbb{N}$, let $\mathcal{M}_k^n$ be the unique complete, simply connected, $n$–dimensional Riemannian manifold of constant sectional curvature equal to $k$.

The three key cases are already familiar: $\mathcal{M}_{-1}^n = \mathbb{H}^n$, $\mathcal{M}_0^n = \mathbb{E}^n$, and $\mathcal{M}_1^n = \mathbb{S}^n$. In fact, for $k \neq 0$ the space $\mathcal{M}_k^n$ can be obtained by rescaling the metric on $\mathbb{H}^n$ or $\mathbb{S}^n$ by a factor of $1/\sqrt{|k|}$, according to whether $k < 0$ or $k > 0$, respectively. So $\mathcal{M}_1^n$ is the unit $n$–sphere, and for $k > 0$ the model space $\mathcal{M}_k^n$ is just the sphere of radius $1/\sqrt{|k|}$. Qualitatively, the most important thing with whether $k$ is negative, positive, or zero.

DEFINITION 1.1.2. Define $D_k$ to be infinite if $k \leqslant 0$, or $D_k := \pi/\sqrt{k}$ if $k > 0$. This is the *diameter* of $\mathcal{M}_k^n$.

PROPOSITION 1.1.3. *For all $n$ and $k$, $\mathcal{M}_k^n$ is a geodesic metric space, and there is a unique geodesic between points at distance strictly less than $D_k$ from each other.*

Note that the distance hypothesis is necessary: there are infinitely many geodesics between antipodal points on a sphere.

EXERCISE 1.1.4. Show that in $\mathcal{M}_k^2$, if $x, y, z$ are distinct points with $d(x, y) = d(x, z) = r < D_k/2$ and $m$ is the midpoint of $[y, z]$ then $d(x, m) < r$. Conclude that balls in $\mathcal{M}_k^2$ of radius less than $D_k/2$ are convex.

**1.2. Comparison geometry.** We will define a notion of curvature in metric spaces by comparing a given metric space to the constant curvature model spaces.

LEMMA 1.2.1. *Let $k \in \mathbb{R}$. Suppose $d_0, d_1, d_2 \geqslant 0$ such that for all $i$, indices mod 3, we have $|d_{i+1} - d_{i+2}| \leqslant d_i \leqslant d_{i+1} + d_{i+2}$ and $d_0 + d_1 + d_2 < 2D_k$. Then there exists a, unique, up to isometry, geodesic triangle in $\mathcal{M}_k^2$ with side lengths $d_0$, $d_1$, and $d_2$.*

The first condition says that the three distances satisfy all permutations of the triangle inequality. The second is vacuous if $k \leqslant 0$. For $k > 0$ it implies $d_i < D_k$. This is necessary for uniqueness, since $\mathcal{M}_k^2$ is only locally uniquely geodesic.

DEFINITION 1.2.2. Let $k \in \mathbb{R}$, and let $x, y, z$ be three points in a geodesic metric space $X$ such that $d(x, y) + d(y, z) + d(z, x) < 2D_k$. Then there is a unique, up to isometry, geodesic triangle $\Delta(\bar{x}, \bar{y}, \bar{z})$ in $\mathcal{M}_k^2$ such that $d_X(x, y) = d_{\mathcal{M}_k^2}(\bar{x}, \bar{y})$, $d_X(y, z) = d_{\mathcal{M}_k^2}(\bar{y}, \bar{z})$, and $d_X(z, x) = d_{\mathcal{M}_k^2}(\bar{z}, \bar{x})$. Call such a triangle a *comparison triangle* for $x, y, z$.

For any geodesic triangle $\Delta(x, y, z)$ in $X$ with vertices $x, y, z$, and any point $p \in \Delta(x, y, z)$, there is a unique *comparison point* $\bar{p} \in \Delta(\bar{x}, \bar{y}, \bar{z})$ such that the distance from $\bar{p}$ to the endpoints of its side of the triangle is the same as for $p$; that is, if $p$ is on a geodesic between $x$ and $y$, define $\bar{p}$ to be the unique point on $[\bar{x}, \bar{y}]$ satisfying $d_{\mathcal{M}_k^2}(\bar{p}, \bar{x}) = d_X(p, x)$, or, equivalently, $d_{\mathcal{M}_k^2}(\bar{p}, \bar{y}) = d_X(p, y)$,

DEFINITION 1.2.3. A geodesic metric space $X$ is *CAT(k)* if geodesic triangles are no fatter than their comparison triangles in $\mathcal{M}_k^2$. That is, for every geodesic triangle $\Delta(x, y, z)$ in $X$ of perimeter less than $2D_k$, and every pair of points $p, q \in \Delta(x, y, z)$, we have $d_X(p, q) \leqslant d_{\mathcal{M}_k^2}(\bar{p}, \bar{q})$.

'CAT' is an acronym for the names of three pioneers of the field: Cartan, Alexandrov, and Toponogov. One might also suggest a backronym: Compare All Triangles.

The CAT($k$) condition is an *upper* curvature bound.

LEMMA 1.2.4. *If $k < k'$ then $\mathcal{M}_k^2$ is CAT($k'$), but $\mathcal{M}_{k'}^2$ is not CAT($k$).*

In practice it is often convenient to have a locally verifiable condition.

DEFINITION 1.2.5. A space has *curvature at most* $k$, or is *locally CAT(k)*, if every point has a CAT($k$) neighborhood.

THEOREM 1.2.6. *A complete geodesic metric space of curvature $k$ is CAT(k) if and only if it has no isometrically embedded loop of length less* *than $2D_k$. If $k \leqslant 0$, the last condition is equivalent to the space being simply connected.*

Example 1.2.7. A *flat torus* is a space obtained from a rectangle in $\mathbb{E}^2$ by identifying opposite sides and taking the resulting length metric. It has curvature 0, since every point has a neighborhood isometric to a neighborhood in $\mathbb{E}^2$, but it is not CAT(0), because it is not simply connected. Indeed, take a loop parallel to one pair of sides and divide it into three equal length pieces. This is a geodesic triangle in the torus, and it is convex in the torus. The comparison triangle in $\mathbb{E}^2$ is not convex: for points on different sides the Euclidean geodesic cuts through the interior of the triangle, and is shorter than the path that travels around the perimeter.

Exercise 1.2.8. Show that for all $k$, a tree $X$ is CAT($k$).

Exercise 1.2.9. Show that a graph with only finitely many different edge lengths is CAT(1) if and only if it does not contain an isometrically embedded cycle of length less than $2\pi$.

### 1.3. Some consequences of the CAT(k) property.

Exercise 1.3.1. Show that there is a unique geodesic between two points at distance less than $D_k$ in a CAT(k) space. In particular, CAT(0) spaces are uniquely geodesic.

Proposition 1.3.2. *In a CAT(0) space $X$, local geodesics are geodesic. That is, if $\gamma\colon \mathbb{R} \to X$ is a unit speed path such that there exists $\epsilon > 0$ such that for all $t \in \mathbb{R}$ the segment $\gamma|_{(t-\epsilon,t+\epsilon)}$ is geodesic, then $\gamma$ is geodesic.*

Proof. Suppose $\gamma$ is not geodesic. Let:

$$\delta := \inf\{|t-s| \mid \gamma \text{ is not geodesic on } [s,t]\} \geqslant \epsilon > 0$$

Then there are $s < t$ such that $3\delta/2 > t - s \geqslant \delta$ such that $\gamma$ is not geodesic on $[s,t]$. Let $r := (t-s)/2$. We have $t-r < \delta$ and $r-s < \delta$, so $\gamma$ is geodesic on $[s,r]$ and on $[r,t]$. Consider the Euclidean comparison triangle for $\gamma(s)$, $\gamma(r)$, and $\gamma(t)$. Its angle at $\overline{\gamma(r)}$ is not $\pi$, since

$$\begin{aligned}
d(\overline{\gamma(s)},\overline{\gamma(t)}) &= d(\gamma(s),\gamma(t)) \\
&< t-s \\
&= t-r+r-s \\
&= d(\gamma(s),\gamma(r)) + d(\gamma(r),\gamma(t)) \\
&= d(\overline{\gamma(s)},\overline{\gamma(r)}) + d(\overline{\gamma(r)},\overline{\gamma(t)})
\end{aligned}$$

But if its angle at $\overline{\gamma(r)}$ is less than $\pi$ then for all sufficiently small positive $u$ we have:

$$2u > d([\overline{\gamma(r)}, \overline{\gamma(s)}](u), [\overline{\gamma(r)}, \overline{\gamma(t)}](u)) = d([\gamma(r), \gamma(s)](u), [\gamma(r), \gamma(t)](u))$$

This is a contradiction, since for $u < \epsilon/2$ we have that $\gamma$ is geodesic on $[r-u, r+u]$, so $d([\gamma(r), \gamma(s)](u), [\gamma(r), \gamma(t)](u)) = 2u$. $\square$

EXERCISE 1.3.3. A *local isometry* is a map $\phi\colon X \to Y$ such that for every $x \in X$ there is an $r_x > 0$ such that $\phi$ restricted to the ball of radius $r_x$ about $x$ is an isometric embedding. If $\phi\colon X \to Y$ is a local isometry from a geodesic metric space to a CAT(0) metric space, show that $\phi$ is an isometric embedding.

PROPOSITION 1.3.4 (Convexity of the CAT(0) metric). *Let $\alpha\colon [0,1] \to X$ and $\beta\colon [0,1] \to X$ be geodesics in a CAT(0) space, parameterized by constant speed. Then for all $t \in [0,1]$,*

$$d_X(\alpha(t), \beta(t)) \leqslant (1-t)d_X(\alpha(0), \beta(0)) + td_X(\alpha(1), \beta(1))$$

PROOF. Let $\gamma\colon [0,1] \to X$ be a unit speed parameterized geodesic from $\beta(0)$ to $\alpha(1)$. Consider triangles $\Delta(\alpha(0), \alpha(1), \gamma(0))$ and $\Delta(\gamma(0), \gamma(1), \beta(1))$ in $X$ and their comparison triangles in $\mathcal{M}_0^2$. See Figure 1.

For all $t \in [0,1]$, we have:

$$\begin{aligned}
d_X(\alpha(t), \beta(t)) &\leqslant d_X(\alpha(t), \gamma(t)) + d_X(\gamma(t), \beta(t)) && \text{triangle ineq.}\\
&\leqslant d_{\mathcal{M}_0^2}(\bar{\alpha}(t), \bar{\gamma}(t)) + d_{\mathcal{M}_0^2}(\bar{\gamma}(t), \bar{\beta}(t)) && \text{CAT(0) ineq.}\\
&= (1-t)d_{\mathcal{M}_0^2}(\bar{\alpha}(0), \bar{\gamma}(0)) + td_{\mathcal{M}_0^2}(\bar{\gamma}(1), \bar{\beta}(1)) && \text{similar Eucl. triangles}\\
&= (1-t)d_X(\alpha(0), \gamma(0)) + td_X(\gamma(1), \beta(1)) && \text{def. comp. triangle}
\end{aligned}$$

$\square$

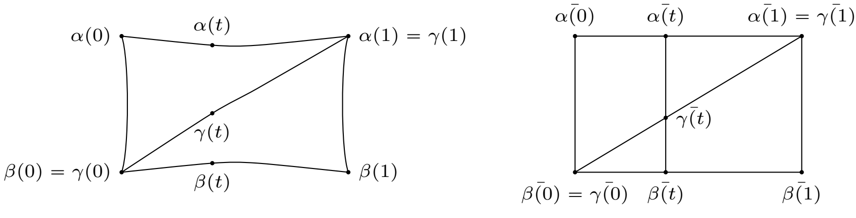



FIGURE 1. Convexity of the CAT(0) metric

PROPOSITION 1.3.5 (Projections). *Let $X$ be a CAT(0) space, and let $C \subset X$ be convex and complete.*

- *$\forall x \in X$ there exists a unique point $\pi(x) \in C$ such that $d(x, \pi(x)) = d(x, C) := \inf_{c \in C} d(x, c)$.*

- *If $x' \in [x, \pi(x)]$, the unique geodesic between $x$ and $\pi(x)$, then $\pi(x') = \pi(x)$.*
- *For all $x \notin C$, and $c \in C$ with $c \neq \pi(x)$, the* Alexandrov angle *of $[\pi(x), x]$ and $[\pi(x), c]$ is at least $\pi/2$, where the Alexandrov angle is defined to be the* $\limsup_{t,t'\to 0}$ *of the Euclidean angle at $\overline{\pi(x)}$ of the comparison triangle for $\pi(x)$, $[\pi(x), x](t)$, $[\pi(x), c](t')$.*
- *The map $\pi\colon X \to C$ is distance non-increasing.*

Proof. In this proof $[a, b]\colon [0, 1] \to X$ denotes the unique constant speed geodesic between $a$ and $b$.

Given $x \in X$, pick a sequence $(c_n) \subset C$ such that $d(x, c_n) \to D := d(x, C)$. For all $\epsilon > 0$ there exists $N$ such that for all $n > N$, $d(x, c_n) < D+\epsilon$. Assume that $\epsilon << D$, let $N$ be the corresponding bound, and let $m, n > N$. Since $C$ is convex, the geodesic $[c_m, c_n]$ is contained in $C$, so it does not enter $B_D(x)$. On the other hand, by convexity of the metric, $[c_m, c_n] \subset B_{D+\epsilon}(x)$. Now consider the comparison triangle in $\mathbb{E}^2 = \mathcal{M}^2_0$. The side $[\bar{c}_m, \bar{c}_n]$ is contained in the annulus $\overline{B_{D+\epsilon}(\bar{x})} - B_D(\bar{x})$. Now, a geodesic segment in a Euclidean annulus of inner diameter $D$ and outer diameter $D+\epsilon$ has length at most $2\sqrt{2D\epsilon + \epsilon^2} \overset{\epsilon\to 0}{\to} 0$. Therefore $(c_n)$ is a Cauchy sequence, so it converges to some point $c \in C$, since $C$ is complete. Set $\pi(x) := c$. By construction $d(x, \pi(x)) = D = d(x, C)$. Exercise 1.1.4 shows that $\pi(x)$ is unique, since if there were two points in $C$ at the same distance, then the midpoint of the segment between them would be a point of $C$ strictly closer to $x$.

The second point follows since if $x' \in [x, \pi(x)]$ then:

$$d(x', C) \leqslant d(x', \pi(x)) = d(x, \pi(x)) - d(x, x') = d(x, C) - d(x, x') \leqslant d(x', C)$$

Thirdly, suppose $x \in X - C$ and $\pi(x) \neq c \in C$. Since $C$ is convex, $[\pi(x), c] \subset C$. Let $t, t' \in [0, 1]$, and consider the comparison triangle for $[\pi(x), x](t)$, $\pi(x)$, $[\pi(x), c](t')$. Its Euclidean angle at $\overline{\pi(x)}$ is at least $\pi/2$, because otherwise we would have, for all sufficiently small $s$, a violation of the fact that $\pi(x)$ is the closest point of $C$ to the point $y := [\pi(x), x](s)$:

$$\begin{aligned}
d(y, \pi(x)) &= d(y, C) \\
&\leqslant d(y, [\pi(x), c]) \\
&\leqslant d(y, [\pi(x), [\pi(x), c](t')]) \\
&\leqslant d(\overline{y}, [\overline{\pi(x)}, \overline{[\pi(x), c](t')}]) \\
&< d(\overline{y}, \overline{\pi(x)}) \quad \text{if angle } < \pi/2 \\
&= d(y, \pi(x))
\end{aligned}$$

Finally, take $x, x' \in X$. Consider $t \in [0,1]$. Let $y := [\pi(x), x](t)$ and $z := [\pi(x'), x'](t)$. Consider the Euclidean quadrilateral $Q$ obtained by identifying the comparison triangle for $\pi(x)$, $\pi(x')$, $y$ and the comparison triangle for $\pi(x')$, $z$, and $y$ along the side $[\overline{y}, \overline{\pi(x')}]$. The angle of $Q$ at $\overline{\pi(x)}$ is at least $\pi/2$. The angle of $Q$ at $\overline{\pi(x')}$ is the sum of the angles of the comparison triangles. But the Alexandrov angle of $[\pi(x'), \pi(x)]$ and $[\pi(x'), x']$ at $\pi(x')$ is bounded above by this same sum, and bounded below by $\pi/2$, so the angle of $Q$ at $\overline{\pi(x')}$ is at least $\pi/2$. Thus:

$$d_X(\pi(x), \pi(x')) = d_{\mathbb{E}^2}(\overline{\pi(x)}, \overline{\pi(x')}) \overset{\text{shape } Q}{\leqslant} d_{\mathbb{E}^2}(\bar{y}, \bar{z}) = d_X(y, z) \qquad \square$$

THEOREM 1.3.6 (Flat strip theorem [**6**, Theorem II.2.13]). *Let $\gamma, \gamma' \colon R \to X$ be geodesics in a CAT(0) space. Suppose that there exists $D$ such that for all $t \in \mathbb{R}$, $d(\gamma(t), \gamma'(t)) \leqslant D$. Then there exists $D' \leqslant D$ such that the convex hull of $\gamma$ and $\gamma'$ in $X$ is isometric to a strip $\mathbb{R} \times [0, D']$ in $\mathbb{E}^2$.*

SKETCH. Assume $\gamma$ and $\gamma'$ are parameterized by unit speed, and that $\gamma'(0) \in \pi_\gamma^{-1}(\gamma(0))$. Consider $f \colon t \mapsto d(\gamma(t), \gamma'(t))$. This is a convex, periodic function, so it is constant $D' \leqslant D$.

For $s < t \in \mathbb{R}$, consider the quadrilateral $Q$ with sides $[\gamma(s), \gamma(t)] = \gamma_{[s,t]}$, $[\gamma(t), \gamma'(t)]$, $[\gamma'(t), \gamma'(s)] = \gamma'_{[s,t]}$, $[\gamma'(s), \gamma(s)]$. Consider the Euclidean quadrilateral $\bar{Q}$ obtained by identifying two comparison triangles obtained from splitting $Q$ across a diagonal. We claim that $\bar{Q}$ has right angles, and the theorem follows. $\square$

PROPOSITION 1.3.7. *Let $X$ be a complete CAT(k) space. Suppose $Y \subset X$ such that:*

$$r_Y := \inf\{r \mid \exists x, Y \subset B_r(x)\} < D_k/2$$

*Then there exists a unique 'center' $c_Y$ such that $Y \subset \overline{B_{r_Y}(c_Y)}$.*

PROOF. Let $(x_n) \subset X$ and $(r_n) \subset \mathbb{R}$ be sequences with $Y \subset B_{r_n}(x_n)$ and $r_n \overset{n\to\infty}{\to} r_Y$. Fix a basepoint $o \in \mathbb{E}^2$. For $\epsilon > 0$, choose $R' < r_Y < R < D_k/2$ such that a geodesic segment in $\mathcal{M}_k^2$ contained in annulus $A$ with inner radius $R'$ and outer radius $R$ has length less than $\epsilon$. There exists $N$ such that for all $n > N$ we have $R' < r_Y \leqslant r_n < R$. Pick $n, n' > N$ and let $m$ be the midpoint of a geodesic segment $[x_n, x_{n'}]$. For $y \in Y$ let $\Delta_y$ be a comparison triangle in $\mathcal{M}_k^2$ for $y, x_n x_{n'}$ with $\bar{y} = o$, and let $\bar{m}_y$ be the comparison point for $m$ in $\Delta_y$.

Since balls of radius less than $D_k/2$ in $\mathcal{M}_k^2$ are convex, the geodesic $[x_n, x_{n'}]$ is contained in $B_R(o)$, and if both subsegments $[\bar{x}_n, \bar{m}_y]$ and $[\bar{m}_y, \bar{x}_{n'}]$ enter $B_{R'}(o)$, then $\bar{m}_y \in B_{R'}(o)$. It is not the case that all of the $m_y$ are contained in $B_{R'}(o)$, because this would imply $d_X(y, m) < R'$ for all $y \in Y$, which contradicts $R' < r_Y$. Thus, for some $y \in Y$, at least one of the geodesic $[\bar{x}_n, \bar{m}_y]$ and $[\bar{m}_y, \bar{x}_{n'}]$ is contained in $A$, so has length less than $\epsilon$. But then $d_X(x_n, x_{n'}) \leqslant d_{\mathcal{M}^2_k}(\bar{x}_n, \bar{x}_{n'}) < 2\epsilon$. This shows that $(x_n)$ is a Cauchy sequence. Since $X$ is complete, $(x_n)$ converges to some point $c_Y$, which then satisfies the proposition.

Uniqueness of the center follows from Exercise 1.1.4. □

COROLLARY 1.3.8. *If $X$ is a complete CAT(0) space and $G < \operatorname{Isom}(X)$ has a bounded orbit then $G$ fixes a point.*

PROOF. If $x \in X$ such that $Gx$ is bounded then $Gx$ has a unique center $c$. Since $G$ fixes $Gx$, it fixes $c$. □

COROLLARY 1.3.9. *If $G$ acts geometrically on a complete CAT(0) space then $G$ has only finitely many conjugacy classes of finite subgroup.*

PROOF. Since the action is cocompact, there is a compact set $K$ that is a fundamental domain for the action. By proper discontinuity, there are only finitely many elements $g \in G$ such that $gK \cap K \neq \varnothing$. In particular, the subset $F$ of $G$ consisting of elements that fix a point in $K$ is finite. If $H$ is a finite subgroup of $G$ then its orbits are finite, hence, bounded, so the previous result says $H$ fixes a point $x$. Since $K$ is a fundamental domain, there exists $g \in G$ such that $gx \in K$, and $gHg^{-1}$ fixes $gx$. Thus, every finite subgroup of $G$ is conjugate into the finite set $F$. □

## 1.4. Isometries of CAT(0) spaces.

DEFINITION 1.4.1. Let $\phi$ be an isometry of a CAT(0) space $X$.

- The *displacement function of* $\phi$ is $d_\phi\colon X \to \mathbb{R} : x \mapsto d(x, \phi x)$.
- The *translation length of* $\phi$ is $|\phi| = \inf_{x\in X} d_\phi(x)$.
- The *minset of* $\phi$ is the (possibly empty) set $\operatorname{Min}(\phi) := \{x \in X \mid d_\phi(x) = |\phi|$.
- $\phi$ is *hyperbolic* if $|\phi| > 0$ and is realized, that is, there exists $x \in X$ such that $d(x, \phi x) = |\phi|$.
- $\phi$ is *elliptic* if $|\phi| = 0$ and is realized, so $\phi$ has a fixed point.
- $\phi$ is *parabolic* if $|\phi|$ is not realized.
- $\phi$ is *semi-simple* if $|\phi|$ is realized. A subgroup of $\operatorname{Isom}(X)$ is *semi-simple* if all of its elements are semi-simple.

THEOREM 1.4.2. *Let $X$ be a CAT(0) space.*

(1) $\phi \in \operatorname{Isom}(X)$ *is hyperbolic if and only if $\phi$ has an 'axis', a geodesic $\gamma\colon \mathbb{R} \to X$ along which $\phi$ acts by translation: $\forall t \in \mathbb{R}$, $\phi\gamma(t) = \gamma(t + |\phi|)$.*

(2) *If $X$ is complete then $\phi$ is hyperbolic if and only if there exists $n \neq 0$ such that $\phi^n$ is hyperbolic.*
(3) *If $\phi$ is hyperbolic then all axes of $\phi$ are parallel and their union is* $\mathrm{Min}(\phi)$.
(4) *If $\phi$ is hyperbolic then $\mathrm{Min}(\phi)$ is isometric to $Y \times \mathbb{R}$ and $\phi|_{\mathrm{Min}(\phi)} \colon (y,t) \mapsto (y, t + |\phi|)$, so $\{y\} \times \mathbb{R}$ is an axis of $\phi$ for all $y \in Y$.*
(5) *If $\psi \in \mathrm{Isom}(X)$ commutes with $\phi$ then it preserves $\mathrm{Min}(\phi)$. If, in addition, $\phi$ is hyperbolic, then $\psi$ preserves the product structure of* $\mathrm{Min}(\phi)$.

Proof. Suppose $\phi$ acts by translation along a geodesic $\gamma$ by distance $\tau$. A geodesic is convex and complete, so by Proposition 1.3.5 the closest point projection map $\pi_\gamma \colon X \to \gamma$ is distance non-increasing. Let $x \in X$ be a point off $\gamma$. Assume $\gamma$ is parameterized with unit speed and so that $\pi_\gamma(x) = \gamma(0)$. Then:

$$d(x, \phi(x)) \geqslant d(\pi_\gamma(x), \pi_\gamma(\phi(x))) = d(\gamma(0), \phi(\pi_\gamma(x))) = d(\gamma(0), \gamma(\tau)) = \tau$$

Thus, $\phi$ achieves its translation length along $\gamma$.

Conversely, suppose that $\phi$ is hyperbolic, so $|\phi| > 0$ and there exists $x \in \mathrm{Min}(\phi)$. Since $d(\phi(x), \phi(\phi(x))) = d(x, \phi(x)) = |\phi|$, $\mathrm{Min}(\phi)$ is $\phi$–invariant. By convexity of the metric, $\mathrm{Min}(\phi)$ is convex. Therefore, $Y := \bigcup_{n \in \mathbb{Z}} [\phi^n(x), \phi^{n+1}(x)] \subset \mathrm{Min}(\phi)$. By construction, $Y$ is a concatenation of geodesic segments. Let $m$ be the midpoint of $[x, \phi(x)]$. Then $\phi(m)$ is the midpoint of $[\phi(x), \phi^2(x)]$, and $d(m, \phi(m)) = |\phi|$. But $|\phi| = d(m, \phi(m)) \leqslant d(m, \phi(x)) + d(\phi(x), \phi(m)) = |\phi|/2| + |\phi|/2 = |\phi|$, so $[m, \phi(x)] + [\phi(x), \phi(m)]$ is the geodesic from $m$ to $\phi(m)$. This shows that $Y$ is locally geodesic, so, by Proposition 1.3.2, it is a geodesic. This proves (1).

Suppose $\gamma$ and $\gamma'$ are two different axes for $\phi$. Then $f(t) := d(\gamma(t), \gamma'(t))$ is a convex periodic function, so it is constant. This proves (3).

Since $\mathrm{Min}(\phi)$ is nonempty and convex, the induced length metric on $\min(\phi)$ makes it a CAT(0) space. Choose an axis $\gamma$ for $\phi$ in $\mathrm{Min}(\phi)$, and consider closest point projection $\pi_\gamma \colon \mathrm{Min}(\phi) \to \gamma$. Let $Y := \pi_\gamma^{-1}(\gamma(0))$. A consequence of the Flat Strip Theorem is that $\mathrm{Min}(\phi)$ is isometric to $Y \times \mathbb{R}$. This proves (4).

Let $\phi$ and $\psi$ be commuting isometries, and suppose $x \in \mathrm{Min}(\phi)$. Then $d(\phi(\psi(x)), \psi(x)) = d(\psi(\phi(x)), \psi(x)) = d(\phi(x), x) = |\phi|$, so $\psi$ preserves $\mathrm{Min}(\phi)$.

Suppose $\phi$ is hyperbolic with axis $\gamma$ and $\psi$ commutes with $\phi$. Then $\psi(\gamma)$ is a geodesic in $\mathrm{Min}(\phi)$, and $\phi(\psi(\gamma(t))) = \psi(\phi(\gamma(t))) = \psi(\gamma(t + |\phi|))$, so $\psi(\gamma)$ is an axis for $\phi$. This proves (5).

If $\phi$ is hyperbolic then by (1) it has an axis upon which it acts by translation by $|\phi|$. For any nonzero power, $\phi^n$ acts on the same axis by translation by $|n||\phi|$, so $\phi^n$ is hyperbolic. Conversely, if $\phi^n$ is hyperbolic then $\mathrm{Min}(\phi^n) = Y \times \mathbb{R}$. Since $\phi$ commutes with $\phi^n$, it preserves $\mathrm{Min}(\phi^n)$ and its product structure. Consider the restriction of $\phi$ to each of the factors. On $\mathbb{R}$, $\phi$ acts by nontrivial translation, since if it fixed $\mathbb{R}$ pointwise then so would $\phi^n$. Since $\phi^n$ acts trivially on $Y$, $\phi|_Y$ has orbits of size at most $n$, so, by Corollary 1.3.8, $\phi$ fixes a point $y \in Y$. This means that $\phi$ acts by translation on the geodesic $\{y\} \times \mathbb{R}$, so $\phi$ is hyperbolic. □

Exercise 1.4.3. Consider $\mathbb{R}^n$ with the Euclidean metric. Every isometry is of the form $f\colon v \mapsto Av + b$ where $A$ is an orthogonal matrix and $b \in \mathbb{R}^n$. Show that either $f$ fixes a point in $\mathbb{R}^n$ or there is a line along which $f$ acts by nontrivial translation. Conclude that $\mathrm{Isom}(\mathbb{E}^n)$ is semi-simple.

The following theorem shows that there is a link between algebra and geometry for groups acting on CAT(0) spaces: when the group has an Abelian subgroup $\mathbb{Z}^n$, there is a corresponding $n$–dimensional flat subspace.

Theorem 1.4.4 (The flat torus theorem [**6**, Theorem II.7.1]). *Suppose $A \cong \mathbb{Z}^n$ acts properly by semi-simple isometries on a CAT(0) space. Then:*

(1) $\mathrm{Min}(A) = \cap_{\alpha\in A}\mathrm{Min}(\alpha)$ *is nonempty and splits as a product $Y \times \mathbb{E}^n$.*
(2) *Every element $\alpha \in A$ leaves* $\mathrm{Min}(A)$ *invariant and acts by the identity on $Y$ and by translation on $\mathbb{E}^n$.*
(3) *$\forall y \in Y$ the quotient of the $n$–flat $\{y\} \times \mathbb{E}^n$ by the $A$–action is an $n$–torus.*
(4) *For all elements $\beta$ in the normalizer of $A$ in* $\mathrm{Isom}(X)$, *$\beta$ preserves* $\mathrm{Min}(A)$ *and respects the product structure.*
(5) *If $G$ is a subgroup of the normalizer of $A$ in* $\mathrm{Isom}(X)$ *then $G$ has a finite index subgroup $G'$ that centralizes $A$. Moreover, if $G$ is finitely generated then $G$ has a finite index subgroup that contains $A$ as a direct factor.*

Theorem 1.4.5 (Solvable subgroup theorem [**6**, Theorem II.7.8]). *If $G$ acts properly and cocompactly by isometries on a CAT(0), then every virtually solvable subgroup of $G$ is finitely generated and virtually Abelian.*

Definition 1.4.6. A finitely generated subgroup $H$ of a finitely generated group $G$ is said to be *undistorted* if the inclusion of $H$ into $G$ is a quasiisometric embedding, with respect to any choice of finite generating sets of $H$ and $G$.

Exercise 1.4.7. Show that the subgroup $H := \langle a\rangle$ of $G := \langle a, b \mid b^{-1}ab = a^2\rangle$ is distorted (ie is not undistorted). Hint: it suffices to show that with respect to generating set $S := \{a, b\}$ of $G$, the word length $|a^n|_S$ grows sublinearly in $n$. Show that in this example it groups like a logarithm.

More generally, the *Baumslag-Solitar group* $BS(m,n) := \langle a, b \mid b^{-1}a^m b = a^n\rangle$ is said to be *unimodular* when $|m| = |n|$, and $\langle a\rangle$ is distorted whenever $BS(m,n)$ is nonunimodular.

Corollary 1.4.8. *If $G$ acts geometrically on a CAT(0) space then solvable subgroups of $G$ are undistorted.*

Proof. The Solvable Subgroup Theorem says a solvable subgroup $H$ of $G$ is finitely generated and virtually Abelian, so it has a finite index subgroup $A \cong \mathbb{Z}^n$ for some $n$. If $n = 0$ then $H$ is finite, hence undistorted. Otherwise, the Flat Torus Theorem says $H \hookrightarrow G$ is conjugate by a quasiisometry to an isometric embedding $\mathbb{E}^n \hookrightarrow X$, so $H \hookrightarrow G$ is a quasiisometric embedding. □

Corollary 1.4.9. *A group acting geometrically on a CAT(0) space contains no subgroup isomorphic to a nonunimodular Baumslag-Solitar group.*

**1.5. $\mathcal{M}_k$–polyhedral complexes.** This section is concerned with the question of when a space built from CAT(k) pieces is itself a CAT(k) space.

Definition 1.5.1. For $k \in \mathbb{R}$, an $\mathcal{M}_k$–polyhedral complex is a space $X$ made from a disjoint union $\coprod_{i\in I} P_i$ of convex polytopes $P_i \in \mathcal{M}_k^{n_i}$ by identifying some polytopes isometrically along faces.

Theorem 1.5.2. *A connected $\mathcal{M}_k$–polyhedral complex made from finitely many isometry types of polytope is a complete geodesic metric space.*

To see that some condition is necessary, take two vertices $x$ and $y$ and connect them by an edge of length $1/n$ for each $n \in \mathbb{N}$. This is an $\mathcal{M}_0$–polyhedral complex. For all $\epsilon > 0$ there is an edge between $x$ and $y$ of length less than $\epsilon$, so $d(x, y) = 0$. So this polyhedral complex is not even a metric space.

It is clear that a point in the interior of some polytope in an $\mathcal{M}_k$–polyhedral complex has a CAT($k$) neighborhood. It turns out that the only potential problems come at the vertices.

Definition 1.5.3. An $\mathcal{M}_k$–polyhedral complex satisfies the *link condition* if the link of every vertex is a CAT(1) space.

Theorem 1.5.4. *A $\mathcal{M}_k$–polyhedral complex with finitely many isometry types of polytopes has curvature at most $k$ if and only if it satisfies the link condition.*

THEOREM 1.5.5. *A $\mathcal{M}_k$–polyhedral complex with finitely many isometry types of polytopes and satisfying the link condition is CAT(k) if and only if it does not contain an isometrically embedded circle of length less than $2D_k$. If $k \leqslant 0$ this is equivalent to the complex being simply connected.*

EXAMPLE 1.5.6. Take three Euclidean squares, say of side length 1, and glue them around a vertex $v$. The link of $v$ is a triangle with edge lengths $\pi/2$, so it is a loop of length $3\pi/2 < 2D_0 = 2\pi$. This is a $\mathcal{M}_0$–polyhedral complex that does not satisfy the link condition. It does not have curvature at most 0. Every point except $v$ has a CAT(0) neighborhood, but $v$ does not.

If you do the same construction for 4 squares then the resulting space is CAT(0), as the link of $v$ in this case would be a circle of length $2\pi$. In fact, this space is isometric to a Euclidean square of side length 2.

If you glue together more then 4 squares then the length of the link of $v$ is greater than $2\pi$, and this should be thought of as $v$ having strict negative curvature. However, every other point is only locally non-positively curved, so the resulting space is CAT(0), but no better.

Suppose that $X$ is a $\mathcal{M}_k$–polyhedral complex with finitely many isometry types of polytopes and no isometrically embedded loop of length less than $2D_k$. Then we can check whether $X$ is CAT(k) by an induction on dimension, as follows: Since there are finitely many isometry types of cells there is some maximum dimension $n$ of polytope that is used to build $X$. We know that $X$ is CAT(k) if and only if every vertex satisfies the link condition. The link of every vertex is an $\mathcal{M}_1$–polyhedral complex of dimension at most $n-1$. Furthermore, since there were only finitely many isometry types of polytopes in $X$, there are only finitely many isometry types of polytopes used in the links. Thus, $X$ is CAT($k$) if and only if the link of every vertex itself satisfies the link condition and has no isometrically embedded loop of length less than $2D_1 = 2\pi$. If no link has a short loop, then $X$ is CAT($k$) if and only if the link of every vertex of the link of every vertex of $X$ is CAT(1). But the dimension of the $\mathcal{M}_1$–polyhedral complexes we are interested in drops at each iteration, and Exercise 1.2.9 says that once the dimension reaches 1 the space is CAT(1) if and only if there are no short loops.

## 2. Construction of the Davis complex

The construction of the Davis complex $\Sigma(W, S)$ of a Coxeter system $(W, S)$ proceeds in three steps: the first is a formal construction of a simplicial complex, the second is a recellulation of the complex to have a coarser cell structure, the third is to metrize the recellulation. This will produce a $\mathcal{M}_0$–polyhedral complex. Then we need to check that the result is CAT(0).

### 2.1. The formal construction.

DEFINITION 2.1.1. A *partially ordered set (poset)* is a set $X$ and a relation $\leqslant$ satisfying the following conditions:

- $x \leqslant x$
- $x \leqslant y \,\&\, y \leqslant x \implies x = y$
- $x \leqslant y \,\&\, y \leqslant z \implies x \leqslant z$

EXAMPLE 2.1.2. Let $X$ be a set. The power set $\mathcal{P}(X)$ of all subsets of $X$ is partially ordered with respect to inclusion.

DEFINITION 2.1.3. An $n$–chain in $(X, \leqslant)$ is a totally ordered subset of $X$ of size $n$, that is, it is a collection of $n$ distinct element $x_1, \dots, x_n \in X$ such that $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n$.

PROPOSITION 2.1.4. *Given a partially ordered set $(X, \leqslant)$, there is a simplicial complex $Y$ whose $n$–simplices are in bijection with $(n+1)$–chains in $(X, \leqslant)$, and $(X, \leqslant)$ is isomorphic to the poset of cells of $Y$ ordered by inclusion.*

EXAMPLE 2.1.5. Let $X$ be the nonempty subsets of $\{r, s, t\}$, partially ordered by inclusion. The poset and corresponding simplicial complex are shown in Figure 2. Notice that the simplicial complex is a barycentric subdivision of a single simplex. The 'recellulation' step will be to forget the subdivision.

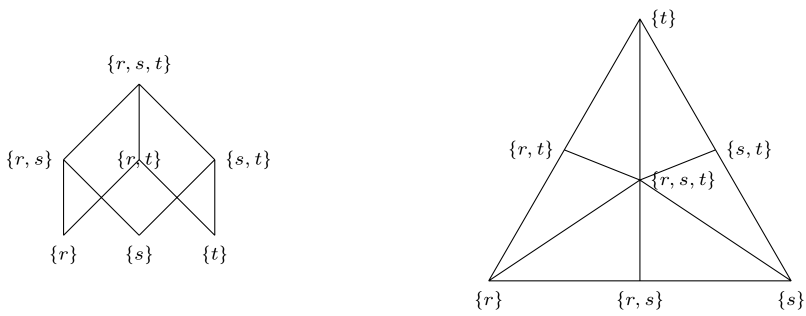



FIGURE 2. Poset and simplicial complex

There will actually be three posets of interest for a Coxeter system $(W, S)$, all defined in terms of spherical subsets of the generators, and all partially ordered by inclusion.

- $\mathcal{S} := \{T \subset S \mid T \text{ is spherical}\} = \{T \subset S \mid W_T \text{ is finite}\}$

- $\mathcal{S}_{>\varnothing} := \{T \neq \varnothing \mid T \text{ is spherical}\}$
- $W\mathcal{S} := \bigcup_{T \text{ spherical}} W/W_T$.

The corresponding simplicial complexes are, respectively:

- The *chamber* $K = K(W,S)$.
- The *nerve* or *link* $L = L(W,S)$.
- The *Davis complex* $\Sigma = \Sigma(W,S)$.

The action of $W$ on $W\mathcal{S}$ by left multiplication extends to a simplicial action of $W$ on $\Sigma$.

LEMMA 2.1.6. *Every simplex of $\Sigma$ is a translate of a unique simplex of $K$.*

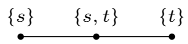
PROOF. A simplex in $\Sigma$ corresponds to a chain $w_0W_{T_0} \subset \cdots \subset w_nW_{T_n}$ in $W\mathcal{S}$, where $w_i \in W$ and $T_i \in \mathcal{S}$. By Corollary 4.0.3, this implies that for $i < j$, $T_i \subset T_j$ and $w_i^{-1}w_j \in W_{T_j}$. In particular, for $i = 0$ and $j > 0$, $w_0^{-1}w_j \in W_{T_j}$, so $w_j \in w_0W_{T_j}$, which means $w_jW_{T_j} = w_0W_{T_j}$. Thus, the chain $w_0W_{T_0} \subset \cdots \subset w_nW_{T_n}$ is the same as $w_0W_{T_0} \subset \cdots \subset w_0W_{T_n}$, which is the $w_0$ translate of the chain $W_{T_0} \subset \cdots \subset W_{T_n}$ in $\mathcal{S}$, which corresponds to a simplex in $K$. $\square$

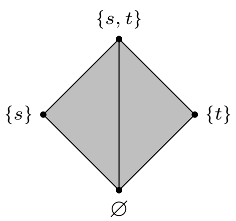
COROLLARY 2.1.7. *$K$ is the fundamental domain for the action of $W$ on $\Sigma$.*

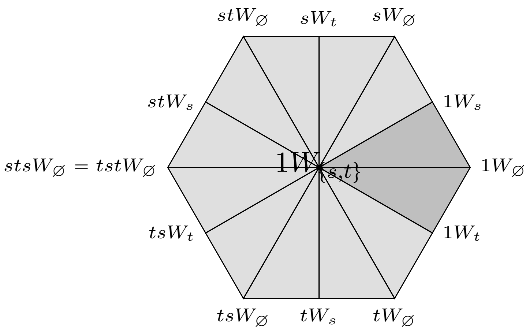
EXERCISE 2.1.8. Show that $\mathcal{U}(W,K)$ is $W$–equivariantly homeomorphic to $\Sigma$.

EXAMPLE 2.1.9. Consider $\mathcal{D}_3 = \langle s,t \mid s^2, t^2, (st)^3\rangle$.

$$\mathcal{S} = \{\varnothing, \{s\}, \{t\}, \{s,t\}\}$$

$L =$ 


$K =$ 

$\Sigma =$ 


Check in this example that $s$ and $t$ act by reflections.

**2.2. Recellulation and metrization.** The definition of $\Sigma$ in the previous section has too many triangles. We simplify the construction as follows. Instead of taking a vertex for each spherical coset we take a cell, and determine attaching maps by inclusion. Specifically, for $T \in \mathcal{S}$ and $g \in W$ take a cell of dimension $|T|$ for $gW_T$. The vertices are then in bijection with cosets of $W_\varnothing$, that is, with elements of $W$. The 1–cells are of the form $gW_s = \{g, gs\}$, attached to vertices $g$ and $gs$. Notice that the 1–skeleton is exactly $Cay(W, S)$. A 2–element spherical subset $\{s, t\}$ generates a dihedral group $\mathcal{D}_{m_{st}}$. It's Cayley graph with respect to $\{s, t\}$ is a cycle of length $2m_{st}$. To the copy of this cycle in $Cay(W, S)$ we attach a 2–cell for $W_{\{st\}}$. Distinct cosets of $W_{\{st\}}$ correspond to distinct cycles in $Cay(W, S)$, so we attached one 2–cell to each distinct $\{s, t\}$–cycle in $Cay(W, S)$. This makes the 2–skeleton of $\Sigma$ a copy of the Cayley 2–complex for the Coxeter presentation of $W$. We continue inductively to construct $\Sigma$ by adding cells one dimension at a time to the existing skeleton. To do this properly we add, for each coset $gW_T$ for $T$ spherical of size $n$, an $n$–ball. We must specify an attaching map from the boundary of the $n$–ball to $\Sigma^{(n-1)}$. In this way we build a complex $\Sigma$ for which the $\Sigma$ of the previous section is the barycentric subdivision. We actually want a polyhedral complex, not just a topological one, so we will specify a metric making the cell corresponding to $gW_T$ a convex Euclidean polytope, and so that the attaching maps are isometries on faces.

Suppose $T \subset S$ is spherical with $|T| = n$. By Theorem 3.0.7 and Theorem 4.0.3, $W_T$ is conjugate (via the square root of its cosine matrix) into $\mathrm{O}(n)$, and acts on $\mathbb{R}^n$ with fundamental domain a simplicial cone, with the elements of $T$ acting as reflections in the codimension 1 faces. A point in the interior of this fundamental domain is *generic point*. For any generic point $x$, the convex hull of finite set $W_T.x$ is a convex Euclidean polytope, and its combinatorial type is independent of the choice of $x$. Let us organize the choice of $x$: let $\bar{d} \in (0, \infty)^T = (d_t \mid t \in T)$. Then there is a unique point in the interior of the fundamental simplicial cone that for all $t \in T$ is distance $d_t$ from the codimension 1 face that is the fixed set of the reflection by $t$. Let $\Sigma_T(\bar{d})$ be the convex Euclidean polytope obtained by taking the convex hull of the $W_T$ orbit of the this point. Call $\Sigma_T(\bar{d})$ a *Coxeter cell* for $T$.

We want to glue Coxeter cells isometrically over shared faces to get a polyhedral complex. We do this by choosing $\bar{d}$ uniformly: choose $\bar{d} \in (0, \infty)^S$, and for spherical $T \subset S$ restrict the choice to $T$: $\bar{d}_T := (d_t \mid t \in T)$, so that if $T$ and $T'$ are spherical then the faces of $\Sigma_T(\bar{d}_T)$ and $\Sigma_{T'}(\bar{d}_{T'})$ corresponding to $T \cap T'$ are both isometric to $\Sigma_{T \cap T'}(\bar{d}_{T \cap T'})$.

PROPOSITION 2.2.1 ([**11**, Proposition 7.3.4]). *For any Coxeter system $(W, S)$ and choice of $\bar{d} \in (0, \infty)^S$ the Davis complex $\Sigma = \Sigma(W, S)$ admits the structure of a Euclidean polyhedral complex such that:*

- – *Its vertex set is in bijection with $W$.*
  – *Its 1–skeleton isomorphic to $Cay(W, S)$.*
  – *Its 2–skeleton is isomorphic to the Cayley complex for the Coxeter presentation of $(W, S)$.*
- *Each $n$–cell corresponds to a coset $gW_T$ where $T$ is spherical of size $n$, and the cell is a Coxeter cell $\Sigma_T(\bar{d}_T)$.*
- *The link of every vertex is $L(W, S)$.*
- *The poset of cells of $\Sigma$ is isomorphic to $W\mathcal{S}$.*

The choice of $\bar{d}$ will not matter in the proof that $\Sigma$ is CAT(0). We will assume, unless specified otherwise, that we have chosen $d_s = 1/2$ for all $s \in S$. The flexibility to make different choices will be used in Section 4 to construct CAT(-1) metrics.

COROLLARY 2.2.2. *$\Sigma$ is simply connected.*

PROOF. The Cayley complex of a presentation is always simply connected [**11**, Proposition 2.2.3]. The fundamental group of a cell complex only depends on its 2–skeleton. □

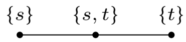
COROLLARY 2.2.3. *$\Sigma$ is a geodesic metric space on which $W$ acts geometrically.*

PROOF. $\Sigma$ is connected $\mathcal{M}_0$–polyhedral complex with finitely many isometry types of cells, so it is a proper geodesic metric space. $W\backslash\Sigma \cong K$ is compact. By construction, $W \curvearrowright \Sigma$ is combinatorial, and $\Sigma$ is locally finite and finite dimensional, so $W \curvearrowright \Sigma$ is properly discontinuous if and only if the cell stabilizers are finite. By construction, cell stabilizers are conjugates of spherical special subgroups, which are finite. □

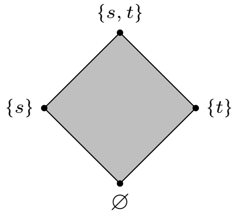
THEOREM 2.2.4. *If $(W, S)$ is a Coxeter system such that every proper special subgroup is finite then $W$ is a simplicial geometric reflection group.*

PROOF. If $W$ is finite then it is a spherical reflection group. If not then the nerve is the boundary of an $|S|-1$ simplex. The fundamental chamber $K$ is a cone on the nerve, so it is an $|S|-1$ simplex. $K$ is a fundamental domain for $W \curvearrowright \Sigma$, by Corollary 2.1.7. Thus, $W$ is a simplicial Coxeter group. (Recall Definition 3.2.11 and Exercise 2.1.8.) By Lannér's Theorem, Theorem 3.2.12, $W$ is a geometric reflection group. $\square$

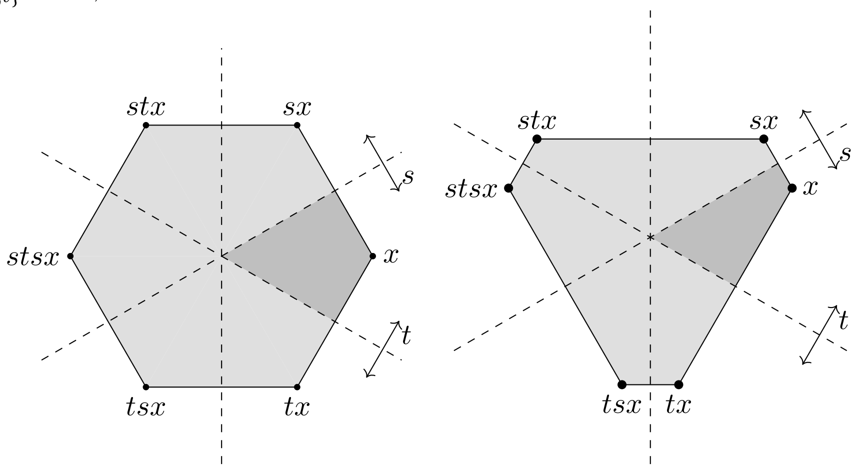
**2.3. Examples of Davis complexes.**

EXAMPLE 2.3.1. We first revisit Example 2.1.9, for which $W = \mathcal{D}_3 = \langle s, t \mid s^2, t^2, (st)^3 \rangle$. Recall $\mathcal{S} = \{\varnothing, \{s\}, \{t\}, \{s, t\}\}$, $L =$ 

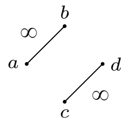
, and

$K =$ 

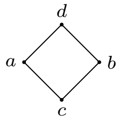
. Below we see $\Sigma(1/2, 1/2)$ and $\Sigma(1/4, 1)$ (not at the same scale). There are both valid choices of Coxeter cell for $W_{\{s,t\}}$. Since in this case the entire group $W$ is spherical, $\Sigma$ is just one closed 2–cell. Note, in each case, the shaded copy of $K$: each vertex of the shaded region has stabilizer $W_T$, where $T$ is the corresponding vertex of $K$, so $x$ corresponds to $\varnothing$ with stabilizer $W_\varnothing = \{1\}$, the origin corresponds to $\{s, t\}$ with stabilizer $W_{\{s,t\}} = W$, etc.




As seen in this first example, the Davis complex for a spherical Coxeter groups is not so exciting on its own; it is simply a ball that we declare has a metric that makes it a Euclidean polytope. What is interesting is that in the non-spherical case the Davis complex is constructed using the spherical subgroups as building blocks.

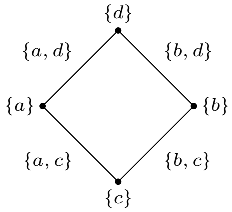
Example 2.3.2. $W = \mathcal{D}_\infty \times \mathcal{D}_\infty$ defined by Coxeter graph $\Gamma =$ 

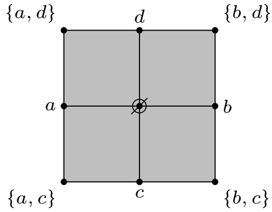
.

The presentation graph is $\Upsilon =$ 

.

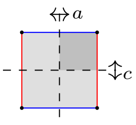
The spherical subsets are all those that contain at most one of $a$ and $b$ and at most one of $c$ and $d$. $\mathcal{S} = \{\varnothing, \{a\}, \{b\}, \{c\}, \{d\}, \{a,c\}, \{a,d\}, \{b,c\}, \{b,d\}\}$.

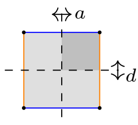
Since the largest spherical subsets have size 2 the Davis complex will be 2–dimensional, and the link will be 1–dimensional. Since $\Upsilon = L^{(1)}$, this means $\Upsilon = L$. That is: $L =$ 

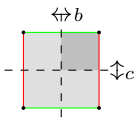
and $K =$ 

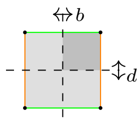
.

The four maximal spherical subsets give isometric Coxeter cells:

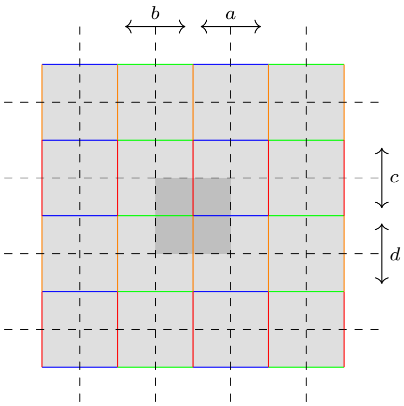
$\Sigma_{\{a,c\}} =$ 

$\Sigma_{\{a,d\}} =$ 

$\Sigma_{\{b,c\}} =$ 

$\Sigma_{\{b,d\}} =$ 


The link tells us to put one of these squares at each vertex, with incident edges of the same color identified, to make a link that is a circle consisting of 4 arcs. We do this around every vertex to get a square tiling of $\mathbb{E}^2$.

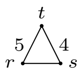


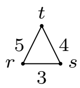

The squares with colored boundaries are copies of one of the four Coxeter 2–cells. The dashed lines are walls across which an element of $W$ acts by reflection. The walls for the four generators are labelled. For instance, the wall across which $a$ reflects is transverse to a set of blue edges. The other set of blue edges in the picture is transverse to a wall that is fixed by $bab$.

This example is a Euclidean reflection group. The dashed reflection lines give the tessellation of $\mathbb{E}^2$ by squares whose fundamental domain is the shaded copy of $K$, with the generators acting by reflections in the sides of $K$. Here we see that the Davis complex is dual to the Euclidean reflection tessellation.

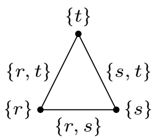
EXERCISE 2.3.3. Construct the Davis complexes for the three $\mathbb{E}^2$ triangle reflection groups, and compare to Figure 2 of Chapter 2.

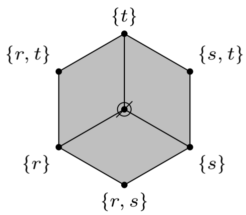
EXAMPLE 2.3.4. Consider the Coxeter system defined by Coxeter graph:

$\Gamma =$ 

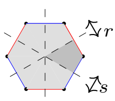

Its presentation graph is: $\Upsilon =$ 

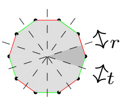

All proper subsets of $S$ are spherical: $\mathcal{S} = \{\varnothing, \{r\}, \{s\}, \{t\}, \{r,s\}, \{s,t\}, \{r,t\}\}$.

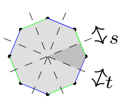
$\Upsilon \cong L^{(1)} = L =$ 

and $K =$ 

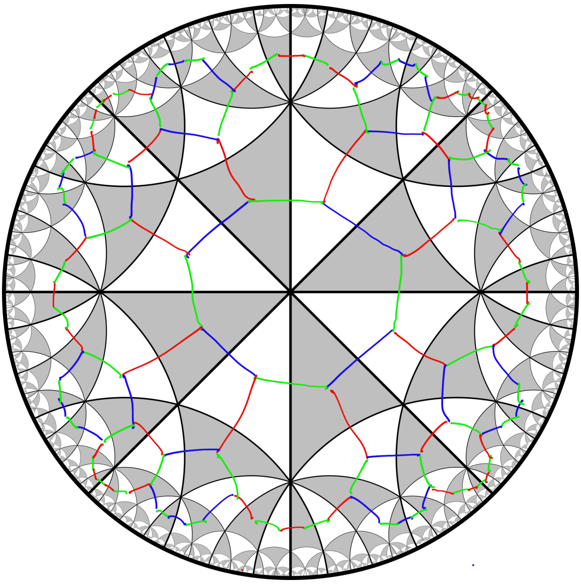

The three maximal spherical subsets give Coxeter cells:

$\Sigma_{\{r,s\}} =$ 

$\Sigma_{\{r,t\}} =$ 

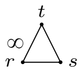
$\Sigma_{\{s,t\}} =$ 

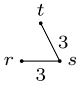

To build the Davis complex $\Sigma$ we put one of each of these three pieces around each vertex, with incident edges of the same color identified. This is example is again a geometric reflection group, a hyperbolic triangle group, so the Davis complex is dual to the tessellation of $\mathbb{H}^2$ by triangles:



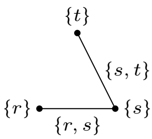
Let us see a non-geometric example.

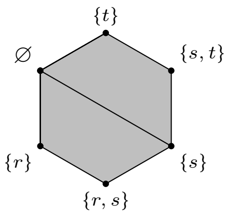
Example 2.3.5. Consider the Coxeter system defined by Coxeter graph:

$\Gamma =$ 

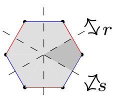

Its presentation graph is: $\Upsilon =$ 

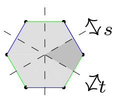

A subset of $S$ is spherical if it does not contain both $r$ and $t$: $\mathcal{S} = \{\varnothing, \{r\}, \{s\}, \{t\}, \{r,s\}, \{s,t\}\}$.

$\Upsilon \cong L^{(1)} = L =$ 

and $K =$ 


The two maximal spherical subsets give Coxeter cells:

$\Sigma_{\{r,s\}} =$ 

$\Sigma_{\{s,t\}} =$ 


In this example the link tells us that at each vertex there is one copy $\Sigma_{r,s}$ and one copy of $\Sigma_{s,t}$, and they are glued together across the face $\Sigma_{\{s\}}$ (a blue edge). Since $W_{\{r,t\}} \cong \mathcal{D}_\infty$ the figure for $K$ says the red-green geodesic $\Sigma_{\{r,t\}}$ is a boundary of $\Sigma$ (because there is no 2-cell in $K$ connecting $\{r\}$ and $\{t\}$). This example is a tree of hexagons, a piece of which is shown in Figure 3. The blue edges are all interior edges, and the red and green edges are boundary edges. Recall that this group does act on a tessellation of $\mathbb{H}^2$

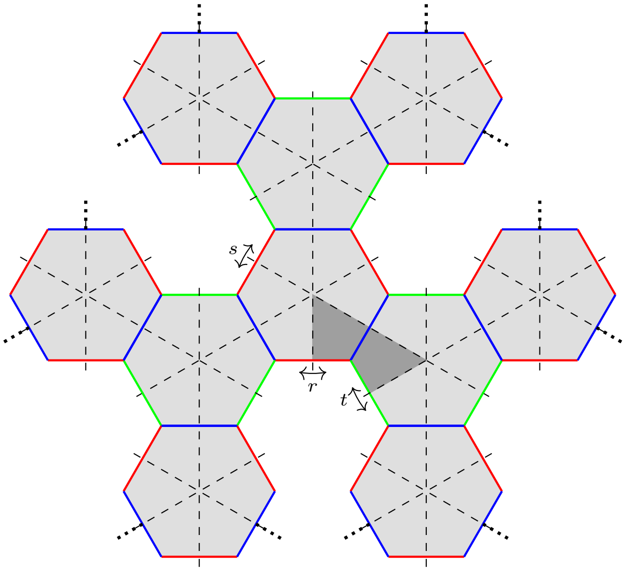



FIGURE 3. $\Sigma$ for $\Delta(3,3,\infty)$ is a tree of hexagons, with fundamental domain $K$ shaded.

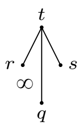
by *ideal* hyperbolic triangles, as seen in Figure 10 of Chapter 2, but this action is not cocompact, since the fundamental domain is not compact.

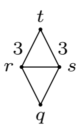
Alternatively there is the Coxeter complex, for which the action is cocompact but not proper. For this example it would amount to collapsing the red and green edges, giving a tree of triangles with all edges blue and all vertices infinite valence. We saw a similar example in Figure 6 of Chapter 3, which was a tree of squares instead of a tree of triangles.

**This example is therefore the first one in which we get something genuinely new from $\Sigma$: we get a geometric action of $W$, where we did not have one from either of the classical constructions.**

Finally, let us a try an example of dimension $> 2$:

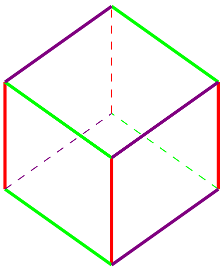
EXAMPLE 2.3.6. Consider the Coxeter system defined by Coxeter graph:

$\Gamma =$ 

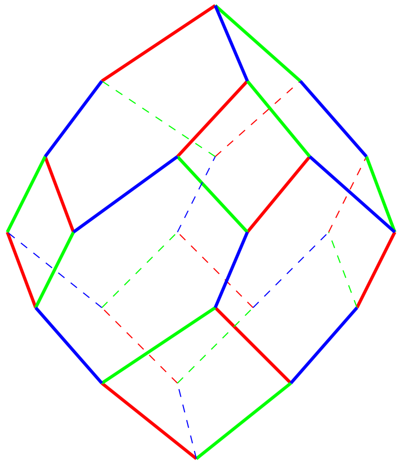

Its presentation graph is: $\Upsilon =$ 


A subset of $S$ is spherical if it does not contain both $q$ and $t$. Both of the triangles in $\Upsilon$ are spherical subgroups, so the nerve $L$ is $\Upsilon$ with the two triangles filled by 2–simplices.

The two maximal spherical subsets $\{q, r, s\}$ and $\{r, s, t\}$ give Coxeter cells as in Figure 4. $\Sigma_{\{q,r,s\}}$ is a cube. $\Sigma_{\{r,s,t\}}$ is known as the *order 4 permutohedron.* Its 1–skeleton is the Cayley graph of the symmetric group on 4 elements corresponding to the generating set consisting of transpositions $r = (12)$ (red), $s = (34)$ (green), and $t = (23)$ (blue).



(A) $\Sigma_{\{q,r,s\}}$



(B) $\Sigma_{\{r,s,t\}}$

FIGURE 4. Coxeter cells for Example 2.3.6.

The nerve tells us to construct $\Sigma$ by gluing $\Sigma_{\{q,r,s\}}$ to $\Sigma_{\{r,s,t\}}$ along the face $\Sigma_{\{r,s\}}$, that is, glue cubes to permutohedra along red-green squares. The blue and violet edges are not glued to anything, they remain on the boundary. We get that $\Sigma$ is a fattened tree of valence 6, where the fat vertices are permutohedra and the fat edges are cubes. Edge paths that alternate blue-violet are geodesics on the surface of the fattened tree that are translates of $\Sigma_{\{q,t\}} \cong \mathbb{R}$, since $W_{\{q,t\}} \cong \mathcal{D}_\infty$.

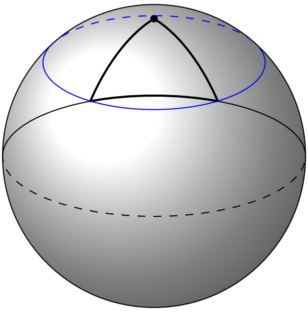
EXERCISE 2.3.7. The Coxeter group $W$ of Example 2.3.6 acts geometrically on a fattened tree $\Sigma$. It follows from standard results in Geometric Group Theory that $W$ has a finite index free subgroup. Find one, and argue that it is free using Theorem 1.3.5.

EXERCISE 2.3.8. By construction, a (conjugate of a) generator acts by Euclidean reflection on each Coxeter cell that it preserves, so there is a codimension 1 fixed subspace in that cell. Show that for each generator $s$ the union of such fixed subspaces is a 'topological wall' separating $\Sigma$ into two complementary connected components. Show that the edges met by this topological wall are the wall $\Omega_s$ of Section 1.2 of Chapter 2.

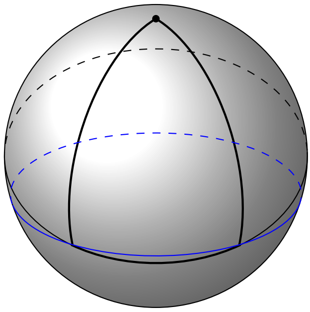
**2.4. The Davis complex is CAT(0).** In this section we will show that the Davis complex is CAT(0). The main technical tool is the following:

Lemma 2.4.1 (Moussong's Lemma). *Suppose $L$ is an $\mathcal{M}_1$–polyhedral complex with all edges of length at least $\pi/2$. Then $L$ is CAT(1) if and only if $L$ is a metric flag complex.*

This will require some work. 'Metric flag' means $L$ is determined by its 1–skeleton, see Definition 2.4.3. The crux of the proof, depicted in Figure 5, is that spherical simplices start to bulge once their edge lengths are longer than $\pi/2$, see Lemma 2.4.2, and this forces geodesics to follow the 1–skeleton. But if edges have length at least $\pi/2$ then a loop in the 1–skeleton of length less than $2\pi$ must be a triangle, and the flag condition implies such a triangle is filled by a simplex, so there are no short geodesic loops.



(A)



(B)

Figure 5. The distance from a vertex of a simplex to the opposite face is convex when the simplex is small, but concave when the simplex is big. Points equidistant from the North Pole are lines of latitude (blue), while sides of the simplex are arcs of great circles.

Lemma 2.4.2. *Suppose a convex spherical simplex has vertices $\{v_0, \ldots, v_n\}$ with $\ell_{ij} := d(v_i, v_j) \geqslant \pi/2$ when $i \neq j$. Then the distance from $v_0$ to the opposite face is realized by* $\min_{i>0} d(v_0, v_i)$.

Proof. Consider a spherical 2–simplex with edges of length at least $\pi/2$. Assume, renumbering if necessary, that $d(v_0, v_1) \leqslant d(v_0, v_2)$. Up to isometry we may assume $v_0 = (0, 0, 1) \in \mathbb{S}^2 \subset \mathbb{R}^3$, and that $v_1 = (a, b, c)$ and $v_2 = (x, y, z)$ with $z \leqslant c \leqslant 0$. The face opposite $v_0$ is the unique spherical geodesic from $v_1$ to $v_2$ given by $\gamma\colon t \mapsto \frac{(1-t)v_1+tv_2}{|(1-t)v_1+tv_2|}$

$$d(v_0, \gamma(t)) = \pi/2 + \sin^{-1}\left(\left|\frac{(1-t)c + tz}{|(1-t)v_1 + tv_2|}\right|\right) \geqslant \pi/2 + \sin^{-1}(|c|) = d(v_0, v_1)$$

Furthermore, either $c = z = 0$ and $\gamma$ runs along the equator, so that $\gamma$ is equidistant from $v_0$, or every point in the interior of $\gamma$ is strictly farther from $v_0$ than $v_1$.

Now proceed by induction. Let $x$ be a point in the codimension 1 face $F(v_1, \ldots, v_n)$ containing $\{v_1, \ldots, v_n\}$ that is not contained in the codimension 2 face $F(x_1, \ldots, x_{n-1})$ containing $\{v_1, \ldots, v_{n-1}\}$. Let $y$ be the first point of $F(x_1, \ldots, x_{n-1})$ on the spherical geodesic starting at $v_n$ and passing through $x$. By the induction hypothesis, $d(v_0, y) \geqslant \min_{0<i<n} d(v_0, v_i)$. By the 2–dimensional case:

$$d(v_0, x) \geqslant \min\{d(v_0, y), d(v_0, v_n)\} \geqslant \min_{0<i\leqslant n} d(v_0, v_i) \qquad \square$$

DEFINITION 2.4.3. A simplicial spherical complex $L$ is a *metric flag complex* if $L$ is metrically determined by its 1–skeleton; that is, if some set of edges of $L$ can be the 1–skeleton of a spherical simplex, then there is a spherical simplex of $L$ whose 1–skeleton is that set of edges.

EXAMPLE 2.4.4. Suppose $a \geqslant b \geqslant c > 0$ are three numbers that can be the three side lengths of a triangle, that is, all permutations satisfy the strict triangle inequality. We claim they are the three sides of a spherical simplex if and only if $a + b + c < 2\pi$. So a triangle graph with edge lengths $a$, $b$, $c$ is a metric flag complex if and only if $a + b + c \geqslant 2\pi$.

The claim relies on a different characterization of simplices, which we will establish in Lemma 2.4.8: there exists a spherical 2–simplex with side lengths $a$, $b$, $c$ between 0 and $\pi$ if and only if $C^* := \begin{pmatrix} 1 & \cos a & \cos b \\ \cos a & 1 & \cos c \\ \cos b & \cos c & 1 \end{pmatrix}$ is positive definite.

We check positive definiteness via Sylvester's Criterion by checking that the principal minors are positive. The first two are easy. The third is the determinant, which, after some manipulation, is equal to $(\cos a - \cos(b + c))(\cos(b - c) - \cos a)$. We will show both factors are positive.

Consider the second factor. It is non-positive when $\cos(b - c) \leqslant \cos a$, but $b - c$ and $a$ are in $[0, \pi)$, so this is true if and only if $b - c < a \leqslant b - c$, which is a contradiction.

Consider the first factor. It is non-positive when $\cos a \geqslant \cos(b + c)$, which could be true for $a \geqslant b + c$ or for $b + c \geqslant 2\pi - a$. The first case implies $b + c \leqslant a < b + c$, and the second implies $a + b + c \geqslant 2\pi$, both of which are contradictions.

2.4.1. *Polar duals.* The edge length hypothesis in Moussong's Lemma turns out to be dual to the 'non-obtuse' hypothesis on dihedral angles from Chapter 2.

LEMMA 2.4.5. *Let $U = \{\mathbf{u}_1, \mathbf{u}_2, \ldots\}$ and $V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots\}$ be finite sets of unit vectors in $\mathbb{R}^{n+1}$. For $\mathbf{u} \in U$, consider the halfspace bounded by $\mathbf{u}^\perp$*

*containing* $\mathbf{u}$. *Let* $K$ *be their intersection. Let* $K'$ *be the intersection of linear halfspaces containing* $V$. *Let* $\sigma := K \cap \mathbb{S}^n$ *and* $\sigma' := K' \cap \mathbb{S}^n$.

*Then* $\sigma$ *is a convex* $n$*–dimensional spherical polytope if and only if* $U$ *spans* $\mathbb{R}^{n+1}$ *and is contained in one complementary component of a linear hyperplane. Similarly,* $\sigma'$ *is a convex* $n$*–dimensional spherical polytope if and only if* $V$ *spans* $\mathbb{R}^{n+1}$ *and is contained in one complementary component of a linear hyperplane*

*A unit vector* $\mathbf{v}$ *is a vertex of* $\sigma$ *if and only if* $< \mathbf{u}, \mathbf{v} > \geqslant 0$ *for all* $\mathbf{u} \in U$ *and* $\langle U \cap \mathbf{v}^\perp \rangle$ *has dimension* $n$.

*A unit vector* $\mathbf{u}$ *is the inward pointing unit normal vector to a codimension 1 face of* $\sigma'$ *if and only if* $< \mathbf{u}, \mathbf{v} > \geqslant 0$ *for all* $\mathbf{v} \in V$ *and* $\langle V \cap \mathbf{u}^\perp \rangle$ *has dimension* $n$.

Proof. $K$ is a finite intersection of linear halfspaces, so it is a convex cone. Thus, any two points in $\sigma$ can be joined by a geodesic in $\sigma$, and this geodesic is unique unless the two points are antipodes in $\mathbb{S}^n$.

$$\begin{aligned} \sigma \text{ is convex in } \mathbb{S}^n &\iff \sigma \text{ does not contain antipodal points of } \mathbb{S}^n \\ &\iff K \text{ does not contain a line} \\ &\iff \text{the span of } U \text{ has positive codimension} \end{aligned}$$

The condition that $U$ is contained in a component of a hyperplane is equivalent to the existence of a unit vector $\mathbf{w}$ such that $< \mathbf{w}, \mathbf{u} >$ is strictly positive for all $\mathbf{u} \in U$. Since $U$ is finite, this is equivalent to the existence of $\epsilon > 0$ such that $< \mathbf{w}, \mathbf{u} > \geqslant \epsilon$, which is equivalent to $\mathbf{w}$ being in the interior of $\sigma$. So the condition is equivalent to $\sigma$ having nonempty interior, thus, being $n$–dimensional.

The proof that $\sigma'$ is an convex $n$–dimensional spherical polytope is similar, with the roles of the two properties exchanged: the existence of $\mathbf{w} \in \mathbb{S}^n$ such that $< \mathbf{w}, \mathbf{v} > \geqslant \epsilon > 0$ for all $\mathbf{v} \in V$ implies that any convex combination of points of $V$ is contained in the positive $\mathbf{w}$ halfspace, so $\sigma'$ is convex, and the spanning condition implies $\sigma'$ is $n$–dimensional.

The condition $< \mathbf{u}, \mathbf{v} > \geqslant 0$ for all $\mathbf{u} \in U$ is equivalent to $\mathbf{v} \in \sigma$. A vector $\mathbf{w}$ is in $\langle U \cap \mathbf{v}^\perp \rangle^\perp$ if and only if for all sufficiently small $\epsilon > 0$ and all $\mathbf{u} \in U$ we have both $\langle \mathbf{v} + \epsilon \mathbf{w}, \mathbf{u} \rangle \geqslant 0$ and $\langle \mathbf{v} - \epsilon \mathbf{w}, \mathbf{u} \rangle \geqslant 0$. Thus, $\mathbf{v}$ spans an extremal ray of $K$, or, equivalently, is a vertex of $\sigma$, if and only if $\langle v \rangle = \langle U \cap \mathbf{v}^\perp \rangle^\perp$.

The condition $< \mathbf{u}, \mathbf{v} > \geqslant 0$ for all $\mathbf{v} \in V$ says the halfspace bounded by $\mathbf{u}^\perp$ containing $\mathbf{u}$ contains $V$. When $\langle V \cap \mathbf{u}^\perp \rangle$ is nontrivial, $K' \cap \langle V \cap \mathbf{u}^\perp \rangle$ is a face of $K'$ of dimension equal to the dimension of $\langle V \cap \mathbf{u}^\perp \rangle$. □

DEFINITION 2.4.6. If $P$ is a convex spherical polytope of dimension $n$ with inward pointing unit normals $U = \{\mathbf{u}_1, \mathbf{u}_2, \dots\}$ and vertices $V = \{\mathbf{v}_1, \mathbf{v}_2, \dots\}$, its *polar dual* is the convex spherical polytope $P^*$ of dimension $n$ with inward pointing unit normals $V$ and vertices $U$.

The previous lemma implies that $P^*$ is in fact a convex spherical polytope of dimension $n$, and that $P$ and $P^*$ are combinatorially dual. It is then clear from the definition that $P^{**} = P$.

COROLLARY 2.4.7. *A convex spherical polytope $P$ is a simplex if and only if $P^*$ is a simplex.*

Let $P$ be a convex spherical polytope whose inward pointing unit normals are $\{\mathbf{u}_i\}$ and whose vertices are $\{\mathbf{v}_i\}$.

Now we can give analogues of Lemma 3.2.6, Lemma 4.1.2, and Proposition 4.1.3 from Chapter 2.

LEMMA 2.4.8. *Let $V = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subset \mathbb{S}^{n-1} \subset \mathbb{R}^n$. Suppose that $\ell_{ij}$ is the spherical distance between $\mathbf{v}_i$ and $\mathbf{v}_j$. Then $V$ is the set of vertices of a convex spherical simplex with edges of length $\ell_{ij}$ if and only if the Gram matrix of $V$ is positive definite.*

PROOF. By Proposition 4.1.3, the Gram matrix of $V$ is positive definite if and only if $V$ is the set of inward pointing unit normals of a spherical simplex $P$, which is true if and only if $V$ is the set of vertices of a spherical simplex $P^*$. □

COROLLARY 2.4.9. *A spherical simplex is determined up to isometry by the Gram matrix of its vertices, or, equivalently, by its edge lengths.*

COROLLARY 2.4.10. *Let $C = (c_{ij})$ be a symmetric matrix with 1's on the diagonal and off-diagonal entries $c_{ij} \in (-1, 1)$. Let $\ell_{ij} := \cos^{-1} c_{ij}$. There exists a spherical simplex with edges of length $\ell_{ij}$ if and only if $C$ is positive definite.*

PROOF. If $C$ is positive definite then $C$ is the Gram matrix of the unit vectors $V$ given be the columns of $\sqrt{C}$. By the previous lemma, these are the vertices of a simplex, which by construction has edges of lengths $\ell_{ij}$. The other direction is also implied by the previous lemma. □

If $P$ is a convex spherical polytope with inward pointing unit normals $U = \{\mathbf{u}_1, \dots\}$ and vertices $V = \{\mathbf{v}_1, \dots\}$, let $C(P)$ be the Gram matrix of $U$. Recall this is the matrix $(\langle \mathbf{u}_i, \mathbf{u}_j \rangle) = (\langle \cos(\angle \mathbf{u}_i, \mathbf{u}_j))$. Recall that if the codimension 1 faces normal to $\mathbf{u}_i$ and $\mathbf{u}_j$ intersect in a codimension 2 face then the dihedral angle of $P$ at their intersection is $\theta_{ij} = \pi - \angle \mathbf{u}_i, \mathbf{u}_j$, so $-\cos\theta_{ij} = \cos(\angle \mathbf{u}_i, \mathbf{u}_j)$. Let $C^*(P)$ be the Gram matrix of $V$, which is $(\langle \mathbf{v}_i, \mathbf{v}_j\rangle) = (\cos(\angle \mathbf{v}_i, \mathbf{v}_j)) = (\cos \ell_{ij})$ where $\ell_{ij}$ is the spherical distance between $\mathbf{v}_i$ and $\mathbf{v}_j$. Evidently, $C(P^*) = C^*(P)$.

LEMMA 2.4.11. *If $P$ is a convex spherical polytope such that all of its edges have length at least $\pi/2$ then $P$ is a simplex.*

PROOF. By duality, there is an edge between vertices $\mathbf{v}_i$ and $\mathbf{v}_j$ in $P$ if and only if the corresponding codimension 1 faces of $P^*$ intersect in a codimension 2 face. All edges of $P$ have length at least $\pi/2$ when the off-diagonal entries of $C^*(P)$ are non-positive. But $C^*(P) = C(P^*)$, so the dihedral angles between codimension 1 faces are $\theta_{ij}$ such that $\cos\theta_{ij}$ for $i \neq j$ is non-negative, hence $\theta_{ij} \leqslant \pi/2$. Then Lemma 3.2.6 says $P^*$ is a simplex, so $P$ is as well. □

2.4.2. *Proof of Moussong's Lemma.* Given Lemma 2.4.11, one direction of Moussong's Lemma is easy:

EXERCISE 2.4.12. Prove that a CAT(1) spherical simplicial complex is a metric flag complex.

For the other direction we first state an alternative characterization of CAT(1) spaces, then give three auxiliary lemmas.

LEMMA 2.4.13 (Bowditch's Lemma). *If $X$ is a compact metric space of curvature at most 1 and every closed rectifiable curve of length strictly less than $2\pi$ is* shrinkable *then $X$ is CAT(1).*

Here, 'shrinkable' means that the curve is homotopic to a strictly shorter curve through a homotopy of closed rectifiable curves of non-increasing length. For example, consider a flat cylinder of circumference $2\pi$ with one boundary component glued to the boundary of a unit hemisphere. The remaining boundary curve is shrinkable: it can be homotoped along the cylinder through constant length curves until one reaches the hemisphere, at which point it can be homotoped through strictly shorter curves to a point. For a non-example, consider $\mathbb{S}^2$ with an open disc of radius less than $\pi/2$ removed. The boundary circle is not shrinkable: it can be homotoped to a point, but only by first increasing its length to $2\pi$.

LEMMA 2.4.14. *The property of being a spherical simplicial complex with all edges of length at least $\pi/2$ is inherited by links.*

PROOF. Let $L$ be a spherical simplicial complex with all edges of length at least $\pi/2$. Let $v_0$, $v_1$, and $v_2$ be the vertices of a 2–simplex $\sigma$ in $L$. Then there are vertices $\hat{v}_1$ and $\hat{v}_2$ in $lk(v_0)$ connected by an edge $\hat{\sigma}$ whose length is the angle $\theta_0 < \pi$ of $\sigma$ at $v_0$. Since $\pi/2 \leqslant \ell_{ij} < \pi$, we have $\cos \ell_{ij} \leqslant 0$ and $\sin \ell_{ij} > 0$, so the spherical law of cosines gives $\cos \theta_0 = \frac{\cos \ell_{12} - \cos \ell_{01} \cos \ell_{02}}{\sin \ell_{01} \sin \ell_{02}} \leqslant 0$, so $\theta_0 \geqslant \pi/2$. ☐

LEMMA 2.4.15. *The property of being a metric flag complex is inherited by links.*

PROOF. Let $L$ be a metric flag complex. Let $v_0$ be an arbitrary vertex of $L$. Since $L$ is simplicial, the link of $v_0$ in $L$ only depends simplices whose vertices are all connected to $v_0$ by edges. Let $\{v_1, \dots, v_k\}$ be the vertices of $L$ that are adjacent to $v_0$.

Define $\ell_{ij}$ to be the length of the edge between $v_i$ and $v_j$, if such an edge exists in $L$. Set $\ell_{ii} = 0$ and $\ell_{ij} = \pi$ if $v_i$ and $v_j$ are distinct vertices not connected by an edge. Define a symmetric bilinear form $B$ on $\mathbb{R}^{k+1}$ by taking coordinate vectors $\mathbf{e}_0, \dots, \mathbf{e}_k$, setting $B(e_i, e_j) = \cos \ell_{ij}$, and extending bilinearly. Let $V := \mathbf{e}_0^{\perp_B}$. Let $p\colon \mathbb{R}^{k+1} \to V : \mathbf{w} \mapsto \mathbf{w} - B(\mathbf{w}, \mathbf{e}_0)\mathbf{e}_0$. Observe, for $i \neq 0$, that $\sin \ell_{0i} > 0$ and $B(p(\mathbf{e}_i), p(\mathbf{e}_i)) = 1 - \cos^2 \ell_{0i} = \sin^2 \ell_{0i}$, so we can define $\hat{\mathbf{e}}_i := \frac{p(\mathbf{e}_i)}{\sin \ell_{0i}}$. Let $\hat{v}_i$ be the vertex of the link $lk(v_0, L)$ corresponding to the edge between $v_0$ and $v_i$.

We have:

$$B(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j) = \frac{\cos \ell_{ij} - \cos \ell_{0i} \cos \ell_{0j}}{\sin \ell_{0i} \sin \ell_{0j}} \tag{14}$$

If $v_0$, $v_i$, and $v_j$ span a simplex in $L$ then $\ell_{ij} < \pi$ and the angle of this simplex at $v_0$ is some $0 < \hat{\ell}_{ij} < \pi$. The Spherical Law of Cosines and (14) give:

$$B(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j) = \cos \hat{\ell}_{ij} \tag{15}$$

Suppose that $\hat{I} \subset \{1, \dots, k\}$ is a set of indices whose corresponding vertices $\hat{v}_i$ in $lk(v_0, L)$ are pairwise connected by edges. Let $I := \{0\} \cup \hat{I}$. By construction $B(\mathbf{e}_0, \mathbf{e}_0) = 1$ and $\mathbf{e}_0$ is $B$–orthogonal to $\hat{\mathbf{e}}_i$ for all $i \in \hat{I}$, so:

$$B \text{ is pos. def. on } \langle \hat{\mathbf{e}}_i \mid i \in \hat{I} \rangle \iff B \text{ is pos. def. on } \langle \mathbf{e}_i \mid i \in I \rangle \tag{16}$$

We have:

$$
\begin{aligned}
&\{\hat{v}_i \mid i \in \hat{I}\} \text{ can span a simplex with edge lengths } \hat{\ell}_{ij} \\
&\quad \overset{2.4.10}{\iff} (\cos \hat{\ell}_{ij}) \text{ for } i, j \in \hat{I} \text{ is positive definite} \\
&\quad \overset{(15)}{\iff} (B(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j)) \text{ for } i, j \in \hat{I} \text{ is positive definite} \\
&\quad \iff B \text{ is positive definite on } \langle \hat{\mathbf{e}}_i \mid i \in \hat{I} \rangle \\
&\quad \overset{(16)}{\iff} B \text{ is positive definite on } \langle \mathbf{e}_i \mid i \in I \rangle \\
&\quad \iff (B(\mathbf{e}_i, \mathbf{e}_j)) \text{ for } i, j \in I \text{ is positive definite} \\
&\quad \iff (\cos \ell_{ij}) \text{ for } i, j \in I \text{ is positive definite} \\
&\quad \overset{2.4.10}{\iff} \{v_i \mid i \in I\} \text{ can span a simplex with edge lengths } \ell_{ij} \\
&\quad \overset{L \text{ is m. flag}}{\iff} \{v_i \mid i \in I\} \text{ do span a simplex in } L \text{ with edge lengths } \ell_{ij} \\
&\quad \iff \{\hat{v}_i \mid i \in \hat{I}\} \text{ do span a simplex in } lk(v_0, L) \text{ with edge lengths } \hat{\ell}_{ij} \square
\end{aligned}
$$

Proof of Moussong's Lemma. Let $L$ be a spherical complex with all edges of length at least $\pi/2$. By Lemma 2.4.11, $L$ is simplicial. By Exercise 2.4.12, if $L$ is CAT(1) it is a metric flag complex.

Conversely, suppose $L$ is a metric flag complex with all edges of length at least $\pi/2$.

Recall that the "link condition" is that the link of every vertex is a CAT(1) space, and this is equivalent to the spherical complex being of curvature at most 1. The "girth condition" is that the complex does not contain an isometrically embedded loop of length less than $2\pi$.

Suppose $L$ is not CAT(1). Either it is of curvature at most 1 and the girth condition fails, or it is not of curvature at most 1, so the link condition fails. Thus, some link is not a CAT(1) space. So either every link has curvature at most 1 but one of them fails the girth condition, or some link is not of curvature at most 1, so fails the link condition. Repeat the argument for iterated links. Since the dimension drops each time we pass to a link, eventually the links are 1–dimensional, in which case they fail to be CAT(1) only if the girth condition fails.

Since the properties of being metric flag and having edges of length at least $\pi/2$ pass to links, every iterated link of $L$ has these properties. Thus, $L$ fails to be CAT(1) if and only if there is some iterated link $L'$ of $L$ that is a metric flag complex with edges of length at least $\pi/2$ and curvature at most 1 and fails the girth condition.

Supposing such an $L'$ exists, we apply Bowditch's Lemma and derive a contradiction. Bowditch's Lemma implies that $L'$ contains a non-shrinkable loop $\gamma$ of length less than $2\pi$. Since $L'$ is compact and locally contractible, we may assume $\gamma$ is locally geodesic with minimal length among all non-trivial, non-shrinkable loops.

Let $v$ be a vertex of $L'$. Since edges of $L'$ have length at least $\pi/2$ and less than $\pi$, we have the following:

- The closed ball of radius $\pi/2$ about any vertex is isometric to the spherical cone on its link.
- The closed ball of radius $\pi/2$ about a vertex $v$ is contained in the union of closed cells containing $v$, the closed star of $v$.
- $L'$ is covered by open balls of radius $\pi/2$ around its vertices.

We further claim that the covering of $L'$ by balls of radius $\pi/2$ around its vertices has nerve $L'^{(1)}$, meaning two such balls intersect if and only if there is an edge of $L'$ connecting their centers. Edges of $L'$ have length less than $\pi$, so it is clear that the nerve contains a subgraph isomorphic to $L'^{(1)}$. Suppose $v$ and $v'$ are vertices of $L'$ such that $d(v, v') < \pi$, so they are adjacent in the nerve of the cover. Suppose $v$ and $v'$ are not contained in a common simplex of $L'$. Since $\overline{B_{\pi/2}v}$ is contained in the closed star of $v$, and similarly for $v'$, the closed stars intersect. This implies that there are simplices $\sigma$ and $\sigma'$ of $L'$ such that $\sigma$ contains $v$, $\sigma'$ contains $v'$, and $\sigma \cap \sigma'$ contains a point in $B_{\pi/2}v \cap B_{\pi/2}v'$. By Lemma 2.4.2, the distance from $v$ to $\sigma \cap \sigma'$ is at least $\pi/2$, as is the distance from $v'$ to $\sigma \cap \sigma'$, which is a contradiction. Thus, every pair of vertices $v$ and $v'$ of $L'$ at distance less than $\pi$ are contained in a common simplex, hence, are adjacent in $L'^{(1)}$.

Now we claim that $\gamma$ can be homotoped to be contained in $L'^{(1)}$. Specifically, let $\delta$ be an arc of $\gamma$ whose interior is contained in the ball of radius $\pi/2$ about a vertex $v$ of $L'$, and whose endpoints are at distance $\pi/2$ from $v$. We claim that there is a length non-increasing homotopy of $\delta$ rel endpoints to a curve $\delta'$ that goes through $v$.

Assuming the claim, we finish the proof. Apply the claim to each open ball along the length of $\gamma$. Chaining together the homotopies gives a length non-increasing homotopy of $\gamma$ to a curve $\gamma'$. The new $\gamma'$ is locally geodesic and non-shrinkable and has length less than $2\pi$. Consider vertices $v$ and $v'$ of $L$ such that $\gamma'$ contains a point $x \in B_{\pi/2}v \cap B_{\pi/2}v'$. Then there is an arc of $\gamma'$ containing $x$ that goes through $v$ and $v'$. But $x$, $v$, and $v'$ are contained in a common simplex, so the unique locally geodesic path from $v$ to $v'$ is the edge between them.

This shows that $\gamma'$ is contained in $L'(1)$. But $\gamma'$ is a loop of length less than $2\pi$, while edges of $L'$ have length at least $\pi/2$, so $\gamma'$ consists of only 3 edges. Since $L'$ is a metric flag complex, a triangle in $L'^{(1)}$ of length less than $2\pi$ is the boundary of a simplex of $L'^{(1)}$. This gives a contradiction, since the boundary of a 2–simplex is shrinkable, but $\gamma'$ is not.

It remains to prove the claim. Suppose that $\delta$ is an arc of $\gamma$ whose interior is contained in $B_{\pi/2}v$, and whose endpoints are at distance $\pi/2$ from $v$. Consider the surface $S$ that is the union of geodesic segments of length $\pi/2$ starting at $v$ and passing through a point of $\delta$. The surface $S$ is a finite sequence of convex spherical 2–simplices $S_1, \ldots, S_n$ that all have common vertex $v$, and such that $S_i$ and $S_{i+1}$ share an edge of length $\pi/2$. Choose an isometry from $S_1$ to a simplex in $\mathbb{S}^2$ sending $v$ to the North pole. Then there is a unique isometry sending $S_2$ into $S^2$ that agrees with the previous map on $S_1 \cap S_2$ and so that $S_1$ and $S_2$ have disjoint interiors. Continuing in this way, we produce a local isometry from $S$ into the Northern hemisphere of $S^2$. See Figure 6, where $\delta$ is violet and $S$ is the yellow/orange ruled surface.

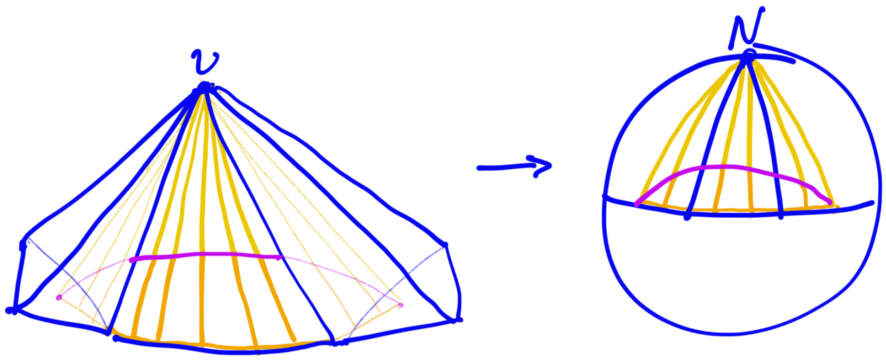



FIGURE 6. Developing an interior arc.

The image of $\delta$ under this map is a local geodesic whose endpoints are on the equator and whose interior is strictly contained in the Northern hemisphere. This implies that the endpoints of this path are antipodes, since otherwise there is a unique local geodesic between them: the equatorial geodesic. If they are antipodal, then there is a whole circle's worth of geodesics between them, and in particular there is a constant length homotopy rel endpoints to a geodesic through the North pole. Pulling this homotopy back to $S$ gives the desired constant length homotopy of $\delta$ to a path through $v$. $\square$

2.4.3. *The Davis complex is CAT(0).*

THEOREM 2.4.16. *Let* $(W, S)$ *be a Coxeter system. Its Davis complex* $\Sigma = \Sigma(W, S)$ *is CAT(0).*

PROOF. $\Sigma$ is a simply connected $\mathcal{M}_0$–polyhedral complex with finitely many isometry types of cells, and the link of every vertex of $\Sigma$ is $L =$

$L(W, S)$, the nerve of $(W, S)$. By Theorem 1.5.5, $\Sigma$ is CAT(0) if and only if it satisfies that link condition, which in this case means that $L$ is CAT(1). This follows from Moussong's Lemma, once we verify that $L$ has edges of length at least $\pi/2$ and is metric flag.

For $s, t \in S$, let $m_{st}$ be the order of $st$.

An edge in $L$ corresponds to a 2–element spherical subset $T = \{s, t\} \in \mathcal{S}$, and the length of this edge is the dihedral angle of $\Sigma_T$ between $\Sigma_s$ and $\Sigma_t$, which is $\pi - \pi/m_{st} \geqslant \pi/2$. So edges in $L$ have length at least $\pi/2$.

Now suppose $T \subset S$ such that the set of vertices $\{\Sigma_t \mid t \in T\}$ in $L$ are pairwise connected by edges. For $s, t \in T$, let $c_{st} = \cos \ell_{st} = -\cos \pi/m_{st}$, so that $C_T = (c_{st})$ is the cosine matrix for $(W_T, T)$.

$$\begin{aligned} &\text{There exists a spherical simplex with edge lengths } \ell_{st} \\ &\overset{2.4.10}{\iff} C_T \text{ is positive definite} \\ &\overset{3.0.7}{\iff} T \in \mathcal{S} \\ &\iff \Sigma_T \text{ is a cell of } \Sigma \\ &\iff \{\Sigma_t \mid t \in T\} \text{ spans a cell in } L \\ &\iff \{\Sigma_t \mid t \in T\} \text{ spans a simplex in } L \end{aligned}$$

This shows that $L$ is a metric flag complex. □

THEOREM 2.4.17. *Let $(W, S)$ be a Coxeter system. Every finite subgroup of $W$ is conjugate into a spherical special subgroup.*

PROOF. $W$ acts geometrically on its Davis complex $\Sigma$, which is CAT(0). A finite subgroup $G$ of $W$ fixes a point $x$ of $\Sigma$, by Corollary 1.3.8. If $x$ is vertex let $\sigma := x$. Otherwise let $\sigma$ be the cell containing $x$ in its interior. The action of $W$ on $\Sigma$ is combinatorial, so $G$ fixes $\sigma$. Cell stabilizers are conjugates of spherical special subgroups. □

EXERCISE 2.4.18. Show the topological walls of Exercise 2.3.8 are convex in $\Sigma$. Recall the dual walls $\Omega_s$ are convex in the combinatorial metric by Corollary 1.2.25 of Chapter 2.

## 3. Classification of virtually solvable subgroups

We have shown that Coxeter groups act geometrically on CAT(0) spaces, their Davis complexes, so the general results of Section 1.4 say that every virtually solvable subgroup of a Coxeter group $W$ is finitely generated, virtually Abelian, and undistorted. In this section we mention a finer result that classifies such groups in terms of the Coxeter system $(W, S)$.

DEFINITION 3.0.1. An *affine* subset $T \subset S$ is one such that $(W_T, T)$ is a Euclidean reflection group. Equivalently, if $\Gamma$ is the Coxeter graph of $(W, S)$, then $\Gamma_T$ is a disjoint union of graphs from Table 2.4.

EXERCISE 3.0.2. Show that if $T' \subset T \subset S$ and $T$ is affine then $\Gamma_{T'}$ is a disjoint union of spherical and affine pieces. If, in addition, $\Gamma_T$ is connected, show that $T'$ is spherical.

If $T$ is affine and $W_T$ acts geometrically on $\mathbb{E}^n$ then $W_T$ contains a finite index subgroup $\mathbb{Z}^n$ consisting of translations.

On the other hand, CAT(0) groups do not contain infinite torsion subgroups, by Chapter 3 Theorem 1.0.1, so for every non-spherical subset $T \subset S$, $W_T$ contains a $\mathbb{Z}$ subgroup.

Theorem 3.0.4 says combinations of these are essentially the only sources of free Abelian subgroups.

DEFINITION 3.0.3. Let $(W, S)$ be a Coxeter system with Coxeter graph $\Gamma$. Let $T \subset S$. Suppose that $\Gamma_T = \coprod_{i=1}^{m} \Gamma_{T_i}$ is the decomposition of $\Gamma_T$ into its connected components. For $i = 1 \dots m$ we define a free Abelian subgroup $A_i$ of $W_{T_i}$:

- If $T_i$ is spherical let $A_i = \{1\}$.
- If $T_i$ is affine let $A_i$ be the translation subgroup of $W_{T_i}$.
- If $T_i$ is neither affine nor spherical let $A_i$ be any $\mathbb{Z}$ subgroup of $W_{T_i}$.

The free Abelian subgroup $A := \prod_{i=1}^{m} A_i$ is a *standard free Abelian subgroup.*

By definition, $\Gamma_T$ is a full subgraph, so not only are the $\Gamma_{T_i}$ disjoint from each other in $\Gamma$, they are not even adjacent to one another, since if there is an edge of $\Gamma$ between vertices of $\Gamma_{T_i}, \Gamma_{T_j} \subset \Gamma_T$ then the edge belongs to $\Gamma_T$, so $\Gamma_{T_i}$ and $\Gamma_{T_j}$ are not connected components of $\Gamma_T$. Of course, the phrasing this way is exactly so that the subgroups $W_{T_i}$ commute with one another.

THEOREM 3.0.4 (Krammer [**19**, Theorem 6.8.2]). *Let $(W, S)$ be a Coxeter system. Up to conjugation, every free Abelian subgroup of $W$ has a finite index subgroup that is a subgroup of a standard free Abelian subgroup.*

Combined with the Solvable Subgroup Theorem, this gives:

COROLLARY 3.0.5. *Let $(W, S)$ be a Coxeter system. Up to conjugation, every virtually solvable subgroup of $W$ has a finite index subgroup that is a subgroup of a standard free Abelian subgroup.*

EXAMPLE 3.0.6. Consider the graph $\Gamma$ of Figure 7. We claim that the maximal rank of a free Abelian subgroup of $W$ is 7. First, find affine sets by finding full subgraphs of $\Gamma$ that match examples from Table 2.4.

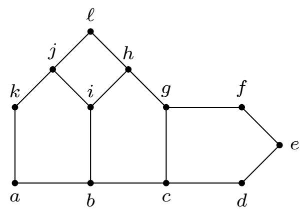



FIGURE 7. The maximal rank free Abelian subgroup is $\mathbb{Z}^7$.

None of graphs with labels appears in $\Gamma$. There are no full subgraphs of types $\tilde{E}_6$, $\tilde{E}_7$, or $\tilde{E}_8$. There are full copies of $\tilde{D}_5$ consisting of $\{i, h, \ell, g, c, f\}$ and $\{a, b, c, d, i, g\}$. There is an $\tilde{A}_3$, the square, 3 copies of $\tilde{A}_4$, the 3 pentagons, and an $\tilde{A}_7$ given by $\{a, b, c, g, h, \ell, j, k\}$.

Note that perimeter is not a full subgraph, as $c$ and $g$ are adjacent in $\Gamma$ but not consecutive on the perimeter.

Each of the subgraphs mentioned above contributes a maximal standard free Abelian subgroup, since in each case the complement of the 1–neighborhood of the defining subgraph is spherical. Thus, the maximum rank free Abelian subgroup is the standard $\mathbb{Z}^7$ coming from $\tilde{A}_7$.

EXERCISE 3.0.7. Show that if $G$ has a finite index subgroup $A \cong \mathbb{Z}^r$ and $H < G$ then $H$ has a finite index subgroup isomorphic to $\mathbb{Z}^q$ for some $q \leqslant r$, with $q = r$ if and only if $[G : H] < \infty$.

EXERCISE 3.0.8. Show that a virtually Abelian group has a finite index normal Abelian subgroup.

EXERCISE 3.0.9. Give an example of a torsion free, nonAbelian, virtually Abelian group. Conclude, despite the previous exercise, that not every virtually Abelian group can be expressed as a semi-direct product $\mathbb{Z}^r \rtimes F$ for some $r$ and some finite $F$.

A key fact used in Theorem 3.0.4 is:

LEMMA 3.0.10 ([**19**, Corollary 6.3.10][**11**, Lemma 12.7.1]). *Suppose* $(W, S)$ *is irreducible and not spherical or affine. If* $w \in W$ *is an element that is not conjugate into any proper special subgroup then* $\langle w \rangle \cong \mathbb{Z}$ *has finite index in its centralizer.*

For example, a Coxeter element (a power of all the generators, recall Definition 1.3.11) is not conjugate into any proper special subgroup. Theorem 1.3.12 implies this is true for powers too, in the irreducible non-spherical case:

COROLLARY 3.0.11. *If $(W, S)$ is irreducible and not spherical or affine then every nontrivial power of every Coxeter element generates an infinite cyclic subgroup with finite index in its centralizer.*

COROLLARY 3.0.12. *If $(W, S)$ be an irreducible Coxeter system that is neither spherical nor affine then $W$ is not virtually Abelian.*

PROOF. Suppose that $W$ has a finite index subgroup $A \cong \mathbb{Z}^r$. Let $w$ be a Coxeter element. Some non-trivial power $w^n$ is contained in $A$. Since $A$ is Abelian, $A$ is contained in the centralizer of $\langle w^n \rangle$, but by Corollary 3.0.11, that centralizer is virtually cyclic, so $r = 1$.

By [**17**], every proper special subgroup of $W$ has infinite index. Since $W$ is virtually cyclic, infinite index is equivalent to finite, so $W$ is a geometric reflection group, by Theorem 2.2.4. The only virtually infinite cyclic geoemetric reflection group is $\mathcal{D}_\infty$, which is affine, contrary to hypothesis. □

COROLLARY 3.0.13. *Let $(W, S)$ be a Coxeter system with Coxeter graph $\Gamma$. Then $W$ is virtually Abelian if and only if $\Gamma$ is a disjoint union of spherical and affine subgraphs. Furthermore, $W$ is virtually $\mathbb{Z}^r$ if and only if the ranks of the translation subgroups of the affine factors of $\Gamma$ sum to $r$.*

PROOF. Subgroups of virtually Abelian groups are virtually Abelian, so, by Corollary 3.0.12, a virtually Abelian Coxeter group can have only spherical and affine irreducible components.

The product of a group that is virtually $\mathbb{Z}^m$ with a group that is virtually $\mathbb{Z}^n$ is virtually $\mathbb{Z}^{m+n}$. □

## 4. When is the Davis complex CAT(-1)?

Now let's be greedy. We have shown that the Davis complex of a Coxeter system always admits a CAT(0) metric, but maybe we can get more? Does it admit a CAT(-1) metric? The answer is 'not always', but it turns out that the 'obvious' obstruction is the only one. It is convenient to simultaneously consider another generalization of negative curvature.

### 4.1. Gromov hyperbolicity.

DEFINITION 4.1.1. A geodesic metric space $X$ is *(Gromov) hyperbolic* if there exists a $\delta$ such that every geodesic triangle in $X$ is $\delta$–thin, in the sense that each side is contained in the $\delta$–neighborhood of the other two.

A space is *$\delta$–hyperbolic* is it hyperbolic with the given $\delta$ as a sufficient thinness bound.

EXERCISE 4.1.2. Show that $\mathbb{H}^2$ is (Gromov) hyperbolic. Conclude that any CAT(-1) space is too.

The difference between being hyperbolic and being CAT(-1) is that the CAT(-1) condition applies at all scales, while hyperbolicity is coarser: we essentially don't care what happens at scales smaller than $\delta$. For example, $\mathbb{R}$ is CAT(-1), hence hyperbolic. If we instead consider $\mathbb{R} \times [0,1]$, this is CAT(0) but not CAT(-1); it contains small Euclidean triangles. It is hyperbolic though; take $\delta > 1$. Then the $\delta$–neighborhood of a point $(r,s)$ contains $\{r\} \times [0,1]$. If $(r_0,s_0)$, $(r_1,s_1)$, and $(r_2,s_2)$ are three points then the $\delta$–neighborhood of any two sides contains $[\min r_i, \max r_i] \times [0,1]$, which contains the third side.

You can make a similar argument with $\mathbb{R} \times \mathbb{S}^1$ for an example of a space that is hyperbolic but not CAT(0).

EXAMPLE 4.1.3. $\mathbb{E}^2$ is not hyperbolic. For an equilateral triangle of side length $s$, the distance from the midpoint of one side to either of the other two is $s\sqrt{3}/4$. There are arbitrarily large equilateral triangles, so there is no uniform $\delta$ for which all triangles are $\delta$–thin.

THEOREM 4.1.4 (eg [**6**, Theorem III.H.1.9]). *Hyperbolicity is invariant under quasiisometry.*

This robustness is a main attraction of the hyperbolic definition. For instance, compare the following two definitions:

DEFINITION 4.1.5. A group is a *CAT(k) group* if it acts geometrically on some CAT(k) space.

DEFINITION 4.1.6. A group is a *hyperbolic group* if some, equivalently, every, geodesic metric space on which it acts geometrically is hyperbolic.

The latter makes sense because of Theorem 4.1.4: any two spaces on which the group acts geometrically are quasiisometric, so either both are hyperbolic or both are non-hyperbolic. There is no need to worry about finding the right space to witness hyperbolicity. Conversely, if $G \curvearrowright X$ geometrically and $X$ is not CAT(k) then we cannot conclude that $G$ is not a CAT(k) group; $X$ just is not the space to certify that property of $G$.

As mentioned above, it is easy to construct examples of spaces that are hyperbolic but not CAT(0). It is not easy to construct examples of groups that are hyperbolic but not CAT(0). In fact, it is a long-open question whether every hyperbolic group admits a geometric action on some CAT(0) space.

EXAMPLE 4.1.7.
- Finite groups are hyperbolic: choose a Cayley graph and take $\delta$ larger than its diameter.
- Groups that act geometrically on a CAT(-1) space are hyperbolic.

- If $G = H \times F$ where $H$ is hyperbolic and $F$ is finite then $G$ is hyperbolic. More generally, if $G$ has a finite index hyperbolic subgroup $H$ then $G$ is also hyperbolic, since $G$ and $H$ are quasiisometric.

A fundamental result about hyperbolic groups is as follows:

THEOREM 4.1.8. *An infinite order element of a hyperbolic group is undistorted and has virtually cyclic centralizer.*

COROLLARY 4.1.9. *A hyperbolic group cannot contain a Baumslag-Solitar group as a subgroup.*

Recall that the notion of distortion was defined in Definition 1.4.6. Baumslag-Solitar groups were defined nearby, and in Corollary 1.4.9 it was shown that a group acting geometrically on a CAT(0) space contains no nonunimodular Baumslag-Solitar subgroup, because of distortion.

EXERCISE 4.1.10. Given that the generators $a$ and $b$ of a Baumslag-Solitar group are infinite order and not commensurable, show that if the group is unimodular then there is a power of $a$ whose centralizer is not virtually cyclic.

**4.2. Moussong's Theorem.** Our detour into hyperbolicity tells us that for a Coxeter system $(W, S)$, $W$ cannot act geometrically on a CAT(-1) space if $\mathbb{Z}^2 < W$. This is the 'obvious' obstruction to the existence of a CAT(-1) metric on $\Sigma$. Moussong's Theorem says it is the only obstruction:

THEOREM 4.2.1 (Moussong's Theorem [**11**, Corollary 12.6.3]). *Let $(W, S)$ be a Coxeter system with Coxeter graph $\Gamma$. The following are equivalent:*

(1) *The Davis complex $\Sigma = \Sigma(W, S)$ admits a CAT(-1) metric.*
(2) *$W$ is hyperbolic.*
(3) *$W$ does not contain $\mathbb{Z}^2$.*
(4) *$S$ does not contain a subset $T$ of either of the following types:*
    - *$T$ is irreducible affine with $|T| > 2$.*
    - *$\Gamma_T = \Gamma_{T'} \coprod \Gamma_{T''}$ with $T'$ and $T''$ both nonspherical.*

Condition (4) is known as *Moussong's condition.*

PROOF SKETCH. CAT(-1) spaces are hyperbolic, and hyperbolicity is invariant under quasiisometry, by Theorem 4.1.4. $W$ acts geometrically on $\Sigma$, by Corollary 2.2.3, so $W$ and $\Sigma$ are quasiisometric by Theorem 0.0.5. Thus, (1) $\implies$ (2).

(2) $\implies$ (3) is a special case of Corollary 4.1.9, since $\mathbb{Z}^2 \cong BS(1, 1)$.

By Theorem 3.0.4, $\mathbb{Z}^2 < W$ if and only if $W$ contains a standard free Abelian subgroup $\mathbb{Z}^r$ for $r \geqslant 2$. A standard free Abelian subgroup arises either as a subgroup of an irreducible affine subgroup, or as a subgroup of a product of infinite special subgroups. These are the two possibilities described by (4), so (3) $\iff$ (4).

The main work is to show (4) $\implies$ (1). The technical result that we will skip is that Moussong's condition implies that the nerve $L = L(W, S)$ is *extra large*, which implies that it does not contain geodesic loops of length less than *or equal to* $2\pi$. The idea is to repeat the CAT(0) construction for $\Sigma$, but using CAT(-1) polytopes.

Recall that if $T \subset S$ is spherical then $W_T$ admits an orthogonal action on $\mathbb{R}^n$, and we defined $\Sigma_T(\bar{d}) = \Sigma_T^{\mathbb{E}}(\bar{d})$ to be the convex hull, in $\mathbb{E}^{|T|}$, of the $W_T$ orbit of a point defined by $\bar{d}$. However, the stabilizer of the origin in the Poincaré ball model of $\mathbb{H}^{|T|}$ is also the orthogonal group, so we could do the same thing in $\mathbb{H}^{|T|}$: specify a point $x$ via $\bar{d}$ using hyperbolic distances, act on it by the finite group $W_T$, and take the hyperbolic convex hull $\Sigma_T^{\mathbb{H}}(\bar{d})$ of the resulting finite set. The combinatorial type of $\Sigma_T^{\mathbb{H}}(\bar{d})$ does not depend on the choice of $\bar{d}$, and if $\bar{d}$ if very close to $\bar{0}$ then $\Sigma_T^{\mathbb{H}}(\bar{d})$ will be metrically close to $\Sigma_T^{\mathbb{E}}(\bar{d})$. The key difference is that the dihedral angles of $\Sigma_T^{\mathbb{E}}(\bar{d})$ do not depend on $\bar{d}$, but the dihedral angles of $\Sigma_T^{\mathbb{H}}(\bar{d})$ do: they are close to those of $\Sigma_T^{\mathbb{E}}(\bar{d})$ when $\bar{d}$ is close to $\bar{0}$, but they go to 0 as all coordinates of $\bar{d}$ go to $\infty$.

We don't like small dihedral angles; to go from $\Sigma$ being an $\mathcal{M}_{-1}$–polyhedral complex to it being CAT(-1) we need to prove a link condition, so no short loops in $L$. The dihedral angles of the cells become the edge length in the link, so small dihedral angles may give short loops in the link. The obvious thing to do is to take $\bar{d}$ very close to $\bar{0}$ to try to prevent this, since we know the link condition is satisfied if we use the Euclidean metric. The problem is that if the link in the Euclidean metric has a loop of length exactly $2\pi$ then the link condition is satisfied, but in the hyperbolic metric the edges in the link will be slightly shorter, so the corresponding loop will have length $< 2\pi$ and the link condition fails. The 'extra large' property prevents this; there are no loops of length exactly $2\pi$, which, since $L$ is finite, means there is a gap between $2\pi$ and the length of the shortest loop. Then the claim is that for $\bar{d}$ close enough to $\bar{0}$ the link with respect to the hyperbolic metric is close enough to the Euclidean one that there are still no loops of length $< 2\pi$. $\square$

COROLLARY 4.2.2. *Suppose* $(W, S)$ *has no rank 3 spherical subgroups, so that* $L = L(W, S) = L^{(1)} = \Upsilon$, *the presentation graph of* $(W, S)$. *The following are equivalent:*

(1) $W$ *is hyperbolic.*
(2) $L$ *is extra large.*

(3) $\Upsilon$ *does not contain a full subgraph of the following forms:*
   - *A triangle labelled $a$, $b$, $c$ with $\frac{1}{a} + \frac{1}{b} + \frac{1}{c} = 1$.*
   - *An unlabelled square.*

PROOF. $L$ and $\Upsilon$ are combinatorially the same, by hypothesis. An edge labelled $m$ in $\Upsilon$ (with $m = 2$ if the edge is unlabelled), corresponds to an edge of length $\pi - \pi/m$ in $L$. Since all of the labels are at least 2, $L$ contains a loop of length at most $2\pi$ only in the two cases of an unlabelled square and a triangle whose reciprocal labels sum to at least 1. However, if they sum to strictly greater than 1 this would give a rank 3 spherical subgroup, so only the equality case needs to be ruled out. □

COROLLARY 4.2.3. *Let $(W, S)$ be a right-angled Coxeter group. This means its presentation graph $\Upsilon$ is unlabelled, or, equivalently, its Coxeter graph $\Gamma$ has all edges labelled $\infty$. Then $W$ is hyperbolic if and only if $\Upsilon$ does not contain a full subgraph that is a square.*

PROOF. Consult Table 2.4 to see that the only irreducible affine graph with all edges labelled $\infty$ is a single edge, so the first part of Moussong's condition cannot occur.

The spherical subsets $T$ of $S$ are those such that $\Gamma_T$ contains no edges. Suppose $\Gamma_T = \Gamma_{T'} \coprod \Gamma_{T''}$ with $T'$ and $T''$ nonspherical. Then $\Gamma_{T'}$ and $\Gamma_{T''}$ each contain an edge, and it suffices to assume that single edge is the whole graph, so that $\Gamma_T$ is a full subgraph that is a disjoint union of two edges. But then $\Upsilon_T$ is a square. Thus, if $\Upsilon$ has no squares then Moussong's condition is satisfied. □

COROLLARY 4.2.4. *Let $\Upsilon_0$ be any finite graph. Subdivide edges until there are no loops, bigons, triangles, or squares. Call the resulting graph $\Upsilon$. Then $\Upsilon$ is the presentation graph of a finitely generated 2–dimensional hyperbolic right-angled Coxeter group.*

## 5. Free and surface subgroups

Given a Coxeter system $(W, S)$, when does $W$ contain a free subgroup? a subgroup isomorphic to the fundamental group of a closed surface?

### 5.1. Free subgroups.

PROPOSITION 5.1.1. *If $(W, S)$ is a Coxeter system then the following are equivalent:*

(1) *$F_2$ is not a subgroup of $W$.*
(2) *$W$ is virtually Abelian.*

(3) *The Coxeter graph $\Gamma$ is a disjoint union of spherical and affine subgraphs; that is, of graphs appearing in Table 2.2 and Table 2.4.*

PROOF. By Corollary 4.0.4 Coxeter groups are linear, so by Theorem 1.0.1 they satisfy the Tits Alternative: every finitely generated subgroup, in particular $W$ itself, is either virtually solvable or contains a free subgroup of rank 2. By Corollary 3.0.5 if $W$ is virtually solvable then it is virtually a standard free Abelian subgroup. Thus, (1) $\iff$ (2).

(2) $\iff$ (3) is Corollary 3.0.12. □

## 5.2. Splittings.

### 5.2.1. *Background.*

DEFINITION 5.2.1. An *amalgamated product* of groups $A_1, \dots, A_n$ over a subgroup $C$ is the group obtained by gluing together the $A_i$ along a copy of a common subgroup $C$. Given $\phi_i \colon C \hookrightarrow A_i$ the amalgamated product $\ast_C A_i$ is obtained from the free product of the $A_i$ by taking the quotient by the equivalence relation generated by the condition:

$$\forall c \in C, \forall 1 \leqslant i, j \leqslant n, \ \phi_i(c) = \phi_j(c)$$

When $C = \{1\}$ the amalgamated product over $C$ is simply the free product.

DEFINITION 5.2.2. An *HNN extension* of a group $A$ over a subgroup $C$ is the group obtained by gluing $A$ to itself along two, not necessarily distinct, copies of $C$. Given $\phi \colon C \hookrightarrow A$ and $\psi \colon C \hookrightarrow A$:

$$A{\ast_C} := A \ast \langle t \rangle / \sim, \qquad \text{where } \forall c \in C, \ t^{-1}\phi(c)t = \psi(c)$$

The new generator $t$ in the resulting group is an infinite order element known as the *stable letter*.

DEFINITION 5.2.3. An *amalgam* is a group that is either an amalgamated product or HNN extension. A *splitting* of a group is an expression of the group as an amalgam.

The study of group amalgams and their corresponding actions on trees is called 'Bass-Serre Theory'.

In Algebraic Topology, van Kampen's Theorem says that when sufficiently nice spaces are glued together along subspaces, the fundamental group of the result is an amalgam of the fundamental groups of the spaces you started with.

DEFINITION 5.2.4. The *set of ends* Ends(X) of a topological space $X$ is the inverse limit $\varprojlim_{K \subset X \text{ compact}} \{\text{unbounded components of } X - K\}$.

For proper spaces we might as well use closed balls in place of compact sets, so points in Ends(X) correspond to chains of nested set that are unbounded complementary components of increasingly large balls about some basepoint. Two such chains are different ends if eventually their constituents are separated by some ball.

EXAMPLE 5.2.5. A bounded set has 0 ends, because balls do not have any unbounded components.

In $\mathbb{R}^2$, the complement of any ball consists of exactly 1 unbounded subset, and no two such sets are separated by a ball, so $\mathbb{R}^2$ has one end.

Consider $\mathbb{R}$. Each ball about the origin has two unbounded components $(-\infty, r)$ and $(r, \infty)$. Two such sets are separated by a ball if and only if one of them is unbounded in the negative direction and the other is unbounded in the positive direction. Therefore, Ends($\mathbb{R}$) has two points, which we identify with $\pm\infty$.

The ends of a space should be thought of as the topologically distinct ways of going off to infinity. It is a fact that a quasiisometry induces a bijection between sets of ends, so it makes sense, for a finitely generated group $G$, to define Ends($G$) as the set of ends of any Cayley graph of $G$.

The *number of ends* of a space or group is the cardinality of its set of ends. For spaces this number can be a non-negative integer or infinity.

EXERCISE 5.2.6. Let $G$ be a finitely generated group with more than two ends. Show $G$ has infinitely many ends.

THEOREM 5.2.7. *If $G$ is a finitely generated group, there are the following possibilities:*

- *$G$ has 0 ends, which occurs if and only if $G$ is finite.*
- *$G$ has 1 end.*
- *$G$ has 2 ends, which occurs if and only if $G$ is virtually $\mathbb{Z}$.*
- *$G$ has infinitely many ends.*

THEOREM 5.2.8 (Stallings's Theorem). *A finitely generated group splits as an amalgam over a finite subgroup if and only if it has more than one end.*

One consequence of Stallings's Theorem and some basic Bass-Serre Theory is:

LEMMA 5.2.9. *If $G$ splits as an amalgam over a finite subgroup and $H < G$ is a 1–ended subgroup then $H$ is conjugate into one of the factors of the splitting.*

On might try to split inductively: If $G$ has more than one end, split it as an amalgam over a finite subgroup, and then look at the factor groups. If they have more than one end, split over a finite subgroup, etc. This process potentially goes on forever, but if it terminates then the group is called *accessible*, and the final stage expresses $G$ as an iterated amalgam over finite groups where the factor groups that are left at the end are either 0 ended (finite) or 1 ended.

THEOREM 5.2.10 (Dunwoody's Theorem). *Finitely presented groups are accessible.*

For an accessible group, the terminal splitting over finite subgroups, resulting in finite or 1–ended factors, is called the *Dunwoody-Stallings decomposition.*

THEOREM 5.2.11. *An infinite, finitely presented group is virtually free if and only if its Dunwoody-Stallings decomposition consists of finite groups; that is, it has no 1–ended factors.*

Using Bass-Serre Theory, we can further refine which groups occur as virtually free groups:

THEOREM 5.2.12. *Let $G$ be a finitely presented group with no 1–ended factor in its Dunwoody-Stallings decomposition. Then one of the following holds:*

(1) *$G$ is finite $\iff$ $G$ has 0 ends $\iff$ $G$ is virtually trivial.*
(2) *$G$ has infinitely many ends $\iff$ $G$ is virtually $F$ for $F$ a non-Abelian free group.*
(3) *$G$ has 2 ends $\iff$ $G$ is virtually $\mathbb{Z}$ $\iff$ One of the following is true:*
    (a) *$G \cong \mathbb{Z} \rtimes H$ for $H$ finite.*
    (b) *$G \cong A *_C B$ for $A$ and $B$ finite with $[A : C] = [B : C] = 2$.*

5.2.2. *Splittings of Coxeter groups.* Recall that connected components of the presentation graph $\Upsilon$ of a Coxeter system correspond to free factors in a free splitting of $W$. We could read this fact off directly from the presentation, according to Exercise 2.0.3. Let us generalize:

PROPOSITION 5.2.13. *Let $(W, S)$ be a Coxeter system with nerve $L$, Suppose there is a spherical set $T_0 \subset \mathcal{S}$ such that the simplex $\sigma_{T_0} \subset L$ corresponding to $T_0$ disconnects $L$. Let $T_i$ be the vertex set of the $i$–th component of $L - \sigma_{T_0}$. Then $W$ splits as an amalgamated product $\circledast_{W_{T_0}} W_{T_i \cup T_0}$ over the finite group $W_{T_0}$.*

The set $L - \sigma_T$ is called a *punctured nerve.*

Proof. Let $\phi_i \colon W_{T_0} \hookrightarrow W_{T_i \cup T_0}$ be the injection determined by the inclusion $T_0 \subset T_i$. The amalgamated product $*_{W_{T_0}} W_{T_i \cup T_0}$ has a defining presentation with generators the disjoint union of the $T_i \cup T_0$. (That is, $T_i$ plus a separate copy of $T_0$ for each $i$.) The relations are the relations of each $W_{T_i \cup T_0}$, together with additional relations $\phi_i(w) = \phi_j(w)$ for each $w \in W_T$. The additional relations simply say to identify the distinct copies of $T_0$ in this presentation, so this presentation Tietze reduces to the original Coxeter presentation of $W$. □

In fact, this is the only way splittings over finite groups occur:

Theorem 5.2.14 ([**11**, Theorem 8.7.2]). *Let $(W, S)$ be a Coxeter system. $W$ splits over a finite subgroup, or, equivalently, $W$ has more than one end, if and only if there exists a spherical $T \subset S$ such that $L - \sigma_T$ is not connected. (Here, $T = \varnothing$ is possible, in which case take $\sigma_T = \varnothing$.)*

This gives an algorithm for computing the Dunwoody-Stallings decomposition: Check if $W$ is finite. If not, check if there exists spherical $T_0$ such that $L - \sigma_{T_0}$ is not connected. If not, $W$ has 1 end. If so, let $T_i$, $i = 1, \dots, n$ be the vertex sets of the components, so that $*_{W_{T_0}} W_{T_i \cup T_0}$. For each $W_{T_i \cup T_0}$, repeat.

Theorem 5.2.15. *Let $(W, S)$ be a Coxeter system. $W$ is virtually free if and only if the above procedure does not uncover any 1–ended pieces.*

The special case of being virtually infinite cyclic can be recognized more explicitly:

Theorem 5.2.16. *Let $(W, S)$ be a Coxeter system with Coxeter graph $\Gamma$ and presentation graph $\Upsilon$. $W$ is 2–ended if and only there is a spherical $T_0$ such that the following equivalent conditions hold:*

- $\Gamma = \Gamma_{T_0} \sqcup \overset{s_1}{\bullet}\overset{\infty}{\text{———}}\overset{s_2}{\bullet}$
- *$\Upsilon$ is the suspension of $\Upsilon_{T_0}$, that is $S = T_0 \cup \{s_1, s_2\}$, and for each $i$ the vertex $s_i$ is connected to every vertex in $\Upsilon_{T_0}$, but there is no edge between $s_1$ and $s_2$.*

*Thus, $W$ is a product of $\mathcal{D}_\infty = W_{\{s_1, s_2\}}$ and the finite group $W_{T_0}$.*

Corollary 5.2.17. *The only irreducible virtually $\mathbb{Z}$ Coxeter system is:*
$\bullet\overset{\infty}{\text{————}}\bullet$

Proof. $\mathcal{D}_\infty$ has a $\mathbb{Z}$ of index 2, so the product of $\mathcal{D}_\infty$ with a finite group has a finite index $\mathbb{Z}$ subgroup. This is equivalent to being 2–ended.

Conversely, suppose $W$ is 2–ended. Item item 3a of Theorem 5.2.12 cannot occur because $W$ is generated by torsion elements, so it does not surject onto $\mathbb{Z}$. Thus item 3b of Theorem 5.2.12 together with Theorem 5.2.14 implies that $W = W_{T_1 \cup T_0} *_{W_{T_0}} W_{T_2 \cup T_0}$ with $T_0 \in \mathcal{S}$, no edges between $T_1$ and $T_2$, and $[W_{T_i \cup T_0} : W_{T_0}] = 2$ for $i \in \{1, 2\}$.

Consider the longest element $\Delta_{T_0}$ of $W_{T_0}$. For $i \in \{1, 2\}$, let $s_i \in T_i$. Then $s_i \Delta_{T_0}$ is the longest element of $s_i W_{T_0}$. If $[W_{T_i \cup T_0} : W_{T_0}] = 2$ then $W_{T_i \cup T_0} = W_{T_0} \sqcup s_i W_{T_0}$, so $T_i = \{s_i\}$ and $s_i \Delta_{T_0}$ is the longest element of $W_{T_i \cup T_0}$. By Lemma 4.0.11 $s_i$ commutes with all $t \in T_0$. □

THEOREM 5.2.18. *Let $(W, S)$ be a Coxeter system. $W$ is virtually a nonAbelian free group if and only if it is infinite, is virtually free, as in Theorem 5.2.15, and is not virtually $\mathbb{Z}$, as in Theorem 5.2.16.*

**5.3. Surface subgroups.** A well-known question of Gromov asks if every 1–ended hyperbolic group contains a surface subgroup, meaning, the a subgroup isomorphic to the fundamental group of a closed surface.

Let us explain the ‘1–ended’ hypothesis. Recall that a finitely presented group $G$ either has some 1–ended factor in its Dunwoody-Stallings decomposition, or all groups in the decomposition are finite and $G$ is virtually free. In the latter case, $G$ has no 1–ended subgroups, so it cannot contain a surface subgroup, since the plane is 1–ended. It is also true that a group is hyperbolic if and only if all of the 1–ended factors in its Dunwoody-Stallings decomposition are hyperbolic. Thus, there is no loss in restricting the question from hyperbolic non-(virtually free) groups to hyperbolic 1–ended groups.

THEOREM 5.3.1 (Gordon, Long, and Reid [**15**, Theorem 1.1]). *A Coxeter group contains a surface subgroup if and only if it is not virtually free.*

Note that by Moussong's Theorem Theorem 4.2.1 a Coxeter group is nonhyperbolic only if it contains $\mathbb{Z}^2$, which is the fundamental group of a torus, so it is the hyperbolic case that is interesting.

Before proving the theorem we have two technical conditions that will be used in an induction.

DEFINITION 5.3.2. Say that $(W, S)$ is *2–spherical* if every pair $\{s, t\} \subset S$ is spherical.

LEMMA 5.3.3 (cf [**15**, Theorem 2.3]). *If $(W, S)$ is 2–spherical then either it is spherical or it contains a surface subgroup.*

PROOF. Suppose $W$ is infinite, and let $T \subset$ be a minimal nonspherical subset. Since $W_T < W$, it suffices to find a surface subgroup in $W_T$, so we may assume $T = S$. Thus, every proper special subgroup of $W$ is spherical.

By Theorem 2.2.4, $W$ is a simplicial geometric reflection group. By hypothesis, $W$ is infinite, so it is not a spherical reflection group. Also, since it is 2–spherical it is not virtually cyclic. If $W$ contains $\mathbb{Z}^2$ then we are done, $W$ contains a subgroup isomorphic to the fundamental group of a torus, so we can exclude the Euclidean reflection groups. Thus, it suffices to consider the case that $W$ is a simplicial hyperbolic reflection group.

In dimension 2 this means that $(W, S)$ is a hyperbolic triangle group. By Theorem 1.0.1, $W$ has a finite index torsion-free subgroup, $G$. The restriction of the action of $W$ on $\mathbb{H}^2$ to $G$ is still geometric, since $G$ has finite index, so $G \cong \pi_1(G \backslash \mathbb{H}^2)$ where $G \backslash \mathbb{H}^2$ a closed surface.

In dimension greater than 2 there are only finitely many simplicial hyperbolic reflection groups, as in Table 2.3. The discussion following [**15**, Theorem 2.3] shows that in each case one can take a 2–dimensional face of the simplicial fundamental domain and see that this face tiles a copy of $\mathbb{H}^2$ with a subgroup of $W$ acting cocompactly. Passing to a torsion-free finite index subgroup of this subgroup gives the desired surface group. $\square$

Definition 5.3.4. A *chordal graph* is a finite simplicial graph that belongs to the smallest class $\mathcal{C}$ of graphs that contain complete graphs and are closed under amalgamation over complete subgraphs. This means, if $A$ and $B$ are in $\mathcal{C}$ and $K$ is a complete graph that occurs as a subgraph of both $A$ and $B$, then the graph $A \cup_K B$ obtained from the disjoint union of $A$ and $B$ by identifying the two copies of $K$ is also in $\mathcal{C}$.

Theorem 5.3.5 (Dirac's chordal graph theorem). *A finite connected simplicial graph is either chordal or it contains a full $n$–cycle for some $n \geqslant 4$.*

Proof of Theorem 5.3.1. Let $(W, S)$ be a Coxeter system with presentation graph $\Upsilon$.

Case 1: $\Upsilon$ contains a full $n$–cycle for $n \geqslant 4$. Let $T \subset S$ be the vertices such that $\Upsilon_T$ is the $n$–cycle. Then $\Upsilon_T$ is the presentation graph for $(W_T, T)$. If all of the edges of $\Upsilon_T$ are unlabelled (= labelled by 2's) then $W_T \cong \mathcal{C}_\infty \times \mathcal{C}_\infty$ has a finite index $\mathbb{Z}^2$, so $W$ contains a surface subgroup. Otherwise, if the edges of $\Upsilon_T$ are labelled $m_1, \ldots, m_n$, then $\sum_{i=1}^n 1/m_i < n/2 \leqslant n-2$, so, by Proposition 2.4.1 $(W_T, T)$ acts on $\mathbb{H}^2$ as a reflection group with fundamental domain an $n$–gon. Pass to a finite index torsion-free subgroup to get a surface subgroup of $W_T$, hence of $W$.

Case 2: $\Upsilon$ is a complete graph. In this case $(W, S)$ is 2–spherical, so it is either finite or contains a surface subgroup, by Lemma 5.3.3.

Case 3: $\Upsilon$ is not complete and does not contain an $n$–cycle for $n \geqslant 4$. By Theorem 5.3.5, $\Upsilon$ is chordal, but it is not complete, so it is a non-trivial amalgam: $\Upsilon = A \cup_K B$ where $A$ and $B$ are full subgraphs with $A \cap B = K$

complete. Let $T_0$ be the vertex set of $K$, and $T_1$ and $T_2$ the vertex sets of $A$ and $B$, respectively. Since $\Upsilon_{T_0} = K$ is complete, by the previous case either $W_{T_0}$ is finite or it contains a surface subgroup. If it contains a surface subgroup then so does $W$. If it is finite then $T_0$ is the vertex set of a simplex $\sigma_{T_0} \in L(W,S)$, and $L - \sigma_{T_0}$ is disconnected, into the subcomplex with vertex set $T_1 - T_0$ and the subcomplex with vertex set $T_2 - T_0$. This gives us a splitting of $W$ as an amalgamated product over a finite subgroup, $W = W_{T_1} *_{W_{T_0}} W_{T_1}$, by Proposition 5.2.13. By Lemma 5.2.9, if there is a surface subgroup in $W$ then it is conjugate into either $W_{T_1}$ or $W_{T_2}$.

It either $W_{T_1}$ or $W_{T_2}$ is infinite repeat the argument. The number of generators goes down with each step, so either at some point this process produces a subgroup that contains a surface group, or it terminates in finite groups, in which case we have expressed $W$ as an amalgam of finite groups, so $W$ is virtually free. $\square$

CHAPTER 6

# Right-angled Coxeter groups

In this chapter we specialize to right-angled Coxeter groups. This means that all of exponents in the Coxeter presentation $m_{st}$ are either 2 or infinity. Consequently, the Davis complex for such a Coxeter group can be chosen, by taking $\bar{d} = (1/2, 1/2, \ldots, 1/2)$, to be a cube complex.

What's the point? In CAT(0) cube complexes there is a naturally defined system of walls with properties very similar to the case of reflection systems (but without the hypothesis of a reflection group action). We have seen for reflection systems how the combinatorics of walls leads to geometric consequences; for example, an edge path being geodesic if and only if it crosses each wall at most once. It turns out that there is a duality between (discrete) wall spaces and CAT(0) cube complexes, so, in a sense, CAT(0) cube complexes are the natural way to combinatorialize all wall spaces, and the geometry is reflected by the *combinatorial metric* on the CAT(0) cube complex, that is, by the edge-length metric on the 1–skeleton. It must be stressed that the cube complex is not CAT(0) with respect to the combinatorial metric. One interpretation of what we're going to do is that for finite dimensional CAT(0) cube complexes there are strong connections between the CAT(0) metric, the combinatorial metric, and the wall metric. We will see an example, Example 1.4.1, where a fairly simple CAT(0) polygonal complex has an obvious wall structure, but the CAT(0) cube complex encoding that wall structure is much more complicated than the polygonal complex, so for non-cubical examples there is not the same close connection between the three metrics.

We will explore all of these topics in Section 1. Here is a motivating question. Suppose $G$ is a CAT(0) group, which, recalling Definition 4.1.5, means that $G$ acts geometrically on some CAT(0) space $X$. Let $H$ be a finitely generated subgroup of $G$. When is $H$ a CAT(0) group? This is an open question. It is not clear that if $H$ is a CAT(0) group then the defining geometric action on a CAT(0) space should be related to the action of $H$ on $X$, but that would be a natural place to start. Specifically, we would have a positive answer if $H$ acts geometrically on a CAT(0) subspace of $X$. Try the following: Let $x \in X$ and let $H.x$ be the $H$–orbit of $x$. The action of $H$ on $H.x$ is cocompact, but $H.x$ is not connected, so it is certainly not CAT(0). Enlarge it to make it CAT(0) by passing to the convex hull $\mathcal{H}(H.x)$, the smallest convex subset of $X$ containing $H.x$. Convex subsets of CAT(0) spaces are CAT(0), and $H$ fixes $H.x$, so it acts on $\mathcal{H}(H.x)$. Proper discontinuity is inherited from $G \curvearrowright X$. However, we need to worry about cocompactness. How much did it cost to enlarge $H.x$ to $\mathcal{H}(H.x)$? Is there some bound $R$ such that $\mathcal{H}(H.x) \subset N_R(H.x)$? We will compute such a bound in Theorem 1.5.9 when $X$ is a finite dimensional CAT(0) cube complex and $H.x$ is *quasiconvex.*

DEFINITION 0.0.1. A subset $Y$ of a geodesic metric space $X$ is $Q$–*quasiconvex* if every geodesic between points of $Y$ is contained in the $Q$–neighborhood of $X$. It is *quasiconvex* if there exists a $Q$ such that it is $Q$–quasiconvex.

So, quasiconvexity allows an additive error in the definition of convexity. This property has been used extensively in the case of hyperbolic groups, in the sense of Definition 4.1.1. There it is a very natural property because, in a sense, additive errors are baked into the definition of hyperbolicity. For example, it is not hard to see that a quasiconvex subset of a hyperbolic space is hyperbolic, and that quasiconvexity is invariant under quasiisometries between hyperbolic spaces. In particular, it makes sense to say a subgroup $H$ of a hyperbolic group $G$ is or is not quasiconvex, because the property does not depend on which geometric model of $G$ is chosen. The situation is different in CAT(0) spaces. The CAT(0) property is sensitive to small perturbations of the metric, and quasiconvexity is not a robust property. In Section 2 we introduce further generalizations of convexity, including the Morse Property, which is quasiisometry invariant, hence a well-defined subgroup property. In Section 3 we show that a refined version of the Morse property has a nice interpretation in right-angled Coxeter groups. There is much more to say about right-angled Coxeter groups, but this is as far as I can make it in one semester.

## 1. Combinatorics of cube complexes

There are two natural metrics to consider on a cube complex: the piecewise Euclidean metric coming from thinking of the whole space will all of the cells filled in, or the combinatorial metric on the vertex set thinking only of the 1–skeleton as defining edge paths. We have previously focused on the piecewise Euclidean metric, particularly in the CAT(0) case. Now we will look at the combinatorial metric.

Figure 1 has a preview of the circle of ideas we will explore. The point is, the CAT(0) condition is geometric. We want to bring in some combinatorics. There are two different viewpoints on how to do this: median graphs and wall spaces. We can push both of these to an even further to an algebraic/combinatorial level of abstraction, median algebras and pocsets, respectively. The circular nature of Figure 1 is not important. It is possible to prove other of the implications directly.

To get started we make some definitions:

Given a graph, one can turn it into a cube complex:

DEFINITION 1.0.1. Let $\Gamma$ be a simple graph. Let $\mathrm{Cube}(\Gamma)$ be the cube complex whose 1–skeleton is $\Gamma$, and where an $n$–cube is attached wherever there is an isometrically embedded subgraph of $\Gamma$ isomorphic to the 1–skeleton of an $n$–cube.

EXAMPLE 1.0.2. $\mathrm{Cube}(K_{2,3}) \cong \mathbb{S}^2$

EXERCISE 1.0.3. If $X$ is a finite dimensional CAT(0) cube complex then $\mathrm{Cube}(X^{(1)}) = X$.

DEFINITION 1.0.4. A *wall structure* on a set is a nonempty collection of nonempty subsets that is closed under taking complements, denoted $*$.

The chosen sets are called *halfspaces*, and a complementary pair of halfspaces $\{A, A^*\}$ is called a *wall*.

A wall structure is *discrete* if there are only finitely many walls separating any pair of points. That is, if $x$ and $y$ are points in the underlying space there are only finitely many walls $\{A, A^*\}$ with $x \in A$ and $y \in A^*$.

If $\{H_1, H_1^*\}$ and $\{H_2, H_2^*\}$ are walls then say they are *transverse* if all of $H_1 \cap H_2$, $H_1 \cap H_2^*$, $H_1^* \cap H_2$, and $H_1^* \cap H_2^*$ are all nonempty. Otherwise they are *nested*.

For example, the halfspaces of a reflection system form a discrete wall structure **on the vertex set** of the underlying graph. In the case of reflection systems there were further hypotheses about a group action and how it related to the wall structure. For wall structures we do not need to have any group action at all.

We will look at wall structures on the vertex set of a graph. The definition is very abstract. There is no requirement that it has anything to do with the graph structure, so we cannot possibly hope for a result, as in reflection systems, that the graph distance between two points is the same as the wall distance. Sageev [**22**] essentially[1] showed how to build a new

[1]Sageev actually worked in a specific case of wall structure involving a group-subgroup pair, but the construction can be made to work more generally, as observed by Niblo and Chatterji [**8**].

ambient space for the walls in which the wall metric is closely related to the geometry of the space, and retaining the natural relationships between the walls. To do this, on first builds a graph $\Gamma$ and the cubes it to get a cube complex $\mathrm{Cube}(\Gamma)$ that is CAT(0) and in which the original walls are represented as hyperplanes.

This gives an equivalence between wall structures and CAT(0) cube complexes.

Shortly later two other equivalence were added. Roller [**21**] went even more abstract and proved equivalences of pocsets, median algebras, and CAT(0) cube complexes, while Chepoi [**9**] proved that median graphs are exactly the 1–skeleta of CAT(0) cube complexes.

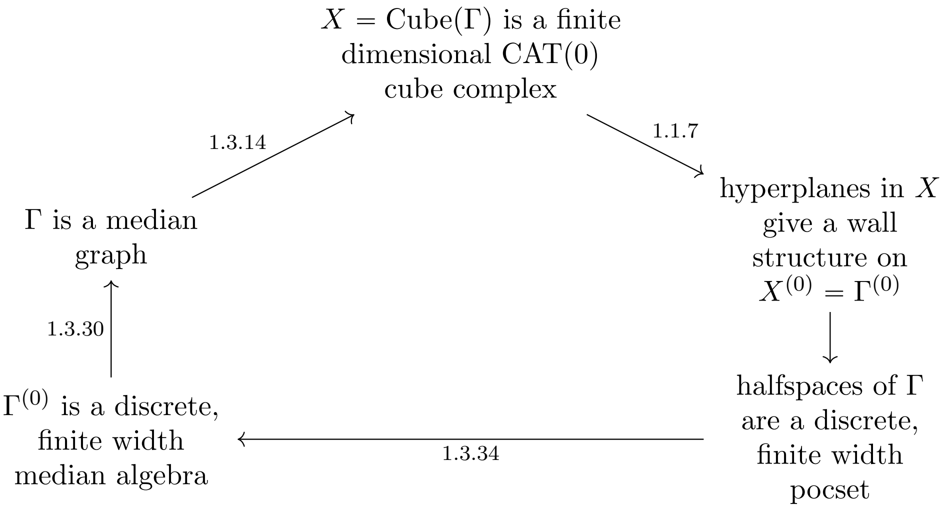
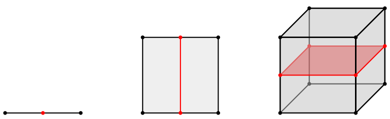



FIGURE 1. The various reformulations of CAT(0) cube complex

## 1.1. Hyperplanes in cube complexes.

DEFINITION 1.1.1. If $[-1/2, 1/2] \times \cdots \times [-1/2, 1/2]$ is a parameterized cube, a *midcube* is a subset such that one coordinate is set to 0.

Here are some cubes with midcubes:



DEFINITION 1.1.2. Define an equivalence relation $[\cdot]_X$ on the edges of a cube complex $X$ by declaring two edges to be equivalent if they are equal or if they are opposite edges of a square. If $\sigma$ is a single cube then a $[\cdot]_\sigma$–equivalence class consists of edges of $\sigma$ transverse to a midcube of $\sigma$.

The *hyperplane* transverse to the $[\cdot]_X$–equivalence class $\mathcal{E}$ is the cube complex built from cubes isometric to the midcubes dual to $[\cdot]_\sigma$–equivalence classes of $\sigma \cap \mathcal{E}$ for each cube $\sigma$ of $X$. These cubes are glued together in the same way that their corresponding midcubes are.

Some cube complexes with hyperplanes are shown in Figure 2. Note that the examples of Figure 2 are non-positively curved, but the hyperplanes do not give the walls of a wall structure on the 0–skeleton: the second example has a hyperplane that is not embedded, resulting in more than 2 complementary components, and the third is a Möbius strip in which the indicated hyperplane does not separate the space.

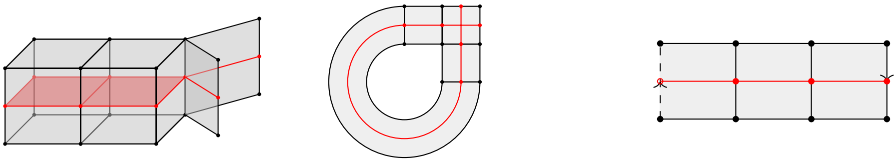


FIGURE 2. Hyperplanes in non-positively curved cube complexes

To clarify, what does it mean that 'the hyperplane is not embedded' in the second example of Figure 2? The point is that 'the hyperplane' is not a subset of $X$, it is an abstract cube complex that comes with a natural map into $X$, and that map is not an embedding. (In the example, 'the hyperplane' is a segment of length 5.) This distinction in language is sometimes ignored, as when we say for the second and third examples that the hyperplane does not separate the space into two components, but we really mean that the *image* of the hyperplane in $X$ does not separate $X$ into 2 components. We will see that in CAT(0) cube complexes there is no confusion, because hyperplanes are embedded.

DEFINITION 1.1.3. If $H$ is a hyperplane in a CAT(0) cube complex, its *carrier* $N(H)$ is the union of closed cubes intersecting $H$.

EXERCISE 1.1.4. Let $X$ be a finite dimensional CAT(0) cube complex, and let $H$ be a hyperplane in $X$.

- Show that $H$ lifts to a closed, convex hyperplane in the universal cover $\widetilde{N(H)}$ of its carrier.
- If $p\colon \widetilde{N(H)} \to N(H)$ is the universal covering map, show that $\widetilde{N(H)} \xrightarrow{p} N(H) \hookrightarrow X$ is a local isometry.
- Show $H$ is a closed, convex set in $X$. (Recall Exercise 1.3.3.)

EXERCISE 1.1.5. Show that a hyperplane in a finite dimensional CAT(0) cube complex has precisely two complementary connected components. Hint:

If $H$ is transverse to an edge $e$ in $X^{(1)}$, $m$ is the midpoint of $e$, and $\pi_e\colon X \to e$ is closest point projection, show that $H = \pi_e^{-1}(m)$.

COROLLARY 1.1.6. *Let $X$ be a finite dimensional CAT(0) cube complex. An edge path in $X^{(1)}$ is a geodesic in the combinatorial metric if and only if it crosses each hyperplane at most once.*

COROLLARY 1.1.7. *Let $X$ be a finite dimensional CAT(0) cube complex. The hyperplanes determine a discrete wall structure on the vertices of $X$.*

**1.2. Pocsets.** This is from [**23**]. Look there for more references.

DEFINITION 1.2.1. A *pocset* is a partially ordered set $(\mathcal{P}, <)$ with a involution $*$ (complemement) such that:

- $\forall A \in \mathcal{P}$, $A$ and $A^*$ are incomparable.
- $\forall A, B \in \mathcal{P}$, $A < B \implies B^* < A^*$.

For example, the halfspaces of a wall system, partially ordered by inclusion, form a pocset.

Define transverse, nested, and discrete similarly to the wall system case:

DEFINITION 1.2.2. $A, B \in \mathcal{P}$ are *nested* if $A < B$ or $B < A$ or $A < B^*$ or $B^* < A$, and *transverse* otherwise, which is the case when $A \cap B$, $A^* \cap B$, $A \cap B^*$, and $A^* \cap B^*$ are all nonempty.

DEFINITION 1.2.3. If $A < B$, then the *interval* from $A$ to $B$ is $[A, B] := \{C \in \mathcal{P} \mid A < C < B\}$.

DEFINITION 1.2.4. A pocset is *discrete* if its intervals are finite, and it has *finite width* if there is a bound on the size of a pairwise transverse set.

DEFINITION 1.2.5. An *ultrafilter* $\alpha$ on a pocset $(\mathcal{P}, <, *)$ is a subset of $\mathcal{P}$ satisfying two conditions:

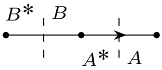
(choice) $\forall A \in \mathcal{P}$, exactly one of $A$ or $A^*$ is in $\alpha$.

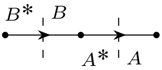
(consistency) If $A \in \alpha$ and $A < B$ then $B \in \alpha$.

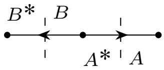
An ultrafilter satisfies the descending chain condition (DCC) if every descending chain $A_1 > A_2 > \cdots$ with $A_i \in \alpha$ eventually terminates.

EXAMPLE 1.2.6. Consider a tree. Think of the midpoint of each edge as a wall dividing the vertex set of the tree into two complementary components. Take the set of such components partially ordered by inclusion with $*$ as the complement operation. This gives a pocset.

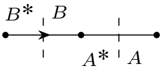
The 'choice' condition of an ultrafilter says that for each edge we must choose one side or the other. Visualize this by orienting each edge with arrow pointing toward the halfspace we have chosen. 'Consistency' says if

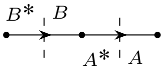
I have already made this choice 

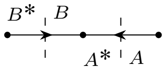
then I am forced to extend by this choice 

because this one 

has $A \in \alpha$ and $A < B$ but $B \notin \alpha$.

On the other hand, having 

does not force the choice orientation of the $\{A, A^*\}$ edge; both 

and 

are consistent.

In fact, ultrafilters on tree posets as above can be completely described as a choice of orientation of the edges such that every vertex has at most one outgoing edge. It is possible that some vertex has no outgoing edges. If such a vertex $v$ exists then it is unique, because for every other vertex $w$ the orientations of edges along the geodesic from $v$ to $w$ all point toward $v$.

Exercise 1.2.7. For an ultrafilter on a tree pocset as in the above example, show that the DCC ultrafilters are in bijection with vertices of the tree.

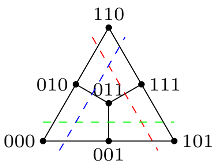
Example 1.2.8. Consider the integer lattice $\mathbb{Z}^2 \subset \mathbb{R}^2$. Consider the 'walls' to be vertical and horizontal lines in $\mathbb{R}^2$ at half-integer coordinates, each of which separates $\mathbb{Z}^2$ into two complementary halves. We get a pocset consisting of these halfspaces ordered by inclusion.

The set of ultrafilters can be described by pairs $(x, y)$ where $x, y \in \mathbb{Z} \cup \{\infty, -\infty\}$ as follows: If $x \in \mathbb{Z}$ then for all vertical walls select the halfspace containing the vertical line at coordinate $x$. If $x = \infty$ then for all vertical walls select the right halfspace. If $x = -\infty$ then for all vertical walls select the left halfspace. Similarly, the $y$ coordinate determines the choice of halfspace for the horizontal walls. The ultrafilter described in this way is DCC if and only if both of $x$ and $y$ are in $\mathbb{Z}$, so, as in the tree case, the DCC ultrafilters correspond to the vertices of $\mathbb{Z}^2$.

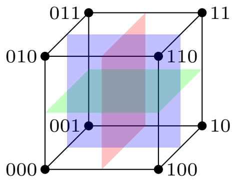
The next example shows not every DCC ultrafilter comes from a vertex of the original graph.

Example 1.2.9. Consider the following graph. The dashed 'walls' each subdivide the graph into two connected components. Take these as the elements of a pocset, ordered by inclusion, with $*$ being complement. We pick a 0–side and a 1–side for each wall, and label the vertices in the order red, green, blue for the side of the respective wall.

In this example all the walls are transverse, so all $2^3$ possible choices of halfspaces are valid ultrafilters. Most of these correspond to the vertices of the original graph, but in this example there is one DCC ultrafilter, '100', that is new. In a preview of what's to come, the DCC ultrafilters can be arranged into a graph $\Gamma$ in which two DCC ultrafilters are adjacent if they differ on exactly one wall. In this example, $\mathrm{Cube}(\Gamma)$ is a single 3–cube, and we see the original walls represented as midcubes that partition the vertices in the same way as the original walls.




(A) A graph with walls




(B) Its dual cube complex

The extra vertex is important; If it were missing then the resulting cube complex would only be three squares incident at a common vertex, which is not CAT(0).

Now we will make precise the construction of a graph from a pocset as illustrated in the previous example.

Definition 1.2.10. Given a a discrete, finite width pocset $(\mathcal{P}, <, *)$, define a graph $\Gamma(\mathcal{P})$ by taking a vertex for each DCC ultrafilter, and connecting two ultrafilters by an edge if they differ on a single pair $\{A, A^*\}$.

We will show:

Theorem 1.2.11. *If $(\mathcal{P}, <, *)$ is a discrete, finite width pocset then $\Gamma(\mathcal{P})$ is a nonempty, connected graph.*

It is possible to show directly using properties of ultrafilters that $\mathrm{Cube}(\Gamma(\mathcal{P}))$ is a CAT(0) cube complex. Instead, we will show that $\Gamma(\mathcal{P})$ is a median graph, and that whenever $\Gamma$ is a median graph $\mathrm{Cube}(\Gamma)$ is a CAT(0) cube complex. The benefit of doing it this way is that it lets us port geometric properties of median graphs to the cube complex setting.

Lemma 1.2.12. *If $(\mathcal{P}, <, *)$ is a discrete, finite width pocset then DCC ultrafilters exist.*

Proof. Let $\mathcal{T} \subset \mathcal{P}$ be a maximal pairwise transverse subset. It is finite by the finite width hypothesis. Define $\omega$ to contain $\mathcal{T}$, and extend to the rest of $\mathcal{P}$ as follows. For any $B \in \mathcal{P}$ such that $B$ and $B^*$ are not in $\mathcal{T}$, there exists $A \in \mathcal{T}$ such that $A$ and $B$ are nested. Otherwise we could add $B$ to $\mathcal{T}$ to get a larger transverse set. Add $B$ to $\omega$ if $A < B$ or $B^* < A$, and add $B^*$ to $\omega$ if $B < A$ or $A < B^*$. One of these is true since $A$ and $B$ are nested. This is well defined, because if $A' \in \mathcal{T}$ is also nested with $B$ then transversality of $A$ and $A'$ prohibits $A < B < A'$ and $A' < B^* < A$, so if $A$ says $B$ should be in $\omega$ then $A'$ agrees. Now $\omega$ satisfies the 'choice' condition.

Suppose that $B < C$ and $B \in \omega$. There exists $A \in \mathcal{T}$ nested with $B$ such that either $A \leqslant B < C$ or $C^* < B^* < A$. In both of these cases we defined $C \in \omega$. Thus, $\omega$ satisfies the 'consistency' condition; it is an ultrafilter.

Now suppose $B_1 > B_2 > \dots$ is a descending chain in $\omega$. For each $A_j$ in $\mathcal{T}$ define:

$$n_j := \begin{cases} 1 & \text{if } B_1 \text{ and } A_j \text{ are transverse} \\ \#[A_j, B_1] & \text{if } A_j < B_1 \\ \#[B_1^*, A_j] & \text{if } B_1^* < A_j \end{cases}$$

The second and third cases are the two possibilities for $B_1$ being nested with $A_j$ when $B_1 \in \omega$. These numbers are finite since the pocset is discrete. Let $n := \max_j n_j$, which exists since the pocset is finite width.

Now consider any element $B_i \notin \mathcal{T}$ in the chain. By definition of $\mathcal{T}$, $B_i$ is nested with some $A_j \in \mathcal{T}$, and since $B_i \in \omega$ the possibilities are $A_j < B_i$ or $B_i^* < A_j$. In the first case, we have $B_1 > B_2 > \dots > B_i > A_j$, and in the second case we have $B_1^* < B_2^* < \dots < B_i^* < A_j$. In both cases this implies $i \leqslant n_j - 1 \leqslant n - 1$, so the chain contains an initial segment of non–$\mathcal{T}$ elements of length at most $n-1$. It is then possible that $B_n \in \mathcal{T}$. By the way we constructed $\omega$, there is no $B \in \mathcal{P}$ such that for some $A \in \mathcal{T}$ we have $B > A$ and $B \in \omega$, so if there is such a $B_n$ then it is the terminal element of the chain. Thus, this chain (in fact, any descending chain starting with $B_1$) has length at most $n$. $\square$

LEMMA 1.2.13. *If $\alpha$ is a DCC ultrafilter with $A \in \alpha$ then there is a DCC ultrafilter that differs from $\alpha$ only on the pair $\{A, A^*\}$ if and only if $A$ is minimal in $\alpha$.*

PROOF. Suppose $\omega$ differs from $\alpha$ only on $\{A, A^*\}$, with $A \in \alpha$. First we check that $\omega$ is an ultrafilter. The 'choice' condition is obviously satisfied by $\omega$.

If $B \lneq A$ then $B \in \alpha \cap \omega$, but then $B < A$ with $B \in \omega$ and $A^* \in \omega$ means $\omega$ is not consistent, so minimality is necessary for consistency.

To see that minimality is also sufficient we only have to check that $A$ minimal and $A^* \lneq B$ implies $B \in \omega$, since consistency for pairs not involving $A^*$ is inherited from $\alpha$. We have $B^* \lneq A$, so by minimality of $A$, $B^* \notin \alpha$, so $B \in \alpha$. Since $\alpha$ and $\omega$ agree off $\{A, A^*\}$, $B \in \omega$.

Finally, the DCC condition is inherited from $\alpha$: Notice that $A$ being minimal in $\alpha$ implies $A^*$ is minimal in $\omega$. If $B_1 > B_2 > \dots$ is a descending chain in $\omega$ then either it contains $A^*$, which, being minimal, must be the last element of the chain, or it does not contain $A^*$, so it is also a chain in $\alpha$ and is therefore finite by the DCC for $\alpha$. □

LEMMA 1.2.14. *If $(\mathcal{P}, <, *)$ is a finite width pocset and $\alpha$ and $\omega$ are two DCC ultrafilters then* $\mathrm{Diff}(\alpha, \omega) := \{A \in \mathcal{P} \mid A \in \alpha$ *and* $A \notin \omega\}$ *is finite.*

PROOF. Descending chains in Diff $=$ Diff$(\alpha, \omega)$ are finite by the DCC for $\alpha$, and ascending chains in Diff are finite by applying $*$ and the DCC for $\omega$, so chains in Diff are finite. Thus, every chain can be extended to a maximal finite chain. If Diff is infinite then finite width implies there are arbitrarily long chains. Consider a sequence of maximal descending chains of unbounded length. Consider the first element of each chain. No pair of such elements can be nested, because that would either violate consistency of $\alpha$, or it would allow us to extend a maximal chain. Thus, the set of first elements is pairwise transverse. By finite width, this set is of bounded size, so at least one of these elements, call it $A_1$, occurs as the first element of infinitely many of the chains. Pass to the subsequence of chains consisting of those that start with $A_1$. Repeat for the second element of each chain. Again, there cannot be any nesting relationships between these elements, because it would either contradict consistency of $\alpha$ or maximality of the chains. Since there are still infinitely many chains, some $A_2$ occurs as the second element of infinitely many chains. Pass to the subsequence of chains consisting of those that start with $A_1 > A_2$. Continuing in this way, there is an infinite descending chain $A_1 > A_2 > \ldots$ in Diff, which is a contradiction. □

Now we can prove the graph is connected:

PROOF OF THEOREM 1.2.11. Let $(\mathcal{P}, <, *)$ be a discrete, finite width pocset and $\Gamma(\mathcal{P})$ its graph. The graph is nonempty by Lemma 1.2.12. Take any two vertices, corresponding to DCC ultrafilters $\alpha$ and $\omega$. By Lemma 1.2.14, $\mathrm{Diff}(\alpha, \omega) := \{A \in \mathcal{P} \mid A \in \alpha$ and $A \notin \omega\}$ is finite, so it has minimal elements. Suppose $A \in \mathrm{Diff}(\alpha, \omega)$ and $B < A$ for some $B \in \alpha$. Then $B$ cannot be in $\omega$, by consistency, so $B$ is also in $\mathrm{Diff}(\alpha, \omega)$. This shows that minimality in $\mathrm{Diff}(\alpha, \omega)$ implies minimality in $\alpha$. Thus, by Lemma 1.2.13, for any minimal element $A_0 \in \mathrm{Diff}(\alpha, \omega)$ there is a DCC ultrafilter $\alpha_1$ that differs from $\alpha_0 := \alpha$ only on $\{A_0, A_0^*\}$. We have $\mathrm{Diff}(\alpha_1, \omega) = \mathrm{Diff}(\alpha_0, \omega) - \{A_0\}$. Repeat. We get a sequence $\alpha = \alpha_0, \alpha_1, \alpha_2, \ldots$ of adjacent DCC ultrafilters in $\Gamma(\mathcal{P})$ such that $\mathrm{Diff}(\alpha_i, \omega)$ decreases in size by 1 with each step, so for $n = \#\,\mathrm{Diff}(\alpha, \omega)$ we have $\alpha_n = \omega$. We have constructed a path from $\alpha$ to $\omega$ in $\Gamma(\mathcal{P})$. In fact, it is easy to argue that this path is geodesic. □

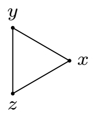
**1.3. Median graphs and median algebras.** The history of median algebras is long and complicated by the fact that the ideas were rediscovered several times in different settings. Two references are Birkhoff and Kiss [**4**], who work in the setting of distributive lattices, and Isbell [**18**], who axiomatizes median algebras as a distinct algebraic structure. Bowditch [**5**] has extensive notes. We will develop some of the geometry of median graphs for application to right-angled Coxeter groups. We will also look at the equivalence between median graphs and median algebras, because there are some (rather complicated) results known for median algebras that have geometric interpretations in median graphs.

1.3.1. *Median graphs.*

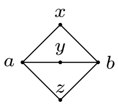
DEFINITION 1.3.1. A *modular graph* is a connected simple graph in which for every three vertices $x, y, z$ there exists a *median* vertex $m(x, y, z) \in [x, y] \cap [y, z] \cap [z, x]$, where the *interval* $[x, y]$ is the union of geodesics from $x$ to $y$.

A *median graph* is a modular graph in which every three vertices have a unique median.

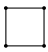
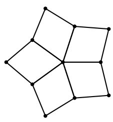
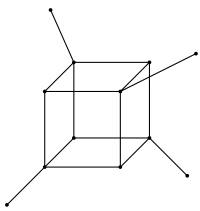
EXERCISE 1.3.2. Show that any tree is a median graph.

EXAMPLE 1.3.3. A triangle 

is not modular; $[x, y] \cap [y, z] \cap [z, x] = \varnothing$.

The graph $K_{2,3} =$ 

is modular but not median; $[x, y] \cap [y, z] \cap [z, x] = \{a, b\}$.

Here are some median graphs:   

EXERCISE 1.3.4. For the median graph examples in Example 1.3.3, convince yourself that they are median. Show that for any vertices $w, x, y, z$ that $m(w, x, m(y, x, z)) = m(m(w, x, y), x, z)$.

EXERCISE 1.3.5. Show that modular graphs are bipartite. Hint: Pick a base vertex $v$ and show that two vertices equidistant from $v$ cannot be adjacent. Find a graph that is bipartite but not modular.

COROLLARY 1.3.6. *Modular graphs contain no triangles. Median graphs contain no $K_{2,3}$.*

There are actually full characterizations of modular and median graph in terms of subgraph structure. We will state a condition for modular graphs in Theorem 1.3.10, and prove a characterization of median graphs in Theorem 1.3.11. Let us introduce some shorthand notation:

DEFINITION 1.3.7. Write $x_0 \to x_1 \to \cdots \to x_n$ if there exists a geodesic from $x_0$ to $x_n$ that passes through $x_1$, $x_2$, ... in that order.

The point is that in writing $x_0 \to x_1 \to x_2$ we suppress the fact that there could be many geodesics from $x_0$ to $x_1$ and many geodesics from $x_1$ to $x_2$. They are interchangeable, in the sense of the next lemma, whose proof is trivial.

LEMMA 1.3.8. *If $\alpha + \beta + \gamma$ is a geodesic and $\beta'$ is a geodesic with the same endpoints as $\beta$ then $\alpha + \beta' + \gamma$ is a geodesic.*

COROLLARY 1.3.9. *If $x_0 \to \cdots \to x_n$ and for some $0 \leqslant i < j \leqslant n$ we have $x_i \to y_1 \to \cdots \to y_m \to x_j$ then:*

$$x_0 \to x_1 \to \cdots \to x_i \to y_1 \to \cdots \to y_m \to x_j \to x_{j+1} \to \cdots \to x_n$$

THEOREM 1.3.10 (see [**9**, Lemma 4.1]). *A graph is modular if and only if it has no triangles and satisfies the following geodesic bigon condition: Suppose $x$ and $y$ are neighbors of $w$ with $w \to x \to z$ and $w \to y \to z$. Then there exists a common neighbor $v$ of $x$ and $y$ such that $w \to x \to v \to z$ and $w \to y \to v \to z$.*

It is clear that the geodesic bigon condition is necessary: take $v$ to be a median of $x, y, z$.

THEOREM 1.3.11. *A modular graph is median if and only if it does not contain $K_{2,3}$ as a subgraph.*

We need a lemma first.

LEMMA 1.3.12. *In a modular graph, $w \in [x, y] \cap [x, z]$ implies $m(w, y, z) \subset m(x, y, z)$.*

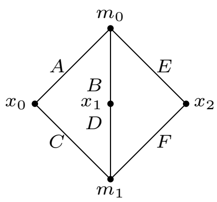
PROOF. Let $m$ be a median point for $w, y, z$. We have $x \to w \to y$ and $w \to m \to y$, so $x \to w \to m \to y$. Similarly, $m \in [x, z]$. But $m \in [y, z]$, so $m \in [x, y] \cap [y, z] \cap [z, x]$. □

PROOF OF THEOREM 1.3.11. We have seen that $K_{2,3}$ cannot be a subgraph of a median graph. We prove the other direction: Suppose we have a modular graph that is not median. Then there exist triples $x_0, x_1, x_2$ with more than one median. Among all such triples, choose one such that

$\sum d(x_i, x_{i+1})$ is minimal. Let $m_0$ and $m_1$ be two medians of $x_0, x_1, x_2$. For each $i$, choose a median $y_i$ for $x_i$, $m_0$, $m_1$.

For each $i \neq k \in \{0,1,2\}$ and $j \in \{0,1\}$ we have $x_i \to m_j \to x_k$ and $x_i \to y_i \to m_j$, so, by Lemma 1.3.8, $x_i \to y_i \to m_j \to y_k \to x_k$. So $m_0$ and $m_1$ are also medians for the triple $y_0$, $y_1$, $y_2$. This contradicts minimality of $\sum d(x_i, x_{i+1})$ unless $x_i = y_i$, so each $x_i$ sits on a geodesic between $m_0$ and $m_1$.

Consider the schematic diagram, where capital letters indicate distances:

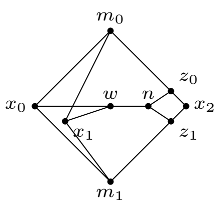



(17)

We have $A + B = d(x_0, x_1) = C + D$ and $A + C = d(m_0, m_1) = B + D$, which implies $A = D$ and $B = C$. Similarly, with $x_0$ and $x_2$ we find $A = F$ and $C = E$. The same computation for $x_1$ and $x_2$ gives $B = F$ and $D = E$. Thus, $A = D = E = C = B = F$. If $D = 1$ then since modular graphs contain no triangles, there is no edge from $x_i$ to $x_j$ when $i \neq j$, and we have found a $K_{2,3}$ subgraph, so we are done.

In the case $D > 1$ there exist vertices $z_i \neq m_i$ at distance 1 from $x_2$ with $x_2 \to z_i \to m_i$. By Lemma 1.3.12 $m_i$ is a median for $x_0$, $x_1$, $z_i$. For $i \in 0, 1$ let $n_i$ be a median for $x_i$, $z_0$, $z_1$. By our choice of $z_0$ and $z_1$, neither $x_2$ nor $z_1$ lies on a geodesic from $z_0$ to $x_0$ or $x_1$, and vice versa, so the $n_i$ are distinct from $z_0$, $z_1$, and $x_2$. However, $d(z_0, z_1) = 2$, so both $n_i$ are adjacent to each $z_j$. Modular graphs have no triangles, so there are no edges between $n_0$, $n_1$, and $x_2$. If $n_0$ and $n_1$ are distinct then $n_0$, $n_1$, $x_2$, $z_0$, and $z_1$ are the vertices of a $K_{2,3}$, so we are done. Suppose not, so suppose $n := n_0 = n_1$.

Let $w$ be a median for $n$, $x_0$, $x_1$. By Lemma 1.3.12, $w$ is a median for $z_0$, $x_0$, $x_1$. We would like this to be a different median that $m_0$.




(18)

We have $z_1 \to n_1 \to x_1$ and $n_1 \to w \to x_1$, so $z_1 \to n_1 \to w \to x_1$. This cannot be the case for $m_0$: if $z_1 \to m_0 \to x_1$ then, since $z_1 \to x_2 \to m_0$, we would have $z_1 \to x_2 \to m_2 \to x_1$, which implies $d(z_1, x_1) = 2D + 1$. That is wrong; $d(z_1, x_1) = 2D - 1$ via $z_1 \to m_1 \to x_1$, so $w \neq m_0$. We have shown that the triple $x_0$, $x_1$, $z_0$ has distinct medians $m_0$ and $w$. This contradicts the choice of $x_0$, $x_1$, $x_2$, since:

$$\sum d(x_i, x_{i+1}) = 6D > 6D - 2 = d(x_0, z_0) + d(x_1, z_0) + d(x_0, x_1) \qquad \square$$

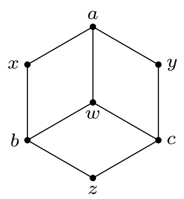
1.3.2. *Cubings of median graphs are CAT(0).*

Exercise 1.3.13. Show that if $\Gamma$ is a modular graph then $\mathrm{Cube}(\Gamma)$ is simply connected. To do this, take an arbitrary loop $\gamma$ in $\Gamma$ and show it can be filled in $\mathrm{Cube}(\Gamma)^{(2)}$. Hint: Consider $\alpha + e + \beta$, where $e$ is an edge of $\gamma$ and $\alpha$ and $\beta$ are geodesics between $\gamma(0)$ and either end of $e$. Apply Exercise 1.3.5 and Theorem 1.3.10 to this loop.

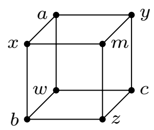
Theorem 1.3.14 (Chepoi [**9**, Theorem 6.1]). *If* $\Gamma$ *is a median graph then* $\mathrm{Cube}(\Gamma)$ *is a CAT(0) cube complex. Conversely, the 1–skeleton of a CAT(0) cube complex is a median graph.*

We will not prove this theorem. The direction claimed by Figure 1 is the easier direction: show that $\mathrm{Cube}(\Gamma)$ is CAT(0). Exercise 1.3.13 establishes that $\mathrm{Cube}(\Gamma)$ is simply connected, so we only need to verify the link condition. The following is the lowest dimensional case, and contains the essential ingredients of verifying the link condition.

Proposition 1.3.15. *Suppose a median graph* $\Gamma$ *has the following subgraph* $\Lambda$*:*




*Then there exists a vertex* $m$ *such that the following subgraph* $\Lambda'$ *is isometrically embedded:*




The point is that the first picture looks like it will give three squares that yield a short loop in the link of $w$ in $K(\Gamma)$, so if $K(\Gamma)$ is going to be CAT(0) there needs to be a 3–cube filling that loop in the link.

Proof. First, $\Lambda$ is isometrically embedded: By Exercise 1.3.5, modular graphs are bipartite, so there are no odd-length cycles. Thus, the parity of the distance between two vertices in $\Lambda$ is the same as in $\Gamma$. In particular, vertices at distance 2 from each other in $\Lambda$ are at distance 2 in $\Gamma$, and vertices at distance 3 from each other in $\Lambda$ are at distance either 1 or 3 from each other in $\Gamma$. Up to symmetry, to verify $\Lambda$ is isometrically embedded we only have to check that $a$ and $z$, which are at distance 3 in $\Lambda$, are not adjacent in $\Gamma$. An edge between $a$ and $z$ would mean that $a$, $b$, $c$, $w$ and $z$ span a $K_{2,3}$ subgraph of $\Gamma$, which is forbidden.

Notice that $\Lambda$ itself is not a median graph; $x$, $y$, $z$ have no median in $\Lambda$. Let $m := m(x, y, z)$ be their median in $\Gamma$. Since $d(x, y) = d(y, z) = d(z, x) = 2$ and $m \notin \Lambda$, $m$ is a common neighbor of $x$, $y$, and $z$, so $\Gamma$ contains the desired subgraph. No-odd-cycles implies $m$ is not a neighbor of $a$, $b$, or $c$, and is not at distance 2 from $w$. Finally, if $m$ were a neighbor of $w$ then $w$, $x$, $y$, $a$, $m$ would be the vertices of a $K_{2,3}$ subgraph, which also forbidden, so $\Lambda'$ is isometrically embedded in $\Gamma$. □

1.3.3. *Some geometry of median graphs.*

PROPOSITION 1.3.16. *Intervals in median graphs are convex: For any interval $[y, z]$ and any $w, x \in [y, z]$ we have $[w, x] \subset [y, z]$.*

This is not obvious. It is clear that the convex hull of $y$ and $z$ should contain geodesics between $y$ and $z$, but we also need geodesics between points of geodesics between $y$ and $z$, etc. We need to show that those were already part of $[y, z]$.

PROOF. Given $y$ and $z$, suppose that there exist $w, x \in [y, z]$ such that $[w, x] \not\subset [y, z]$. We may assume that among all such pairs $w$, $x$ we choose one such that $d(w, x)$ is minimal. Note that $w \neq x$, since otherwise $[x, w] = \{x\} = \{w\} \subset [y, z]$.

Let $v \in [w, x] - [y, z]$. Define $a := m(v, x, y)$, $b := m(v, x, z)$, and $c := m(v, a, b)$.

From the definitions of the points, we have $y \to a \to x$ and $x \to b \to z$ and $y \to x \to z$ and $a \to c \to b$. Apply Lemma 1.3.8:

$$y \to a \to x \quad \& \quad y \to x \to z \implies y \to a \to x \to z$$

$$x \to b \to z \quad \& \quad y \to a \to x \to z \implies y \to a \to x \to b \to z$$

$$a \to c \to b \quad \& \quad y \to a \to x \to b \to z \implies y \to a \to c \to b \to z$$

Similarly, $w \to v \to x$ and $v \to a \to x$ and $v \to c \to a$, so:

$$w \to v \to x \quad \& \quad v \to a \to x \implies w \to v \to a \to x$$

$$v \to c \to a \quad \& \quad w \to v \to a \to x \implies w \to v \to c \to a \to x$$

So $v \in [w, c]$ and $c$ is strictly closer than $x$ to $w$, unless $c = x = a$. This would contradict minimality of $d(w, x)$, so $c = x = a$. A symmetric argument shows $b = x$ as well.

Now, we have $x = a \in [y, v]$ and $x = b \in [z, v]$ and $x \in [y, z]$, so $x$ is a median of $y, z, v$.

Run the same argument with the roles of $x$ and $w$ swapped to see that $w$ is also a median of $y, z, v$. By uniqueness of medians, $w = x$, which is a contradiction. □

LEMMA 1.3.17. *Given three vertices $x, y, z$ of a median graph, the median $m = m(x, y, z)$ is the unique closest point of $[y, z]$ to $x$.*

PROOF. Suppose there exists $w$ with $y \to w \to z$ and $d(w, z) \leqslant d(m, z)$.

$$\begin{aligned} d(x,y) + d(x,z) &\leqslant d(x,w) + d(w,y) + d(x,w) + d(w,z) \\ &= d(y,z) + 2d(x,w) \\ &\leqslant d(y,z) + 2d(x,m) \\ &= d(x,m) + d(m,y) + d(x,m) + d(m,z) \\ &= d(x,y) + d(x,z) \end{aligned}$$

So all of the inequalities are equalities, which implies $d(x, y) = d(x, w) + d(w, y)$ and $d(x, y) = d(x, w)+d(w, z)$. Thus, $w \in [x, y]\cap[y, z]\cap[z, x] = \{m\}$, by uniqueness of medians. □

PROPOSITION 1.3.18. *In a median graph, $w \in [x, y] \cap [x, z]$ and $v \in [y, z]$ implies $w \in [x, v]$.*

In words this is ‘convexity of betweenness’: if $w$ is between $x$ and $y$ and $w$ is between $x$ and $z$ then $w$ is between $x$ and anything in the interval of $y$ and $z$.

PROOF. We have $x \to w \to y$ and $w \to m(w, y, z) \to y$, so $x \to w \to m(w, y, z) \to y$. If we also had $x \to m(w, y, z) \to v$, then it would follow that $x \to w \to m(w, y, z) \to v$, as desired, so it suffices to prove the lemma in the case $w = m(w, y, z)$.

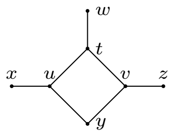
By Lemma 1.3.12, $m(x, y, z) = m(w, y, z)$, so $w = m(x, y, z)$. By Lemma 1.3.17, $w = m(x, y, z)$ is the unique closest point of $[y, z]$ to $x$. By Proposition 1.3.16, $[w, v] \subset [y, z]$, so $w$ is closer to $x$ than any other point in $[w, v]$. But by Lemma 1.3.17, the closest point of $[w, v]$ to $x$ is $m(x, w, v)$, so $w = m(x, w, v) \in [x, v]$. □

We will use this result in the next subsection:

LEMMA 1.3.19. *Let $w, x, y, z$ be vertices in a median graph.*

$$m(m(x, w, y), w, z) = m(x, w, m(y, w, z))$$

PROOF. Let $u := m(w, x, y)$ and $v := m(w, y, z)$ and $t := m(u, v, w)$.




We have $u \in [x, w] \cap [x, y]$ and $v \in [w, y]$, so, by Proposition 1.3.18, $u \in [x, v]$. Then $u \in [x, w] \cap [x, v]$, so, by Lemma 1.3.12, $m(v, w, x) = m(u, v, w) = t$.

The symmetric argument with $z$ and $v$ swapped with $x$ and $u$ gives $m(u, w, z) = t$. Thus:

$$\begin{aligned} m(m(x, w, y), w, z) &= m(u, w, z) \\ &= t \\ &= m(x, w, v) \\ &= m(x, w, m(y, w, z)) \end{aligned}$$ □

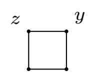
DEFINITION 1.3.20. A subset $Y \subset X$ is *gated* if for every $x \in X$ there exists a *gate for* $x$, $\omega(x) \in Y$, such that for every $y \in Y$ there exists a geodesic from $x$ to $y$ through $\omega(x)$. The map $x \mapsto \omega(x)$ is the *gate map.*

PROPOSITION 1.3.21. *Convex subgraphs of a median graph are gated.*

PROOF. Let $C$ be a convex subgraph of a median graph. Let $x$ be arbitrary. Let $y$ be a closest point of $C$ to $x$, and let $z \in C$ be arbitrary. Consider $m = m(x, y, z)$. It lies on $[x, y]$, so it is closer to $x$ than $y$, unless $m = y$. It also lies on $[y, z]$, which is contained in $C$ since $C$ is convex, so it cannot be strictly closer to $x$ than $y$. Thus $m = y$. Now, $m$ also lies on $[x, z]$, so it is strictly closer to $x$ than $z$ if $z \neq m = y$. This shows that $y$ is the unique closest point of $C$ to $x$, and that as $z$ varies throughout $C$, there is always a geodesic from $x$ to $z$ that goes through $y$, so $y$ is the gate for $x$. □

COROLLARY 1.3.22. *For vertices $y$ and $z$ in a median graph, $[y, z]$ is gated and the gate map is $x \mapsto \omega(x) = m(x, y, z)$.*

PROOF. If $w \in [y, z]$ then by Proposition 1.3.18, $m(x, y, z) \in [x, w]$. □

The 'gate' terminology suggests a property stronger than the definition. The gate for a point is not the only entry into the gated set. Consider the simple example 

. There is an interval $[w, x] = \{w, x\}$. The gate for $z$ on this interval is $w$. For every point in $[w, x]$ there is a geodesic from $z$ to that point through $w$. There are also geodesics that do not go through the gate, like $z \to y \to x$. So one should not imagine that the gate is the *only* way into $[w, x]$, only that there is no shortcut that can be obtained by avoiding the gate. It does follow immediately from the definition that the gate for a point is the unique closest point of the gated set, so the gate map is closest point projection.

PROPOSITION 1.3.23 (Lipschitz projection to convex sets). *Let $C$ be a convex subgraph of a median graph. The gate map $\omega$ of $C$ is Lipschitz: for all $x$ and $y$ we have $d(\omega(x), \omega(z)) \leqslant d(x, y)$.*

PROOF. It is enough to show the proposition when $d(x, y) = 1$, then extend along geodesics. So suppose $d(x, y) = 1$.

Suppose $d(x, C) \neq d(y, C)$. Without loss of generality, assume $d(x, C) < d(y, C)$. Then $1 + d(x, C) = d(y, C) \leqslant d(y, \omega(x)) \leqslant 1 + d(x, C)$, so these are equalities and $\omega(x)$ realizes the distance from $y$ to $C$, so $\omega(y) = \omega(x)$.

Suppose $d(x, C) = d(y, C)$. The gate property says:

$$d(y, \omega(y)) + d(\omega(y), \omega(x)) = d(y, \omega(x))$$

But $d(y, \omega(x)) \leqslant d(y, x) + d(x, \omega(x))$. Thus:

$$d(\omega(x), \omega(y)) \leqslant d(x, y) + d(x, \omega(x)) - d(y, \omega(y)) = 1 \qquad \square$$

1.3.4. *Median algebras.*

DEFINITION 1.3.24. A *median algebra* is a set $M$ and a symmetric ternary operation $\langle\cdot\rangle\colon M^3 \to M$ satisfying, for all $v, w, x, y, z \in M$:

- M (majority) $\langle x, x, y\rangle = x$
- A (associativity) $\langle\langle x, w, y\rangle, w, z\rangle = \langle x, w, \langle y, w, z\rangle\rangle$
- D (distributivity) $\langle\langle x, w, y\rangle, v, z\rangle = \langle x, \langle w, v, z\rangle, \langle y, v, z\rangle\rangle$

By setting $v = w$, (M) and (D) imply (A). It is also true that (M) and (A) imply (D), but this takes some work, eg [**5**, Theorem 3.2.1]. We take it as a fact, the point being that to prove something is a median algebra it is enough to prove (M) and (A), but (D) gives us a new tool.

DEFINITION 1.3.25. If $(M, \langle\cdot\rangle)$ is a median algebra and $x, y \in M$, the *median interval* between $x$ and $y$ is $[x, y]_M := \{z \in M \mid \langle x, y, z\rangle = z\}$.

It follows immediately from the definitions that:

$$\langle x, y, z\rangle = [x, y]_M \cap [y, z]_M \cap [z, x]_M \tag{19}$$

DEFINITION 1.3.26. A median algebra is *discrete* if all median intervals are finite.

EXAMPLE 1.3.27. $\mathbb{Z}$ and $\mathbb{R}$ are median algebras with the usual (statistical) median and the usual intervals. $\mathbb{Z}$ is a discrete median algebra; $\mathbb{R}$ is not discrete.

PROPOSITION 1.3.28. *The vertex set of a median graph with the median operation in the graph is a median algebra whose median intervals agree with the graph intervals.*

Notice this gives us new results for median graphs, since we can now apply Condition (D) of Definition 1.3.24:

LEMMA 1.3.29. *In a median graph,* $\pi_{[w,x]} \circ \pi_{[y,z]} = \pi_{[\pi_{[w,x]}(y),\pi_{[w,x]}(z)]}$.

PROOF OF PROPOSITION 1.3.28. We set $\langle x, y, z\rangle := m(x, y, z) = [x, y] \cap [y, z] \cap [z, x]$. This is symmetric, and $[x, x] = \{x\}$, so $\langle x, x, y\rangle = [x, x] \cap [x, y] = \{x\}$. Property (A) was Lemma 1.3.19.

$$\begin{aligned}
[x, y]_M &= \{z \mid \langle x, y, z\rangle = z\} \\
&= \{z \mid z = m(x, y, z)\} \\
&= \{z \mid z = [x, y] \cap [y, z] \cap [z, x]\} \\
&= \{z \mid z \in [x, y]\} \\
&= [x, y]
\end{aligned}$$

□

The converse of Proposition 1.3.28 is true for discrete median algebras:

PROPOSITION 1.3.30. *If* $(M, \langle\cdot\rangle)$ *is a discrete median algebra then there is a median graph whose vertex set is* $M$*, the graph median is* $\langle\cdot\rangle$*, and the graph intervals are the median intervals.*

To prove this we need some facts about median algebras analogous to properties of median graphs.

LEMMA 1.3.31. $x \in [y, z]_M$ *implies* $[y, x]_M \subset [y, z]_M$ *and* $[x, z]_M \subset [y, z]_M$

PROOF. By the definition of median intervals, $x \in [y, z]_M$ and $w \in [y, z]_M$ means $x = \langle x, y, z\rangle$ and $w = \langle w, x, y\rangle$. Then:

$$\begin{aligned}
\langle w, y, z\rangle &= \langle\langle w, x, y\rangle, y, z\rangle \\
&= \langle\langle w, y, x\rangle, y, z\rangle \\
&= \langle w, y, \langle x, y, z\rangle\rangle \\
&= \langle w, y, x\rangle \\
&= \langle w, x, y\rangle \\
&= w
\end{aligned}$$

So $w \in [y, z]_M$. The other claim is similar. □

LEMMA 1.3.32. *Suppose* $x \in [y,z]_M$. *If* $x \neq z$ *then* $z \notin [y,x]_M$. *If* $x \neq y$ *then* $y \notin [x,z]_M$.

PROOF. Suppose $z \neq x \in [y,z]_M$. Then $\langle z,y,x\rangle = x \neq z$, so $z \notin [y,x]_M$. The other claim is similar. $\square$

LEMMA 1.3.33. *Let* $(M, \langle\cdot\rangle)$ *be a median algebra. For any* $y,z \in M$ *define* $\pi_{[y,z]_M} : M \to [y,z]_M : x \mapsto \langle x,y,z\rangle$. *Then* $\pi_{[w,x]_M}([y,z]_M) = [\pi_{[w,x]_M}(y), \pi_{[w,x]_M}(z)]_M$.

PROOF. For any $u \in [y,z]_M$ we have:

$$\begin{aligned}
\pi_{[w,x]_M}(u) &= \langle w,x,u\rangle \\
&= \langle w,x,\langle u,y,z\rangle\rangle \\
&= \langle u,\langle w,x,y\rangle,\langle w,x,z\rangle\rangle \\
&= \langle u,\pi_{[w,x]_M}(y),\pi_{[w,x]_M}(z)\rangle \\
&\in [\pi_{[w,x]_M}(y),\pi_{[w,x]_M}(z)]_M
\end{aligned}$$

For any $v \in [\pi_{[w,x]_M}(y), \pi_{[w,x]_M}(z)]_M$, consider $u := \langle v,y,z\rangle \in [y,z]_M$.

$$\begin{aligned}
\pi_{[w,x]_M}(u) &= \langle w,x,u\rangle \\
&= \langle w,x,\langle v,y,z\rangle\rangle \\
&= \langle v,\langle w,x,y\rangle,\langle w,x,z\rangle\rangle \\
&= v
\end{aligned}$$

$\square$

PROOF OF PROPOSITION 1.3.30. Define a *chain* in $M$ to be a sequence $x_0, x_1, \ldots, x_n$ such that $[x_i, x_{i+1}]_M = \{x_i, x_{i+1}\}$.

We claim that for any two points $y,z \in M$ there exists a chain from $y$ to $z$ contained in $[y,z]_M$. This is proved by induction on interval size. If $\#[y,z]_M = 2$ then $y,z$ is already a chain. Suppose the claim is true for all intervals of size at most $n \geqslant 2$, and suppose $\#[y,z]_M = n+1 \geqslant 3$. Then there exists $x \in [y,z]_M - \{y,z\}$. By Lemma 1.3.32, $[y,x]_M$ and $[x,z]_M$ are strictly smaller than $[y,z]_M$, so by the induction hypothesis there exist chains from $y$ to $x$ in $[y,x]_M$, and from $x$ to $z$ in $[x,z]_M$. By Lemma 1.3.31, $[y,x]_M \cup [x,z]_M \subset [y,z]_M$, so the claim is proved.

Define a graph by taking the vertices to be the set $M$, and adding an edge between $x$ and $y$ if and only if $[x,y]_M = \{x,y\}$. Chains in $M$ correspond to vertices along an edge path in the graph, so the existence of chains between any pair of elements shows the graph is connected.

Consider a geodesic in the graph from $y$ to $z$. Its vertices give a minimal length chain from $y$ to $z$. By Lemma 1.3.33, the projection to $[y,z]_M$ gives a chain. By minimality the projection is bijection between the chain and its image.

Suppose there exists a first vertex $x$ on the geodesic that is not in $[y, z]_M$. Let $w$ be its predecessor. Since $w$ and $x$ are adjacent, $[w, x]_M = \{w, x\}$. By bijectivity of the projection $w \neq \langle x, y, z\rangle$, and since $w$ and $x$ are adjacent, so are $w$ and $\langle x, y, z\rangle$. Now observe:

$$\begin{aligned} w &\neq \langle x, y, z\rangle \\ &= \langle\langle x, w, x\rangle, y, z\rangle \\ &= \langle x, \langle w, y, z\rangle, \langle x, y, z\rangle\rangle \\ &= \langle x, w, \langle x, y, z\rangle\rangle \\ &\in [w, x]_M \cap [w, \langle x, y, z\rangle]_M \cap [x, \langle x, y, z\rangle]_M \\ &= \{w, x\} \cap \{w, \langle x, y, z\rangle\} \cap [x, \langle x, y, z\rangle]_M \\ &\subset \{w\} \end{aligned}$$

That is a contradiction, so $[y, z] \subset [y, z]_M$.

Conversely, suppose $x \in [y, z]_M$. Take any minimal length chain $y = w_0, \ldots, w_n = z$. Let $i$ be the first index such that $\langle y, x, w_i\rangle = x$. Such an $i$ exists because $x \in [y, z]_M$, so $\langle y, x, z = w_n\rangle = x$. Similarly, let $j$ be the last index such that $\langle x, z, w_j\rangle = x$. By Lemma 1.3.33, $w_0, \ldots, w_i$ projects via $\pi_{[y,x]_M}$ to a chain in $[y, x]_M \subset [y, z]_M$ from $y$ to $x$, and $w_j, \ldots, w_n$ projects via $\pi_{[x,z]_M}$ to a chain in $[x, z]_M \subset [y, z]_M$ from $x$ to $z$. Therefore, we have that there exists a chain in $[y, z]_M$ from $y$ to $z$ through $x$, given by:

$$y = \pi_{[y,x]_M}(w_0), \ldots, \pi_{[y,x]_M}(w_{i-1}), x, \pi_{[x,z]_M}(w_{j+1}), \ldots, \pi_{[x,z]_M}(w_n) = z$$

Since $x \in [y, z]_M$ was arbitrary, every vertex in $[y, z]_M$ lies on some minimal length chain from $y$ to $z$. This shows $[y, z]_M \subset [y, z]$.

We have built a simple connected graph whose vertex set is the median algebra, and whose intervals match the median intervals. The fact that it is a median graph now follows from (19). □

In light of these results we can drop the subscript $M$ from the median intervals, and use $\langle\cdot\rangle$ for the median in a median graph.

1.3.5. *Median algebras vs pocsets.*

THEOREM 1.3.34. *The set of DCC ultrafilters of a discrete, finite width pocset form a discrete median algebra, where the median $\mu$ is defined democratically: if $\alpha$, $\beta$, $\gamma$ are ultrafilters then $\mu(\alpha, \beta, \gamma)$ chooses $A \in \mathcal{P}$ if a majority of $\alpha$, $\beta$, and $\gamma$ do.*

PROOF. First we show that $\mu(\alpha, \beta, \gamma)$ as described is a DCC ultrafilter. The 'choice' condition of ultrafilters is satisfied, since for any $A \in \mathcal{P}$ at least two of $\alpha$, $\beta$, and $\gamma$ agree on $A$, hence also on $A^*$. 'Consistency' is satisfied because if, say, $A \in \alpha \cap \beta$ and $A < B$ then consistency of $\alpha$ and $\beta$ implies $B \in \alpha \cap \beta$, so $B \in \mu(\alpha, \beta, \gamma)$.

If $A_1 > A_2 > \cdots$ is a descending chain in $\mu(\alpha, \beta, \gamma)$ then the DCC for $\alpha$ says only finitely many of the $A_i$ are in $\alpha$, so there is a cofinite descending subchain consisting of elements in $\beta \cap \gamma$, but the DCC for $\beta$ then says the remaining chain is finite. Thus, $\mu(\alpha, \beta, \gamma)$ is a DCC ultrafilter.

The map $\mu$ is clearly symmetric, and for $\mu(\alpha, \alpha, \beta)$, $\alpha$ always wins the vote, so $\mu(\alpha, \alpha, \beta) = \alpha$. The associativity median axiom is true because $\mu(\alpha, \delta, \mu(\beta, \delta, \gamma)$ and $\mu(\mu(\alpha, \delta, \beta), \delta, \gamma)$ can both be expressed as follows: For $A \in \mathcal{P}$ the four ultrafilters $\alpha$, $\beta$, $\gamma$, and $\delta$ vote; if there is a majority then majority rules; if there is a tie then the vote of $\delta$ is decisive.

Consider DCC ultrafilters $\alpha$ and $\gamma$. By Lemma 1.2.14, $\operatorname{Diff}(\alpha, \gamma)$ is finite. We claim this gives a bound $\#[\alpha, \gamma] \leqslant 2^{\# \operatorname{Diff}(\alpha, \gamma)}$ on the size of their median interval $[\alpha, \gamma] := \{\beta \mid \mu(\alpha, \beta, \gamma) = \beta\}$. How can $\mu(\alpha, \beta, \gamma) = \beta$ for every $A \in \mathcal{P}$? $\beta$ must be on the winning side of every vote, so if $\alpha$ and $\gamma$ agree on $A$ then $\beta$ has to go along with the majority. When $\alpha$ and $\gamma$ disagree then $\beta$ is free to be the swing vote either way. Thus the number of possible such $\beta$ are bounded above by the number of distinct ways of voting on the elements of $\operatorname{Diff}(\alpha, \gamma)$. □

Proposition 1.3.35. *If $(\mathcal{P}, <, *)$ is a discrete, finite width pocset then the graph of the pocset $\Gamma(\mathcal{P})$ of Definition 1.2.10 is isomorphic to the median graph produced by applying Proposition 1.3.30 to the discrete median algebra of Theorem 1.3.34.*

Proof. In both cases the vertex set of the graph is the set of DCC ultrafilters of $\mathcal{P}$. In the construction of $\Gamma(\mathcal{P})$ two ultrafilters $\alpha \neq \omega$ are connected by an edge if and only if $\operatorname{Diff}(\alpha, \omega) := \{A \in \mathcal{P} \mid A \in \alpha \text{ and } A^* \in \omega\}$ consists of a single element. In the median algebra, we saw in the proof of Theorem 1.3.34 that $\beta \in [\alpha, \omega]$ when $\alpha$, $\beta$, and $\omega$ are unanimous off of $\operatorname{Diff}(\alpha, \omega) \cup \operatorname{Diff}(\alpha, \omega)^*$. Flipping a minimal element of $\operatorname{Diff}(\alpha, \omega)$ gives such a $\beta$, so $\alpha$ and $\omega$ are adjacent in the median graph if and only if $[\alpha, \omega] = \{\alpha, \omega\}$, which is true if and only if $\# \operatorname{Diff}(\alpha, \omega) = 1$. □

In Section 1.3.7 we will show:

Proposition 1.3.36. *If $\Gamma$ is a median graph and $(\mathcal{P}, <, *)$ is the pocset of median halfspaces, then $\Gamma(\mathcal{P})$ is isomorphic to $\Gamma$.*

This says the surprising behavior in Example 1.2.9, where there are DCC ultrafilters that do not arise from vertices of the graph, does not occur for median graphs. Looking back at Example 1.2.9, the new vertex '100' that appeared was necessary precisely to be the median of vertices '000', '101', '110', which did not have a median in the original graph.

Before proving Proposition 1.3.36 we have to explain what we mean by 'median halfspace'.

1.3.6. *Median halfspaces and walls.*

DEFINITION 1.3.37. A *(median) halfspace* $H$ of a median graph $\Gamma$ is a convex subgraph such that the full subgraph $H^*$ on the set of complementary vertices $\Gamma^{(0)} - H^{()}$ is also convex. A *(median) wall* in $\Gamma$ is a pair $\{H, H^*\}$ of complementary halfspaces.

PROPOSITION 1.3.38. *For each edge in a median graph there is a unique wall separating its two vertices.*

COROLLARY 1.3.39. *In a median graph, the graph distance is equal to the wall distance.*

Before proving Proposition 1.3.38 we need an improved version of Proposition 1.3.16.

DEFINITION 1.3.40. The *join* of subsets $A$ and $B$ is $[A, B] := \bigcup_{a \in A,\, b \in B} [a, b]$. This is the union of all geodesics between a point of $A$ and a point of $B$.

LEMMA 1.3.41. *The join of convex subgraphs of a median graph is convex.*

PROOF. Let $A$ and $B$ be convex subgraphs of a median graph, and let $C := [A, B]$. Let $a_1, a_2 \in A$, and let $b_1, b_2 \in B$. Let $c_1 \in [a_1, b_1]$ and $c_2 \in [a_2, b_2]$. Let $c_3 \in [c_1, c_2]$. We must show $c_3 \in C$.

Consider $c_1' := \langle a_1, b_1, c_3 \rangle$. It is still in $[a_1, b_1]$, and by Proposition 1.3.18 we have $c_1 \to c_1' \to c_3$. Similarly, for $c_2' := \langle a_2, b_2, c_3 \rangle$ we have $c_3 \to c_2' \to c_2$. But we also have $c_1 \to c_3 \to c_2$, so we get $c_1 \to c_1' \to c_3 \to c_2' \to c_2$, and we see that $c_3 \in [c_1', c_2']$. Thus, we might as well assume that $c_1 = c_1'$ and $c_2 = c_2'$.

Now define $a_3 := \langle a_1, a_2, c_3\rangle$ and $b_3 := \langle b_1, b_2, c_3\rangle$. By convexity, $a_3 \in A$ and $b_3 \in B$. Compute:

$$\begin{aligned}
\langle a_3, b_3, c_3\rangle &= \langle\langle a_1, a_2, c_3\rangle, \langle b_1, b_2, c_3\rangle, c_3\rangle \\
&= \langle\langle a_1, c_3, a_2\rangle, c_3, \langle b_2, c_3, b_1\rangle\rangle \\
&= \langle a_1, c_3, \langle a_2, c_3, \langle b_2, c_3, b_1\rangle\rangle \\
&= \langle a_1, c_3, \langle\langle a_2, c_3, b_2\rangle, c_3, b_1\rangle\rangle \\
&= \langle a_1, c_3, \langle b_1, c_3, \langle a_2, b_2, c_3\rangle\rangle\rangle \\
&= \langle\langle a_1, c_3, b_1\rangle, c_3, \langle a_2, b_2, c_3\rangle\rangle \\
&= \langle\langle a_1, b_1, c_3\rangle, \langle a_2, b_2, c_3\rangle, c_3\rangle \\
&= \langle c_1, c_2, c_3\rangle \\
&= c_3
\end{aligned}$$

Thus, $c_3$ is on a geodesic between $a_3 \in A$ and $b_3 \in B$, so it is in $C$. □

Proof of Proposition 1.3.38. Let $a$ and $b$ be adjacent vertices. By Zorn's Lemma, there exist maximal convex subgraphs containing $a$ but not $b$. Let $H$ be any such subgraph. Suppose that $H^*$ is not convex. Then there exist $x_1, x_2, x_3$ with $x_1, x_2 \in H^*$ and $x_3 \in [x_1, x_2] \cap H$. By Lemma 1.3.41, $[H, x_1]$ is a convex set properly containing $H$, so by maximality of $H$ it must also contain $b$. Thus, there is an $h_1 \in H$ such that $b \in [h_1, x_1]$. Similarly, there is an $h_2 \in H$ such that $b \in [h_2, x_2]$. Let $h_3 := \langle b, h_1, h_2\rangle$. Since $H$ is convex, $h_3 \in H$. Now, we have $x_1 \to b \to h_1$ and $b \to h_3 \to h_1$, so $x_1 \to b \to h_3 \to h_1$, and in particular, $x_1 \to b \to h_3$. Similarly, $x_2 \to b \to h_3$. By Proposition 1.3.18, $h_3 \to b \to x_3$. This is a contradiction, because $h_3$ and $x_3$ are in the convex set $H$, but $H$ does not contain $b$. Thus, $\{H, H^*\}$ is a wall separating $a$ and $b$.

Suppose there are distinct walls $\{H_1, H_1^*\}$ and $\{H_2, H_2^*\}$ separating $a$ and $b$. Without loss of generality, we assume $a \in H_1 \cap H_2$ and there exists $c \in H_1 \cap H_2^*$. Then convexity implies $\langle a, b, c\rangle \in H_2^* \cap H_1$, since $a, c \in H_1$ and $b, c \in H_2^*$. However, $a$ and $b$ are adjacent, so $[a, b] = \{a, b\}$. Thus, $\langle a, b, c\rangle$ is either equal to $a \notin H_2^*$ or $b \notin H_1$. □

The existence part of the argument of Proposition 1.3.38 can be improved:

Proposition 1.3.42 ([**21**, Theorem 2.8]). *If $A$ and $B$ are disjoint convex subgraphs of a median graph then there exist walls separating them.*

Definition 1.3.43. Given a wall $\{H, H^*\}$ in a median graph, let $\partial H$ be the full subgraph spanned by vertices of $H$ that are adjacent to $H^*$.

LEMMA 1.3.44. *Given a wall $\{H, H^*\}$ in a median graph, $\partial H$ is convex.*

PROOF. Suppose not. Then there exist $x, y \in \partial H$ with $[x, y] \not\subset \partial H$. Assume that $d(x, y)$ is minimal among all such pairs.

By definition of $\partial H$, $x$ has neighbors in $H^*$. In fact, by Proposition 1.3.21, since $H^*$ is convex there is a unique closest point of $H^*$ to $x$, so $x$ has a unique neighbor $a$ in $H^*$, and likewise $y$ has a unique neighbor $b$ in $H^*$.

For all $z \in [x, y]$ we estimate:

$$\begin{aligned} d(z, H^*) &\leqslant d(z, \langle a, b, z\rangle) \\ &= d(\langle x, y, z\rangle, \langle a, b, z\rangle) \\ &\leqslant d(\langle x, y, z\rangle, \langle a, y, z\rangle) + d(\langle a, y, z\rangle, \langle a, b, z\rangle) \\ &\leqslant d(x, a) + d(y, b) \quad \text{by Lemma 1.3.33} \\ &= 2 \end{aligned}$$

Pick a geodesic $\gamma$ from $x$ to $y$ that contains a vertex not in $\partial H$. By minimality of $d(x, y)$, we may assume $\gamma \cap \partial H = \{x, y\}$.

For any $z \in \gamma - \{x, y\}$, let $c := \langle a, b, z\rangle$. By the above estimate, $d(c, z) = 2$. There is a unique vertex $w$ between $c$ and $z$: since $d(c, z) = 2$ and $d(z, H^*) > 1$, $w$ is a closest point of $H$ to $c$.

Consider $\langle x, z, w\rangle$ and $\langle y, z, w\rangle$. Since $d(w, z) = 1$, both of these points are in $\{w, z\}$. They cannot both be $w$, or else $d(x, y) \leqslant d(x, w) + d(w, y) = d(x, z) - 1 + d(z, y) - 1 = d(x, y) - 2$. Suppose $\langle x, z, w\rangle = z$. Then $z \in [x, w]$ with $x, w \in \partial H$, so $d(x, y) \leqslant d(x, w)$ implies $d(y, z) = 1$. But now $d(z, c) = d(z, b) = 2$ realize $d(z, H^*)$, so $b = c$, and then $d(b, w) = d(b, y) = 1$ realize $d(b, H)$, so $w = y$. Similarly, if $\langle y, z, w\rangle = z$ then $w = x$.

So every interior vertex of $\gamma$ is adjacent to either $x$ or $y$. The only possibilities that do not contradict geodesicity are that $\gamma$ has length 2 or 3. But length 2 has the same problem as in the previous paragraph: it implies $a$ and $b$ both realize $d(z, H^*)$, so $a = b$, and it follows that $x = y$. So $\gamma$ has length 3. We will show that this also leads to a contradiction.

Let the vertices of $\gamma$ be $x, z_1, z_2, y$. Consider $\langle x, z_2, b\rangle$. It is in $H$, by convexity, but there is a unique geodesic, of length 2, from $b$ to $z_2$, so it is either $z_2$ or $y$. On the other hand, $z_2$ is on a geodesic from $x$ to $y$. Thus, $\langle x, z_2, b\rangle = z_2$. This implies $4 = d(x, b) = d(x, a) + d(a, b)$, so $d(a, b) = 3$. This gives a contradiction, because $d(z_1, b) \leqslant 3$ via $z_1 \to y \to b$, but $a$ is the gate of $H^*$ for $z_1$, so $d(z_1, b) = d(z_1, a) + d(a, b) = 5$. □

LEMMA 1.3.45. *Given a vertex $v$ in a median graph, there is a bijection between edges at $v$ and halfspaces $H$ with $v \in \partial H$.*

PROOF. By Proposition 1.3.38 every edge incident to $v$ corresponds to a wall for which $v$ is in the boundary of a halfspace. We only need to show that no two of these coincide. Suppose that two different edges belong to the same wall, and take the opposite vertices. Both of these are at distance 1 from $v$, realizing the distance from $v$ to $H^*$. But $H^*$ is convex, so by Proposition 1.3.21 it has a unique closest point to $v$. □

LEMMA 1.3.46. *If $v$ is a vertex in a median graph, define:*

$$\alpha_v := \{\textit{halfspaces containing } v\}$$

*Then a halfspace $H \in \alpha_v$ is minimal if and only if $\{H, H^*\}$ is the wall corresponding to one of the edges incident to $v$, which, by Lemma 1.3.45, is equivalent to $v \in \partial H$.*

PROOF. Let $H \in \alpha_v$. By Lemma 1.3.44, $\partial H$ is convex. If $v \notin \partial H$ then by Proposition 1.3.42 there is a wall $\{A, A^*\}$ separating $\{v\}$ and $\partial H$. Assume $v \in A$ and $\partial H \subset A^*$. Then $A \in \alpha_v$ and $H^* < A^*$, which implies $A < H$, so $H$ is not minimal in $\alpha_v$.

Conversely, if $v \in \partial H$ let $e = [v, w]$ be the edge incident to $v$ separated by $\{H, H^*\}$. If $H \neq A \in \alpha_v$ then by Proposition 1.3.38, $\{A, A^*\}$ does not separate $e$, so $w \in A \cap H^*$, which means $A \nless H$. Thus, $H$ is minimal in $\alpha_v$. □

1.3.7. *Compatibility of some of the constructions.* We can now prove Proposition 1.3.36, which claimed that if $\Gamma$ is a median graph and $(\mathcal{P}, <, *)$ is the pocset of median halfspaces, then $\Gamma(\mathcal{P})$ is isomorphic to $\Gamma$.

PROOF OF PROPOSITION 1.3.36. Let $v$ be a vertex of $\Gamma$. It is easy to see that $\alpha_v := \{H \mid H \text{ is a halfspace containing } v\}$ is a DCC ultrafilter. We will show that $v \mapsto \alpha_v$ induces an isomorphism $\Gamma \to \Gamma(\mathcal{P})$.

First, this map is injective on vertices, since if $v$ and $w$ are distinct vertices then by Proposition 1.3.42 there is a wall $\{H, H^*\}$ separating them, so $\alpha_v$ and $\alpha_w$ disagree on $H$.

More specifically, by Proposition 1.3.38, $v$ and $w$ are adjacent in $\Gamma$ if and only if they are separated by exactly one wall, which is the same as saying $\#\operatorname{Diff}(\alpha_v, \alpha_w) = 1$. By construction, this is equivalent to $\alpha_v$ and $\alpha_w$ being adjacent in $\Gamma(\mathcal{P})$.

This shows $v \to \alpha_v$ is an embedding of $\Gamma$ into $\Gamma(\mathcal{P})$. We still have to see that the map is surjective on vertices. Choose a vertex $v \in \Gamma$, and suppose $\omega$ is a DCC ultrafilter adjacent to $\alpha_v$ in $\Gamma(\mathcal{P})$. By construction, there is a single wall $\{H, H^*\}$ on which $\alpha_v$ and $\omega$ differ, and, assuming $H \in \alpha_v$ and $H^* \in \omega$, we have that $H$ is minimal in $\alpha_v$ and $H^*$ is minimal in $\omega$. By Lemma 1.3.46, $v \in \partial H$. Let $w \in H^*$ be adjacent to $v$. Then

$\mathrm{Diff}(\alpha_v, \alpha_w) = \{H\} = \mathrm{Diff}(\alpha_v, \omega)$, so $\omega = \alpha_w$. Since $\Gamma(\mathcal{P})$ is connected, by Theorem 1.2.11, this implies that every DCC ultrafilter comes from a vertex of $\Gamma$. □

We also note that the wall structure of a median graph, which at this point is formally defined but might be hard to visualize, is exactly the wall structure coming from hyperplanes in $\mathrm{Cube}(\Gamma)$.

Proposition 1.3.47. *Let $\Gamma$ be a median graph. There is a bijection between walls of $\Gamma$ and hyperplanes of $\mathrm{Cube}(\Gamma)$.*

Proof. Recall that a hyperplane of $X := \mathrm{Cube}(\Gamma)$ is dual to an equivalence class of edges, where the equivalence relation $R_X$ on edges is generated by the condition that opposite edges of a square are related.

We can also define an equivalence relation $R_\Gamma$ on edges of $\Gamma$ by saying that two edges are equivalent if is some wall of $\Gamma$ that crosses both.

We will show these two equivalence relations are the same.

Let $e$ be an edge of $\Gamma$. By Proposition 1.3.38, there is a unique wall $\{H, H^*\}$ separating its vertices. Suppose there is an embedded 4–cycle $[v_0, v_1]$, $[v_1, v_2]$, $[v_2, v_3]$, $[v_3, v_0]$ in $\Gamma$ with $e = [v_0, v_1]$. We assume $v_0 \in H$ and $v_1 \in H^*$.

No-triangles implies $\Gamma$ does not contain a diagonal of this 4–cycle, so we have $v_3 \to v_0 \to v_1$. This implies $v_3 \in H$, because if $v_3 \in H^*$ and $v_1 \in H^*$ then convexity of $H^*$ would imply $v_0 \in H^*$. We also have $v_2 \to v_1 \to v_0$, which implies $v_2 \in H^*$, because if $v_2 \in H$ and $v_0 \in H$ then convexity of $H$ would imply $v_1 \in H$. This implies that if two edges are $R_X$-related then they are $R_\Gamma$–related.

Now suppose $\{H, H^*\}$ is a wall and $v_0, v_3 \in \partial H$ are adjacent.

By definition, there exist $v_1, v_2 \in \partial H^*$ adjacent to $v_0$ and $v_3$, respectively, and no-triangles implies $v_1 \neq v_2$. By Lemma 1.3.44, $\partial H^*$ is convex, so the path from $v_1$ to $v_2$ through $v_0$ and $v_3$ is not geodesic, which implies $d(v_1, v_2) < 3$. It cannot be 2 because that would give an odd-length cycle, so it must be 1. Thus, for every edge $[v_0, v_3] \subset \partial H$ there is a corresponding edge $[v_1, v_2] \subset \partial H^*$ such that $v_0$, $v_1$, $v_2$, $v_3$ are the vertices of a square, with the wall transverse to both elements of the pair of opposite sides $[v_0, v_1]$ and $[v_2, v_3]$. So, if two edges $[v, w]$ and $[x, y]$ are $R_\Gamma$–related, with $v, x \in H$ and $w, y \in H^*$, then by convexity of $\partial H$ there is an path in $\partial H$ with vertices $v = v_0, v_1, \ldots, v_n = x$, and, by the above, a path $w = v'_0, v'_1, \ldots, v'_n = y$ in $\partial H^*$ such that for each $i$ there is an embedded 4–cycle $v_i$, $v'_i$, $v'_{i+1}$, $v_{i+1}$. Thus $[v_i, v'_i]$ and $[v_{i+1}, v'_{i+1}]$ are $R_X$–related, which, by transitivity, implies $[v, w]$ and $[x, y]$ are $R_X$–related. □

COROLLARY 1.3.48. *If $X$ is a finite dimensional CAT(0) cube complex and $(\mathcal{P}, <, *)$ is the discrete, finite width pocset of halfspaces of hyperplanes of $X$, then $X$ is isomorphic to the CAT(0) cube complex* $\mathrm{Cube}(\Gamma(\mathcal{P}))$.

PROOF. It is enough to show that $X^{(1)}$ is isomorphic to $\Gamma(\mathcal{P})$, since the cube complex is determined by its 1–skeleton. By Theorem 1.3.14, $X^{(1)}$ is a median graph, and by Proposition 1.3.47 the wall structure in the median sense is the same as the wall structure in the CAT(0) sense, so they give the same pocset of halfspaces on the vertices. Proposition 1.3.36 says going from median graph to pocset of halfspaces back to median graph is an isomorphism. □

**1.4. Cubing a non-cubical complex.** It is immediate from the construction that if a group acts on a wall structure on a graph then it acts on the cube complex of the pocset of halfspaces. The action may not be as nice as one might hope. In particular, it might not be geometric.

EXAMPLE 1.4.1. Consider the Coxeter group with Coxeter graph an unlabelled triangle. We have seen this is a 2–dimensional Euclidean reflection group, so its Davis complex is homeomorphic to the plane. All proper subsets of the generators a spherical. The maximal special subgroup are all isomorphic to $\mathcal{D}_3$, so their Coxeter cell is a Euclidean hexagon, which we can choose to be regular, so we can take the Davis complex to be a regular hexagonal tessellation of $\mathbb{E}^2$. The 1–skeleton is the Cayley graph of the group, which forms a reflection system. In fact, looking back at Example 2.3.1, a wall crosses a hexagon in a pair of opposite edges, so we get three families of parallel walls, at slopes 0, $\pi/3$, and $-\pi/3$. See Figure 4.

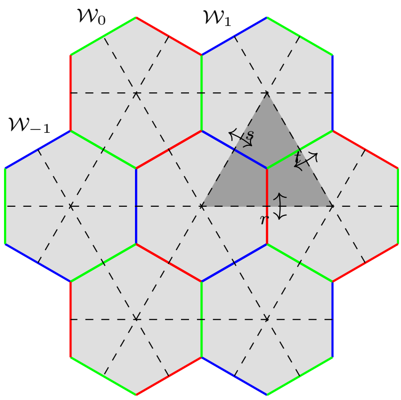



FIGURE 4. $\Sigma$ for $\Delta(3,3,3)$ is a hexagonal tessellation of $\mathbb{E}^2$, with fundamental domain $K$ shaded.

Consider the pocset for this wall structure. Walls are transverse if and only if they have different angle, so the width of the pocset is 3. It is also discrete, since it is build from a wall system on a graph. What does the corresponding cube complex look like?

Any three walls of different angle give a 3–cube. Fix two walls, say $\Omega^r$ and $\Omega^s$, the ones flipped by $r$ and $s$, which have slope 0 and $\pi/3$, respectively. Each of these is transverse to *every* wall of slope $-\pi/3$. Let's number them. There is a wall of slope $-\pi/3$ through $\Omega^r \cap \Omega^s$; call it $\mathcal{W}_0$. Let $\mathcal{W}_1 := \Omega^t$. Number the others sequentially, so that we have a $\mathcal{W}_i$ of slope $-\pi/3$ for all $i \in \mathbb{Z}$.

For all $i \in \mathbb{Z}$, $\Omega^r$, $\Omega^s$, and $\mathcal{W}_i$ are pairwise transverse, so they correspond to a 3–cube in the dual cube complex. All of these 3–cubes have a common face corresponding to the transverse pair $\Omega^r$, $\Omega^s$, so the cube complex is not locally finite.

Since the group action preserves the wall structure, there is an induced group action on the cube complex. However, the induced action is not cocompact. To see this, look at the action on triples of walls $\{\Omega^r, \Omega^s, \mathcal{W}_i\}$. They cut out an equilateral triangle in the plane with side length a linear function of $|i|$. The group acts by isometries, so $\{\Omega^r, \Omega^s, \mathcal{W}_i\}$ and $\{\Omega^r, \Omega^s, \mathcal{W}_j\}$ can only possibly be in the same orbit when $|i| = |j|$. This means that of the $\mathbb{Z}$–many distinct 3–cubes corresponding to $\{\Omega^r, \Omega^s, \mathcal{W}_i\}$, there are at least $\mathbb{N}$–many distinct ones in the quotient, all containing edges incident to a common vertex. So the quotient is not locally finite.

**1.5. Helly, Ramsey, and Dilworth.** In this section we take two theorems from combinatorics and apply them to hyperplanes in CAT(0) cube complexes to see that a quasiconvex set is close to its convex hull. This is from Hagen [**16**].

First, here is a property of median graphs/CAT(0) cube complexes.

THEOREM 1.5.1 (Helly Property). *Let $C_1, \dots, C_n$ be a collection of pairwise intersecting convex subgraphs of a median graph. Then $\bigcap_{i=1}^n C_i \neq \varnothing$.*

PROOF. There is nothing to prove for $n \leqslant 2$, so suppose $n \geqslant 3$. Suppose that the statement is true for collections of size at most $n-1$. Then for each $i$ there exists a point $x_i \in \bigcap_{j\in\{1,\dots,n\}-\{i\}} C_j$.

Consider $[x_i, x_j]$. The point $x_i$ is in every $C_k$ except possibly $C_i$, and similarly the point $x_j$ is in every $C_k$ except possibly $C_j$, so for $k \neq i, j$, $x_i$ and $x_j$ are both in $C_k$. Since $C_k$ is convex, $[x_i, x_j] \subset C_k$. Thus, $[x_i, x_j] \subset \bigcap_{k\in\{1,\dots,n\}-\{i,j\}} C_k$.

Let $m := m(x_1, x_2, x_n) = [x_1, x_2] \cap [x_2, x_n] \cap [x_n, x_1]$. Since $m \in [x_1, x_2]$, $m$ is in every $C_k$ except possibly $C_1$ and $C_2$. Since $m \in [x_2, x_n]$, $m$ is in every $C_k$ except possibly $C_2$ and $C_n$, so it is in $C_1$. Since $m \in [x_n, x_1]$, $m$ is in every $C_k$ except possibly $C_n$ and $C_1$, so it is in $C_2$. $\square$

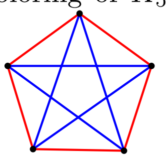
COROLLARY 1.5.2. *If $X$ is a CAT(0) cube complex, collections of halfspaces have the Helly property, as do collections of hyperplanes.*

Now the combinatorics.

THEOREM 1.5.3 (Ramsey's Theorem). *For all $r, b \in \mathbb{N}$ there is a* Ramsey number $\mathrm{Ram}(r, b)$, *a least positive integer such that if $K$ is a complete graph with at least* $\mathrm{Ram}(r, b)$ *vertices, then for every bicoloring (red/blue) of the edges either there is a red clique of size at least $r$ or a blue clique of size at least $b$.*

EXAMPLE 1.5.4. $\mathrm{Ram}(3, 3) = 6$. Suppose $K$ is the complete graph on at least 6 vertices. Let $v_0$ be some vertex, and let $v_1, \ldots, v_5$ be any five of its neighbors. Among the five edges $[v_0, v_1], \ldots, [v_0, v_5]$, at least 3 of them have a common color. Let us assume that edges $[v_0, v_1]$, $[v_0, v_2]$, and $[v_0, v_3]$ are red. If any one of the edges $[v_1, v_2]$, $[v_2, v_3]$, or $[v_1, v_3]$ are red then we would have a red triangle through $v_0$. If all three of them are blue then they form a blue triangle. Thus $\mathrm{Ram}(3, 3) \leqslant 6$.

The following bicoloring of $K_5$ has no monochromatic triangle, which



shows $\mathrm{Ram}(3, 3) > 5$.

The next theorem is about finite partially ordered sets. Recall that a chain in a partially ordered set is a totally ordered subset. An *antichain* is a subset such that no two elements are comparable in the partial order.

THEOREM 1.5.5 (Dilworth's Theorem). *In any finite partially ordered set the size of the largest antichain is equal to the size of the smallest partition of the set into chains.*

Now let $X$ be a finite dimensional CAT(0) cube complex. Let $\mathcal{W}$ be some finite collection of hyperplanes. Let $\overrightarrow{\mathcal{W}} = \{\overrightarrow{w} \mid w \in \mathcal{W}\}$ be a choice of one of the halfspaces for each hyperplane, and for $w \in \mathcal{W}$ let $\overleftarrow{w}$ be the halfspace for $w$ that is not in $\overrightarrow{\mathcal{W}}$. Then $\overrightarrow{\mathcal{W}}$ is partially ordered by inclusion. Let $\mathcal{A} \subset \overrightarrow{\mathcal{W}}$ be an antichain. Consider the complete graph with vertex set $\mathcal{A}$, and color the edges as follows: Color edge $[\overrightarrow{w}_i, \overrightarrow{w}_j]$ *red* if $w_i$ and $w_j$ are crossing hyperplanes. Color edge $[\overrightarrow{w}_i, \overrightarrow{w}_j]$ *blue* if $w_i$ and $w_j$ are disjoint.

Suppose this graph has a red clique. By construction, the vertices of this clique correspond to a collection of pairwise crossing hyperplanes. By the Helly property, these hyperplanes have a common point of intersection, so all of them contain a midcube of some common cube of $X$. A cube has as many midplanes as its dimension, so the size of a red clique is bounded by $\dim X$.

Now suppose this graph has a blue clique $\{\overrightarrow{w}_i\}_{i \in I}$. Suppose $|I| \geqslant 3$, and let $i, j, k \in I$ be distinct. Suppose $w_i \subset \overleftarrow{w}_j$ and $w_k \subset \overrightarrow{w}_j$. Since $\mathcal{A}$ is an antichain, $\overrightarrow{w}_j \not\subset \overrightarrow{w}_i$, so $\overrightarrow{w}_i \subset \overleftarrow{w}_j$. Similarly, $\overrightarrow{w}_j \not\supset \overrightarrow{w}_k$, so $\overrightarrow{w}_k \supset \overleftarrow{w}_j$. This implies $\overrightarrow{w}_i \subset \overrightarrow{w}_k$, which is a contradiction. We conclude that $w_j$ cannot separate $w_i$ from $w_k$.

DEFINITION 1.5.6. A *facing tuple* is a collection of hyperplanes that are disjoint from one another and such that no one of them separates any pair of the others.

PROPOSITION 1.5.7. *Let $X$ be a finite dimensional CAT(0) cube complex, let $\overrightarrow{\mathcal{W}}$ be a finite collection of halfspaces, and let $F$ be the maximum size of a facing tuple in $\mathcal{W}$. Then $\overrightarrow{\mathcal{W}}$ contains a chain of length strictly greater than* $\frac{|\overrightarrow{\mathcal{W}}|}{\mathrm{Ram}(1+\dim X, 1+F)}$.

PROOF. Let $\overrightarrow{\mathcal{W}} = \coprod_{i=1}^{C} \overrightarrow{\mathcal{W}}_i$ be a smallest partition of $\overrightarrow{\mathcal{W}}$ into chains. By Dilworth's Theorem, there is an antichain $\mathcal{A} \subset \overrightarrow{\mathcal{W}}$ of size $C$. Make a bicolored complete graph from $\mathcal{A}$ as above. We have seen that the size of the largest red clique is at most $\dim X$ and the size of the largest blue clique is at most $F$. By Ramsey's Theorem, the graph must be small: $|\mathcal{A}| < \mathrm{Ram}(1+\dim X, 1+F)$. The average size of a chain in the partition of $\overrightarrow{\mathcal{W}}$ is $\frac{|\overrightarrow{\mathcal{W}}|}{C} = \frac{|\overrightarrow{\mathcal{W}}|}{\mathcal{A}} > \frac{|\overrightarrow{\mathcal{W}}|}{\mathrm{Ram}(1+\dim X, 1+F)}$. At least one chain has at least average length. $\square$

COROLLARY 1.5.8. *Let $\gamma$ be a finite geodesic edge path in a finite dimensional CAT(0) cube complex $X$. Let $\mathcal{W}$ be the set of hyperplanes crossed by $\gamma$. For any choice of $\overrightarrow{\mathcal{W}}$, $\overrightarrow{\mathcal{W}}$ contains a chain of length strictly greater than* $\frac{|\gamma|}{\mathrm{Ram}(1+\dim X, 3)}$.

PROOF. Every edge of $\gamma$ crosses a wall, and, since $\gamma$ is a geodesic, it crosses each wall at most once, so $|\gamma| = |\mathcal{W}|$. Suppose $\mathcal{W}$ contains a facing triple $\{w_i, w_j, w_k\}$. Since they are disjoint, $\gamma$ crosses them in a well-defined order, which we assume is alphabetical. But $w_j$ does not separate $w_i$ from $w_k$, because they are a facing triple, so $w_i$ and $w_k$ are on the same size of $w_j$. This is a contradiction, because it means that $\gamma$ first crosses $w_i$, then $w_j$, and finds itself on the opposite side of $w_j$ from $w_k$. To cross $w_k$ it would have to first cross back over $w_j$. This it cannot do, since a geodesic can cross a wall at most once. Thus $F \leqslant 2$.

By Proposition 1.5.7, any choice of $\overrightarrow{\mathcal{W}}$ contains a chain of length greater than $\frac{|\overrightarrow{\mathcal{W}}|}{\mathrm{Ram}(1+\dim X, 3)} = \frac{|\gamma|}{\mathrm{Ram}(1+\dim X, 3)}$. $\square$

THEOREM 1.5.9. *For every $D$ and $Q$, let $R := Q \cdot \mathrm{Ram}(1+D,3)$. If $X$ is a CAT(0) cube complex of dimension at most $D$ and $Z \subset X$ is $Q$–quasiconvex then $\mathcal{H}(Z) \subset N_R(Z)$.*

PROOF. Let $x \in \mathcal{H}(Z)$, so no hyperplane separates $x$ from all of $Z$. Let $y \in Z$ be a closest point of $Z$ to $x$. Let $\gamma$ be a geodesic from $x$ to $y$. By Corollary 1.5.8, there exists a chain $\{w_1, w_2, \dots, w_n\}$ of hyperplanes crossed by $\gamma$, for $n > |\gamma|/\mathrm{Ram}(1+D,3)$. Let $\overrightarrow{w}_i$ be the halfspace of $w_i$ containing $x$. Since no wall separates $x$ from all of $Z$, there exists $z \in Z \cap \overrightarrow{w}_1$.

Consider the median $m = m(x,y,z) = [x,y] \cap [y,z] \cap [x,z]$. Since $x, z \in \overrightarrow{w}_1$, which is convex, $m \in [x,z] \subset \overrightarrow{w}_1$.

Since $m \in [x,y]$, and $y$ is a closest point of $Z$ to $x$, $y$ is also a closest point of $Z$ to $m$, so $d(m,Z) = d(m,y)$. Since $\{w_1, \dots, w_n\}$ is a chain with $m \in \overrightarrow{w}_1$ and $y \in \overleftarrow{w}_n$, $d(m,y) \geqslant n$, so $d(m,Z) \geqslant n$.

Since $y$ and $z$ are both in $Z$, which is $Q$–quasiconvex, $m \in [y,z] \subset N_Q(Z)$.

We now have $Q > d(m,Z) \geqslant n > |\gamma|/\mathrm{Ram}(1+D,3)$. □

## 2. More robust versions of convexity

We mentioned in the previous section that quasiconvexity is a useful property in hyperbolic spaces, but not as well behaved in CAT(0) spaces. Let us see why.

EXAMPLE 2.0.1. Consider the combinatorial metric on the Euclidean plane tessellated by unit squares, making it a CAT(0) square complex. Identify the vertices with the integer lattice. In this metric vertical and horizontal lines are convex geodesics, but geodesics tracking diagonal lines are not even quasiconvex. For example, consider the main diagonal $\{(x,x) \mid x \in \mathbb{Z}\}$. These points lie on a (actually, on many) combinatorial geodesic. But for $a < b$ there is also a geodesic from $(a,a)$ to $(b,b)$ consisting of a vertical line from $(a,a)$ to $(a,b)$ followed by a horizontal line from $(a,b)$ to $(b,b)$, whose maximum distance to the diagonal is $b-a$.

So in the combinatorial metric on a CAT(0) cube complex, even geodesics may fail to be quasiconvex, and there are quasiisometries taking convex geodesic to non-quasiconvex geodesics (rotation, in this example). In the CAT(0) metric geodesics are of course convex, since there is a unique geodesic between any two points, but the latter problem can still occur.

EXERCISE 2.0.2 (logarithmic spiral). Show that $\phi \colon \mathbb{E}^2 \to \mathbb{E}^2$ given in polar coordinates by $(r,\theta) \mapsto (r, \theta + \log(1+r))$ is a quasiisometry. Show that the convex geodesic ray $r \mapsto (r,0)$ is sent by $\phi$ to a non-quasiconvex set.

The essential problem is that convexity is defined in terms of *geodesics*, but geodesics are not, in general, well-behaved under quasiisometries. A quasiisometry sends a geodesic to a quasigeodesic, by definition, but, as in Exercise 2.0.2, a quasigeodesic might be fairly wild without some stronger condition on the space like hyperbolicity.

A solution to making a version of convexity that is well-behaved under quasiisometry is to insist that all quasigeodesics stay close to the subspace. Obviously, one has less control on quasigeodesics with worse quasigeodesic constants, so 'close' has to take those constants into account. This version of quasigeodesic quasiconvexity is usually called the *Morse property.*

DEFINITION 2.0.3. Let $\mu\colon \mathbb{R}^2 \to \mathbb{R}$ be a function. A subset $Z$ of a geodesic metric space $X$ is *$\mu$–Morse* if for every $L$ and $A$, every $(L, A)$–quasigeodesic segment $\gamma$ with endpoints on $Z$ stays in the $\mu(L, A)$–neighborhood of $Z$.

Say $Z$ is *Morse* if there exists a $\mu$ such that $Z$ is $\mu$–Morse.

EXERCISE 2.0.4. In the plane with Euclidean metric, estimate quasigeodesic constants for the map

$$\gamma\colon [0, 3s] \to X : t \mapsto \begin{cases} (0, t) & 0 \leqslant t \leqslant s \\ (t - s, s) & s \leqslant t \leqslant 2s \\ (s, 3s - t) & 2s \leqslant t \leqslant 3s \end{cases}$$

Conclude that the $x$–axis is not Morse.

LEMMA 2.0.5 (The Morse Lemma). *Geodesics in hyperbolic spaces are Morse.*

EXERCISE 2.0.6. Let $\phi\colon X \to Y$ be a quasiisometry between geodesic metric spaces. Let $Z$ be a Morse subset of $X$. Show that $\phi(Z)$ is Morse in $Y$.

A consequence of the exercise is that it makes sense, if $H$ is a subgroup of a finitely generated group $G$, to say that $H$ is a *Morse subgroup*, or not, because this property does not depend on the choice of geometric model for $G$.

Here is a different formulation of the Morse property:

DEFINITION 2.0.7. Let $Z < X$ be a closed set. Let $\pi_Z\colon X \to Z$ be closest point projection. Let $\rho\colon \mathbb{R}^+ \to \mathbb{R}^+$ be a non-decreasing, sublinear function, that is, $\lim_{r\to\infty} \frac{\rho(r)}{r} = 0$. Say $Z$ is *$\rho$–contracting* if for all $x, y \in X$ we have:

$$d(x, y) \leqslant d(x, Z) \implies \operatorname{diam} \pi_Z(x) \cup \pi_Z(y) \leqslant \rho(d(x, Z))$$

Say $Z$ is *sublinearly contracting* if it is $\rho$–contracting for some sublinear function $\rho$. Say $Z$ is *strongly contracting* if there is a constant $C$ such that $Z$ is $\rho$–contracting for the (sublinear!) constant function $\rho \equiv C$.

In words, for any point $x \notin Z$, take the biggest ball possible around $x$ that just barely meets $Z$. Project that ball to $Z$. Asymptotically, the diameter of the projection is negligible compared to the diameter of the ball.

THEOREM 2.0.8 ([**2**]). *If $X$ is a geodesic metric space and $Z < X$ then $Z$ is Morse if and only if it is sublinearly contracting.*

EXAMPLE 2.0.9. Let $Z$ be a geodesic in a tree $X$. Let $x \notin Z$. Let $B = \{y \in X \mid d(x,y) \leqslant d(x,Z)\}$. Then $\pi_Z(B) = \pi_Z(x)$ is a single point, so $Z$ is 0–contracting.

EXERCISE 2.0.10. Consider a geodesic $Z$ in the Poincarè Disc model of $\mathbb{H}^2$, which, up to isometry, we may suppose is the $x$–axis. Consider any point not on the geodesic, which, up to isometry, we may suppose lies on the positive $y$ axis. Show that every hyperbolic ball about $y$ is contain in the Euclidean ball $B$ centered at $(0, 1/2)$ of Euclidean radius $1/2$. Show that $\pi_Z(B)$ is a compact interval, and conclude that $Z$ is strongly contracting.

EXERCISE 2.0.11. Show that given $Q$ and $\delta$ there exists $C$ such that if $Z$ is a $Q$–quasiconvex subset of a $\delta$–hyperbolic space then $Z$ is $C$–contracting.

A $\mu$–Morse set is $Q$–quasiconvex for $Q := \mu(1,0)$, so the exercise together with Theorem 2.0.8 shows that Morse is equivalent to strongly contracting in hyperbolic spaces. That is why our examples have all been strongly contracting. It turns out that the same is true in CAT(0) spaces [**7**] and in finite dimensional CAT(0) cube complexes with respect to the combinatorial metric [**13**]. Thus, for our purposes it will be enough to consider strong contraction.

EXERCISE 2.0.12. Let $G$ be a group generated by a finite set $S$. Let $H$ be a subgroup of $G$. Show that $H$ is finitely generable if and only if the subset of vertices of $Cay(G,S)$ corresponding to elements of $H$ is coarsely connected, in the sense that there is some $R$ such that any two elements of $H$ can be connect by a sequence $h_1, h_2, \ldots, h_n$ in $H$ with $d(h_i, h_{i+1}) \leqslant R$ in $Cay(G,S)$.

EXERCISE 2.0.13. Show that if $H$ is quasiconvex in $Cay(G,S)$ then $H$ is finitely generable.

EXERCISE 2.0.14. Suppose $K < H < G$ with $G$ finitely generable. Show that if $K$ is Morse in $H$ and $H$ is Morse in $G$ then $K$ is Morse in $G$. (You may assume that it suffices to consider discrete quasigeodesics, meaning those whose domain is the integral points of an interval in $\mathbb{R}$.)

EXERCISE 2.0.15.

- Give an example of finitely generable groups $K < H < G$ with $K$ Morse in $G$, $H$ isometrically embedded in $G$, but $K$ not Morse in $G$.
- Give an example of finitely generable groups $K < H < G$ with $H$ Morse in $G$ but $K$ not Morse in $G$.

EXERCISE 2.0.16. Show that strong contraction is equivalent to the following: There exist $A$ and $B$ such that $d(x,y) \leqslant d(x,Z) - A \implies \operatorname{diam} \pi_Z(x) \cup \pi_Z(y) \leqslant B$.

Here are two further variations on the Morse property that will appear in the next section:

DEFINITION 2.0.17. A subspace $Z$ of a geodesic metric space $X$ is *stable* if it both Morse in $X$ and is itself hyperbolic.

'Morse' only describes how $Z$ sits in $X$, not the intrinsic geometry of $Z$. In a hyperbolic space Morse implies quasiconvex implies hyperbolic, so Morse sets are automatically stable. In more general spaces this is no longer true. For example, in $\mathbb{Z}^2 * \mathbb{Z}$ the $\mathbb{Z}^2$ factor is Morse but not hyperbolic, so not stable.

DEFINITION 2.0.18. A subspace $Z$ of a geodesic metric space $X$ is *eccentric* if it is minimally Morse unstable, in the following sense:

- It is Morse.
- It is not stable.
- Given $\mu$ there exists $\epsilon$ such that if $Z' < Z$ is $\mu$–Morse in $X$ and not stable then $d_{Haus}(Z', Z) \leqslant \epsilon$.

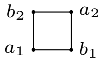
Here the *Hausdorff distance* between two sets is:

$$d_{Haus}(A, B) := \inf\{r \mid A \subset \mathcal{N}_r(B) \text{ and } B \subset \mathcal{N}_r(A)\}$$

## 3. Morse, stable, and eccentric subspaces of RACGs

'Morse' means quasigeodesic segments have to stay close to a subset. A subset $Z$ is not Morse if there are detours in $X$ that allow one to travel between points of $Z$ without taking too much longer than staying in $Z$. We had one example that was the $x$–axis in the plane, where we found that it was possible to travel between points by going up, over, and down, at the cost of being a quasigeodesic with some uniform quasigeodesic constants. This motivates the following construction:

PROPOSITION 3.0.1. *Let $(W, S)$ be a right-angled Coxeter group with presentation graph $\Upsilon$. Suppose $T \subset S$ contains opposite vertices of a full square in $\Upsilon$, but not the whole square. Then $W_T$ is not Morse.*

PROOF. Let $\square$ = 

be a full square in $\Upsilon$ with $a_1, a_2 \in T$ and $b_2 \notin T$. Since the square is full there is no edge in $\Upsilon$ between $a_1$ and $a_2$, so $W_{\{a_1,a_2\}} \cong \mathcal{D}_\infty$ and, by Theorem 2.0.1, special subgroups are convex, so $\Sigma_{\{a_1,a_2\}} = \mathbb{R}$ is a geodesic in $\Sigma$. The same is true for $\{b_1, b_2\}$, and for $\square$ we have that $W_\square \cong \mathcal{D}_\infty \times \mathcal{D}_\infty$ and $\Sigma_\square$ is a convex, square-tessellated copy of $\mathbb{E}^2$ in $\Sigma$.

**Case 1: $b_1 \notin T$.** Then $\Sigma_\square \cap \Sigma_T = \Sigma_{\{a_1,a_2\}}$ is a geodesic, and $\Sigma_{\{b_1,b_2\}}$ is a geodesic that intersects $\Sigma_T$ only at 1. Furthermore, every edge $e$ of $\Sigma_{\{b_1,b_2\}}$ belongs to two squares, one whose transverse sides are labelled $a_1$ and the other whose transverse sides are labelled $a_2$. In both of these the side opposite $e$ has the same label as $e$, so we see a strip of squares with vertical sides all labelled like $e$, and horizontal sides alternating $a_1$, $a_2$. There is a hyperplane dual to $e$ that cuts this strip in half, separating adjacent translates of $\Sigma_T$. Since the wall distance equals the graph distance, this implies $\Sigma_{\{b_1,b_2\}}$ and $\Sigma_T$ are orthogonal.

For each $n \in \mathbb{N}$ we have a path $1 \to (b_2b_1)^n + (b_2b_1)^n \to (b_2b_1)^n(a_1a_2)^n + (b_2b_1)^n(a_1a_2)^n \to (a_1a_2)^n$ from 1 to $(a_1a_2)^n$ that is a $(3, 0)$–quasigeodesic and gets $2n$–far away from $\Sigma_T$. Thus, $\Sigma_T$ is not Morse. This is illustrated in Figure 5.

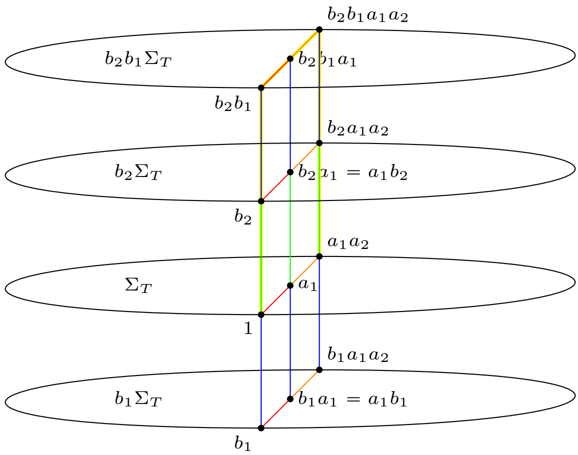



FIGURE 5. Case 1 of non-square complete implies non-Morse

**Case 1: $b_1 \in T$.** The picture is similar except that only every other edge in $\Sigma_{\{b_1,b_2\}}$ is orthogonal to its neighboring cosets of $\Sigma_T$. See Figure 6. $\square$

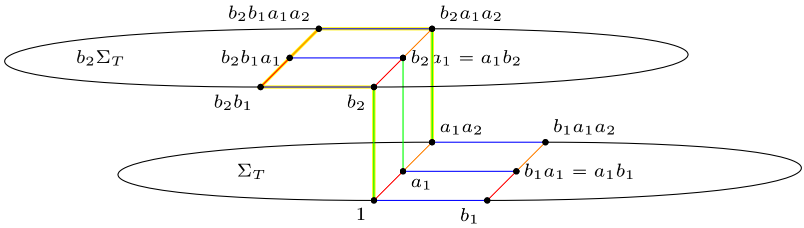



FIGURE 6. Case 2 of non-square complete implies non-Morse

In fact, Proposition 3.0.1 is an 'if and only if', but the other direction will take some work.

DEFINITION 3.0.2. An induced subgroup of $\Upsilon$ is *square complete* if whenever it contains opposite vertices of an induced square it contains the whole square.

THEOREM 3.0.3 ([**25**, Theorem 1.11],[**13**, Proposition 4.9]). *A special subgroup of a right-angled Coxeter group is Morse if and only if its presentation subgraph is square complete.*

DEFINITION 3.0.4. A *grid of hyperplanes* is a pair $(\mathcal{H}\mathcal{V})$ of chains of hyperplanes $\mathcal{H} = \{H_i\}_{i\in I}$ and $\mathcal{V} = \{V_j\}_{j\in J}$ such that every $H_i$ crosses every $V_j$.

One should imagine a grid of squares in the planes, with $\mathcal{V}$ being the vertical hyperplanes and $\mathcal{H}$ being the horizontal hyperplanes.

DEFINITION 3.0.5. If $X$ is a CAT(0) cube complex and $Y$ is subcomplex, let $\mathfrak{H}(Y)$ be the set of hyperplanes of $X$ that intersect $Y$.

THEOREM 3.0.6 (cf [**13**, Proposition 4.5]). *Let $X$ be a finite dimensional CAT(0) cube complex. Consider the combinatorial metric on $X^{(1)}$. Let $Y$ be a gated subgraph of $X^{(1)}$. The following are equivalent:*

(1) *$Y$ is strongly contracting.*
(2) *$Y$ is Morse.*
(3) *There exists $C_3$ such that for every flat square $[0,r]\times[0,r]\subset X$ with $[0,r]\times\{0\}\subset Y$ we have $[0,r]\times\{r\}\subset\mathcal{N}_{C_3}Y$.*
(4) *There exists $C_4$ such that if $(\mathcal{H},\mathcal{V})$ is a grid of hyperplanes with $\mathcal{V}\subset\mathfrak{H}(Y)$ and $\mathcal{H}\cap\mathfrak{H}(Y)=\varnothing$ then* $\min\{\#\mathcal{V},\#\mathcal{H}\}<C_4$.

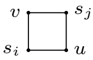
PROOF. We have already mentioned that (2) $\iff$ (1) in CAT(0) cube complexes. Not (3) $\implies$ not (2) is the same argument as Proposition 3.0.1, since big squares that get arbitrarily far from $Y$ would provide efficient detours. Thus, (2) $\implies$ (3).

Suppose we have a grid of hyperplanes as in (4). [**12**, Theorem 2.7] derives from the grid a flat rectangle of width at least $a := \#\mathcal{V} - 2$ and height at least $b := \#\mathcal{H} - 1$ with its base on $Y$. Consider a square of sidelength $r := \min\{a, b\}$ contained in the rectangle, with base on $Y$. Since $\mathcal{H}$ is a chain, the top of the square is separated from the base, and from $Y$, by each horizontal hyperplane that meets it, so the top of the square is not in the $(r-1)$–neighborhood of $Y$. Thus, $C_3 \geqslant r \geqslant \min\{\#\mathcal{V}-2, \#\mathcal{H}-1\}$, or, conversely, $\min\{\#\mathcal{V}, \#\mathcal{H}\} \leqslant C_3 + 2$. Thus, (3) implies (4) for $C_4 := C_3 + 2$.

Now assume (4). By Corollary 1.5.8, there is a constant $D$ depending only on the dimension of $X$ such that from a collection of hyperplanes crossing a geodesic one may extract a chain consisting of at least $1/D$ fraction of them. Define $A := C_4D$ and $B := C_4D$. Suppose $x$ and $x'$ are points of $X$ with $d(x, x') \leqslant d(x, Y) - A$ and $\operatorname{diam} \pi_Y(x) \cup \pi_Y(x') > B$. Since the wall distance is equal to the graph distance, there are at least $A$–many more hyperplanes separating $x$ and $Y$ than there are separating $x$ and $x'$, so there are at least $A$–many hyperplanes separating $\{x, x'\}$ from $Y$. There exists a chain $\mathcal{H}$ in that set consisting of at least $A/D = C_4$ many hyperplanes. There are more than $B$–many hyperplanes separating $\pi_Y(x)$ from $\pi_Y(x')$, so we can choose a chain $\mathcal{V}$ of these of size greater than $B/D = C_4$. Furthermore, $(\mathcal{H}, \mathcal{V})$ is a grid, as follows. Consider a geodesic from $x$ to $\pi_Y(x)$. Each edge crosses a hyperplane that separates $x$ from $Y$, otherwise there would be a closer point of $Y$ to $x$ than $\pi_Y(x)$. So $\mathfrak{H}([x, \pi_Y(x)]) \cap \mathfrak{H}(Y) = \varnothing$. The same is true for $x'$ and $\pi_Y(x')$. So for each $V \in \mathcal{V}$, $\pi_Y(x)$ and $\pi_Y(x')$ are on opposite sides of $V$, but $x$ and $\pi_Y(x)$ are on the same side and $x'$ and $\pi_Y(x')$ are on the same side. Thus, $x$ and $x'$ are on opposite sides. For $H \in \mathcal{H}$, $x$ and $x'$ are on the same side and $\pi_Y(x)$ and $\pi_Y(x')$ are both on the opposite side. Thus, $V$ and $H$ cross, as we have found points in all four possible intersections of complementary halfspaces. This grid contradicts (4), so $d(x, x') \leqslant d(x, Y) - A$ implies $\operatorname{diam} \pi_Y(x) \cup \pi_Y(x') \leqslant B$. By Exercise 2.0.16, this implies (1). ☐

Proof of Theorem 3.0.3. One direction was Proposition 3.0.1. In the other, suppose $\Sigma_T$ is not Morse. By Theorem 3.0.6, there are arbitrarily large square grids $(\mathcal{H}, \mathcal{V})$ with $\mathcal{V} \subset \mathfrak{H}(Y)$ and $\mathcal{H} \cap \mathfrak{H}(Y) = \varnothing$. Take one of size $n$ larger than the size of the largest clique in $\Upsilon$. We may assume the bottom left corner of the square is at the vertex 1. Let $s_1$, $s_2$, …be the labels on the edges dual to $H_i \in \mathcal{H}$. Then we have $s_1 \notin T$ and $s_i \neq s_{i+1}$.

Suppose all of the $s_i$ commute. Then they cannot all be distinct, since there are more of them than the largest clique in $\Upsilon$. But if they all commute and $s_i = s_j$ then the word $s_1 \cdots s_n$ is not geodesic, contradicting that we have a flat square. Thus, there exists a minimal $i$ such that there is an index $j$ for which $s_i$ and $s_j$ do not commute. By minimality of $i$, $s_i$ and $s_k$ do commute for all $k < i$. Consider the hyperplane $\mathcal{W}_{s_i}$ through the edge $[1, s_i]$. Now, a reflection through one midcube in a cube fixes all of the other midcubes, so all $s \in S$ that commute with $s_i$ fix $\mathcal{W}_{s_i}$. Now, $H_i = s_1 s_2 \cdots s_{i-1} \mathcal{W}_{s_i} = \mathcal{W}_{s_i}$, so $\mathcal{W}_{s_i}$ does not cross $\Sigma_T$, so $s_i \notin T$.

Similarly, find edges along the bottom of the square labelled $u$ and $v$ that do not commute. The fact that these all came from a grid means $s_i$ and $u$ commute, $s_i$ and $v$ commute, $s_j$ and $u$ commute, and $s_j$ and $v$ commute. Thus, there is a full square 

$\subset \Upsilon$ such that $u, v \in T$ and $s_i \notin T$. This means $\Upsilon_T$ is not square complete. □

Corollary 3.0.7. *A special subgroup of a right-angled Coxeter group is stable if and only if its presentation subgraph is square complete and contains no square.*

Example 3.0.8 ([**3**]). Consider the right-angled Coxeter group defined by the presentation graph in Figure 7.

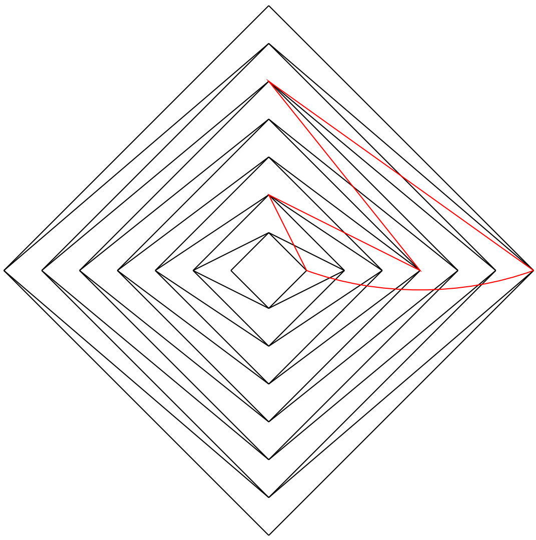


Figure 7. A 'CFS' graph with a stable surface subgroup

There are lots of squares in this picture. In fact, every vertex lies on a square that can be connected to a square containing any other vertex through a sequence of squares that share a diagonal, which makes it something called a 'CFS graph'. $\Sigma$ has lots of intersecting copies of $\mathbb{E}^2$, which we might expect makes it very non-hyperbolic. However, the red subgraph is a 5–cycle that is square complete. Let $T$ be the vertex set of the red subgraph. Then $W_T$ is square complete and square-free, so it is a stable subgroup. It is also an $\mathbb{H}^2$ reflection group, so $\Sigma_T$ is a stable, quasiisometrically embedded copy of $\mathbb{H}^2$ in $\Sigma$.

DEFINITION 3.0.9. Let $\Upsilon$ be a simple graph. A *minsquare* subgraph is a full subgraph of $\Upsilon$ that:

- is square complete,
- contains a square,
- and is minimal with respect to inclusion among subgraphs of $\Upsilon$ satisfying the first two conditions.

THEOREM 3.0.10 ([**14**]). *Let $(W, S)$ be a right-angled Coxeter group with presentation graph $\Upsilon$. Let $\Sigma$ be the Davis complex of $(W, S)$. Every eccentric subspace of $\Sigma$ is at bounded Hausdorff distance from some $g\Sigma_T$ where $\Upsilon_T$ is a minsquare subgraph of $\Upsilon$.*

REMARK. It is not true in general that a Morse subspace is close to a Morse subgroup. Both minimality and non-hyperbolicity are important in this theorem.

COROLLARY 3.0.11. *The set of quasiisometry types of special subgroup defined by a minsquare subgraph of the presentation graph is a quasiisometry invariant for right-angled Coxeter groups.*

This means, if $W_1$ and $W_2$ are right-angled Coxeter groups defined by presentation graphs $\Upsilon_1$ and $\Upsilon_2$, and if $\Upsilon_1$ contains a minsquare subgraph $\Upsilon_1'$ such that there is no minsquare subgraph of $\Upsilon_2$ that generates a group quasiisometric to $W_{\Upsilon_1'}$, then $W_1$ and $W_2$ are not quasiisometric.

# Bibliography


[1] Daniel Allcock, *Infinitely many hyperbolic Coxeter groups through dimension 19*, Geom. Topol. **10** (2006), 737–758, doi:10.2140/gt.2006.10.737. MR 2240904

[2] Goulnara N. Arzhantseva, Christopher H. Cashen, Dominik Gruber, and David Hume, *Characterizations of Morse geodesics via superlinear divergence and sublinear contraction*, Doc. Math. **22** (2017), 1193–1224, https://www.math.uni-bielefeld.de/documenta/vol-22/36.html.

[3] Jason Behrstock, *A counterexample to questions about boundaries, stability, and commensurability*, Beyond hyperbolicity, London Math. Soc. Lecture Note Ser., vol. 454, Cambridge Univ. Press, Cambridge, 2019, pp. 151–159. MR 3966609

[4] Garrett Birkhoff and S. A. Kiss, *A ternary operation in distributive lattices*, Bull. Amer. Math. Soc. **53** (1947), 749–752, doi:10.1090/S0002-9904-1947-08864-9. MR 21540

[5] Brian H. Bowditch, *Median algebras*, July 2022 draft (2022), http://www.warwick.ac.uk/~masgak/papers/median-algebras.pdf.

[6] Martin R. Bridson and André Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der mathematischen Wissenschaften, vol. 319, Springer, Berlin, 1999. MR 1744486 (2000k:53038)

[7] Christopher H. Cashen, *Morse subsets of CAT(0) spaces are strongly contracting*, Geom. Dedicata **204** (2020), no. 1, 311–314, doi:10.1007/s10711-019-00457-x.

[8] Indira Chatterji and Graham Niblo, *From wall spaces to* CAT(0) *cube complexes*, Internat. J. Algebra Comput. **15** (2005), no. 5-6, 875–885, doi:10.1142/S0218196705002669. MR 2197811

[9] Victor Chepoi, *Graphs of some* CAT(0) *complexes*, Adv. in Appl. Math. **24** (2000), no. 2, 125–179, doi:10.1006/aama.1999.0677. MR 1748966

[10] H. S. M. Coxeter, *Discrete groups generated by reflections*, Ann. of Math. (2) **35** (1934), no. 3, 588–621, doi:10.2307/1968753. MR 1503182

[11] Michael W. Davis, *The geometry and topology of Coxeter groups*, London Mathematical Society Monographs Series, vol. 32, Princeton University Press, Princeton, NJ, 2008. MR 2360474

[12] Anthony Genevois, *Algebraic characterisation of relatively hyperbolic special groups*, preprint (2017), arXiv:1709.01258.

[13] Anthony Genevois, *Hyperbolicities in CAT(0) cube complexes*, preprint (2017), arXiv:1709.08843v2.

[14] Anthony Genevois, *Quasi-isometrically rigid subgroups in right-angled Coxeter groups*, preprint (2019), arXiv:1909.04318.

[15] C. McA. Gordon, D. D. Long, and A. W. Reid, *Surface subgroups of Coxeter and Artin groups*, J. Pure Appl. Algebra **189** (2004), no. 1-3, 135–148, doi:10.1016/j.jpaa.2003.10.011. MR 2038569

[16] Mark Hagen, *Large facing tuples and a strengthened sector lemma*, Tunisian J. Math. **4** (2022), 55–86, DOI:10.2140/tunis.2022.4.55.

[17] Tetsuya Hosaka, *Parabolic subgroups of finite index in coxeter groups*, Journal of Pure and Applied Algebra **169** (2002), no. 2, 215–227, DOI:https://doi.org/10.1016/S0022-4049(01)00068-8.

[18] John R. Isbell, *Median algebra*, Trans. Amer. Math. Soc. **260** (1980), no. 2, 319–362, DOI:10.2307/1998007. MR 574784

[19] Daan Krammer, *The conjugacy problem for Coxeter groups*, Groups Geom. Dyn. **3** (2009), no. 1, 71–171, DOI:10.4171/GGD/52. MR 2466021

[20] Roland K. W. Roeder, *Compact hyperbolic tetrahedra with non-obtuse dihedral angles*, Publ. Mat. **50** (2006), no. 1, 211–227, DOI:10.5565/PUBLMAT_50106_12. MR 2325019

[21] Martin A. Roller, *Poc-sets, median algebras and group actions: An extended study of Dunwoody's construction and Sageev's theorem*, Habilitationsschrift, University of Regensburg, 1998, arXiv:1607.07747.

[22] Michah Sageev, *Ends of group pairs and non-positively curved cube complexes*, Proc. London Math. Soc. (3) **71** (1995), no. 3, 585–617, DOI:10.1112/plms/s3-71.3.585. MR MR1347406 (97a:20062)

[23] Michah Sageev, CAT(0) *cube complexes and groups*, Geometric group theory, IAS/Park City Math. Ser., vol. 21, Amer. Math. Soc., Providence, RI, 2014, pp. 7–54, DOI:10.1090/pcms/021/02. MR 3329724

[24] David E. Speyer, *Powers of Coxeter elements in infinite groups are reduced*, Proc. Amer. Math. Soc. **137** (2009), no. 4, 1295–1302, DOI:10.1090/S0002-9939-08-09638-X. MR 2465651

[25] Hung Tran, *On strongly quasiconvex subgroups*, Geom. Topol. **23** (2019), no. 3, 1173–1235, DOI:10.2140/gt.2019.23.1173. MR 3956891