GROMOV'S SIMPLICIAL NORM AND BOUNDED COHOMOLOGY

LVZHOU CHEN

Abstract. These are lecture notes for the course *Gromov's Simplicial Norm and Bounded Cohomology* in Spring 2022 at the University of Texas at Austin.

Contents

Date: May 6, 2022.

1. Introduction to Gromov's simplicial norm

One interesting topic in geometry and topology is to relate geometric quantities of a manifold to topological invariants. One typical problem asks about which manifolds admit Riemannian metrics with negative (positive, or non-positive) sectional (Ricci, or scalar) curvature.

Here we are interested in volumes of closed manifolds. Usually one needs a Riemannian metric to make sense of it, but is it possible to get a topological invariant out of it? The Mostow rigidity (and Gauss–Bonnet in dimension 2) implies that the volume of a hyperbolic closed manifold is determined by its topology. Gromov's *simplicial volume*, as a special case of the *simplicial norm*, is a way to define this invariant in a purely topological way.

Why should one be interested in such an invariant? The following basic problem is an example where one needs a topological notion of volume/area.

Problem 1.1. *Given two orientable connected closed surfaces S, S', what is the largest possible degree* $\deg(f)$ *of a continuous map $f : S \to S'$?*

As we will see below (Lemma 1.11), the simplicial volumes of S and S', denoted $\|S\|_1$ and $\|S'\|_1$, satisfy
$$\|S\|_1 \geq |\deg(f)| \cdot \|S'\|_1$$
for any continuous map f. Intuitively, S needs to have enough area to cover S' for $|\deg(f)|$ times. This provides an upper bound $\|S\|_1/\|S'\|_1$ when $\|S'\| > 0$, or equivalently when S' has genus at least two as we will prove. Moreover, the upper bound obtained this way is actually sharp, and in Section 1.6 we will exactly determine the set of all possibly degrees
$$\deg(S, S') := \{\deg(f) \mid f : S \to S'\}.$$

1.1. The simplicial norm.

Fix $n \in \mathbb{Z}_{\geq 0}$. Given a topological space X, Gromov [Gro82] introduced a semi-norm $\|\cdot\|_1$ on the singular homology $H_n(X;\mathbb{R})$ for each n as a real vector space to measure the size of each homology class. Recall that $H_n(X;\mathbb{R})$ is the homology of the singular chain complex
$$\cdots \xrightarrow{\partial_{n+2}} C_{n+1} \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_n} C_{n_1} \xrightarrow{\partial_{n-1}} \cdots,$$
where $C_n(X;\mathbb{R})$ is the space of singular n-chains, namely the real vector space spanned by the set $S_n(X)$ of all singular n-simplices. As usual, we have the subspaces $B_n \subset Z_n \subset C_n$, where $Z_n := \ker \partial_n$ and $B_n := \operatorname{Im} \partial_{n+1}$ are the spaces of cycles and boundaries respectively. So by definition $H_n(X;\mathbb{R})$ is the quotient Z_n/B_n.

Given the standard basis $S_n(X)$, equip the space $C_n(X;\mathbb{R})$ with the ℓ^1-norm, i.e. $|c|_1 = \sum_{i=1}^k |\lambda_i|$ for any $c = \sum_{i=1}^k \lambda_i c_i$ expressed uniquely as a (finite) linear combination of basis elements $c_i \in S_n(X)$ with coefficients $\lambda_i \in \mathbb{R}$.

Definition 1.2 (Simplicial norm)**.** The restriction of this ℓ^1-norm to Z_n induces a semi-norm on its quotient $H_n(X;\mathbb{R})$, explicitly,
$$\|\sigma\|_1 := \inf_{[c]=\sigma} |c|_1,$$
where the infimum is taken over all cycles $c \in Z_n$ representing the homology class $\sigma \in H_n(X;\mathbb{R})$. This semi-norm is called *Gromov's simplicial norm.*

In words, $\|\sigma\|_1$ is the infimal number of simplices that we need to represent σ.

The following property is immediate from the definition but important.

Proposition 1.3 (Functorial)**.** *For any continuous map $f : X \to Y$, then the induced map $f_* : H_n(X;\mathbb{R}) \to H_n(Y;\mathbb{R})$ is non-increasing with respect to the simplicial norm, i.e.*
$$\|f_*\sigma\|_1 \leq \|\sigma\|_1$$
for any $\sigma \in H_n(X;\mathbb{R})$.

Proof. For any cycle $c = \sum_i \lambda_i c_i \in Z_n(X;\mathbb{R})$ representing σ, the cycle $f_*c = \sum \lambda_i f_* c_i = \sum_i \lambda_i (f \circ c_i)$ represents $f_*\sigma$. Hence by definition

$$\|f_*\sigma\|_1 \le |f_*c|_1 \le \sum_i |\lambda_i| = |c|_1.$$

Since c is arbitrary, taking infimum implies

$$\|f_*\sigma\|_1 \le \|\sigma\|_1.$$

□

Corollary 1.4 (Invariance)**.** *If $f : X \to Y$ is a homotopy equivalence, then $f_* : H_n(X;\mathbb{R}) \to H_n(Y;\mathbb{R})$ is an isometric isomorphism (i.e. an isomorphism that is norm-preserving) with respect to the simplicial norm.*

More generally, if for a map $f : X \to Y$ there is $g : Y \to X$ such that g_*f_* is the identity on $H_n(X;\mathbb{R})$, then f_* is an isometric embedding (i.e. injective and norm-preserving).

Proof. The first part easily follows from the second part by taking g to be a homotopy inverse of f.

For the second part, by functoriality of g and the fact that $g_*f_* = id$, we have $\|\sigma\|_1 = \|(g_*f_*)\sigma\|_1 \le \|f_*\sigma\|_1$. Combining with the functoriality of f, we must have $\|\sigma\|_1 = \|f_*\sigma\|_1$ for any $\sigma \in H_n(X;\mathbb{R})$. Hence f_* is norm-preserving. Injectivity easily follows from the fact that $g_*f_* = id$. □

It is often convenient to consider cycles with rational coefficients since they can be scaled to integral cycles. We can always find a rational homology class arbitrarily close to a given homology class with respect to the simplicial norm; see the lemma below. This follows from the fact that B_n and Z_n are rational subspaces. Here a point $c \in C_n(X;\mathbb{R})$ is rational if $c \in C_n(X;\mathbb{Q})$, and an $\mathbb{R}$-linear subspace is rational if it has a basis consisting of rational points. Any point in a rational subspace V is a limit of rational points in V with respect to the norm $|\cdot|_1$ (think about it). Here B_n and Z_n are rational because the boundary maps $\partial_{k+1} : C_{k+1}(X;\mathbb{R}) \to C_k(X;\mathbb{R})$ are rational linear, i.e. obtained from $C_{k+1}(X;\mathbb{Q}) \to C_k(X;\mathbb{Q})$ by tensoring with $\mathbb{R}$ over $\mathbb{Q}$.

Lemma 1.5. *If $\sigma \in H_n(X;\mathbb{Q})$, then $\|\sigma\|_1 = \inf |c|_1$ where the infimum is taken over all* rational *cycles $c = \sum \lambda_i c_i$ (i.e $\lambda_i \in \mathbb{Q}$ and $\partial c = 0$).*

For a general $\sigma \in H_n(X;\mathbb{R})$ and any $\epsilon > 0$, there is $\sigma' \in H_n(X;\mathbb{Q})$ with $\|\sigma - \sigma'\|_1 \le \epsilon$.

Proof. For the first part, note that $B_n(X;\mathbb{Q})$ is dense in $B_n(X;\mathbb{R})$ with respect to the norm $|\cdot|_1$, since $B_n(X;\mathbb{R})$ is a rational subspace. As $\sigma \in H_n(X;\mathbb{Q})$, it can be represented by some rational cycle c. All other (resp. rational) cycles take the form $c+b$ with $b \in B_n(X;\mathbb{R})$ (resp. $b \in B_n(X;\mathbb{Q})$), so the result follows by density.

The second part is due to the density of $Z_n(X;\mathbb{Q})$ in $Z_n(X;\mathbb{R})$, which holds since $Z_n(X;\mathbb{R})$ is a rational subspace. □

Exercise 1.6. *Recall that $H_0(X;\mathbb{R})$ is isomorphic to the $\mathbb{R}$-vector space with basis corresponding to the path connected components of the space X. For any path component C and a point $c \in C$, thought of as a singular 0-simplex, we have a homology class $\sigma = [c]$. Show that $\|\sigma\|_1 = 1$.*

Remark 1.7. *If A is a subspace of X, then we can define a simplicial (semi-)norm similarly on the relative homology group $H_n(X,A;\mathbb{R})$. Here one can treat $H_n(X,A;\mathbb{R})$ as the homology of the chain complex $C_n(X,A) = C_n(X)/C_n(A)$ (with the induced differentials). These vector spaces are equipped with semi-norms induced from $C_n(X)$ and thus we can define an induced semi-norm on $H_n(X,A;\mathbb{R})$ as before. When A is empty, this agrees with our definition above.*

More generally, one can analogously define simplicial norm for any normed *chain complex; see* [Fri17].

Exercise 1.8. *Concretely, we can think of $H_n(X,A;\mathbb{R}) = Z_n(X,A)/B_n(X,A)$, where $B_n(X,A) = B_n(X) \cup C_n(A)$ and $Z_n(X,A) = \partial_n^{-1} C_{n-1}(A)$, with $C_i(A)$ treated naturally as a subspace of $C_i(X)$*

for both $i = n-1, n$. Show that the semi-norm induced from this quotient agrees with the definition in the remark above.

1.2. **The simplicial volume.** Now we specialize to measure the size of an oriented connected compact manifold M with (possibly empty) boundary ∂M. Let $n = \dim M$. The orientation picks out a generator $[M] \in H_n(M, \partial M; \mathbb{Z}) \cong \mathbb{Z}$, called the *fundamental class*. We think of it as a class in $H_n(M, \partial M; \mathbb{R}) \cong \mathbb{R}$ using the map $H_n(M, \partial M; \mathbb{Z}) \to H_n(M, \partial M; \mathbb{R})$ induced by the standard inclusion $\mathbb{Z} \to \mathbb{R}$. Concretely, if M has a triangulation, then the sum of all n-simplices with compatible orientation is a cycle representing the fundamental class.

Definition 1.9 (Simplicial volume)**.** The simplicial volume of M is $\|[M]\|_1$, which we often abbreviate as $\|M\|_1$. Note that the choice of orientation does not affect the simplicial volume.

If M is non-orientable, then M has an orientable double cover N, and we define $\|M\|_1 := \|N\|_1/2$. If M is disconnected, define $\|M\|_1$ as the sum of volumes of its components.

Exercise 1.10. *If M is orientable and closed, with finitely many components N_i. Show that $\sum_i \|N_i\|_1 = \|\sum_i [N_i]\|_1$, which explains the definition above for the disconnected case.*

Recall that, for any continuous map $f : M^n \to N^n$ between oriented connected closed (occ) manifolds, the degree $\deg(f)$ is the unique integer such that $f_*[M] = \deg(f) \cdot [N]$.

Lemma 1.11. *For any continuous map $f : M^n \to N^n$ between occ manifolds, we have*

$$|\deg(f)| \cdot \|N\|_1 \leq \|M\|_1.$$

Moreover, if f is a (finite) covering map, then equality holds.

Proof. The inequality follows from functoriality (Proposition 1.3) since $\|f_*[M]\|_1 = \|\deg(f) \cdot [N]\|_1 = |\deg(f)| \cdot \|[N]\|_1$.

Let $c = \sum_i \lambda_i c_i$ be a cycle representing the fundamental class $[N]$. Each map $c_i : \Delta^n \to N$ has $d := |\deg(f)|$ lifts $\tilde{c}_i^j$ to M, $j = 1, \cdots, d$. Then $\tilde{c} = \sum_i \sum_{j=1}^d \tilde{c}_i^j$ is a cycle and clearly $f_*[\tilde{c}] = d[c] = |\deg(f)| \cdot [N] = \pm f_*[M]$. Hence $[\tilde{c}] = \pm[M]$, and $\|M\|_1 \leq |\deg(f)| \cdot |c|_1$. Since c is arbitrary, minimizing its norm gives the reversed inequality we desire. □

Corollary 1.12. *If an orientable closed connected manifold M admits a selfmap $f : M \to M$ with $|\deg(f)| > 1$, then $\|M\|_1 = 0$.*

Exercise 1.13. *Extend the lemma and corollary above to the case of manifolds with boundary.*

Example 1.14.

(1) For any sphere S^n, $n \geq 1$, we have $\|S^n\|_1 = 0$.
(2) For the n-torus $T^n = (S^1)^n$, $n \geq 1$ we have $\|T^n\|_1 = 0$.
(3) More generally, if $M = S^1 \times N$ for a closed manifold N, then $\|M\|_1 = 0$.

These properties of the simplicial volume help us understand the simplicial norm of certain homology classes.

Lemma 1.15. *For $n \geq 1$, if a homology class $\sigma \in H_n(X; \mathbb{R})$ is represented by a sphere, i.e. there is a map $f : S^n \to X$ with $f_*[S^n] = \sigma$, then $\|\sigma\|_1 = 0$.*

Proof. By functoriality and the fact that spheres (of dimension at least one) have zero simplicial volume, $\|\sigma\|_1 = \|f_*[S^n]\|_1 \leq \|S^n\|_1 = 0$. Thus $\|\sigma\|_1 = 0$. □

Corollary 1.16. *For any X, the simplicial norm $\|\cdot\|_1$ vanishes on $H_1(X; \mathbb{R})$.*

Proof. Basically, every 1-cycle is a bunch of circles and thus this should follow from Lemma 1.15. To make it precise, we use the approximation by rational cycles from Lemma 1.5 to reduce the problem to integral cycles, which is a standard trick in these topics.

By the second part of Lemma 1.5, it suffices to show that $\|\sigma\|_1 = 0$ for all rational homology classes $\sigma \in H_n(X;\mathbb{R})$. Any such σ is represented by some rational cycle c, and up to scaling, it suffices to consider the case where c is integral, i.e. $c = \sum_i n_i c_i$ for some $n_i \in \mathbb{Z} \setminus \{0\}$. Up to changing the orientation on c_i we may assume $n_i > 0$.

Now create n_i disjoint oriented segments for each c_i for all i. The fact that $\partial c = 0$ implies that we can pair the boundary points of these segments so that the endpoint of a segment s is always paired with the starting point of some segment s' so that the corresponding paths glue up in X respecting the orientations. The end result is a closed oriented 1-manifold, i.e. a disjoint union of finitely many oriented circles S^1_k indexed by k. In other words, there is a map $\varphi : \sqcup_k S^1_k \to X$ such that $\varphi_* \sum_k [S^1_k] = \sigma$. Hence by Lemma 1.15 and the triangle inequality,

$$\|\sigma\|_1 \leq \sum_k \|\varphi_*[S^1_k]\|_1 = 0,$$

so $\|\sigma\|_1 = 0$ as desired. $\square$

1.3. **Volumes of surfaces.** In this section we aim to obtain the first nontrivial examples. We have seen that the simplicial norm is boring on H_0 and vanishes on H_1. Interesting examples emerge in H_2. For orientable connected closed surfaces, we have seen in Example 1.14 that the simplicial volume vanishes when the genus is zero or one. For surfaces of higher genus, the simplicial volume is nonzero and is proportional to the Euler characteristic.

Theorem 1.17. *For any orientable connected closed surface S of genus at least two, we have $\|S\|_1 = -2\chi(S)$.*

Remark 1.18. *Note that by Gauss–Bonnet, for any hyperbolic metric, S has area $-2\pi\chi(S) = \pi\|S\|_1$, so the simplicial volume is proportional to the hyperbolic volume. The factor π is the area of the ideal hyperbolic triangle, or equivalently, the supremum of areas of all hyperbolic triangles (ideal or not). We will generalize this to higher dimension, which is referred to as Gromov's proportionality theorem.*

To combine the results for all genera, it is convenient to introduce the following χ^- notation.

Notation 1.19. For an orientable connected compact surface S, let $\chi^-(S) = \chi(S)$ if $\chi(S) \leq 0$ and let $\chi^-(S) = 0$ otherwise, i.e. we adjust $\chi(S)$ to 0 when S is a sphere or a disk. For a general orientable compact surface $S = \sqcup \Sigma_i$ with components, let $\chi^-(S) := \sum \chi^-(\Sigma_i)$. In other words, $\chi^-(S)$ is the Euler characteristic of S after deleting all components homeomorphic to spheres or disks.

Then the following theorem easily follows from Theorem 1.17 and the case of the sphere and torus.

Theorem 1.20. *For any orientable closed surface S, we have $\|S\|_1 = -2\chi^-(S)$.*

We will prove Theorem 1.17 by establishing inequalities in both directions, which involve two different kinds of ideas.

The strategy for proving $\|S\|_1 \leq -2\chi(S)$ is to construct a sequence of cycles representing the fundamental class $[S]$ approaching the optimal value. As we explained earlier, one concrete way to represent the fundamental class is to triangulate S and take the formal sum of triangles with compatible orientations.

Suppose S has a triangulation with v vertices, e edges and f faces, we know $\chi(S) = v - e + f$. The cycle described above has norm f.

Lemma 1.21. *We have $2e = 3f$, so $\chi(S) = v - \frac{f}{2}$ and $f = 2v - 2\chi(S)$. Hence $\|S\|_1 \leq 2v - 2\chi(S)$.*

Proof. Each triangle has 3 edges, each of which is shared by two triangles. $\square$

So this is close to be optimal except for the error $2v$. The best one can do here is to take a triangulation with $v = 1$, which exists.

Exercise 1.22. *For any occ surface S, there is a triangulation with a single vertex.*

The bounded error can be remedied by taking finite covers.

Lemma 1.23. *If S has genus at least one, then $\|S\|_1 \leq -2\chi(S)$.*

Proof. For any such S and any $d \in \mathbb{Z}_+$, there is a degree d cover $f : S' \to S$. Then by taking a triangulation on S' with a single vertex, we have $\|S'\| \leq$ by Lemma 1.21. Note that both χ and $\|\cdot\|_1$ are multiplicative, i.e. $d\chi(S) = \chi(S')$ and $d\|S\|_1 = \|S'\|_1$ (by Lemma 1.11). Thus we obtain

$$\|S\|_1 = \frac{\|S'\|_1}{d} \leq \frac{2 - 2\chi(S')}{d} = \frac{2 - 2d\chi(S)}{d} = \frac{2}{d} - 2\chi(S).$$

Taking $d \to \infty$, we obtain the desired inequality. $\square$

We considered above all possible ways of representing (resp. a multiple of) the fundamental class using triangulations (resp. of a finite cover).

The reversed inequality uses a technique called "straightening", which involves hyperbolic geometry. Roughly speaking, every hyperbolic triangle has area no greater than π, and the hyperbolic area of the surface is $-2\pi\chi(S)$, so one needs at least $-2\chi(S)$ triangles to cover the entire surface once. So we just need an argument to straighten an arbitrary singular cycle representing the fundamental class into one only involving hyperbolic triangles. We will explain this in more detail (Section 1.5) after a crash course on hyperbolic geometry (1.4).

1.4. **Some hyperbolic geometry.**

We give a quick introduction/review of some hyperbolic geometry, mainly to describe geodesics and isometries. A more detailed treatment can be found in [BP92], [Thu97], or any standard textbook/notes on hyperbolic geometry.

The n-dimensional hyperbolic space $\mathbb{H}^n$ ($n \geq 2$) is the unique (up to isometry) simply connected complete Riemannian manifolds with constant sectional curvature -1. There are several models for $\mathbb{H}^n$, providing different ways to view the space.

1.4.1. *The hyperboloid model.*

Consider the bilinear form $\langle x, y\rangle = x_1y_1 + \cdots + x_ny_n - x_{n+1}y_{n+1}$ on $\mathbb{R}^{n+1}$ for any $x, y \in \mathbb{R}^{n+1}$. The set $H := \{x \mid \langle x, x\rangle = -1\}$ is a hyperboloid of two sheets. The restriction of $\langle\cdot,\cdot\rangle$ on either component gives a complete Riemannian metric of constant curvature -1. The upper sheet H_+ (i.e. with $x_{n+1} > 0$) is the hyperboloid model of $\mathbb{H}^n$.

With this model, the isometry group $\mathrm{Isom}(\mathbb{H}^n)$ is identified with $\mathrm{O}^+(n, 1)$, the group of linear transformations preserving the bilinear form $\langle\cdot,\cdot\rangle$ and stabilizing the upper sheet. The isometry group acts simply transitively on the orthonormal frame bundle, i.e. given any two points $x, y \in \mathbb{H}^n$ and two orthonormal bases (i.e. two orthonormal frames) at the two points, there is a unique isometry taking the frame at x to the frame at y.

An advantage of this model is that, any k-dimensional totally geodesic subspace of $\mathbb{H}^n$ is the intersection with some linear subspace of dimension $k + 1$. In particular, bi-infinite geodesics are intersections with planes through the origin.

The linearity provides a way to take *convex combinations* of points. More precisely, given k points $p_1, \cdots, p_k \in \mathbb{H}^n = H_+ \subset \mathbb{R}^{n+1}$, any coefficients $\lambda_1, \cdots, \lambda_k \geq 0$ with $\sum_{i=1}^k \lambda_i = 1$ uniquely determine a point $p(\lambda_1, \cdots, \lambda_k) \in \mathbb{H}^n$ as the intersection of H_+ with the segment connecting the origin with $\sum_{i=1}^k \lambda_i p_i \in \mathbb{R}^{n+1}$. Apparently $p(\lambda_1, \cdots, \lambda_k)$ depends continuously on the coefficients λ_i, so this defines a continuous map $p : \Delta^{k-1} \to \mathbb{H}^n$, where $\Delta^{k-1} = \{(\lambda_1, \cdots, \lambda_k) \in \mathbb{R}^k \mid \lambda_i \geq 0, \sum \lambda_i = 1\}$ is the standard $(k-1)$-simplex. This gives a way to straighten singular simplices in $\mathbb{H}^n$: For any $c : \Delta^{k-1} \to \mathbb{H}^n$, let $p_1, \cdots, p_k$ be the image of the k vertices in the natural order, and let $\widetilde{\mathrm{str}}(c)$ be the map p defined above. This operation has the following properties which we record for reference later:

Lemma 1.24.

(1) If c_1 and c_2 agree on some face of Δ^{k-1}, then so do $\widetilde{\mathrm{str}}(c_1)$ and $\widetilde{\mathrm{str}}(c_2)$.
(2) For any isometry $g \in \mathrm{Isom}\mathbb{H}^n$, we have $\widetilde{\mathrm{str}}(g \circ c) = g \circ \widetilde{\mathrm{str}}(c)$.

Proof. The first part follows from the construction. The second part holds since the isometries in this model are linear maps, which commute with both taking convex combinations and scaling. $\square$

The straightening operation that we will introduce in Section 1.5 relies on this construction.

1.4.2. *The Poincaré ball model.* In this model, $\mathbb{H}^n$ is identified with the open unit disk $\mathbb{D}^n \subset \mathbb{R}^n$ with the metric $\frac{4ds^2}{(1-\|x\|_2^2)^2}$ (at any $x \in \mathbb{D}^n$), where ds^2 is the Euclidean metric. So the metric gets more distorted in this model compared to the Euclidean one when x is closer to the boundary.

In this model, geodesics are circular arcs perpendicular to the boundary sphere. The isometry group consists of *Möbius transformations* that preserve the unit disk.

Definition 1.25. A self-diffeomorphism f of $S^n = \mathbb{R}^n \cup \{\infty\}$ is a Möbius transformation if one of the following equivalent descriptions holds:

(1) f is a composition of inversions and reflections;
(2) f is conformal (i.e. angle preserving);
(3) f takes round spheres and hyperplanes to round spheres and hyperplanes;
(4) f is a Euclidean similarity possibly composed with an inversion, i.e. $f(x) = \lambda A i(x) + b$, where i is either the identity or an inversion, $A \in O(n)$, $\lambda > 0$, and $b \in \mathbb{R}^n$.

Here an *inversion* with respect to a round sphere $S(p,r)$ in R^n centered at p of radius r is the map on $S^n = \mathbb{R}^n \cup \{\infty\}$ given by $i(x) = p + \frac{x-p}{\|x-p\|} \cdot \frac{r^2}{\|x-p\|}$, which fixes $S(p,r)$ pointwise, swaps p and ∞, and preserves all rays from p so that $\|x-p\| \cdot \|i(x)-p\| = r^2$. The equivalence in the definition above (when $n \geq 3$) essentially follows from Liouville's theorem:

Theorem 1.26 (Liouville)**.** *A conformal diffeomorphisms f between two open subsets of $\mathbb{R}^n$ with $n \geq 3$ takes the form $f(x) = \lambda A i(x) + b$, where i is either the identity or an inversion, $A \in O(n)$, $\lambda > 0$, and $b \in \mathbb{R}^n$.*

See [BP92, Theorem A.3.7] for a detailed proof. A more geometric argument can be found here: `https://lamington.wordpress.com/2013/10/28/liouville-illiouminated/`.

This is a good model to talk about the boundary at infinity of $\mathbb{H}^n$, denoted $\partial\mathbb{H}^n$. Although $\partial\mathbb{H}^n$ is not part of the hyperbolic space $\mathbb{H}^n$, it compactifies the space (i.e. $\overline{\mathbb{H}}^n := \mathbb{H}^n \cup \partial\mathbb{H}^n$) and helps us understand geodesics and isometries. In this model, the boundary is exactly the unit sphere S^{n-1} and the compactification $\overline{\mathbb{H}}^n$ is homeomorphic to the closed unit disk $\overline{\mathbb{D}}^n$. Each bi-infinite geodesic naturally has two endpoints on the boundary which uniquely determines the geodesic. Each isometry extends to a homeomorphism on $\overline{\mathbb{H}}^n$ and in particular acts on the boundary (and the action is 2-transitive).

1.4.3. *The upper-half space model.* This model identifies $\mathbb{H}^n$ with the open upper-half space $\{x \in \mathbb{R}^n \mid x_n > 0\}$ equipped with the metric $\frac{ds^2}{x_n^2}$, where ds^2 is the Euclidean metric. Geodesics in this model are vertical lines and circular arcs perpendicular to the hyperplane $\{x_n = 0\}$. Isometries are Möbius transformations preserving the upper-half space. The boundary can be seen as the union of the hyperplane $\{x_n = 0\}$ with ∞.

1.4.4. *Isometries.* There is a classification of orientation-preserving isometries. Any isometry extends to a continuous homeomorphism on $\overline{\mathbb{H}}^n$, which is topologically a closed ball. Thus by Brouwer's fixed point theorem, each isometry must fix some point in $\overline{\mathbb{H}}^n$. Given $g \neq id \in \mathrm{Isom}^+(\mathbb{H}^n)$, there are three mutually exclusive cases:

(1) g is *elliptic* if it fixes some point in $\mathbb{H}^n$. Up to conjugation, we may assume that g fixes the origin 0 in the disk model, in which case g is conjugate to an orthogonal transformation in $\mathrm{SO}(n)$, determined by its action on the unit tangent space at 0.
(2) g is *parabolic* if it has no fixed point in $\mathbb{H}^n$ and has a unique fixed point in $\partial\mathbb{H}^n$. Up to conjugation, we may assume that g fixes ∞ in the upper-half space model, in which case we can deduce that g is conjugate to a horizontal (i.e. preserving x_n) translation on $\mathbb{R}^n$.
(3) g is *hyperbolic* if it has no fixed point in $\mathbb{H}^n$ and has two fixed points in $\partial\mathbb{H}^n$. Up to conjugation, we may assume that g fixes ∞ and 0 in the upper-half space model. Then g is the composition of a scaling with an orthogonal transformation in $\mathrm{SO}(n-1) \subset \mathrm{SO}(n)$ centered at the origin fixing the x_n axis. This axis is the unique bi-infinite geodesic preserved by g, called the *axis* of g.

1.5. **Straightening.** Let M^n be a hyperbolic manifold. We introduce a linear map $\mathrm{str} : C_k(M;\mathbb{R}) \to C_k(M;\mathbb{R})$, called *straightening*. Fix the universal covering map $\pi : \mathbb{H}^n \to M$. For any singular simplex $c : \Delta^k \to M$, pick any lift $\tilde{c} : \Delta^k \to \mathbb{H}^n$, which we straighten to $\widetilde{\mathrm{str}}(\tilde{c}) : \Delta^k \to \mathbb{H}^n$ using the construction described in Section 1.4.1. Define the straightening of c as $\mathrm{str}(c) := \pi\,\widetilde{\mathrm{str}}(\tilde{c}) : \Delta^k \to M$, which is independent of the choice of the lift since $\widetilde{\mathrm{str}}$ commutes with isometries. This extends to a linear map $\mathrm{str} : C_k(M;\mathbb{R}) \to C_k(M;\mathbb{R})$.

Lemma 1.27. *We have $|\mathrm{str}(c)|_1 \le |c|_1$ for any chain c and $\partial\mathrm{str} = \mathrm{str}\partial$. Moreover,* str *induces the identity map on $H_*(M;\mathbb{R})$.*

Proof. The straightening map does not increase the number of simplices, so $|\mathrm{str}(c)|_1 \le |c|_1$ for any chain c. It commutes with the boundary map by construction.

For any singular simplex $c : \Delta^k \to \mathbb{H}^n$, there is an obvious linear homotopy in $\mathbb{R}^{n+1}$ from c to $\widetilde{\mathrm{str}}(c)$, which scales to one on $\mathbb{H}^n$ and projects down to a homotopy on M. Based on this, one can build a chain homotopy between str and id, or directly observed that $[\mathrm{str}(c)] = [c]$ for any cycle c. □

It follows that, when computing the simplicial norm of any homology class $\sigma \in H_*(M;\mathbb{R})$ it suffices to look at cycles consisting of (straight) hyperbolic simplices. A key fact about hyperbolic simplices is that their volume has a uniform upper bound only depending on the dimension, in contrast with Euclidean simplices.

Lemma 1.28 ([BP92, Theorem C.2.1 and Lemma C.2.3])**.** *For each $n \ge 2$, let v_n be the supremum of volumes of all hyperbolic n-simplices (possibly with vertices at infinity). Then $v_2 = \pi$ and $v_n \le \frac{\pi}{(n-1)!}$.*

Remark 1.29. *A theorem of Haagerup–Munkholm*[HM81] *shows that v_n is achieved uniquely by the regular*[1] *hyperbolic ideal n-simplex.*

Proof. Any hyperbolic simplex has volume no more than some ideal hyperbolic simplex. In fact, for any hyperbolic simplex with vertices $p_0, \dots, p_n \in \mathbb{H}^n$, choose a point p in its interior. The geodesic rays from p to p_i determines a point $p_i' \in \partial\mathbb{H}^n$ for each i. The ideal hyperbolic simplex with vertices $p_0', \cdots, p_n'$ contains the starting one.

When $n = 2$, all ideal hyperbolic triangles are conjugate up to an isometry and have the same area π, so $v_2 = \pi$. The area can be computed explicitly in the upper-half space model, say for the ideal triangle with vertices $-1, 1, \infty$, or can be seen by Gauss–Bonnet.

For $n \ge 3$, we show $v_n \le \frac{v_{n-1}}{n-1}$ by a nice computation in the upper-half space model following [BP92, Lemma C.2.3], which implies the bound $\frac{\pi}{(n-1)!}$ by induction. Let σ be any ideal hyperbolic n-simplex in the upper-half space model and put one of its vertex as ∞. The remaining n vertices form

[1] completely symmetric in the sense that the isometry group is the symmetric group on the $n+1$ vertices

an ideal hyperbolic $(n-1)$-simplex τ sitting on a totally geodesic subspace X of dimension $n-1$ not containing ∞ as a boundary point. So X is a round hemisphere centered at some $p \in \mathbb{R}^{n-1} \subset \partial \mathbb{H}^n$. Up to an isometry we may assume $p = 0$ and the sphere has radius 1. So X is the upper unit hemisphere. The vertical projection to $\mathbb{R}^{n-1}$ restricts to a homeomorphism from X to the unit closed ball in $\mathbb{R}^{n-1}$, and let s be the inverse. Explicitly, $s(x) = (x, h(x))$ where $h(x) = \sqrt{1 - \|x\|^2}$.

Let τ_0 be the image of τ under this projection, which is a Euclidean simplex with vertices on the unit sphere. Then

$$\begin{aligned}\mathrm{vol}(\sigma) &= \int_{\tau_0} \int_{h(x)}^{\infty} \frac{dydx}{y^n} \\ &= \frac{1}{n-1} \int_{\tau_0} \frac{dx}{h(x)^{n-1}}.\end{aligned}$$

It suffices to show that

$$\int_{\tau_0} \frac{dx}{h(x)^{n-1}} \leq \mathrm{vol}(\tau).$$

Note that s gives a way to parameterize τ using τ_0, so we just need to show that the pullback of the volume form is dominates $\frac{dx}{h(x)^{n-1}}$ for each x in the unit ball, where we think of dx as the standard volume form on $\mathbb{R}^{n-1}$.

For any x, let ω be the hyperbolic volume form restricted to $T_{s(x)}X$. Note that ω evaluates to $\frac{1}{h(x)^{n-1}}$ for any orthonormal basis of $T_{s(x)}X$. To find out its pullback, choose an orthonormal basis of $T_x\mathbb{R}^{n-1}$. If $x = 0$, then $Dh = 0$ so s_* is the identity map and $s^*\omega = \frac{dx}{h(x)^{n-1}}$. If $x \neq 0$, we may choose the orthonormal basis so that one of them is $e_1 = \frac{x}{\|x\|}$. Then $Dh = 0$ in all directions perpendicular to e_1 and $Dh(e_1) = \frac{\|x\|}{h(x)}$, so s_* is the identity map in the subspace perpendicular to e_1, and it takes this orthonormal basis to an orthogonal basis where all elements have length 1 except that $\|s_*(e_1)\|^2 = 1 + \frac{\|x\|^2}{1-\|x\|^2} = \frac{1}{1-\|x\|^2} = \frac{1}{h(x)^2}$. Thus $s^*\omega = \frac{1}{h(x)} \cdot \frac{dx}{h(x)^{n-1}} \geq \frac{dx}{h(x)^{n-1}}$ since $h(x) \leq 1$. This verifies that the pullback of the volume form by s dominates $\frac{dx}{h(x)^{n-1}}$, and hence

$$\int_{\tau_0} \frac{dx}{h(x)^{n-1}} \leq \mathrm{vol}(\tau) \leq v_{n-1}.$$

□

Lemma 1.30. *Let M^n be an oriented closed connected hyperbolic manifold. Then $\|M\|_1 \geq \frac{\mathrm{vol}(M)}{v_n}$.*

Proof. Let $c = \sum \lambda_i c_i$ be a cycle representing $[M]$. By Lemma 1.27, we may assume that c consists of straight hyperbolic simplices without increasing $\|c\|_1$. Let vol be the volume form. Then we have

$$\mathrm{vol}(M) = \langle [M], \mathrm{vol} \rangle = \left\langle \sum \lambda_i c_i, \mathrm{vol} \right\rangle \leq \sum |\lambda_i| \cdot \max \mathrm{vol}(c_i) \leq |c|_1 \cdot v_n.$$

Sine c is arbitrary, we conclude that $\mathrm{vol}(M) \leq v_n \cdot \|M\|_1$. □

Remark 1.31. *Conceptually, we obtained this lower bound by some sort of ℓ^1–ℓ^∞ duality, where we used a cocycle (the volume form here) that is bounded on all straight hyperbolic simplices. This suggests the use of bounded cocycles and bounded cohomology as a dual theory to better understand simplicial norms.*

Restricting to the case $n = 2$, we can now finish the proof of Theorem 1.17.

Proof of Theorem 1.17. The above lemma for $M = S$ a hyperbolic surface, we have

$$\|S\|_1 \geq \frac{\mathrm{area}(S)}{v_2} = \frac{-2\pi\chi(S)}{\pi} = -2\chi(S)$$

by Gauss–Bonnet. Combining with Lemma 1.23, we conclude $\|S\|_1 = -2\chi(S)$. □

The bound in Lemma 1.30 is also sharp in higher dimensions; see Theorem 1.36. This geometric argument also works for closed manifolds with varying negative curvature, so one can deduce that any such manifold has positive simplicial volume. More generally, we have the following conjecture attributed to Gromov [Gro82, p.11]:

Conjecture 1.32. *Any closed manifold of non-positive curvature and negative Ricci curvature has* $\|M\|_1 > 0$.

This is still open. See [CW19] for a recent partial positive answer that uses a slightly different straightening and uniformly bounds the Jacobian of the straightened maps.

Similarly, this argument can be used to prove positivity of simplicial norms of other homology classes.

Proposition 1.33. *The simplicial norm is an honest norm (instead of a semi-norm) on* $H_k(M;\mathbb{R})$ *for any* $k \geq 2$ *and hyperbolic closed orientable manifold* M*. That is,* $\|\sigma\|_1 > 0$ *for any* $\sigma \neq 0 \in H_k(M;\mathbb{R})$.

Proof. For any $\sigma \neq 0$, there is $\sigma^* \in H^k(M;\mathbb{R})$ that pairs nontrivially with σ. Represent σ^* by a differential k-form ω. Up to scaling we may assume that $|\omega(V)| \leq 1$ for all k orthogonal vectors V of norm 1 at any point on M. As a result, the restriction of ω on any straight hyperbolic k-simplex is bounded by the volume form. Thus for any straightened chain $\sum \lambda_i c_i$ representing σ, we have

$$|\langle \sigma, \omega\rangle| \leq \max \operatorname{vol}(c_i) \cdot \sum \|\lambda_i\| \leq v_k \cdot |c|_1.$$

It follows that

$$\|\sigma\|_1 \geq \frac{|\langle \sigma, \omega\rangle|}{v_k} > 0.$$

□

1.6. **Application: degrees of maps between surfaces.** Now we apply the calculation in Theorem 1.20 to solve Problem 1.1 as an application.

Theorem 1.34. *Given closed orientable connected surfaces* S *and* S' *that are not spheres, the set* $\deg(S,S')$ *of all possible degrees of maps* $f : S \to S'$ *is*

$$\deg(S,S') = \{d : |d \cdot \chi(S')| \leq |\chi(S)|\}.$$

Proof. By the degree inequality Lemma 1.11, we have $\deg(f) \cdot \|S'\|_1 \leq \|S\|_1$ for any map $f : S' \to S$. Since spheres are excluded, we know $\|S\|_1 = -2\chi(S)$ and $\|S'\|_1 = -2\chi(S')$. It follows that $|d \cdot \chi(S')| \leq |\chi(S)|$ for any $d \in \deg(S,S')$.

It suffices to show that any such d can be realized as $\deg(f)$ for some $f : S \to S'$. We restrict our attention to those $d > 0$ since $d = 0$ is realized by a constant map and those negative values can be obtained by composing with a orientation-reversing homeomorphism. Given any $d > 0$ with $|d \cdot \chi(S')| \leq |\chi(S)|$, there is a degree d cover $p : S'_d \to S'$. Then $|\chi(S'_d)| = |d\chi(S')| \leq |\chi(S)|$, i.e. S'_d has genus no more than that of S. So there is a map $g : S \to S'_d$ with $\deg(g) = 1$ by pinching part of S to a point. Thus the composition $f = gp$ has $\deg(f) = \deg(g) \cdot \deg(p) = d$. □

Exercise 1.35.

(1) *If* $S = S^2$*, show that* $\deg(S,S') = \{0\}$ *unless* $S' = S^2$*, in which case* $\deg(S,S') = \mathbb{Z}$.
(2) *If* $S' = S^2$*, show that* $\deg(S,S') = \mathbb{Z}$.

1.7. **Gromov's proportionality.** The goal of this section is to prove the following theorem, showing that the lower bound of $\|M\|_1$ in Lemma 1.30 is sharp:

Theorem 1.36 (Gromov's Proportionality)**.** *Let* M^n *be an oriented closed connected hyperbolic manifold. Then* $\|M\|_1 = \frac{\operatorname{vol}(M)}{v_n}$.

Note that if for some $\epsilon > 0$ we have a cycle $c = \sum \lambda_i c_i$ with $\lambda_i > 0$, each c_i straight hyperbolic and positively oriented with $\mathrm{vol}(c_i) \geq v_n - \epsilon$, then c represents $\lambda[M]$ for some $\lambda > 0$ and

$$\lambda \mathrm{vol}(M) = \langle \lambda[M], \mathrm{vol}\rangle = \sum \lambda_i \mathrm{vol}(c_i) \geq \min\{\mathrm{vol}(c_i)\} \cdot \sum \lambda_i \geq (v_n - \epsilon)|c|_1 \geq \lambda(v_n - \epsilon)\|M\|_1. \tag{1.1}$$

Thus it suffices to show the existence of such a cycle for any $\epsilon > 0$.

There is no obvious construction of such a cycle. However, if we were allowed to have *infinitely* many terms in a cycle, there is a natural construction. A formal way to make this work is to define Gromov's norm using *measure homology* instead of singular homology. That is, on the set $S_k(X)$ of singular k-simplices in X, let $\mathcal{C}_k(X)$ be the space of signed measures with compact support and bounded total variation, equipped with the total variation norm. This gives a chain complex with boundary maps defined in the usual way and induces a semi-norm on its homology (called measure homology). Zastrow and Hansen both showed that measure homology coincides with singular homology for CW complexes [Han98, Zas98], and Löh showed that the isomorphism is isometric [LÖ6], thus this is an equivalent way to define Gromov's norm as Thurston originally claimed [Thu, Chapter 6].

The advantage is that this gives more room to construct the desired cycle in the measure sense. In fact, any singular chain $\sum \lambda_i c_i$ can be thought of as the signed measure $\sum \lambda_i \delta_{c_i}$, where δ_{c_i} is the Dirac mass at c_i. So in this way $C_k(X)$ is a subspace of $\mathcal{C}_k(X)$.

Instead of giving a rigorous introduction of measure homology and proving the equivalence, we briefly describe this measure cycle and then approximate it using an honest singular cycle to finish the proof as in [Cal, Section 3.1]. For any $\epsilon > 0$, let Δ be a straight positively oriented n-simplex in $\mathbb{H}^n$ so that $\mathrm{vol}(\Delta) > v_n - \epsilon$. Let $D(\Delta)$ be the space of all isometric orientation-preserving embeddings of Δ in $\mathbb{H}^n$, which can be identified as $\mathrm{Isom}^+(\mathbb{H}^n)$ and comes with the Haar measure. Then $D(\Delta)/\pi_1(M)$ can be thought of as all isometric positively oriented copies of Δ in M, with the induced measure. This makes it into a measure chain $\mathrm{smear}(\Delta)$ in $\mathcal{C}_n(M)$. Similarly, let $\bar{\Delta}$ be the reflection of Δ across some face. The same construction provides a measure chain $\mathrm{smear}(\bar{\Delta})$ in $\mathcal{C}_n(M)$. The key fact is that $\mathrm{smear}(\Delta) - \mathrm{smear}(\bar{\Delta})$ is a measure *cycle*, since every face of Δ is the face of its reflection across this face, and this reflection is a translate of $\bar{\Delta}$. Each copy of $\bar{\Delta}$ has the opposite orientation, but the negative sign "corrects" it.

For the approximation, we need Δ to be chosen with a stronger property in the beginning. It is a fact that, for any $\epsilon > 0$ and any $C > 0$, there is Δ such that $\mathrm{vol}(\Delta') > v_n - \epsilon$ for any Δ' obtained from Δ by moving each vertex by a distance no more than C.

Proof of Theorem 1.36. Let $F \subset \mathbb{H}^n$ be a compact fundamental domain of M and fix $p \in F$. For any $\epsilon > 0$ and $C = \mathrm{diam} F$, let Δ be chosen as above. For any $(n+1)$-tuple $\vec{g} = (g_0, \cdots, g_n) \in \pi_1(M)^{n+1}$, let $c_{\vec{g}}$ be the straight hyperbolic n-simplex with vertices $(g_0 p, \cdots, g_n p)$, which is an approximation of those positively-oriented isometric embeddings of Δ or $\bar{\Delta}$ whose vertices lie in $g_0 F, \cdots, g_n F$ respectively, if exist. Let $\lambda_{\vec{g}}$ (resp. $\bar{\lambda}_{\vec{g}}$) be the measure of such embeddings of Δ (resp. $\bar{\Delta}$). Note that both $\lambda_{\vec{g}}$ and $\bar{\lambda}_{\vec{g}}$ are $\pi_1(M)$ invariant under the diagonal action of $\pi_1(M)$ on $\pi_1(M)^{n+1}$. Let $\pi : \mathbb{H}^n \to M$ be the universal covering map. Then set the following chains in $C_n(M; \mathbb{R})$

$$c = \sum_{\vec{g} \in \pi_1(M)^{n+1}/\pi_1(M)} \lambda_{\vec{g}} \pi c_{\vec{g}} \quad \text{and} \quad \bar{c} = \sum_{\vec{g} \in \pi_1(M)^{n+1}/\pi_1(M)} \bar{\lambda}_{\vec{g}} \pi c_{\vec{g}}$$

to approximate $\mathrm{smear}(\Delta)$ and $\mathrm{smear}(\bar{\Delta})$ respectively.

Claim 1.37.

1. *$\lambda_{\vec{g}} = 0$ for all but* finitely many *$\vec{g} \in \pi_1(M)^{n+1}/\pi_1(M)$, i.e. c is a singular chain, and so is c'.*
2. *If $\lambda_{\vec{g}} > 0$ (resp. $\bar{\lambda}_{\vec{g}} > 0$), then* $\mathrm{vol}(c_{\vec{g}}) > v_n - \epsilon$.
3. *For ϵ small enough, if $\lambda_{\vec{g}} > 0$ (resp. $\bar{\lambda}_{\vec{g}} > 0$), then $c_{\vec{g}}$ is positively oriented (resp. negatively oriented).*

(4) $c - \bar{c}$ *is a* cycle.

Proof.

(1) Up to the $\pi_1(M)$ action, we may assume $g_0 p = p \in F$. Then the space of isometric embeddings of Δ with the first vertex inside F is compact, so the remaining n vertices can only lie in finitely many possible gF.
(2) This means that $c_{\vec{g}}$ can be obtained from some embedding of Δ (or $\bar{\Delta}$) by moving each vertex by distance no more than $C = \text{diam}F$. So the assertion follows from our choice of Δ.
(3) No Δ can be close to $\bar{\Delta}$ when they have large enough volume.
(4) Any face τ of $c_{\vec{g}}$ is a straight $(n-1)$-simplex with vertices $(h_1 p, \cdots, h_n p)$ for some $h_i \in \pi_1(M)$. If a face of Δ has an orientation-preserving isomeric embedding with vertices in $h_1 F, \cdots, h_n F$, then this contributes to the coefficient of τ in ∂c by the measure of the total measure of all such isometric embeddings of Δ. For each such embedding, its reflection across τ is an orientation-preserving isometric embedding of Δ, so this has the same contribution to the coefficient of τ in $\partial \bar{c}$. Since this holds for all faces τ, we observe that $\partial(c - \bar{c}) = 0$.

□

By the claim above, all nontrivial terms in c' are negatively oriented, which can be made positively oriented by adding a negative sign. Thus the cycle $c - c'$ can be expressed as a positive linear combination of positively oriented straight simplices of volume at lest $v_n - \epsilon$, this implies $\text{vol}(M) \geq (v_n - \epsilon)\|M\|_1$ by equation (1.1). Since ϵ can be arbitrarily small, this proves $\frac{\text{vol}(M)}{v_n} \geq \|M\|_1$ and completes the proof in combination with Lemma 1.30. □

The smearing operation in the measure homology setup does something more [Thu, Chapter 6]. If M and N are oriented closed oriented Riemannian manifolds with isometric universal cover X, given a singular k-simplex $c : \Delta \to M$, we can similarly do the "smearing" operation by first lifting it to the universal cover, taking all translates of it, and projecting to N. Equipped with the natural measure from $\text{Isom}^+(X)$, this gives a chain in $\mathcal{C}_k(N)$. This extends to a linear map $\text{smear}_{M,N} : \mathcal{C}_k(M) \to \mathcal{C}_k(N)$. This has the property that $\text{smear}_{M,N}[M] = \frac{\text{vol}(M)}{\text{vol}(N)} \cdot [N]$ and implies

$$\frac{\|N\|_1}{\text{vol}(N)} \leq \frac{\|M\|_1}{\text{vol}(M)}.$$

Flipping the roles of M and N implies the following more general version of Gromov's proportionality:

Theorem 1.38 (Gromov's Proportionality, general case)**.** *Suppose M and N are orientable closed Riemannian manifolds with isometric universal covers, then*

$$\frac{\|M\|_1}{\text{vol}(M)} = \frac{\|N\|_1}{\text{vol}(N)}.$$

It is unclear in general though how this constant proportion is related to the geometry of the universal cover.

2. Bounded cohomology

Bounded cohomology is a dual theory of homology with simplicial norm, and we had a glimpse of its power in the estimate of simplicial volume of hyperbolic manifolds. It is also interesting on its own, providing new invariants sometimes quite different from ordinary cohomology. We will first focus on (bounded) cohomology of groups for concreteness. We will see later in Gromov's mapping theorem that the bounded cohomology of a connected space X is canonically isomorphic to the bounded cohomology of $\pi_1(X)$, so we are not losing any information.

2.1. **Group cohomology.** We start with ordinary cohomology of groups as a warm-up. A good detailed reference on this topic is Brown's book [Bro82].

2.1.1. *$K(G,1)$ spaces.* Given a group G, a $K(G,1)$ *space* X_G is a connected *aspherical* CW complex with $\pi_1(X_G) = G$. Here being aspherical means that all higher homotopy groups of X_G vanish. Equivalently (by Whitehead's theorem), the universal cover of X_G is contractible. Such a space is unique up to homotopy equivalence, and is also called the *Eilenberg–Maclane space* or the *classifying space* of the group G (as a discrete group). The uniqueness follows from the universal property below:

Lemma 2.1. *Suppose X is a $K(G,1)$ space. Let Y be a connected CW complex with $\pi_1(Y) = H$. Any homomorphism $\varphi : H \to G$ is induced by a continuous map $f : Y \to X_G$, which is unique up to homotopy.*

Corollary 2.2. *Suppose X and Y are both $K(G,1)$ spaces. Then any map $f : X \to Y$ inducing an isomorphism $f_* : G \to G$ is a homotopy equivalence.*

Proof. Let $g_* : G \to G$ be the inverse of f_*. By the universal property, it is induced by a map $g : Y \to X$. Then $gf : X \to Y$ induces the identity map on $\pi_1(X)$ and hence is homotopic to the identity map on X by uniqueness. Similarly fg is homotopic to id_Y, so g is a homotopy inverse of f. □

One way to build a $K(G,1)$ space is to start with a 2-complex with fundamental group G and add higher dimensional cells to kill all higher homotopy groups inductively. A similar idea can be used to prove the universal property above. See details in [Hat02] for instance.

One can use the $K(G,1)$ space to give a topological definition of the group (co)homology.

Definition 2.3. Given a ring R, the homology $H_*(G;R)$ (resp. cohomology $H^*(G;R)$) of G with R coefficients is $H_*(X_G;R)$ (resp. $H^*(X_G;R)$).

Example 2.4.

(1) *For $G = \mathbb{Z}$, we can take $X_G = S^1$. We know $H_k(G;\mathbb{Z}) = H_k(X_G;\mathbb{Z}) \cong \mathbb{Z}$ for $k = 0, 1$. Since it is one-dimensional, we have $H_k(G;\mathbb{Z}) = H_k(X_G;\mathbb{Z}) = 0$ for $k \geq 2$.*
(2) *For a free group $G = F_n$ with $n \geq 1$, we can take X_G as a graph (e.g. a wedge of circles). We have $H_0(G;\mathbb{Z}) \cong \mathbb{Z}$, $H_1(G;\mathbb{Z}) = \mathbb{Z}^n$, and $H_k(G;\mathbb{Z}) = 0$ for $k \geq 2$ by dimension restrictions.*
(3) *$H_0(G;R) = R$ since $K(G,1)$ space is connected by definition.*
(4) *$H_1(G;R) = \mathrm{Ab}(G) \otimes_{\mathbb{Z}} R$.*
(5) *$H^1(G;R) = \mathrm{Hom}(G,R)$.*
(6) *For a cyclic group $G = \mathbb{Z}/m$, we can choose an infinite lens space as the $K(G,1)$ space, which gives $H_k(\mathbb{Z}/m;\mathbb{Z}) \cong \mathbb{Z}/m$ for all k odd, $H_k(\mathbb{Z}/m;\mathbb{Z}) = 0$ for $k > 0$ and even. See* [Hat02, Example 2.43] *for details in a similar computation.*

It follows from the last example that no $K(\mathbb{Z}/m,1)$ space can be finite dimensional. As an application of this fact, we have:

Proposition 2.5. *For any nontrivial finite group G, there is no free action on $\mathbb{R}^n$.*

Proof. Suppose G acts freely. Let $H \cong \mathbb{Z}/m$ be a nontrivial cyclic subgroup. Then H also acts freely on $\mathbb{R}^n$. The action is also properly discontinuous since H is finite. Thus $\mathbb{R}^n/H$ is a $K(H,1)$ space as $\mathbb{R}^n$ is contractible. This is a contradiction. □

2.1.2. *Aside: The co-Hopfian property.* The uniqueness of $K(G,1)$ spaces (up to homotopy) has interesting consequences in the context of closed manifolds.

Lemma 2.6. *If M and N are both connected aspherical n-manifolds (without boundary) with isomorphic fundamental group, then M and N are either both compact or both non-compact.*

Proof. Note that a connected n-manifold M (without boundary) is compact if and only if $H_n(M;\mathbb{Z}/2)\cong\mathbb{Z}/2$. Since M and N are $K(G,1)$ spaces for the same G, they are homotopy equivalent and $H_n(M;\mathbb{Z}/2)\cong H_n(N;\mathbb{Z}/2)$, so they have the same compactness. □

This is related to the *co-Hopfian* property of fundamental groups

Definition 2.7. A group G is co-Hopfian if every injective homomorphism $h: G\to G$ is an isomorphism.

Exercise 2.8. *Show that the following groups are not co-Hopfian.*

(1) $G=\mathbb{Z}$.
(2) G is a free group or free abelian group.
(3) G is a free product $H\star K$ where K is not co-Hopfian.

Lemma 2.9. *If M is a closed connected aspherical manifold, then any subgroup of $\pi_1(M)$ that is isomorphic to $\pi_1(M)$ must have finite index. In particular, if $\pi_1(M)$ is not co-Hopfian and M is orientable, then M has a self-map f with $|\deg(f)|>1$.*

Proof. Let H be a subgroup of $G=\pi_1(M)$ that is isomorphic to G. Let $\pi: M'\to M$ be the covering map corresponding to the inclusion $H\to G$. As M is aspherical, so is M'. Since M is closed, M' must be closed by Lemma 2.6, so π is a finite cover.

The isomorphism $\pi_1(M)\to H$ can be realized as a homotopy equivalence $\varphi: M\to M'$, which necessarily has $|\deg\varphi|=1$. So the composition $f=\pi\varphi$ is a self-map with $|\deg f|=|\deg\pi|>1$ if H is a proper subgroup. □

Corollary 2.10. *If M is an orientable closed connected aspherical manifold with $\|M\|_1>0$, then $\pi_1(M)$ is co-Hopfian.*

Proof. As $\|M\|_1>0$, M cannot have a self-map f with $|\deg f|>1$ by the degree inequality. □

Asphericity is often deduced from geometry of the manifold:

Lemma 2.11. *A complete Riemannian manifold M with non-positive sectional curvature is aspherical. Moreover, the universal cover of M is diffeomorphic to the Euclidean space.*

Proof. For any $p\in M$, the exponential map $T_pM\to M$ is a covering map by the Cartan–Hadamard theorem. Hence T_pM is one realization of the universal cover. □

Corollary 2.12. *If M is occ with negative sectional curvature, then $\pi_1(M)$ is co-Hopfian.*

Proof. Negative curvature implies $\|M\|_1>0$ by a straightening argument, so the assumptions of Corollary 2.10 are met. □

Example 2.13. *Let S be an occ surface of genus $g>1$. Then $\pi_1(S)$ is co-Hopfian.*

One can also prove this using the fact that infinite-index subgroups of $\pi_1(S)$ are free groups.

Exercise 2.14. *It is crucial to assume M to be aspherical in Corollary 2.10. Let $N=M\#(S^2\times S^1)$ be a connected sum, where M is an occ hyperbolic 3-manifold. Show that $\pi_1(N)$ is not co-Hopfian although N is occ and has $\|N\|_1>0$.*

2.1.3. *The bar complex.* There is also an explicit and purely algebraic definition of group (co)-homology, which we will see to be equivalent. Consider the following *bar complex* with coefficients in R. Let $C_n(G;R)$ be the free R-module with basis consisting of n-tuples $(g_1,\cdots,g_n)\in G^n$, and let $C^n(G,R)=\mathrm{Hom}_R(C_n(G;R),R)$, where each element assigns a value in R to each element of G^n. The differential $\partial: C_n(G;R)\to C_{n-1}(G;R)$ is determined by

(2.1)
$$\partial(g_1,\cdots,g_n):=(g_2,\cdots,g_n)+\sum_{i=1}^{n-1}(-1)^i(g_1,\cdots g_{i-1},g_ig_{i+1},g_{i+2},\cdots,g_n)+(-1)^n(g_1,\cdots,g_{n-1}).$$

It is standard to check that $\partial^2=0$ (which also follows from the topological explanation below), and one can define $H_*(G;R)$ as the homology of this chain complex. Similarly, take $\delta: C^{n-1}(G;R)\to C^n(G;R)$ as the dual of ∂ as usual, we can define $H^*(G;R)$ using the cochains $C^n(G;R)$.

This weird-looking differential (2.1) in the bar complex comes from the following topological interpretation. Let EG be the simplicial complex where n-simplices correspond to elements in G^{n+1} for each $n\geq 0$, where the faces of $(g_0,\cdots,g_n)$ are $(g_0,\cdots,\hat{g}_i,\cdots,g_n)$ (i.e. omitting g_i), $i=0,\cdots,n$.Then EG is contractible since each simplicial map can be coned off, say by adding a fixed element $g\in G$ as the last coordinate. The group G acts by $g(g_0,\cdots,g_n)=(gg_0,\cdots,gg_n)$, which is a free action. Thus the quotient $BG:=EG/G$ is a $K(G,1)$ space.

One can think of an n-simplex as an equivalence class $[(g_0,\cdots,g_n)]$, where $(g_0,\cdots,g_n)\sim(g_0',\cdots,g_n')$ if $(g_0',\cdots,g_n')=(gg_0,\cdots,gg_n)$ for some $g\in G$. This leads to a way to define group (co)homology using the so-called homogeneous coordinates.

Instead, we represent each equivalence class $[(g_0,\cdots,g_n)]$ by an n-tuple $(g_0^{-1}g_1,g_1^{-1}g_2,\cdots,g_{n-1}^{-1}g_n)$, which is independent of the representative $(g_0,\cdots,g_n)$. This is called the *inhomogeneous coordinate*. Then $(g_1,\cdots,g_n)$ in inhomogeneous coordinate corresponds to the equivalence class $[(id,g_1,g_1g_2,\cdots,g_1g_2\cdots g_n)]$ in homogeneous coordinate. Geometrically, one can think of the homogeneous coordinate as marking on vertices and the inhomogeneous coordinate as marking on edges. Under this identification, the boundary map in inhomogeneous coordinates is exactly the differential (2.1). This justifies that the (co)homology defined using bar complex is exactly the (simplicial) homology of the $K(G,1)$ space BG, and thus agrees with our topological definition.

Under this setup, a k-cochain is an assignment $f: G^k\to R$, labeling each k-tuple in inhomogeneous coordinate an element in R and extending to an R-linear map on the space of k-chains.

Example 2.15. *A* 0*-cochain is a constant in* R*, and* $\delta: C^0(G;R)\to C^1(G;R)$ *is the zero map. Hence every* 0*-chchain is a cocycle and* $H^0(G;R)=R$*.*

For a 1*-cochain* $f: G\to R$*, its coboundary is determined by* $(\delta f)(g_1,g_2)=f\partial(g_1,g_2)=f(g_2-g_1)=f(g_2)-f(g_1g_2)+f(g_1)$*. So it is a cocycle if and only if* f *is a homomorphism to the abelian group* R *(forgetting the ring structure). This shows that* $H^1(G;R)=\mathrm{Hom}(G,R)$*.*

2.2. **Bounded cohomology of groups.** Roughly speaking,, in bounded cohomology, we consider cochains that are bounded functions instead of arbitrary functions on G^k. For this to make sense, we need to measure the size of elements in R. For simplicity, we consider $R=\mathbb{R}$ or $\mathbb{Z}$, equipped with the usual absolute value.

Then a k-cochain f in inhomogeneous coordinates as a map $f: G^k\to R$ is bounded if

$$|f|_\infty:=\sup|f(g_1,\cdots,g_k)|<\infty.$$

Equivalently, $|f|_\infty$ is the sup norm as a linear map on the space of k-chains.

For each k, let $C_b^k(G,R)$ be the subspace of $C^k(G;R)$ consisting of bounded cochains. Then the coboundary restricts to map $\delta_k: C_b^k(G;R)\to C_b^{k+1}(G;R)$ since $|\delta_k f|_\infty\leq(k+1)|f|_\infty$. This gives rise to a new chain complex.

Definition 2.16. (Bounded cohomology) Let $Z_b^n(G;R):=\ker\delta_n$ and $B_b^n(G;R):=\operatorname{Im}\delta_{n-1}$. The n-th *bounded cohomology* of G is $H_b^n(G;R):=Z_b^n(G;R)/B_b^n(G;R)$.

The norm $|\cdot|_\infty$ on $C_b^n(G;R)$ restricts to $Z_b^n(G;R)$ and induces a semi norm $\|\cdot\|_\infty$ on $H_b^n(G;R)$. Explicitly, for any $\alpha\in H_b^n(G;R)$,

$$\|\alpha\|_\infty=\inf_{[f]=\alpha}|f|_\infty.$$

It is natural to ask how bounded cohomology differs from ordinary cohomology. There is a natural map connecting them, by treating a bounded cocycle as an ordinary cocycle.

Definition 2.17. (Comparison map) The inclusion $C_b^n(G;R)\to C^n(G;R)$ induces a homomorphism $c: H_b^n(G;R)\to H^n(G;R)$ called the *comparison map*.

Example 2.18. *As shown in Example 2.15, for degree $n = 0$, a 0-cochain is a constant map and thus bounded. Thus $H^0_b(G; R) = H^0(G; R) = R$.*

For degree $n = 1$, a 1-cochain $f : G \to R$ is a cocycle if and only if it is a homomorphism. However, a homomorphism to $R = \mathbb{Z}, \mathbb{R}$ is always unbounded except the trivial one. Thus $H^1_b(G; R) = 0$ and the comparison map is not surjective in general.

2.3. **Quasimorphisms.** Nontrivial bounded cohomology classes emerge in degree $n = 2$. Here is the idea: For any $(n-1)$-cochain f, δf is always a cocycle, but it might happen that δf is bounded while f is not, in which case δf is a bounded n-cocycle potentially nontrivial. We will focus on $R = \mathbb{R}$ throughout this section and thus often omit the coefficient.

When $n = 2$, as we calculated in Example 2.15, for any $f : G \to \mathbb{R}$, we have $(\delta f)(g, h) = f(g) + f(h) - f(gh)$. So δf is bounded exactly when f is a quasimorphism, defined as follows.

Definition 2.19. (Quasimorphism) A map $\varphi : G \to \mathbb{R}$ is a *quasimorphism* if

$$D(\varphi) := \sup_{g,h \in G} |\varphi(g) + \varphi(h) - \varphi(gh)| < \infty.$$

The number $D(\varphi)$ is called the *defect* of φ. Quasimorphisms can be thought of as homomorphisms with bounded error, measured by the defect.

A quasimorphism φ is *homogeneous* if $\varphi(g^n) = n\varphi(g)$ for all $g \in G$ and all $n \in \mathbb{Z}$, i.e. φ restricts to a homomorphism on every cyclic subgroup.

Note that quasimorphisms on G form an $\mathbb{R}$-vector space, which we denote as $\widehat{Q}(G)$. Homogeneous quasimorphisms form a linear subspace, denoted as $Q(G)$. Clearly, homomorphisms to $\mathbb{R}$ are homogeneous quasimorphisms, i.e. $H^1(G) \subset Q(G)$. Also note that any bounded function $\varphi : G \to \mathbb{R}$ is trivially a quasimorphism. This gives another subspace $C^1_b(G) = C^1_b(G; \mathbb{R})$ of $\widehat{Q}(G)$.

Every quasimorphism can be made homogeneous by the following process.

Definition 2.20. (Homogenization) For any $\varphi \in \widehat{Q}(G)$, the homogenization $\bar{\varphi}$ is defined as

$$\bar{\varphi}(g) := \lim_{+\infty} \frac{\varphi(g^n)}{n}.$$

Lemma 2.21. *The homogenization $\bar{\varphi}$ is a well-defined homogeneous quasimorphism, and $\bar{\varphi} - \varphi$ is a bounded function on G. Quantitatively, we have $|\varphi - \bar{\varphi}|_\infty \le D(\varphi)$.*

In the proof of this lemma, we need the following fact about sub-additive sequence.

Lemma 2.22. *Let a_n be a real-valued sequence that is sub-additive, i.e. $a_{m+n} \le a_m + a_n$ for all $m, n \ge 1$. Then we have*

$$\lim_{n \to +\infty} \frac{a_n}{n} = \inf_{n \ge 1} \frac{a_n}{n}.$$

In particular, the limit exists iff $\frac{a_n}{n}$ is bounded below.

Proof. Clearly we have $\liminf \frac{a_n}{n} \ge \inf_{n \ge 1} \frac{a_n}{n}$. So it suffices to show that $\limsup \frac{a_n}{n} \le \inf_{n \ge 1} \frac{a_n}{n}$.

Fix $m \ge 1$ and express any n as $n = qm + r$ with $0 < r \le m$. Let $B = \max_{0 < r \le m} a_r$. By sub-additivity and induction, we have $a_n \le qa_m + a_r \le qa_m + B$. Thus

$$\frac{a_n}{n} \le \frac{qa_m + B}{qm + r} = \frac{a_m + \frac{B}{q}}{m + \frac{r}{q}}.$$

Hence as $n \to \infty$, we have $q \to \infty$ and

$$\limsup \frac{a_n}{n} \le \frac{a_m}{m}.$$

Since m is arbitrary, we obtain

$$\limsup \frac{a_n}{n} \le \inf_{m \ge 1} \frac{a_m}{m}.$$

$\square$

There is an analogous result for sup-additive sequences, replacing inf by sup, which can be deduced by considering the sequence $-a_n$.

Now we prove Lemma 2.21.

Proof of Lemma 2.21. Fixing $g \in G$, we show that the limit defining $\bar{\varphi}(g)$ exists. Although the sequence $\varphi(g^n)$ is not sup- or sub-additive, a small modification does the job.

Let $\varphi_+(g) := \varphi(g) + D(\varphi)$ and let $\varphi_-(g) := \varphi(g) - D(\varphi)$. By definition, we have $|\varphi(g) + \varphi(h) - \varphi(gh)| \leq D(\varphi)$, which implies

$$\varphi(gh) - D(\varphi) \leq \varphi(g) + \varphi(h) \leq \varphi(gh) + D(\varphi),$$

and hence

$$\varphi_+(gh) = \varphi(gh) - D(\varphi) + 2D(\varphi) \leq \varphi(g) + \varphi(h) - 2D(\varphi) = \varphi_+(g) + \varphi_+(h).$$

So φ_+ is sub-additive and similarly φ_- is sup-additive.

It follows that

$$\varphi_-(g) \leq \frac{\varphi_-(g^n)}{n} \leq \frac{\varphi_+(g^n)}{n} \leq \varphi_+(g)$$

for all $n \geq 1$. So $\frac{\varphi_+(g^n)}{n}$ is bounded below and $\frac{\varphi_-(g^n)}{n}$ is bounded above. Thus both have finite limit by Lemma 2.22 and its sup-additive analogue. The limits agree since $\varphi_+(g^n) - \varphi_-(g^n) = 2D(\varphi)$. As $\varphi_-(g^n)\varphi(g^n) \leq \varphi_+(g^n)$, we have

$$\bar{\varphi}(g) = \lim_{n\to+\infty} \frac{\varphi_-(g^n)}{n} = \lim_{n\to+\infty} \frac{\varphi(g^n)}{n} = \lim_{n\to+\infty} \frac{\varphi_+(g^n)}{n}.$$

This shows that $\bar{\varphi}$ is a well-defined function.

This also gives a way to bound the difference $\varphi - \bar{\varphi}$. As

$$\lim_{n\to+\infty} \frac{\varphi_+(g^n)}{n} = \inf_{n\geq 1} \frac{\varphi_+(g^n)}{n} \leq \varphi_+(g) = \varphi(g) + D(\varphi),$$

we get

$$\bar{\varphi}(g) \leq \varphi(g) + D(\varphi).$$

Similarly using φ_- we have $\bar{\varphi}(g) \geq \varphi(g) - D(\varphi)$. As g is arbitrary, we conclude that $|\varphi - \bar{\varphi}|_\infty \leq D(\varphi)$. In particular, $\bar{\varphi}$ is a quasimorphism as the sum of φ and the bounded function $\bar{\varphi} - \varphi$.

Finally, it remains to check that $\bar{\varphi}$ is homogeneous. For every $k \geq 1$, we have

$$\bar{\varphi}(g^k) = \lim_{n\to+\infty} \frac{\varphi(g^{kn})}{n} = k \cdot \lim_{n\to+\infty} \frac{\varphi(g^{kn})}{kn} = k\bar{\varphi}(g).$$

For $k = 0$, $\bar{\varphi}(id) = \lim \frac{\varphi(id)}{n} = 0$. So it suffices to show that $\bar{\varphi}(g) + \bar{\varphi}(g^{-1}) = 0$ to deal with $k < 0$. Indeed, this easily follows from $|\varphi(g^n) + \varphi(g^{-n})| \leq |\varphi(id)| + D(\varphi)$. □

Remark 2.23. *The triangle inequality and the bound $|\varphi - \bar{\varphi}|_\infty \leq D(\varphi)$ implies that $D(\bar{\varphi}) \leq 4\, D(\varphi)$. A more involved argument shows that $D(\bar{\varphi}) \leq 2\, D(\varphi)$.*

Every quasimorphism uniquely decomposes as the sum of a homogeneous quasimorphism (namely, its homogenization) and a bounded function.

Lemma 2.24. *We have $\widehat{Q}(G) = C_b^1(G) \oplus Q(G)$. That is, $C_b^1(G) \cap Q(G) = 0$ and $\widehat{Q}(G) = C_b^1(G) + Q(G)$.*

Proof. If $\varphi \in C_b^1(G) \cap Q(G)$, then $|\varphi(g)| = |\frac{\varphi(g^n)}{n}| \leq \frac{|\varphi|_\infty}{|n|}$. As n is arbitrary, we must have $\varphi(g) = 0$ for all g.

For any $\varphi \in \widehat{Q}(G)$, we have $\varphi = \bar{\varphi} + (\varphi - \bar{\varphi})$, where the homogenization $\varphi \in Q(G)$ and the difference $\varphi - \bar{\varphi}$ is a bounded function by Lemma 2.21. □

Proposition 2.25. *We have the following exact sequence*

$$0 \to H^1(G;\mathbb{R}) \to Q(G) \xrightarrow{\delta} H^2_b(G;\mathbb{R}) \xrightarrow{c} H^2(G;\mathbb{R}).$$

In particular, the kernel of the comparison map $c : H^2_b(G;\mathbb{R}) \to H^2(G;\mathbb{R})$ is isomorphic to the quotient $Q(G)/H^1(G)$, which can be thought of as the space of "interesting" homogeneous quasimorphisms.

Proof. We just check that $\ker\delta = H^1(G)$ (as homomorphisms to $\mathbb{R}$) and $\ker c = \operatorname{Im}\delta$.

As we calculated in Example 2.15, $(\delta\varphi)(g,h) = \varphi(g)+\varphi(h)-\varphi(gh)$, so clearly $H^1(G) \subset \ker\delta$. Conversely, if $[\delta\varphi] = 0 \in H^2_b(G)$, then $\delta\varphi = \delta f$ for some bounded function $f : G \to \mathbb{R}$. This implies that $\varphi - f$ is a homomorphism. Then we get two homogeneous quasimorphisms $\varphi - f$ and φ that differ by a bounded function f. By Lemma 2.24, we must have $f = 0$ and thus φ is a homomorphism. Hence $\ker\delta = H^1(G)$.

$\operatorname{Im}\delta \subset \ker c$ holds by definition. Suppose $\alpha \in \ker c$, i.e. $\alpha = [\delta f]$ for some function $f : G \to \mathbb{R}$. Then δf must be bounded, so f is a quasimorphism. Since f and $\bar f$ differ by a bounded function, we have $\alpha = [\delta f] = [\delta \bar f]$. As $\bar f \in Q(G)$ we conclude that $\alpha \in \operatorname{Im}\delta$. Thus $\operatorname{Im}\delta = \ker c$. □

This gives a way to characterize the kernel of the comparison map in degree two.

Corollary 2.26. *The kernel of the comparison map $c : H^2_b(G;\mathbb{R}) \to H^2(G;\mathbb{R})$ is identified with $Q(G)/H^1(G;\mathbb{R})$.*

Exercise 2.27. *Prove the following variant of the exact sequence in Proposition 2.25*

$$0 \to C^1_b(G) + H^1(G) \to \widehat{Q}(G) \xrightarrow{\delta} H^2_b(G;\mathbb{R}) \xrightarrow{c} H^2(G;\mathbb{R}).$$

We will often use the following basic estimate, which is immediate from the definition and induction.

Lemma 2.28. *Let φ be a quasimorphism. For $g = g_1 \cdots g_n$, we have*

$$|\varphi(g) - \sum_i \varphi(g_i)| \le (n-1)D(\varphi).$$

In particular,

$$|\varphi(g)| \le \sum_i |\varphi(g_i)| + (n-1)D(\varphi).$$

Homogeneous quasimorphisms have the following nice properties:

Lemma 2.29. *Let φ be a homogeneous quasimorphism.*

(1) If g and h commute, then $\varphi(gh) = \varphi(g)+\varphi(h)$. So φ restricts to homomorphisms on abelian subgroups.

(2) φ is conjugation-invariant, i.e. $\varphi(g) = \varphi(hgh^{-1})$ for all $g,h \in G$.

Proof.

(1) For any $n \in \mathbb{Z}_+$, we have $(gh)^n = g^n h^n$, so

$$n|\varphi(g)+\varphi(h)-\varphi(gh)| = |\varphi(g^n)+\varphi(h^n)-\varphi((gh)^n)| = |\varphi(g^n)+\varphi(h^n)-\varphi((g^nh^n)| \le D(\varphi).$$

Letting $n \to \infty$ we see $\varphi(g)+\varphi(h)-\varphi(gh) = 0$.

(2) For any $n \in \mathbb{Z}_+$, we have

$$n|\varphi(g)-\varphi(hgh^{-1})| \le |\varphi(g^n)-\varphi(hg^nh^{-1})| = |[\varphi(h)+\varphi(g^n)+\varphi(h^{-1})] - \varphi(hg^nh^{-1})| \le 2D(\varphi),$$

where we used the fact that $\varphi(h)+\varphi(h^{-1}) = 0$. Letting $n \to \infty$ we get $\varphi(g) = \varphi(hgh^{-1})$.

□

Corollary 2.30. *If G is abelian, then $Q(G) = H^1(G;\mathbb{R})$.*

Exercise 2.31. *Let φ be a homogeneous quasimorphism. Prove that $|\varphi([g,h])| \le D(\varphi)$ for all g, h, where $[g,h] = ghg^{-1}h^{-1}$ is the commutator.*

Remark 2.32. *Actually Bavard* [Bav91, Lemma 3.6] *showed that $D(\varphi) = \sup_{g,h} |\varphi([g,h])|$. See also* [Cal09, Lemma 2.24].

Definition 2.33. A group G is *perfect* if it has trivial abelianization, or equivalently, G agrees with its commutator subgroup, i.e. each $g \in G$ is a product of commutators.

We say G is *uniformly perfect*, if there is a uniform $n \in \mathbb{Z}_+$ such that each $g \in G$ is a product of at most n commutators.

Corollary 2.34. *If G is uniformly perfect, then $Q(G) = H^1(G) = 0$.*

Proof. Suppose for $n \in \mathbb{Z}_+$, each $g \in G$ is a product of at most n commutators. Then for any homogeneous quasimorphism $\varphi \in Q(G)$, and $g = [a_1, b_1] \cdots [a_k, b_k]$ with $k \le n$ and $a_i, b_i \in G$, we have

$$|\varphi(g)| \le |\sum_i \varphi([a_i, b_i])| + (k-1)D(\varphi) \le (2k-1)D(\varphi) \le (2n-1)D(\varphi)$$

by Exercise 2.31. This shows that φ is a bounded function. Hence $\varphi = 0$ by Lemma 2.24. □

Lemma 2.35. *For any surjective homomorphism $f : G \to H$, the pullback map $f^* : Q(H) \to Q(G)$ is injective and defect-preserving.*

Proof. The pullback is given by $(f^*\varphi)(g) = \varphi(f(g))$, so injectivity is immediate. As for the defect, we have

$$D(f^*\varphi) = \sup_{g_1, g_2 \in G} |\varphi(f(g_1)f(g_2)) - \varphi(f(g_1)) - \varphi(f(g_2))| = \sup_{h_1, h_2 \in H} |\varphi(h_1 h_2) - \varphi(h_1) - \varphi(h_2)| = D(\varphi).$$

□

This gives a way to obstruct homomorphisms from groups with few quasimorphisms to those with lots of quasimorphisms.

2.4. **de Rham quasimorphisms.** Let ω be a 1-form on a connected closed hyperbolic manifold M^n. Fix a based point $p \in M$ and let $G = \pi_1(M, p)$. Then any $g \in G$ is uniquely represented by an oriented geodesic loop ℓ_g based at p. The de Rham quasimorphism (due to Barge–Ghys [BG88]) associated to ω is

$$\varphi_\omega(g) := \int_{\ell_g} \omega.$$

Lemma 2.36. *φ_ω defined above is indeed a quasimorphism.*

Proof. We need to bound $\varphi_\omega(g) + \varphi_\omega(h) - \varphi_\omega(gh)$ for any $g, h \in G$.

Let $\widetilde{M} \cong \mathbb{H}^n$ be the universal cover and $\tilde{p}$ be a lift of p. Denote by $\tilde{\omega}$ the pullback of ω on $\widetilde{M}$. Then for any $g \in G$, the unique lift of ℓ_g starting at $\tilde{p}$ is the geodesic $\tilde{\ell}_g$ from $\tilde{p}$ to $g\tilde{p}$. So $\varphi_\omega(g) = \int_{\tilde{\ell}_g} \tilde{\omega}$.

Given $g, h \in G$, we have an oriented geodesic triangle Δ with sides $\tilde{\ell}_g$, $g\tilde{\ell}_h$ and $\tilde{\ell}_{gh}$, where the induced orientation of Δ is opposite to the orientation on $\tilde{\ell}_{gh}$. Thus

$$|\varphi_\omega(g) + \varphi_\omega(h) - \varphi_\omega(gh)| = \left| \int_{\partial\Delta} \tilde{\omega} \right| = \left| \int_\Delta d\tilde{\omega} \right| \le \|d\tilde{\omega}\|_{\widetilde{M}} \cdot \mathrm{area}(\Delta) \le \pi \|d\omega\|_M,$$

where $\|d\omega\|$ is the supremum of $|d\omega(v)|$ over all orthonormal 2-frames v on M, which is finite by compactness. Note that we also used the fact that the area of a hyperbolic triangle is uniformly bounded by π. □

Also note that φ_ω is a homomorphism if ω is closed, as the integral over Δ. When ω is exact, then $\varphi_\omega \equiv 0$.

There is also a nice description of the homogenization $\bar{\varphi}_\omega$. For any $g \in G$, there is a unique closed geodesic loop L_g (which is length minimizing in the free homotopy class) representing the conjugacy of g, and $\bar{\varphi}_\omega(g) = \int_{L_g} \omega$. The reason is that, changing the base point p only varies φ_ω by a bounded amount and does not affect the homogenization. So one can move p so that ℓ_g agrees with L_g.

2.5. **Quasimorphisms on free groups.** Various kinds of quasimorphisms were constructed on free groups. Brooks constructed lots quasimorphisms that imply $Q(F_n)$ is infinite-dimensional. A recent construction by Rolli [Rol09] gives a simpler way to prove this, so we will start with his construction.

2.5.1. *Rolli's construction.* Let $\ell^\infty(\mathbb{Z}_+)$ be the space of $\mathbb{R}$-valued bounded functions on the set $\mathbb{Z}_+$, which is an infinite dimensional space. For any $f \in \ell^\infty(\mathbb{Z}_+)$, extend it uniquely to an odd function $f : \mathbb{Z} \to \mathbb{R}$, i.e. $f(-n) = -f(n)$ for all n.

Given $f, g \in \ell^\infty(\mathbb{Z}_+)$, extended as above, define a quasimorphism $\varphi_{f,g}$ on $F_2 = \langle a, b\rangle$ as follows. For any element $w \in F_2$, express it as a reduced word in the generators $w = a_1^{m_1} b_1^{n_1} \cdots a_k^{m_k} b_k^{n_k}$ with $k \geq 1$ and each $m_i, n_i \in \mathbb{Z} \setminus \{0\}$ except that m_1 or n_k could be 0. Define $\varphi_{f,g}(w) = \sum_{i=1}^k f(m_i) + \sum_{i=1}^k g(n_i)$.

It is straightforward to check that

$$|\varphi(u) + \varphi(v) - \varphi(uv)| \leq \max\{3|f|_\infty, 3|g|_\infty\},$$

so φ is a quasimorphism.

The homogenization $\bar{\varphi}_{f,g}$ can be described as follows. For any $w \in F_2$, it has a reduced expression $w = uvu^{-1}$ where v is the unique cyclically reduced word in the conjugacy class of w. Then $\bar{\varphi}_{f,g}(w) = \bar{\varphi}_{f,g}(v) = \varphi_{f,g}(v)$ if v is not a power of the generator a or b, and clearly $\bar{\varphi}_{f,g}(a) = \bar{\varphi}_{f,g}(b) = 0$.

Lemma 2.37. *We have an embedding $\ell^\infty(\mathbb{Z}_+) \to Q(F_2)$ sending f to $\bar{\varphi}_{f,f}$. Moreover, the image intersects $H^1(F_2)$ trivially.*

Proof. By the description above, for any $n \in \mathbb{Z}_+$, we have $\bar{\varphi}_{f,f}(a^n b^n) = 2f(n)$. Thus $\bar{\varphi}_{f,f} = 0$ if and only if $f = 0$. The intersection with $H^1(F_2)$ is trivial since $\bar{\varphi}_{f,f}$ vanishes on the two generators. ☐

Theorem 2.38. *Both $Q(F_2)$ and $H_b^2(F_2; \mathbb{R})$ are infinite-dimensional.*

Proof. Since $\ell^\infty(\mathbb{Z}_+)$ is infinite-dimensional, the embedding above shows that $Q(F_2)$ is infinite-dimensional. The same space embeds in $Q(F_2)/H^1(F_2) \cong H_b^2(F_2; \mathbb{R})$, so it is also infinite-dimensional. ☐

Corollary 2.39. *If G surjects F_2, then $Q(G)$ is infinite-dimensional.*

Proof. $Q(F_2)$ embeds $Q(G)$ by Lemma 2.35. ☐

Note that this applies to all non-abelian free groups and closed hyperbolic surface groups.

Exercise 2.40. *Generalize Rolli's construction to a free product $G = A \star B$. What kind of functions on A, B do you need for the construction to work out?*

2.5.2. *Brooks quasimorphisms.* Now we turn to the quasimorphisms constructed by Brooks, which can be generalized to groups acting on δ-hyperbolic groups. There are two versions, big counting quasimorphisms and little counting quasimorphisms. We follow the exposition in [Cal09, Section 2.3.2].

Let G be a free group generated by S. Fix a reduced word σ. For any $g \in G$, define the big counting function $C_\sigma(g)$ as the number of copies of σ that appear as subwords in the reduced word representing g. Similarly, define the little counting function $c_\sigma(g)$ as the maximal number n such that there are n *disjoint* copies of σ in the reduced expression of g.

Example 2.41.

(1) For $S = \{a, b\}$, $\sigma = ab$, and $g = aba^{-1}baba$, we have $C_\sigma(g) = c_\sigma(g) = 2$.
(2) For $S = \{a, b\}$, $\sigma = aba$, and $g = ababa$, we have $C_\sigma(g) = 2$ but $c_\sigma(g) = 1$. Similarly for $h = abababa$, we have $C_\sigma(h) = 3$ while $c_\sigma(h) = 2$.

For the discussion below, we will use the following observations.

Lemma 2.42. *(1) $C_\sigma(g^{-1}) = C_{\sigma^{-1}}(g)$ and similarly for c_σ.*
(2) σ and σ^{-1} cannot overlap as subwords.

Definition 2.43 (Counting quasimorphisms)**.** The *big counting quasimorphism* associated to σ is $H_\sigma(g) := C_\sigma(g) - C_\sigma(g^{-1})$. The *little counting quasimorphism* associated to σ is $h_\sigma(g) := c_\sigma(g) - c_\sigma(g^{-1})$.

Lemma 2.44. *Clearly from the definition, $H_\sigma(g^{-1}) = -H_\sigma(g)$ and similarly for h_σ.*

We show below that these are indeed quasimorphisms and we estimate their defect. For a reduced expression $u = u_1u_2$, let $s = 1$ (resp. $s = -1$) if a copy of σ (resp. σ^{-1}) appears in the juncture, and let $s = 0$ if no w or w^{-1} appears.

Lemma 2.45. *With the notation above , we have*

$$0 \le s \cdot (H_\sigma(u) - H_\sigma(u_1) - H_\sigma(u_2)) \le |\sigma| - 1,$$

and $h_\sigma(u) - h_\sigma(u_1) - h_\sigma(u_2) = 0$ or s.

Proof. If σ does not appear in the juncture, then $c_\sigma(u) = c_\sigma(u_1) + c_\sigma(u_2)$ and $C_\sigma(u) = C_\sigma(u_1) + C_\sigma(u_2)$.

If σ appears in the juncture, there are at most $|\sigma| - 1$ subwords of length $|\sigma|$ in the juncture, so $C_\sigma(u) - C_{\sigma^{-1}}(u_1) - C_{\sigma^{-1}}(u_2) \le |\sigma| - 1$.

As for c_σ, clearly $c_\sigma(u_1) + c_\sigma(u_2) \le c_\sigma(u)$. Consider a collection of $c_\sigma(u)$ disjoint copies of σ in u. There is at most one copy that lies in the juncture. The remaining copies are disjoint and either in u_1 or u_2. Thus $c_\sigma(u_1) + c_\sigma(u_2) \ge c_\sigma(u) - 1$. Hence $c_\sigma(u) - c_\sigma(u_1) + c_\sigma(u_2) = 0, 1$.

Combining these observations proves the lemma. □

Lemma 2.46. *We have $D(H_\sigma) \le 3(|\sigma| - 1)$ and $D(h_\sigma) \le 3$. Thus H_σ and h_σ are indeed quasimorphisms.*

Proof. For any $g, h \in G$, there are unique reduced expressions $g = uv^{-1}$, $h = vw^{-1}$ and $gh = uw^{-1}$.

Then

$$\begin{aligned} H_\sigma(g) + H_\sigma(h) - H_\sigma(gh) &= H_\sigma(uv^{-1}) + H_\sigma(vw^{-1}) + H_\sigma(wu^{-1}) \\ &= (H_\sigma(uv^{-1}) - H_\sigma(u) - H_\sigma(v^{-1})) \\ &+ (H_\sigma(vw^{-1}) - H_\sigma(v) - H_\sigma(w^{-1})) \\ &+ (H_\sigma(wu^{-1}) - H_\sigma(w) - H_\sigma(u^{-1})), \end{aligned}$$

where each parenthesis has absolute value is at most $|\sigma| - 1$ by Lemma 2.45. Hence $D(H_\sigma) \le 3(|\sigma| - 1)$. The same method shows that $D(h_\sigma) \le 3$. □

Intuitively, we can think of $g, h, (gh)^{-1}$ as a tripod with legs u, v, w. Whenever there is a copy of σ that appears in one of the three legs, there is a copy of σ^{-1} that appears on the opposite side to cancel out the contribution to the defect. So the contribution only comes from copies of σ and σ^{-1} that lie in the juncture, but there are limited spaces at the junctures, giving the defect bound.

It is straightforward to check that

Lemma 2.47. *The homogenization of H_σ applied to g is the number of copies of σ (minus the number of copies of σ^{-1}) as cyclic subwords in the cyclically reduced word representing the conjugacy class of g.*

Exercise 2.48. *Prove the following.*

(1) *If g is a word shorter than σ, then $H_\sigma(g) = 0$.*
(2) *If g is cyclically reduced word and $\sigma = g^2$, then $\bar{H}_\sigma(g) \geq 1 > 0$, although g is shorter than σ.*
(3) *If $\bar{H}_\sigma(g) > 0$ for some g shorter than σ, show that $\sigma = uvu$ as a reduced expression for some nontrivial subword u.*
(4) *In every conjugacy class, there is some (cyclically reduced) σ such that $\bar{H}_\sigma(g) = 0$ for all g shorter than σ.*

    Hint: Use a lexicographical order on reduced words and take a minimal one in the cyclical reduced words in the given conjugacy class. This is from [Tao16, Lemma 3.1].
(5) *Given the above result, find an infinite sequence of distinct σ_n so that $\{\bar{H}_{\sigma_n}\}_n$ are linearly independent homogeneous quasimorphisms.*

2.6. **The rotation quasimorphism and circle dynamics.** Now we construct a quasimorphism in relation to circle dynamics. Throughout this section, let $T = \mathrm{Homeo}^+(S^1)$, the group of orientation-preserving homeomorphisms on the circle. Every such homeomorphism f lifts to an orientation-preserving homeomorphism $\tilde{f}$ on $\mathbb{R}$, where we think of $\mathbb{R}$ as the universal cover of S^1 via the map $\mathbb{R} \to \mathbb{R}/\mathbb{Z} \cong S^1$. Then $\tilde{f}$ commutes with the deck transformation, i.e. $\tilde{f}(x+1) = \tilde{f}(x) + 1$, and conversely any orientation-preserving homeomorphism on $\mathbb{R}$ descends to one on S^1. This leads to a new group

$$\widehat{T} = \{g \in \mathrm{Homeo}^+(\mathbb{R}) \mid g(x+1) = g(x) + 1\},$$

which has a natural map $\pi : \widehat{T} \to T$. The kernel $\ker \pi = \mathbb{Z}$ is the set of translations by integers as the lifts of a given map are unique up to a deck transformation. Thus we have a central extension

$$1 \to \mathbb{Z} \to \widehat{T} \to T \to 1.$$

We aim to build a quasimorphism $\widetilde{\mathrm{rot}}$ on $\widehat{T}$, which descends to a map $\mathrm{rot} : T \to \mathbb{R}/\mathbb{Z}$ so that $\mathrm{rot}(f)$ captures the dynamical properties of f for any $f \in T = \mathrm{Homeo}^+(S^1)$.

For any $p \in \mathbb{R}$, let $\tau_p(f) = f(p) - p$ for any $f \in \widehat{T}$.

Lemma 2.49. *τ_p is a quasimorphism and $D(\tau_p) \leq 1$.*

Proof. For any $f, g \in \widehat{T}$, we need to estimate $\tau_p(fg) - \tau_p(f) - \tau_p(g) = f(g(p)) - g(p) - (f(p) - p)$. Note that changing g to $g+n$ does not affect this quantity for any $n \in \mathbb{Z}$ and the same for f. Hence we may assume $p \leq g(p) < p+1$, and $p \leq f(p) < f(p)+1$. As f is orientation-preserving, the bound for $g(p)$ implies

$$f(p) \leq f(g(p)) < f(p+1) = f(p) + 1.$$

Combining with the bound for $g(p)$ we obtain

$$(f(p) - p) - 1 = f(p) - (p+1) \leq f(g(p)) - g(p) \leq (f(p)+1) - p = (f(p) - p) + 1.$$

Thus

$$|\tau_p(fg) - \tau_p(f) - \tau_p(g)| = |f(g(p)) - g(p) - (f(p) - p)| \leq 1.$$

Hence $D(\tau_p) \leq 1$. □

Lemma 2.50. *For any $p, q \in \mathbb{R}$, the difference $\tau_p - \tau_q$ is a bounded function on $\widehat{T}$.*

Proof. Note that $\tau_{p+n}(f) = f(p+n) - (p+n) = f(p) - p = \tau_p$, so we may assume $q \leq p < q+1$. Then $f(q) \leq f(p) < f(q) + 1$. Thus

$$-1 \leq f(p) - f(q) - (p - q) \leq 1,$$

i.e. $|\tau_p(f)-\tau_q(f)|\le 1$ for all f. □

It follows that the homogenization $\overline{\tau_p}$ does not depend on p.

Definition 2.51 (Rotation quasimorphism)**.** The *rotation quasimorphism* $\widetilde{\mathrm{rot}}\in Q(\widehat{T})$ is the homogenization of τ_p for any $p\in\mathbb{R}$. As an explicit formula, taking $p=0$, we have $\widetilde{\mathrm{rot}}(f)=\lim_{n\to+\infty}\frac{f^n(0)}{n}$.

$\widetilde{\mathrm{rot}}(f)$ is sometimes called the *translation number* of f. It measures how fast f moves forward on average.

For any $f\in T$, its *rotation number* $\mathrm{rot}(f)\in\mathbb{R}/\mathbb{Z}$ is $\widetilde{\mathrm{rot}}(\tilde{f})$ mod $\mathbb{Z}$ for any lift $\tilde{f}\in\widehat{T}$ of f.

Example 2.52. *For any translation $g(x)=x+\alpha$, $\alpha\in\mathbb{R}$, we have $\widetilde{\mathrm{rot}}(g)=\alpha$.*

We will first show that rationality of the rotation number characterizes the existence of periodic orbits. We say $x\in S^1$ is n-periodic under $f\in T$ if $f^n(x)=x$. If $n=1$, then x is a fixed point. We say f has a periodic orbit if there is some x is n-periodic for some n.

Lemma 2.53. *$g\in\widehat{T}$ has a fixed point if and only if $\widetilde{\mathrm{rot}}(g)=0$.*

Proof. Suppose $g(p)=p$, then $g^n(p)=p$ for all n, i.e. $\tau_p(g^n)=0$. Hence $\widetilde{\mathrm{rot}}(g)=0$.

Suppose g has not fixed point, then either $g(x)<x$ for all $x\in\mathbb{R}$ or $g(x)>x$ for all $x\in\mathbb{R}$. Without loss of generality, suppose $g(x)>x$ for all $x\in\mathbb{R}$. Note $t:=\inf_{x\in\mathbb{R}}(g(x)-x)=\inf_{x\in[0,1]}(g(x)-x)$ as $g(x+1)=g(x)+1$, it must be achieved by some $x_0\in[0,1]$ by compactness. So we have $g(x)-x\ge g(x_0)-x_0=t>0$. Thus $g^n(0)>g^{n-1}(0)+t>\cdots>nt$, and $\widetilde{\mathrm{rot}}(g)=\lim\frac{g^n(0)}{n}\ge t>0$. Thus g must have fixed point if $\widetilde{\mathrm{rot}}(g)=0$. □

Lemma 2.54. *$g\in\widehat{T}$ has $\widetilde{\mathrm{rot}}(g)=\frac{m}{n}$ as a reduced fraction with $n\in\mathbb{Z}_+$ and $m\in\mathbb{Z}\setminus\{0\}$ if and only if there is $p\in\mathbb{R}$ such that $g^n(p)=p+m$.*

Proof. If $g^n(p)=p+m$, then it is easy to see that $\widetilde{\mathrm{rot}}(g)=\frac{m}{n}$ as before.

If $\widetilde{\mathrm{rot}}(g)=\frac{m}{n}$, then $\widetilde{\mathrm{rot}}(g^n)=m$ and $\widetilde{\mathrm{rot}}(g^n-m)=0$. Thus by Lemma 2.53 we know $(g^n-m)(p)=p$ for some $p\in\mathbb{R}$, i.e. $g^n(p)=p+m$. □

Theorem 2.55. *For $f\in T$, f has a periodic orbit if and only if $\mathrm{rot}(f)\in\mathbb{Q}\mod\mathbb{Z}$. Moreover, f has an n-periodic point if and only if $\mathrm{rot}(f)\in\frac{1}{n}\mathbb{Z}\mod\mathbb{Z}$. In particular, f has fixed points if and only if $\mathrm{rot}(f)\equiv 0\mod\mathbb{Z}$.*

Proof. Let $\tilde{f}\in\widehat{T}$ be an arbitrary lift of f. Then $x\in S^1$ is n-periodic under f if and only if it has a lift $p\in\mathbb{R}$ such that $\tilde{f}^n(p)=p+m$ for some $m\in\mathbb{Z}$. So the conclusion follows easily from the previous lemmas. □

Here is the general structure of an element $f\in T$ acting with a fixed point. Let $\mathrm{Fix}(f)$ be the set of fixed points, which is a closed subset of S^1.

Lemma 2.56. *The action of f on each complementary interval I of $\mathrm{Fix}(f)$ is conjugate to a nontrivial translation T on $\mathbb{R}$. That is, there is a homeomorphism $h:I\to\mathbb{R}$ such that $hf=Th$.*

Proof. Choose any $x_0\in I$ and let $x_n=f^n(x_0)$ for all $n\in\mathbb{Z}$. Since $I\cap\mathrm{Fix}(f)=\emptyset$, $f|_I$ is a monotone. Without loss of generality, assume $x_{n+1}>x_n$ for all n (where the order is induced from the orientation on $I\subset S^1$). Let $I_n=[x_n,x_{n+1}]$. Then $\cup_n I_n=I$ since $\{x_n\}$ has no accumulation point inside I (which would be a fixed point of f if otherwise). Note that $f^n(I_0)=I_n$ for all $n\in\mathbb{Z}$.

Choose an arbitrary homeomorphism $h_0:I_0\to[0,1]$ with $h_0(x_0)=0$ and $h_0(x_1)=1$. Define $h:I\to\mathbb{R}$ by setting $h|_{I_n}=T^nh_0f^{-n}$ for all $n\in\mathbb{Z}$, where $T(x)=x+1$ is the unit translation on $\mathbb{R}$. It follows that $h(I_n)=[n,n+1]$ and $h(x_n)=n$ for all $n\in\mathbb{Z}$. It is easy to check by construction that h is a homeomorphism[2] and $hf=Th$. □

[2]If one asks for h that is orientation preserving where $\mathbb{R}$ is equipped with the usual orientation and I has the induced orientation, then we might need to choose T as $T(x)=x-1$

As a byproduct, we have the following observations.

Lemma 2.57. *If $f \in T$ has fixed points, then $f = [a, b]$ for some $a, b \in T$.*

Proof. Note that the unit translation $T(x) = x + 1$ is a commutator $[u, v]$ of two dilations $u(x) = 2x$ and $v(x) = 2(x+1) - 1$. Thus by the previous lemma, for each complementary interval I of $\mathrm{Fix}(f)$, there are homeomorphisms u_I, v_I on $\bar{I}$ fixing both end points such that $f|_{\bar{I}} = [u_I, v_I]$. Define $u|_{\bar{I}} = u_I$ for each complementary interval I and $u|_{\mathrm{Fix(f)}} = id|_{\mathrm{Fix(f)}}$, and define v similarly. Then by construction, $u, v \in T$ and $[u, v]|_I = f|_I$. It follows that $f = [u, v]$ as desired. □

Proposition 2.58. $T = \mathrm{Homeo}^+(S^1)$ *is uniformly perfect. More precisely, each $f \in T$ is a product of at most two commutators.*

Proof. By the previous lemma, It suffices to show that for any $f \in T$ there is a commutator $[a, b]$ with $a, b \in T$ such that $[a, b]f$ has a fixed point. Indeed, for any $x, y \in S^1$, there is a commutator that takes x to y, which can be done by fixing an arbitrary nontrivial commutator and conjugating it appropriately. □

By Corollary 2.34, we have

Corollary 2.59. $Q(\mathrm{Homeo}^+(S^1)) = 0$.

Let us now go back to the relation between rotation numbers and circle dynamics. Ideally, one would like to understand the dynamics up to conjugacy. However, in circle dynamics, it is often natural to consider a weaker equivalence, generated by *semi-conjugacy*.

Definition 2.60. A map $h : S^1 \to S^1$ is a *semi-conjugacy* between two actions $\rho_1, \rho_2 : G \to T$ of a group G if h is a surjective continuous map of degree one and $h\rho_1(g) = \rho_2(g)h$ for all g. When h is a homeomorphism, then it is called a conjugacy.

Here we are considering a single map $f \in T$, which we can treat as a $\mathbb{Z}$ action on S^1 via $n \mapsto f^n$.

Denjoy's construction is a standard way to cook up a new action on the circle from an existing one that is semi-conjugate but often not conjugate.

Example 2.61 (Denjoy's construction)**.** *Suppose $\rho : G \to T$ is an action that admits a countable G-invariant subset $O \subset S^1$. Note that O always exists if G is a countable group, say by taking a G-orbit. Enumerate O as $O = \{x_n\}_{n\geq 1}$ and choose a sequence of positive integers $\{a_n\}_{n\geq 1}$ such that $\sum_{n\geq 1} a_n < \infty$.*

Replace each x_n by a closed interval I_n of length a_n, we obtain a new circle $Y \cong S^1$ and a natural surjective continuous map $h : Y \to S^1$ of degree one, where h collapses each I_n to x_n. We have a new action ρ' of G on Y: for each $g \in G$, define $\rho'(g)(x) = x$ if $x \notin \cup I_n$ and $\rho'(g) : I_n \to I_m$ is the unique orientation preserving linear homeomorphism if $\rho(g)(x_n) = x_m$. Then $h\rho'(g) = \rho(g)h$ for all $g \in G$.

An action is minimal *if the only closed invariant subsets are either empty or the entire space, or equivalently, every orbit is dense. Note that the action on Y is never* minimal *since $Y \setminus \cup_n \mathrm{int}(I_n)$ is a nontrivial closed invariant subset, where $\mathrm{int}(I_n)$ is the interior of I_n. Hence if the starting action ρ is minimal, then the two actions are semi-conjugate but not conjugate.*

Now we consider the case where the rotation number of $f \in T$ is irrational and show that the action is semi-conjugate to the irrational rigid rotation by $\mathrm{rot}(f)$. This is an old theorem due to Poincaré.

Theorem 2.62 (Poincaré)**.** *If $f \in T$ has $\mathrm{rot}(f) = \alpha \notin \mathbb{Q}/\mathbb{Z}$, then there is a surjective continuous degree one map $h : S^1 \to S^1$ such that $hf = T_\alpha h$, where T_α is the rigid rotation on $S^1 = \mathbb{R}/\mathbb{Z}$ by α. Moreover, if f acts minimally, then h is a homeomorphism.*

To prove this, we need the following lemmas.

Lemma 2.63. *For any $g \in \widehat{T}$, the set $\{g^n(x) - x - n\widetilde{\mathrm{rot}}(g)\}_{n\in\mathbb{Z}}$ is bounded for any $x \in \mathbb{R}$.*

Proof. Recall that $\tau_x(g) = g(x) - x$ and $\widetilde{\mathrm{rot}}$ is the homogenization of τ_x. Then $\tau_x - \widetilde{\mathrm{rot}}$ is a bounded function. In particular, its evaluation on $\{g^n\}_{\mathbb{Z}}$ must be bounded, that is, $\{g^n(x) - x - n\widetilde{\mathrm{rot}}(g)\}_{n\in\mathbb{Z}}$ is a bounded set. □

Remark 2.64. *One can use this property as the definition of $\widetilde{\mathrm{rot}}(g)$.*

Lemma 2.65. *For any $g \in \widehat{T}$ with $\widetilde{\mathrm{rot}}(g) = \alpha$, there is a monotone map $\tilde{h} : \mathbb{R} \to \mathbb{R}$ such that $\tilde{h}(x+1) = \tilde{h}(x) + 1$ and $\tilde{h}g = T_\alpha \tilde{h}$, where T_α is the translation by α.*

Proof. Define $\tilde{h}(x) := \sup_{n\in\mathbb{Z}}\{g^n(x) - n\widetilde{\mathrm{rot}}(g)\}$ for each $x \in \mathbb{R}$. Note that this is a finite number as we are taking the supremum of the bounded set in Lemma 2.63 translated by x.

Clearly $g^n(x+1) - n\widetilde{\mathrm{rot}}(g) = 1 + [g^n(x) - n\widetilde{\mathrm{rot}}(g)]$, so $\tilde{h}(x+1) = \tilde{h}(x) + 1$. Moreover, $\tilde{h}(x) \geq \tilde{h}(y)$ for all $x \geq y$ since $g^n(x) \geq g^n(y)$.

Since $g^n(g(x)) - n\widetilde{\mathrm{rot}}g = \widetilde{\mathrm{rot}}(g) + [g^{n+1}(x) - (n+1)\widetilde{\mathrm{rot}}(g)]$, we have $\tilde{h}(g(x)) = \tilde{h}(x) + \widetilde{\mathrm{rot}}(g) = \tilde{h}(x) + \alpha$. Thus $\tilde{h}g = T_\alpha \tilde{h}$. □

Proof of Theorem 2.62. Fix a lift $g = \tilde{f} \in \widehat{T}$ of f, and let $\tilde{h}$ be the map from Lemma 2.65. It descends to a (degree one) map $h : S^1 \to S^1$ since $\tilde{h}$ commutes with integral translations. Since $\tilde{h}$ is monotone, it is continuous if and only if it is surjective (i.e. it has no jumps). Let $\mathrm{Jump}(\tilde{h})$ be the complement of the image of $\tilde{h}$, which is a countable union of nontrivial intervals, invariant under integral translations and translation by α. Thus they descend to a countable union of intervals on S^1 invariant under T_α.

Since α is irrational, every orbit of T_α is dense, so $\mathrm{Jump}(\tilde{h})$ is empty. Hence $\tilde{h}$ is continuous and surjective, and so is h. Then it is easy to check that h gives the desired semi-conjugacy.

The preimage $\tilde{h}^{-1}(x)$ of any $x \in \mathbb{R}$ is an interval. The union U of the interiors of such intervals is invariant under integral translations and under $\tilde{f}$. So it descends to a union of open intervals on S^1 invariant under f. If every orbit of f is dense, then U must be empty. So in this case, $\tilde{h}$ and h are injective. Hence h is a conjugacy when f acts minimally. □

Remark 2.66. *When f has rational rotation, it is possible that* $\mathrm{Jump}(\tilde{h})$ *is a nontrivial collection of intervals. One can collapse them to get a semi-conjugacy, but the resulting space is not a circle if h has finite image. However, one can first "blow up" the starting circle by Denjoy's construction so that* $\mathrm{Fix}(f)$ *has interiors, then the semi-conjugacy works for the blow-up circle. Thus f is equivalent to the rigid rotation by* $\mathrm{rot}(f)$ *in all cases under the semi-conjugacy relations.*

Exercise 2.67. *Let $f \in T$ be the boundary map of a parabolic element in* $\mathrm{PSL}_2(\mathbb{R})$. *How does $\tilde{h}$ behave in the construction above? What if f is the boundary map of a hyperbolic element?*

For the more general question of understanding G actions on S^1 for groups G other than $\mathbb{Z}$, we will discuss it later using the bounded Euler class living in $H^2_b(G;\mathbb{Z})$, which is closely related to the rotation number.

2.7. **Amenable groups.** We discuss basic properties of amenable groups and their relations to bounded cohomology. There are many equivalent definitions of amenability. Here we start with two most common ones and focus on the case of discrete groups (instead of the more general setting of locally compact groups). Let $L^\infty(G) = C^1_b(G)$ be the space of bounded functions on G equipped with the sup norm, which admits a G-action, where $(gf)(h) = f(g^{-1}h)$ for all $h \in G$ and $f \in L^\infty(G)$.

Definition 2.68 (Amenable)**.** A discrete group G is *amenable* if there is an invariant mean m, which is a linear functional $m : L^\infty(G) \to \mathbb{R}$ such that

(1) $m(f) \geq 0$ if $f \geq 0$,
(2) $m(1_G) = 1$, where 1_G is the function taking constant value 1, and

(3) $m(gf) = m(f)$ for all $g \in G$ and $f \in L^\infty(G)$.

Equivalently, one can think of this as a (left) G-invariant *finitely additive* (non-negative) measure μ on G with total mass 1, where $\mu(A) = m(1_A)$ for all $A \subset G$. The other direction of the equivalence is given by $m(f) = \int_G f \, d\mu$.

Example 2.69.

(1) *Any finite group G is amenable, where $m(f) = \frac{1}{|G|}\sum_{g\in G} f(g)$.*

(2) *The free group $G = F_2$ is not amenable. Suppose there is a G-invariant finitely additive measure μ. Then $\mu(\{id\}) = 0$ since otherwise $\mu(G)$ won't be finite. We express $G \setminus \{id\} = X_a \sqcup X_{a^{-1}} \sqcup X_b \sqcup X_{b^{-1}}$, where X_* is the set of elements starting with $*$. Note that $a(X_b \sqcup X_{b^{-1}}) \subset X_a$, so $\mu(X_b \sqcup X_{b^{-1}}) \le \mu(X_a)$. Similarly $\mu(X_a \sqcup X_{a^{-1}}) \le \mu(X_b)$. Thus $\mu(X_a) \le \mu(X_b) \le \mu(X_a) - \mu(X_{b^{-1}})$. So we must have $\mu(X_{b^{-1}}) = 0$. For the same reason, we must have $\mu(X_a) = \mu(X_{a^{-1}}) = \mu(X_b) = 0$, which implies $\mu(G) = 0$, contradicting that $\mu(G) = 1$.*

Amenability interacts nicely with subgroups and quotients.

Lemma 2.70.

(1) *If G is amenable, then every subgroup $H \le G$ is amenable.*

(2) *If G is amenable, then every quotient group $Q = G/N$ is amenable.*

(3) *If for a normal subgroup $N \lhd G$, both N and $Q = G/N$ are amenable, then G is amenable.*

Proof.

(1) Let μ be a left G-invariant finitely additive measure with $\mu(G) = 1$. Let $\{Hg_\lambda\}_{\lambda\in\Lambda}$ be the right cosets. For any $A \subset H$, define $\nu(A) := \mu(\cup_{\lambda\in\Lambda} Ag_\lambda)$. Then clearly ν is a finitely additive measure on H with $\nu(H) = 1$, and left G-invariance of μ implies left H-invariance of ν.

(2) Let μ be a G-invariant finitely additive measure with $\mu(G) = 1$. For any $A \in Q$, define $\nu(A) = \mu(\pi^{-1}(A))$, where $\pi : G \to Q$ is the quotient map. Then clearly ν is a finitely additive measure on Q with $\nu(Q) = 1$. To see the Q-invariance, for any $q \in Q$, let $g \in \pi^{-1}(q)$, and then $\pi^{-1}(qA) = g\pi^{-1}(A)$, so $\nu(qA) = \mu(g\pi^{-1}(A)) = \mu(\pi^{-1}(A)) = \nu(A)$ by G-invariance of μ.

(3) Let μ_N and μ_Q be the invariant measures witnessing the amenability of N and Q respectively. For any $A \subset G$, define a function $f_A : Q \to \mathbb{R}$ by $f_A(gN) = \mu_N(N \cap g^{-1}A)$. This is a bounded function on Q so we can define $\mu(A) := \int_Q f_A \, d\mu_Q$. It is easy to check that μ is a finitely additive measure with $\mu(G) = 1$. To see its invariance, note that $f_{hA}(hgN) = \mu_N(N \cap g^{-1}A) = f_A(gN)$, so $\bar{h}^{-1} f_{hA} = f_A$, where $\bar{h}$ is the image of $h \in G$ in Q. So f_{hA} and f_A have the same integral and hence $\mu(hA) = \mu(A)$ for all $h \in G$.

□

Proposition 2.71. *$\mathbb{Z}$ is amenable.*

Proof. Let $\mathcal{P}(\mathbb{Z})$ be the power set of $\mathbb{Z}$. Finitely additive probability measures on $\mathbb{Z}$ is a closed subset P of $[0,1]^{\mathcal{P}(\mathbb{Z})}$, equipped with the product topology. For any $\epsilon > 0$, consider the subset $P_\epsilon \subset P$ of ϵ-almost invariant measures, i.e. those $\mu \in P$ satisfying $|\mu(zA) - \mu(A)| \le \epsilon$ for all $A \subset \mathbb{Z}$ and a chosen generator $z \in \mathbb{Z}$. This is a closed subset and thus compact since $[0,1]^{\mathcal{P}(\mathbb{Z})}$ is compact by Tychonoff's theorem. The intersection $\cap_{\epsilon>0} P_\epsilon$ consists of $\mathbb{Z}$-invariant measures, which is nonempty as long as we show $P_\epsilon \ne \emptyset$ for each $\epsilon > 0$.

For each $N \in \mathbb{Z}_+$, let $\mu_N(A) = \frac{|A\cap[-N,N]|}{2N+1}$. Then $\mu_N \in P$ and $|\mu_N(gA) - \mu_N(A)| \le \frac{2}{2N+1}$. So $\mu_N \in P_\epsilon$ if N is large enough, verifying that P_ϵ is nonempty. This completes the proof. □

Remark 2.72. *The subset $[-N, N]$ of $\mathbb{Z}$ has the property that its boundary is small compared to the size of the entire subset, and the ratio goes to zero as N goes to infinity. This is an example of* *a* Følner sequence. *The argument above generalizes to show that any group with a Følner sequence is amenable, which is actually an equivalent definition.*

Exercise 2.73. *Show that the direct limit of amenable groups is amenable.*

Proposition 2.74. *Solvable groups are amenable.*

Proof. Note that we have shown that all cyclic groups are amenable. By Lemma 2.70 (3) and induction, all finitely generated abelian groups are amenable. Any abelian group is the direct limit of finitely generated ones, so they are amenable by the exercise above.

For any solvable group, we can see by induction and Lemma 2.70 (3) that its derived subgroups are amenable, based on the fact abelian groups are amenable. □

2.8. **Vanishing results.** Amenable groups have trivial bounded cohomology and in particular no interesting homogeneous quasimorphisms. We give a direct proof of the latter result.

Proposition 2.75. *If G is amenable, then $Q(G) = H^1(G)$.*

Proof. Given any $\varphi \in Q(G)$, we aim to show that φ is a homomorphism. Note that $\varphi(gh) - \varphi(g) - \varphi(h)$ is a bounded function in variables $g, h \in G$. So for any fixed g, the one-variable function $f_g : G \to \mathbb{R}$ defined by $f_g(h) = \varphi(gh) - \varphi(h)$ is bounded. As G is amenable, let $f(g) = m(f_g) \in \mathbb{R}$. This provides a function $f : G \to \mathbb{R}$, which we claim to be a homomorphism and is bounded away from φ. Then $\varphi = f$ by Lemma 2.24.

Indeed, the function $f_g - \varphi(g)$ is bounded by $-D(\varphi)$ and $D(\varphi)$ from below and above, so its mean $m(f_g) - \varphi(g) = f(g) - \varphi(g)$ has absolute value bounded by $D(\varphi)$, which holds for all $g \in G$. So $|f - \varphi| \leq D(\varphi)$.

To show that f is a homomorphism, note that $f(g_1g_2)$ is the mean of $f_{g_1g_2}$ and

$$\begin{aligned} f_{g_1g_2}(h) &= \varphi(g_1g_2h) - \varphi(h) \\ &= [\varphi(g_1g_2h) - \varphi(g_2h)] + [\varphi(g_2h) - \varphi(h)] \\ &= f_{g_1}(g_2h) + f_{g_2}(h) \\ &= (g_2^{-1}f_{g_1})(h) + f_{g_2}(h). \end{aligned}$$

So

$$m(f_{g_1g_2}) = m(g_2^{-1}f_{g_1}) + m(f_{g_2}) = m(f_{g_1}) + m(f_{g_2})$$

by the invariance of the mean m. That is, $f(g_1g_2) = f(g_1) + f(g_2)$. So f is a homomorphism as desired. □

Instead of directly proving vanishing results for bounded cohomology of amenable groups, we deduce it from the following exact sequence for amenable covers.

Theorem 2.76 (Johnson, Trauber, Gromov)**.** *Suppose there is a short exact sequence of groups*

$$1 \to K \to G \to A \to 1,$$

where A is amenable. Then there the induced map $H_b^n(G) \to H_b^n(K)^A$ is an isometric isomorphism with respect to the standard semi norm, where $H_b^n(K)^A$ denotes the A invariant part. In particular (by taking $K = 1$ and $G = A$), $H_b^n(A) = 0$ for all n.

Proof. Here we give a topological proof, interpreting $H_b^n(G)$ as the bounded singular cohomology of a $K(G,1)$ space X_G; More details about this equivalence will be explained in the next section.

Let $\pi : X_K \to X_G$ be the covering space of X_G corresponding to the normal subgroup K. Then A acts on X_K as deck transformations. Then $C_b^n(X_G)$ naturally corresponds to $C_b^n(X_K)^A$ by pullback, the A-invariant cochains on X_K. This correspondence commutes with the coboundary and induces a map $\pi^* : H_b^n(X_G) \to H_b^n(X_K)$, which is non-increasing with respect to the sup norm, and the image clearly lies in $H_b^n(X_K)^A$.

Let m be an invariant mean on A. This gives a map going backwards $m : C_b^n(X_K) \to C_b^n(X_K)^A$ as follows. For any $\alpha \in C_b^n(X_K)$ and singular simplex $\Delta \to X_K$, we have a bounded function f on A defined as $f(a) = \alpha(a\Delta)$. Define $m(\alpha)(\Delta) = m(f)$. The A-invariance implies that $m(\alpha)$ is A-invariant. It is straightforward to check that m is linear, commutes with the coboundary, and $m \circ \pi = id$. Thus it induces a map $m^* : H_b^n(X_K) \to H_b^n(X_G)$ so that $m^* \circ \pi^* = id$. This shows that π^* is injective. Moreover, both m^* and π^* are non-increasing with respect to the sup norm, so $m^* \circ \pi^* = id$ also implies that π^* is an isometric embedding.

To see that the image of π^* is $H_b^n(X_K)^A$, it suffices to show that any class $\sigma \in H_b^n(X_K)^A$ is represented by a cocycle in $C_b^n(X_K)^A$. Let $\alpha \in C_b^n(X_K)$ be a cocycle representing σ. Since σ is A-invariant, we know that, for any $a \in A$, we have $a\alpha - \alpha = \delta f_a$ for some $f_a \in C_b^{n-1}(X_K)$. For every singular $(n-1)$-simplex $\Delta \to X_K$, define $f(\Delta)$ as the mean of $f_a(\Delta)$ as a bounded function over $a \in A$. It is bounded because α is. Then it follows by averaging that $m(\alpha) - \alpha = \delta f$, so $[m(\alpha)] = [\alpha] = \sigma$ and $m(\alpha) \in C_b^n(X_K)^A$ as desired. □

There is also a result for surjection with amenable kernel.

Theorem 2.77 (Gromov)**.** *Suppose there is a short exact sequence of groups*

$$1 \to A \to G \to H \to 1,$$

where A is amenable. Then the induced map $H_b^n(H) \to H_b^n(G)$ is an isometric isomorphism.

In degree two, this generalizes to the (left) exactness of H_b^2.

Theorem 2.78 (Bouarich)**.** *Suppose there is an exact sequence of groups*

$$K \xrightarrow{i} G \xrightarrow{\pi} H \to 1.$$

Then the following induced sequence is exact:

$$0 \to H_b^2(H) \to H_b^2(G) \to H_b^2(K).$$

Similarly,

$$0 \to Q(H) \to Q(G) \to Q(K).$$

Proof. We prove this for homogeneous quasimorphisms. This is to show the exactness of $0 \to Q(H) \to Q(G) \to Q(K)$. The case of H_b^2 is similar and can be found in [Cal09, Section 2.7.2].

The injectivity $\pi^* : Q(H) \to Q(G)$ follows from the fact that this map is defect non-decreasing, which is apparent by the surjectivity of $G \to H$.

The remaining nontrivial part is to show that $\ker i^* \subset \operatorname{Im} \pi^*$. Suppose $\varphi \in Q(G)$ vanishes on the image of K. We may assume $K = \ker \pi$. Then we aim to show that $\varphi(gk) = \varphi(g)$ for all $g \in G$ and $k \in K$. Note that for any $n \in \mathbb{Z}$, we have $\pi((gk)^n) = \pi(g^n)$, so $(gk)^n \cdot g^{-n} \in K$ and $\varphi((gk)^n \cdot g^{-n}) = 0$. Hence

$$|\varphi((gk)^n) + \varphi(g^{-n})| \le D(\varphi),$$

that is $|\varphi(gk) - \varphi(g)| \le D(\varphi)/n$. Taking $n \to \infty$ we conclude $\varphi(gk) = \varphi(g)$ for all $g \in G$ and $k \in K$ as desired. □

2.9. **Bounded cohomology of topological spaces.** Given a topological space X, its singular cohomology is defined by cochains $C^n(X; \mathbb{R})$, which are functions on the space of singular simplices. The *bounded singular cohomology* $H_b^n(X; \mathbb{R})$ is defined in the same way replacing each $C^n(X; \mathbb{R})$ by the subspace $C_b^n(X; \mathbb{R})$ of *bounded* functions, which is equipped with a sup-norm. The norm induces a semi norm $\|\cdot\|_\infty$ on $H_b^n(X; \mathbb{R})$ as usual, by minimizing the sup norm of bounded cochains representing the given bounded cohomology class.

Then it is straightforward to see that

Lemma 2.79. *Any continuous map $f : X \to Y$ induces a map $f^* : H_b^n(Y) \to H_b^n(X)$ such that $\|f^*(\alpha)\|_\infty \leq \|\alpha\|_\infty$ for all $\alpha \in H_b^n(Y)$.*

In particular, if f is a homotopy equivalence, then f^* is an isometric isomorphism.

Proof. This is analogous to Proposition 1.3 and Corollary 1.4. □

Its relation to bounded group cohomology is given by Gromov's mapping theorem.

Theorem 2.80 (Gromov's mapping theorem)**.** *Let X be a CW complex with $\pi_1(X) = G$. Then the natural map $K(G, 1) \to X$ induces an isometric isomorphism $H_b^n(X) \cong H_b^n(G)$.*

This is astonishing as the bounded cohomology only depends on $\pi_1(X)$. When X itself is a $K(G, 1)$, which up to homotopy we may assume to be the simplicial complex BG that we built explicitly, the theorem asserts that the bounded cohomology via singular simplices on BG agrees with the bounded cohomology via simplicial simplices on BG.

One way to prove this is to establish a theory of strongly injective resolutions along the lines of homological algebra, introduced by Ivanov. We will focus more on applications and skip this. See [Fri17] or [Iva17] for more details.

Combining this with Theorem 2.77 we have

Theorem 2.81 (Gromov)**.** *For any $f : X \to Y$ such that $f_* : \pi_1(X) \to \pi_1(Y)$ is surjective with amenable kernel, the induced map $f^* : H_b^n(Y) \to H_b^n(X)$ is an isometric isomorphism.*

As in the bounded group cohomology, we have a comparison map $c : H_b^n(X; \mathbb{R}) \to H^n(X; \mathbb{R})$.

3. More on Gromov's simplicial norm

Further properties of the simplicial norm can be deduced by a duality principle between homology and bounded cohomology.

3.1. Duality.

There is duality principle relating the norm on bounded cohomology to the simplicial norm. Note that the usual pairing $H^n(X; \mathbb{R}) \times H_n(X; \mathbb{R}) \to \mathbb{R}$ gives a way to pair bounded cohomology classes with homology classes, by composing with the comparison map. The usual ℓ^1-ℓ^∞ duality implies that, for any singular chain $s \in C_n(X)$, we have

$$|s|_1 = \sup_{f \in C_b^n(X)} \frac{\langle f, s \rangle}{|f|_\infty} = \sup_{f \in C_b^n(X), |f|_\infty \leq 1} \langle f, s \rangle.$$

Recall that $\|\sigma\|_1 = \inf_{[s]=\sigma} |s|_1$, which can be thought of as the induced norm on the quotient by $\operatorname{Im} \partial$. The dual space of a quotient consists of the dual functions that vanish on the subspace, in our context those f with $\delta f = 0$, i.e. cocycles. This means, for any $\sigma \in H_n(X; \mathbb{R})$ we have

$$\|\sigma\|_1 = \sup_{f \in C_b^n(X), \delta f = 0, |f|_\infty \leq 1} \langle f, \sigma \rangle.$$

The pairing does not change if we replace f by another cocycle in the same class, so we get

$$\|\sigma\|_1 = \sup_{\alpha \in H_b^n(X), \|\alpha\|_\infty \leq 1} \langle \alpha, \sigma \rangle.$$

This is called the duality principle.

Proposition 3.1 (Duality principle)**.** *For any $\sigma \in H_n(X; \mathbb{R})$, we have*

$$\|\sigma\|_1 = \sup_{\alpha \in H_b^n(X), \|\alpha\|_\infty \leq 1} \langle \alpha, \sigma \rangle.$$

Corollary 3.2. *If $H_b^n(X) = 0$, then $\| \cdot \|_1$ vanishes on $H_n(X)$.*

Example 3.3. *We showed that $H_b^1(X; \mathbb{R}) = 0$, so $\|\cdot\|_1$ vanishes on $H_1(X; \mathbb{R})$ by the duality principle, which we proved directly in Corollary 1.16.*

Recall that we showed $\|M\|_1 = 0$ for $M = S^n, T^n$ by constructing self maps with nonzero degree. In both examples, the fundamental groups are amenable.

Proposition 3.4. *If $\pi_1(M)$ is amenable, then the simplicial volume $\|M\|_1 = 0$.*

Proof. We know amenable groups have trivial bounded cohomology by Theorem 2.76, so by Gromov's mapping theorem 2.80 we know $H_b^n(M) = 0$ for all n. In particular, taking $n = \dim M$, we obtain $\|M\|_1 = 0$ by Corollary 3.2. □

In particular, every simply connected closed manifold has vanishing simplicial volume, which is not obvious from the definition.

Example 3.5. *For a closed hyperbolic n-manifold M, the volume form ω represents a cohomology class in $H^n(M)$. The volume form itself is not a bounded cocycle, since there are large fat simplices with arbitrarily large volume. However, the composition $\omega \circ \mathrm{str}$ is bounded. This gives rise to a bounded cohomology class $\alpha \in H_b^n(M)$ so that its image under the comparison map is $[\omega]$. Moreover, $\|\alpha\|_\infty \le \|\omega \circ \mathrm{str}\|_\infty = \sup \mathrm{vol}(\Delta^n) = v_n$. Hence $\frac{1}{v_n}\|\alpha\|_\infty$ has norm no more than 1. So by the duality principle,*

$$\|M\|_1 \ge \frac{1}{v_n}\langle \alpha, [M]\rangle = \frac{1}{v_n}\langle \omega, [M]\rangle = \frac{\mathrm{vol}(M)}{v_n}.$$

This essential the same straightening argument we used to show $\|M\|_1 \ge \frac{\mathrm{vol}(M)}{v_n}$, written in terms of the duality principle. Since the equality holds by the Proportionality Theorem, we conclude that the bound $\|\alpha\|_\infty \le v_n$ is sharp, i.e. $\|\alpha\|_\infty = v_n$.

3.2. **Additivity.** The goal of this section is to show Gromov's additivity theorem

Theorem 3.6 (Gromov)**.** *For $n \ge 3$, let M be the connected sum of closed orientable n-manifolds M_1 and M_2. Then $\|M\|_1 = \|M_1\|_1 + \|M_2\|_1$.*

Here the assumption on the dimension $n \ge 3$ is important. When $n = 2$, if M_1 and M_2 have genus $g_1 \ge 1$ and $g_2 \ge 1$ respectively, then M has genus $g_1 + g_2$. Then by Theorem 1.20, we know $\|M\|_1 = -2\chi(M) = 4(g_1 + g_2 - 1)$, while $\|M_1\|_1 + \|M_2\|_1 = -2\chi(M_1) - 2\chi(M_2) = 4(g_1 + g_2 - 2)$. So in this case, we have $\|M\|_1 > \|M_1\|_1 + \|M_2\|_1$.

We will deduce Theorem 3.6 from a more general result [Fri17, Theorem 7.6]. Let $n \ge 2$, and let $\{(M_i, \partial M_i)\}_{i\in I}$ be a finite collection of orientable compact connected n-manifolds with boundary. We say $(M, \partial M)$ is obtained by gluing $\{(M_i, \partial M_i)\}_{i\in I}$ together, if there are orientation reversing homeomorphisms $f_j : S_j^+ \to S_j^-$ of $(n-1)$-manifolds $\{S_j^\pm\}_{j\in J}$, such that each $S_j^\pm$ is a component of some ∂M_i with induced orientation, and each component of M_i is identified with at most one $S_j^\pm$. Then ∂M is necessarily the union of those boundary components of M_i's that never appear as some $S_j^\pm$.

We say the gluing is *compatible* if $f_{j*}K_j^+ = K_j^-$, f_{j*} is the induced map on π_1 and $K_j^\pm$ is the kernel of the inclusion map $\pi_1(S_j^\pm) \to \pi_1(M_i)$ if $S_j^\pm$ is a component of ∂M_i. Note that if boundary inclusion $S_j^\pm \to M_i$ is π_1-injective, then the gluing is compatible.

Theorem 3.7. *Let $(M, \partial M)$ be obtained by gluing $\{(M_i, \partial M_i)\}_{i\in I}$ together. If each component of ∂M_i has amenable fundamental group, then*

$$\|M, \partial M\|_1 \le \sum_{i\in I} \|M_i, \partial M_i\|.$$

If in addition the gluing is compatible, then

$$\|M, \partial M\|_1 = \sum_{i\in I} \|M_i, \partial M_i\|.$$

Here are some consequences and examples.

When the boundary components are simply connected, then the assumptions of the theorem (amenability and compatibility) hold automatically.

Corollary 3.8. *For $n \geq 3$, if $(N, \partial N)$ is obtained from a closed orientable n-manifold M by removing an open n-ball. Then $\|N, \partial N\|_1 = \|M\|_1$.*

Proof. In this case, we have $\partial N \cong S^{n-1}$, and M is obtained by gluing $(N, \partial N)$ with (B^n, S^{n-1}). Since $\pi_1(\partial N) = \pi_1(S^{n-1}) = 1$ as $n \geq 3$, the gluing is compatible. So by Theorem 3.7,

$$\|N, \partial N\|_1 = \|M\|_1 - \|B^n, S^{n-1}\|_1 = \|M\|_1.$$

The fact that $\|B^n, S^{n-1}\|_1 = 0$ can be obtained by a self map with degree greater than one, or by the equality above for the case $M = S^n$. □

This fails when $n = 2$ as $\|S\|_1 = -2\chi^-(S)$ for all surfaces possibly with boundary.

Then we can deduce Theorem 3.6.

Proof of Theorem 3.6. Let $(N_i, \partial N_i)$ be M_i with an open n-ball removed, $i = 1, 2$. Then M is obtained by gluing $(N_i, \partial N_i)$ together. As $N_i \cong S^{n-1}$, which is simply connected, the gluing is compatible. Thus $\|M\|_1 = \|N_1\|_1 + \|N_2\|_1 = \|M_1\|_1 + \|M_2\|_1$ by the corollary above. □

Example 3.9.

(1) S^2 can be obtained from gluing two disks along the boundary. The boundary circle has fundamental group $\mathbb{Z}$, which is amenable. Although the boundary is not π_1-injective, the kernel for both copies are the entire $\mathbb{Z} = \pi_1(S^1)$, so the gluing is compatible. It follows that $\|S^2\|_1 = 2\|D^2, S^1\|_1$. Actually we know that both quantities are zero.

(2) A genus g surface S can be obtained by gluing two surfaces S_1 and S_2 both with a single circle boundary, where S_i has genus g_i. When $g_1, g_2 > 0$, the boundary is π_1-injective, so $\|S\|_1 = \|S_1\|_1 + \|S_2\|_1$, which is compatible with the Euler characteristic computation. When $g_2 = 0$ and $g_1 = g$, the gluing is not compatible, and we get a strict inequality $\|S\|_1 < \|S_1\|_1$ as $-\chi(S_1) > -\chi(S)$.

Here is another example, which shows that the equality in Theorem 3.7 fails without compatibility.

Example 3.10. *Let K be a knot in S^3. Let $(M, \partial M)$ be the knot complement, i.e. S^3 with an open tubular neighborhood of K removed. A Dehn filling of $(M, \partial M)$ is a closed manifold N obtained by gluing $(M, \partial M)$ with a solid torus $(S^1 \times D^2, T^2)$. The topology of N depends on the identification between ∂M and T^2, specifically which element of $\pi_1(\partial M)$ is identified with ∂D^2.*

This gluing is not compatible when K is nontrivial, as $\pi_1(\partial M)$ injects $\pi_1(M)$ while the kernel of $\pi_1(T^2) \to \pi_1(S^1 \times D^2)$ is $\mathbb{Z}$. As $\mathbb{Z}^2 = \pi^1(T^2)$ is amenable, we only get an inequality from Theorem 3.7:

$$\|N\|_1 \leq \|M, \partial M\|_1 + \|S^1 \times D^2, T^2\|_1 = \|M, \partial M\|_1.$$

Here $\|S^1 \times D^2, T^2\|_1 = 0$ can be seen by doubling (which is a compatible gluing) and $\|S^1 \times S^2\|_1 = 0$ (for having an S^1 factor). Alternatively, there is a self map of degree greater than one.

Suppose K is a hyperbolic knot (i.e. the interior of M has a hyperbolic structure of finite volume), then a theorem of Thurston shows that for all but finitely many Dehn fillings, N has a hyperbolic structure. Moreover, the hyperbolic volumes of different Dehn fillings N converge nontrivially to the volume of M. By the Proportionality Theorem 1.38, we have $\|N\|_1 = \mathrm{vol}(N)/v_3$, which is not a constant as we vary the Dehn fillings. So we must have a strict inequality $\|N\|_1 < \|M, \partial M\|_1$ (for most Dehn fillings N). In the limit we observe $\mathrm{vol}(M)/v_3 \leq \|M, \partial M\|_1$. Actually this is an equality by smearing in the thick part of M to produce an efficient representative of the relative fundamental class; see [Thu, Lemma 6.5.4].

A compatible gluing with non-amenable fundamental groups on the boundaries also violates the theorem.

Example 3.11. *Let S be a surface of genus at least 2. Let M_φ be the mapping torus associated to a mapping class φ. When φ is pseudo-Anosov, M_φ is hyperbolic, so $\|M_\varphi\|_1 = \mathrm{vol}(M_\varphi)/v_3$. It is known that $\mathrm{vol}(M_\varphi)$ is comparable to the translation length of φ with respect to the Weil-Peterson metric on the Teichmüller space. In particular $\mathrm{vol}(M_{\varphi^n}) \to \infty$ as $n \to \infty$.*

On the other hand, M_φ is obtained from $S \times [0,1]$ by a compatible gluing. However, $\|M_{\varphi^n}\|_1 > \|S \times [0,1]\|_1$ for large n, violating the inequality in Theorem 3.7, and the reason is that $\pi_1(S)$ is non-amenable as it has free subgroups.

Example 3.12. *For a closed 3-manifold M, its simplicial volume can be computed from its prime factors in the prime decomposition. When M is prime, it can be decomposed along π_1-injective tori so that each component admits one of the eight geometries, by the geometrization theorem. Five out of the eight geometry only allows amenable fundamental groups. Among the remaining three, manifolds with $\mathbb{H}^2 \times \mathbb{R}$ or $\widetilde{\mathrm{PSL}}_2$ are Seifert fibered, and the S^1 fiber (which has amenable fundamental group) forces the simplicial volume to vanish. So only the hyperbolic geometry contributes nontrivially to the simplicial volume of M.*

Proposition 3.13. *If a closed 3-manifold M admits a complete metric with negative sectional curvature, then it admits a complete hyperbolic metric.*

Proof. Negative curvature implies that M is aspherical and the fundamental group $\pi_1(M)$ is δ-hyperbolic. In particular, $\pi_1(M)$ cannot have any $\mathbb{Z}^2$ subgroup. It follows that M is prime and has trivial JSJ decomposition. So itself is geometric. Negative curvature implies that $\|M\|_1 > 0$, hence M can only admit the hyperbolic geometry. □

From now on, in the setup of Theorem 3.7, we assume all components of ∂M_i to have amenable π_1. We will first prove the inequality in Theorem 3.7. Naively, we would like to piece together almost optimal cycles representing fundamental classes $[M_i, \partial M_i]$. The boundary of the cycles restricted on S_j^+ and S_j^- represent the fundamental classes respectively, which cancel out after gluing as homology classes, but not necessarily as cycles. The hope is that this can fixed as $S_j^\pm$ has amenable fundamental group and thus zero simplicial volume. To carry this out, we will use the duality principle in the relative context, between $H_n(X,Y)$ and $H_b^n(X,Y)$, for a subspace $Y \subset X$. A key ingredient is the following theorem:

Theorem 3.14. *Let (X,Y) be a pair of countable CW complexes. Suppose each component of Y has amenable fundamental group. Then the natural map $H_b^n(X,Y) \to H_b^n(X)$ is an isometric isomorphism for all $n \geq 2$.*

Note that cochains in $C_b^n(X,Y)$ are cochains in X that vanish on Y. So intuitively, for any $\alpha \in H_b^n(X)$ one can find a cocycle that vanishes on Y using amenability of $\pi_1(Y)$ at an arbitrarily small cost to the norm. See [Fri17, Theorem 5.14] for details.

This allows us to treat any $\alpha \in H_b^n(M, \partial M)$ as a class in $H^n(M, \mathcal{S} \cup \partial M)$, where $\mathcal{S} = \cup \overline{S}_j$ and $\overline{S}_j$ is the image of $S_j^\pm$ in M. Then we can pull back α to each M_i to obtain $\alpha_i \in H^n(M_i, \partial M_i)$.

Lemma 3.15. *For any $\alpha \in H_b^n(M, \partial M)$, with the notation above, we have*

$$\langle \alpha, [M, \partial M] \rangle = \sum_{i \in I} \langle \alpha_i, [M_i, \partial M_i] \rangle.$$

Proof. Let c_i be a relative cycle in $C_n(M_i, \partial M_i)$ representing $[M_i, \partial M_i]$. Identify it with its image in M. Then ∂c_i on each component of ∂M_i represents the fundamental class. Then for the chain $c = \sum c_i \in C_n(M)$, the part of ∂c on each $\overline{S}_j$ is the difference of two fundamental cycles since

$f_j : S_j^+ \to S_j^-$ is orientation reversing, hence it can be written as $\partial c'_j$ for some chain $c'_j \in C_n(\overline{S}_j)$. It follows that $c - c' \in C_n(M, \partial M)$ represents $[M, \partial M]$, where $c' = \sum c'_j$.

Representing α as a cocycle vanishing on all $\overline{S}_j$'s, we have $\langle \alpha, c' \rangle = 0$, so

$$\langle \alpha, [M, \partial M] \rangle = \langle \alpha, c - c' \rangle = \langle \alpha, c \rangle = \sum_{i \in I} \langle \alpha_i, c_i \rangle = \sum_{i \in I} \langle \alpha_i, [M_i, \partial M_i] \rangle.$$

□

Proof of the inequality in Theorem 3.7. By the duality principle, we have

$$\|M, \partial M\|_1 = \sup_{\alpha \in H_b^n(M, \partial M), \|\alpha\|_\infty \leq 1} \langle \alpha, [M, \partial M] \rangle.$$

Given any such α, in the notation above, each α_i is obtained by pulling back α, so $\|\alpha_i\|_\infty \leq \|\alpha\|_\infty \leq 1$, so $\|M_i, \partial M_i\|_1 \geq \langle \alpha_i, [M_i, \partial M_i] \rangle$ by the duality principle on M_i. Hence the lemma above, we have

$$\sum_{i \in I} \|M_i, \partial M_i\|_1 \geq \sum_{i \in I} \langle \alpha_i, [M_i, \partial M_i] \rangle = \langle \alpha, [M, \partial M] \rangle.$$

As this holds for all $\alpha \in H_b^n(M, \partial M)$ with $\|\alpha\|_\infty \leq 1$, we deduce that

$$\|M, \partial M\|_1 \leq \sum_{i \in I} \|M_i, \partial M_i\|_1$$

as desired. □

The equality in Theorem 3.7 is obtained by a partial converse to Theorem 3.14 for compatible gluing:

Theorem 3.16. *Suppose M is obtained by compatible gluing as in Theorem 3.7. Given $\varphi_i \in H_b^n(M_i, \partial M_i)$ for all $i \in I$, for any $\epsilon > 0$, there is $\alpha \in H_b^n(M, \partial M)$ such that $\alpha_i = \varphi_i$ for all i and*

$$\|\alpha\|_\infty \leq \max_{i \in I} \{\|\varphi_i\|_\infty\} + \epsilon.$$

Proof. We give a sketch of the proof; Consult [Fri17, Section 9.2] for more details. Let $G = \pi_1(M)$. Then cochains in $H_b^n(M)$ correspond to G-invariant cochains on the universal cover $\widetilde{M}$. The manifold M has the structure of a graph of spaces, where vertices correspond to interiors of M_i's (called vertex spaces) and each edge corresponds to a gluing along some $S_j^\pm$. Denote by Γ the corresponding finite graph.

The universal cover $\widetilde{M}$ has an induced structure of a tree of spaces, where each vertex represents a component of $p^{-1}(M_v)$ for a vertex space M_v, where $p : \widetilde{M} \to M$ is the covering map. This gives a contraction from $\widetilde{M}$ to a tree T, which admits a G action with quotient being the graph Γ. Compatible gluing ensures that each inclusion of M_v in M is π_1-injective by van Kampen. This implies that the restriction of p to each vertex space $\widetilde{M}_v$ in $\widetilde{M}$ is a universal covering map $\widetilde{M}_v \to M_{\bar{v}}$, where $\bar{v}$ is the image of v under the quotient map $T \to \Gamma$.

Each φ_i can be represented by a cocycle f_i on $\widetilde{M}_i$ that is $\pi_1(M_i)$-invariant such that $|f_i|_\infty < \|\varphi\|_i + \epsilon$. So we have cocycles f_v defined on each vertex space $\widetilde{M}_v$ of $\tilde{M}$. The key is to define an extension f of these cocycles to $\widetilde{M}$ so that $|f|_\infty \leq \max_v \{|f_v|_\infty\} = \max_{i \in I} \{|f_i|_\infty\}$.

For each singular simplex Δ in $\widetilde{M}$, we define $f(\Delta)$ as $f_v(\Delta_v)$ if Δ has a "barycenter" v and zero it has not barycenter, where Δ_v is supported on the closure of $\widetilde{M}_v$ as a "projection" of Δ. We say v is a barycenter of Δ if for every pair of vertices x_i, x_j of Δ, all paths connecting them in $\widetilde{M}$ intersects $\widetilde{M}_v$; One can also view this by first projecting to the tree T. The barycenter may not exist but is unique if it exists. The projection Δ_v is chosen rather arbitrarily, with the property that its vertex x'_i agrees with the corresponding vertex x_i of Δ if $x_i \in M_v$, and lies on the unique boundary of

$\widetilde{M_v}$ separating x_i from M_v otherwise. The definition of f guarantees the bound on its norm. The amenability of the boundary components is used to resolve problems on the boundaries of vertex spaces in this construction. □

Now we complete the proof of Theorem 3.7.

Proof of the equality in Theorem 3.7. For any $\epsilon > 0$, by the duality principle, there is $\varphi_i \in H_b^n(M_i, \partial M_i)$ with $\|\varphi_i\|_\infty \le 1$ such that $\langle \varphi_i, [M_i, \partial M_i] \rangle > \|M_i, \partial M_i\|_1 - \epsilon$.

Let α be as in the theorem above for the same ϵ. Then $\|\alpha\|_\infty \le 1 + \epsilon$

$$(1+\epsilon)\|M, \partial M\|_1 \ge \langle \alpha, [M, \partial M] \rangle = \sum_{i \in I} \langle \alpha_i, [M_i, \partial M_i] \rangle = \sum_{i \in I} \langle \varphi_i, [M_i, \partial M_i] \rangle = \sum_{i \in I} \|M_i, \partial M_i\|_1 - \epsilon |I|.$$

Letting $\epsilon \to 0$ gives the inequality in the other direction. □

3.3. **Volume of a product.** Geometrically, the product $M \times N$ of two closed Riemannian manifolds M^m, N^n with the product metric has $\mathrm{vol}(M \times N) = \mathrm{vol}(M) \cdot \mathrm{vol}(N)$. For the simplicial volume, in general it is only known that $\|M \times N\|_1$ is comparable to $\|M\|_1 \cdot \|N\|_1$ by some universal constants depending on the dimensions m, n.

Theorem 3.17 (Gromov). *Let M, N be closed manifolds of dimension m, n. Then*

$$\|M\|_1 \cdot \|N\|_1 \le \|M \times N\|_1 \le \binom{m+n}{m} \|M\|_1 \cdot \|N\|_1.$$

Proof. The product of simplices $\Delta^m \times \Delta^n$ can be triangulated with $\binom{m+n}{n}$ simplices of dimension $m + n$, where simplices correspond to (geodesic) paths from the lower left corner to the upper right corner of a grid made of $m \times n$ squares (so the vertices of a path corresponds to vertices of the simplex). This proves the upper bound.

For the lower bound, let α_M and α_N be cohomology classes of norm at most 1 in $H_b^m(M)$ and $H_b^n(N)$ respectively so that $\langle \alpha_M, M \rangle > \|M\|_1 - \epsilon_1 \ge -$ and $\langle \alpha_N, N \rangle > \|N\|_1 - \epsilon_2 \ge 0$. The cup product $\alpha = \alpha_M \cup \alpha_N$ satisfies $\|\alpha\|_\infty \le \|\alpha_M\|_\infty \cdot \|\alpha_N\|_\infty \le 1$, and

$$\langle \alpha, [M \times N] \rangle = \langle \alpha_M, [M] \rangle \cdot \langle \alpha_N, [N] \rangle \ge (\|M_1\| - \epsilon_1)(\|N\|_1 - \epsilon_2).$$

Hence by the duality principle, we have $\|M \times N\|_1 \ge (\|M_1\| - \epsilon_1)(\|N\|_1 - \epsilon_2)$, where $\epsilon_1, \epsilon_2 \ge 0$ can be arbitrarily small (and actually can be chosen to be zero). This proves the lower bound. □

Remark 3.18. *The same holds for two arbitrary homology classes instead of fundamental classes. There are various improvements of the bounds. For instance, if $m = n = 2$ and one of the two classes is the fundamental class of a closed surface, then there is an equality* [HL20, Theorem E]*:*

$$\|\alpha \times \beta\|_1 = \frac{3}{2} \|\alpha\|_1 \cdot \|\beta\|_1.$$

An immediate consequence generalizes our earlier observation that $\|S^n \times M\|_1 = 0$ (or with S^n replaced by any manifold having a self map of degree greater than one).

Corollary 3.19. $\|M \times N\|_1 = 0$ *if* $\|M\|_1 = 0$.

In the special case where $\pi_1(M)$ is amenable, we can view $M \times N$ as a trivial M-bundle over N where the fibers are amenable. Such genuine bundles in the world of manifolds always have zero volume.

Theorem 3.20. *Let E be a fiber bundle over base B, where the fiber F has $\pi_1(F)$ amenable. Suppose F, E, B are orientable connected closed manifolds and* $\dim F \ge 1$. *Then* $\|E\|_1 = 0$.

Proof. From the long exact sequence of homotopy groups for a fibration, we extract the following exact sequence:

$$\pi_1 F \to \pi_1 E \xrightarrow{p} \to \pi_1 B \to 1.$$

As $\pi_1(F)$ is amenable, Theorem 2.77 and Theorem 2.80 imply that $H_b^n(B) \cong H_b^n(\pi_1 B) \xrightarrow{p^*} H_b^n(\pi_1 E) \cong H_b^n(E)$ is an isometric isomorphism. Let $n = \dim E$. However,

$$\langle E, p^*(\alpha)\rangle = \langle p_*([E]), \alpha\rangle = 0$$

for all $\alpha \in H_b^n(B)$ since the pairing factors through the comparison map $H_b^n(B) \to H^n(B)$, which is a trivial space as $\dim B = \dim E - \dim F < \dim E = n$ as $\dim F \geq 1$. Hence $\|E\|_1 = 0$ by the duality principle. $\square$

This echos with our earlier comment that Seifert fibered 3-manifolds (with $\widetilde{\mathrm{PSL}}_2$ or $\mathbb{H}^2 \times \mathbb{R}$ geometry) have zero simplicial volume so that only hyperbolic pieces contribute nontrivially to the volume of a closed 3-manifold. Although Seifert fibered spaces are not strictly speaking an S^1 fiber bundle over the base surface, but it can be viewed as an S^1 bundle over an orbifold, and the exact sequence above holds with $\pi_1(B)$ interpreted as the orbifold fundamental group.

4. Groups acting on Gromov-hyperbolic spaces

The goal of this section is two fold. One is a proof of Mostow's rigidity as an application of Gromov's proportionality, and the other is Epstein–Fujiwara's construction of quasimorphisms for groups acting nicely on negatively curved spaces[3], which generalizes Brooks' counting quasimorphism and applies to many groups of interest in geometric group theory.

Both require some knowledge of coarse geometry of metric spaces, so we will give a brief introduction, focusing on basics of quasi-isometries, Gromov-hyperbolic spaces, and the Gromov boundary. A good reference for this part is [BH99].

4.1. **Quasi-isometries.**

Definition 4.1. Given constants $L > 0, C \geq 0$, for metric spaces X and Y, a map $f : X \to Y$ is an (L, C)-*quasi-isometric embedding* (or (L, C)-QI embedding for short) if f is L-bi-Lipschitz up to a bounded error C, i.e.

$$\frac{1}{L} d_X(x, y) - C \leq d_Y(f(x), f(y)) \leq L d_X(x, y) + C, \forall x, y \in X.$$

We simply say f is a *quasi-isometric embedding* (QI-embedding) if f is (L, C)-QI embedding for some L and C. If in addition, there is a QI-embedding $g : Y \to X$ such that $f \circ g$ and $g \circ f$ are in bounded distance to id_X and id_Y respectively, then we say f is a *quasi-isometry* with *quasi-inverse* g.

A QI-embedding f is QI if and only if it is *quasi-onto* in the sense that Y is contained in the r-neighborhood of the image of f for some $r > 0$. A quasi-inverse can be chosen by taking each point to a point in the image of f within distance r.

Example 4.2. *Given a group G with a finite generating set S, the* word length *with respect to S is*

$$|g|_S := \inf\{n \mid g = s_1 \cdots s_n, s_i \in S^{\pm}\}.$$

Then we get an associated word-metric $d(g, h) := |g^{-1}h|$. *This is indeed a metric on G as the symmetry follows from the fact that $|g|_S = |g^{-1}|_S$ and the triangle inequality follows from $|gh|_S \leq |g|_S + |h|_S$. The word-metric is left-invariant in the sense that $d(kg, kh) = d(g, h)$ for all $k, g, h, \in G$.*

[3]We skipped this part for time considerations

Let S' be another finite generating set of G. Distinguish the two word-metrics by denoting them as d and d' respectively. The identity map $id : (G, d) \to (G, d')$ is a quasi-isometry. Indeed, $|g|_{S'} \le L_1|g|_S$ for $L_1 = \max_{s\in S} |s|_{S'}$, so

$$d'(g, h) = |g^{-1}h|_{S'} \le L_1|g^{-1}h|_S \le L_1 d(g, h).$$

Similarly $|g|_S \le L_2|g|_{S'}$ for $L_2 = \max_{s'\in S'} |s'|_S$, and

$$d(g, h) =\le L_2 d'(g, h).$$

Hence id is an $(L, 0)$-QI embedding for $L = \max\{L_1, L_2\}$. Any bijective QI embedding is a quasi-isometry.

Therefore, a finitely generated group has a well-defined coarse geometry.

Example 4.3. *$\mathbb{Z}$ is quasi-isometric to $\mathbb{R}$ equipped with the Euclidean metric, via the natural inclusion. A quasi-inverse is a contraction g taking each interval $[n - 1/2, n + 1/2)$ to n for all $n \in \mathbb{Z}$.*

Example 4.4. *Given a closed Riemannian manifold M, let $\widetilde{M}$ be its universal cover with the pullback metric. Then $G = \pi_1(M)$ acts on $\widetilde{M}$ by isometries. For any chosen base point $p \in \widetilde{M}$, the orbit map $f : G \to \widetilde{M}$ given by $f(g) = gp$ is a quasi-isometry, shown by the Švarc–Milnor lemma below. Note that the fundamental group is purely topological but the Riemannian metric on M is an additional structure. The coarse geometry of the fundamental group determines what kind of metric a topological manifold M can carry. For instance, it is known that if $G = \pi_1(M)$ is virtually nilpotent, then M cannot have a metric with negative sectional curvature by considering the growth of the size of balls of increasing radius r in both G and $\widetilde{M}$.*

Lemma 4.5 (Švarc–Milnor Lemma)**.** *If a group G acts properly discontinuously on a proper geodesic metric space X by isometries with compact quotient, then G is finitely generated and any orbit map is a quasi-isometry.*

Here a metric space X is *proper* if every bounded closed subset (e.g. a closed ball) is compact. It is *geodesic* if any two points $x, y \in X$ is connected by a geodesic γ so that the length of γ is $d(x, y)$. The length of a curve $\gamma : [0, t] \to X$ is the supremum of $\sum_{i=0}^{n} d(\gamma(t_{i-1}), \gamma(t_i))$ over all increasing sequences $\{t_i\}_{i=0}^{n}$ with $t_0 = 0$ and $t_n = t$ for all $n \ge 1$. It is a (unit speed global) geodesic if $d(\gamma(s), \gamma(s')) = |s - s'|$ for all $s, s' \in [0, t]$.

The finite generation part of the lemma can be deduced from the following more general fact.

Lemma 4.6. *Suppose G acts on a space X by homeomorphisms such that there is an open subset $U \subset X$ so that $GU = X$. Let $S = \{g \mid gU \cap U \ne \emptyset\}$. If X is connected, then S generates G.*

Proof. Let H be the subgroup generated by S. Let $V = HU$ and $V' = (G \setminus H)U$, both of which are open subsets of X and $V \cup V' = GU = X$. Our goal is to show $G = H$, i.e. $V' = \emptyset$. Suppose not, since X is connected, we must have $V \cap V' \ne \emptyset$. That is, there is $h \in H$ and $g \in G \setminus H$ such that $hU \cap gU \ne \emptyset$. Then $U \cap h^{-1}gU \ne \emptyset$, so by definition $h^{-1}g \in S \subset H$, contradicting $g \notin H$. □

Proof of Lemma 4.5. Let $p \in X$ be any base point. Since X/G is compact, there is a closed ball $B = B(p, r)$ of radius $r > 0$ such that $GB = X$. As X is proper, B is compact, and $S = \{g \mid B \cap gB \ne \emptyset\}$ is finite since the action is properly discontinuous. Hence G is finitely generated by Lemma 4.6. The next step is to compare the word metric d_S to the metric in X restricted to the orbit Gp.

On the one hand, for any $s \in S$, there is $q \in B \cap sB$, so $d(p, sp) \le d(p, q) + d(q, sp) \le 2r$. It follows that $d(p, gp) \le 2r|g|_S$ for all $g \in G$, and $d(gp, hp) \le 2r d_S(g, h)$.

Everything above works for all r large enough. So in the beginning we may choose r so that $GB(p, r/3) = X$. This helps us obtain the comparison in the other direction. For any $g \in G$, let $\gamma : [0, d] \to X$ be a geodesic with $\gamma(0) = p$ and $\gamma(d) = gp$, where $d = d(p, gp)$. Choose a partition of $[0,d]$ with break points $0=t_0<t_1<\cdots<t_n=gp$ so that $|t_i-t_{i-1}|\le r/3$ for all $1\le i\le n$, where $n\le 3d/r+1$. For each $0\le i\le n$, choose g_i so that $\gamma(t_i)\in g_iB(p,r/3)$, where $g_0=id$ and $g_n=g$. Then

$$d(g_ip,g_{i-1}p)\le d(g_ip,\gamma(t_i))+d(\gamma(t_i),\gamma(t_{i-1}))+d(\gamma(t_{i-1}),g_{i-1}p)\le r,$$

so $B(g_ip,r)\cap B(g_{i-1}p,r)\neq\emptyset$. That is, $d_S(g_i,g_{i-1})\le 1$ for all i. Hence $d_S(id,g)\le n\le 3d/r+1=\frac{3}{r}d(p,gp)+1$. As G acts by isometry, we have

$$d_S(g,h)\le\frac{3}{r}d(gp,hp)+1$$

for all $g,h\in G$.

Combining the two directions, we see that $(G,d_S)\to(X,d)$ is an $(2r,r/3)$-QI embedding. As X is contained in the r neighborhood of image, we know this is a quasi isometry. □

4.2. **Gromov-hyperbolic spaces and their boundary.** Gromov-hyperbolic spaces are coarse-geometric generalizations of hyperbolic spaces that are preserved by quasi-isometries. Throughout this section, all metric spaces are assumed to be proper and geodesic.

Definition 4.7. Given $\delta\ge 0$, a proper geodesic metric space is *δ-hyperbolic* if every geodesic triangle is δ-thin: each side is contained in the union of the δ-neighborhoods of the other two sides. We simply say X is *Gromov-hyperbolic* if it is δ-hyperbolic for some δ.

A metric space of finite diameter d is δ-hyperbolic, but it is not of interest.

Example 4.8. *Every simplicial tree (with the simplicial metric) is 0-hyperbolic since every geodesic triangle is a tripod, so that each side is contained in the union of the other two.*

Example 4.9. *The hyperbolic space $\mathbb{H}^n$ is δ-hyperbolic for a uniform δ. It comes down to check an ideal hyperbolic triangle is δ-thin, which can be done by an explicit calculation for a nicely chosen one, say in the upper-half plane model.*

Example 4.10. *If a complete Riemannian manifold X is simply connected and its sectional curvature is no more than $\kappa<0$, then X is δ-hyperbolic for some $\delta=\delta(\kappa)$.*

Definition 4.11. A group G generated by a finite set S is δ-hyperbolic if its Cayley graph with respect to S is δ-hyperbolic.

Example 4.12. *Any free group is 0-hyperbolic with respect to a free basis.*
If $n\ge 2$, then $\mathbb{Z}^n$ (with respect to the standard basis) is not δ-hyperbolic for any δ.

Changing the finite generating set S would change the metric on G by a quasi-isometry, which always take a δ-hyperbolic space to a δ'-hyperbolic space for a suitable δ'. The key is to understand the image of geodesics under a QI-embedding.

Definition 4.13. For an interval I, a curve $c: I\to X$ is a (L,C)-*quasi-geodesic* segment (resp. ray or line) if c is an (L,C)-QI embedding, where $I=[a,b]$ (resp. $I=[0,\infty)$ or $I=\mathbb{R}$) is equipped with the Euclidean metric.

We summarize two key facts about quasi-geodesics in δ-hyperbolic spaces.

Lemma 4.14 (Morse Lemma)**.** *For every pair (L,C) and δ, there is a universal constant $R=R(\delta,L,C)\ge 0$ such that every (L,C)-quasi-geodesic segment with endpoints p,q in a δ-hyperbolic space X has Hausdorff distance at most R to any geodesic segment connecting p,q.*

Lemma 4.15 (Local-to-global Principle)**.** *For every pair (L,C) and δ, there is a universal constant $K=K(\delta,L,C)$ such that if a curve $c:\mathbb{R}\to X$ in a δ-hyperbolic space whose restriction to each interval of length K is an (L,C)-quasi-geodesic, then c is a (global) $(2L,2C)$-quasi-geodesic.*

We will skip the proofs of these facts but look into a few important consequences.

Theorem 4.16. *If $X' \to X$ is an (L,C)-quasi-isometric embedding, where X is δ-hyperbolic, then X' is δ'-hyperbolic, where δ' only depends on L, C, δ. In particular, Gromov-hyperbolicity is QI invariant.*

Proof. Let R be as in the Morse lemma, and we show $\delta' = L(\delta+2R+C)$ works. Let Δ' be a geodesic triangle in X' with sides a', b', c'. Without loss of generality, we show that $c' \subset N_{\delta'}(a') \cup N_{\delta'}(b')$. Let $a = fa'$, $b = fb'$, and $c = fc'$, which are (L,C)-geodesics. By the Morse lemma, there are geodesic segments α, β, γ in X with Hausdorff distance at most R to a, b, c respectively. Then

$$c \subset N_R(\gamma) \subset N_{R+\delta}(\alpha) \cup N_{R+\delta}(\beta) \subset N_{2R+\delta}(a) \cup N_{2R+\delta}(b).$$

That is, for each $c(x)$, without loss of generality, we may assume that there is $a(y)$ so that $d(c(x), a(y)) \le 2R + \delta$. Since f is (L,C) QI embedding, we have

$$\frac{d(c'(x), a'(y))}{L} - C \le d(fc'(x), fa'(y)) = d(c(x), a(y)) \le 2R + \delta.$$

So $d(c'(x), a'(y)) \le L(2R + \delta + C) = \delta'$ as desired. □

Corollary 4.17. *Whether a group is Gromov-hyperbolic does not depend on the choice of the finite generating set.*

Example 4.18. *If $G = \pi_1(M)$ for a closed negatively curved manifold M, then G is δ-hyperbolic for some δ.*

Lemma 4.19. *If $\gamma : [0,\infty) \to X$ is an (L,C)-quasi-geodesic ray with starting point p in a δ-hyperbolic space X, then there is a geodesic ray starting at p within Hausdorff distance R from the Morse lemma.*

Proof. For each $n \in \mathbb{Z}_+$, let $\gamma_n = \gamma|_{[0,n]}$ and c_n be a geodesic connecting $\gamma(0)$ and $\gamma(n)$. By the Morse lemma, c_n is contained in the R-neighborhood of γ_n, where R does not depend on n. Fixing each $m \in \mathbb{Z}_+$, we claim that Arzelà–Ascoli applies to the sequence of geodesic segments $\{c_n|_{[0,m]}\}$ for $n \ge m$. The sequence is equi-continuous since they are (unit speed) geodesics. They lie in the ball of radius of m around p, which is compact.

Hence Arzelà–Ascoli implies that (any subsequence of) $\{c_n|_{[0,m]}\}$ has a subsequence that converges uniformly to a map $\ell_m : [0,m] \to X$. Here ℓ_m must be a geodesic as the limit of geodesics, and it lies in the R-neighborhood of γ since each c_n does. This can be done inductively, so that the subsequence at stage $m+1$ is a subsequence of the subsequence at stage m. It follows that ℓ_m is the restriction of $\ell_{m'}$ to $[0,m]$ for all $m' > m$, hence this uniquely determines a geodesic ray $\ell : [0,\infty) \to X$. It lies in the R-neighborhood of γ since each ℓ_m does.

On the other hand, any fixed $\gamma(t)$ lies in the R-neighborhood of c_n for all $n > t$. Fix m much larger than t, for any $\epsilon > 0$, there is some n such that $c_n([0,m])$ lies in the ϵ-neighborhood of $\ell([0,m])$. Then $\gamma(t)$ lies in the $(R+\epsilon)$-neighborhood of $\ell([0,m])$ if m is chosen large enough so that $d(\gamma(t), c_n(s)) > R$ for all $s > m$. Letting $\epsilon \to 0$, we see that γ lies in the R-neighborhood of ℓ.

Hence γ and ℓ has Hausdorff distance at most R. □

Corollary 4.20. *Any bi-infinite quasi-geodesic line in X has finite Hausdorff distance to a bi-infinite geodesic.*

Proof. Think of the quasi-geodesic line as two quasi-geodesic rays γ_1, γ_2 starting at the same point p. By the lemma above, there are geodesic rays ℓ_1, ℓ_2 starting at p close to γ_1 and γ_2 respectively. Take geodesics c_n connecting $\ell_1(n)$ to $\ell_2(n)$, and take their weak limits using Arzelà–Ascoli as above we find a bi-infinite geodesic close to the union $\ell_1 \cup \ell_2$ and hence close to the starting quasi-geodesic γ. □

Definition 4.21 (Gromov boundary)**.** For a δ-hyperbolic space X, define its *(Gromov) boundary (or ideal boundary)* ∂X as the equivalence class of quasi-geodesic rays, where two quasi-geodesic rays are equivalent if and only if they have finite Hausdorff distance to each other.

It does not affect the definition if we restrict our attention to quasi-geodesic rays starting at a chosen base point. This can be seen by an Arzelà–Ascoli argument similar to the proofs above. We can further restrict to genuine geodesic rays starting at a base point by Lemma 4.19.

The topology on ∂X is given by the following convergence: $\{y_n\}$ converges y in ∂X if there are geodesic rays $c_n \in y_n$ and $c \in y$ so that any subsequence of c_n has a subsequence that converges uniformly on compact sets to c.

Example 4.22. *The Gromov boundary of the hyperbolic space $\mathbb{H}^n$ is the usual ideal boundary $\partial\mathbb{H}^n \cong S^{n-1}$.*

Theorem 4.23 (Boundary map)**.** *If $f : X \to Y$ is a QI-embedding of Gromov-hyperbolic spaces, then f induces a natural continuous injective map $\partial f : \partial X \to \partial Y$. If f is QI, then ∂f is a homeomorphism.*

Proof. QI sends quasi-geodesic rays to quasi-geodesic rays, and two sets with Hausdorff distance D has distance at most $LD + C$ for their image under f if f is (L, C)-QI. This defines the natural map ∂f. If the image of two sets have Hausdorff distance D in Y then the two sets have Hausdorff distance at most $D(L + C)$, which proves injectivity. The map is natural in the sense that, if $g : Y \to Z$ is also QI, then $\partial(gf) = \partial g \circ \partial f$. Moreover, if two QIs are a bounded distance apart, then they induce the same boundary map. Hence if f has a quasi-inverse g, then ∂f and ∂g are inverses.

It remains to check the continuity of ∂f. It suffices to show that if $c_n : [0, \infty) \to X$ is a sequence of geodesic rays converging on compact sets to a geodesic c in X, then the geodesic rays c'_n corresponding to the quasi-geodesics $f \circ c_n$ constructed as in Lemma 4.19 has a subsequence converging to a geodesic ray c' uniformly on compact sets in Y, where $d_H(c', f \circ c) < \infty$ and d_H is the Hausdorff distance.

For each $m \in \mathbb{Z}_+$, as $c_n|_{[0,m]}$ converges uniformly to $c|_{[0,m]}$, so they are in a ϵ-neighborhood of $c|_{[0,m]}$ for n large, hence the images $f(c_n|_{[0,m]})$ has Hausdorff distance at most $L\epsilon + C$ to $f(c|_{[0,m]})$. Moreover, by construction, $c'_n|_{[0,m]}$ has Hausdorff distance at most R to $f(c_n|_{[0,m]})$. It follows that $c'_n|_{[0,m]}$ is $L\epsilon + C + R$ away from $f(c|_{[0,m]})$ and lies in a compact set independent of n. Hence we can apply Arzelà–Ascoli to get a subsequence of c'_n that converges to a geodesic ray c' uniformly on all compact sets, and $c'|_{[0,m]}$ is at most $L\epsilon + C + R$ away from $f(c|_{[0,m]})$ for all m. Therefore, c' represents the same class as $f(c)$ as desired. ☐

4.3. **Mostow's rigidity.** The goal of this section is to explain the Mostow rigidity and a proof using the machinery of boundary maps of quasi-isometries and the simplicial volume.

Theorem 4.24 (Mostow Rigidity)**.** *Let M, N be closed oriented hyperbolic n-manifolds with $n \geq 3$. Suppose $\varphi : M \to N$ is a homotopy equivalence, then there is an isometry $f : M \to N$ homotopic to φ.*

Note that hyperbolic manifolds are $K(\pi, 1)$ spaces, so isomorphism $\pi_1(M) \to \pi_1(N)$ is induced by some homotopy equivalence, and thus by Mostow's rigidity it is induced by an isometry.

Mostow's rigidity shows that, in dimension $n \geq 3$, a topological closed manifold has a unique hyperbolic structure if exists. This is far from being true in dimension two, as there is a big (moduli) space of hyperbolic metrics.

Corollary 4.25. *In dimension $n \geq 3$, any hyperbolic geometric quantity invariant under isometry is a topological invariant among closed hyperbolic n-manifolds. Examples include the hyperbolic volume, injectivity radius and the length spectrum (i.e. the set of lengths of closed geodesics).*

Historically, the topological invariance of the hyperbolic volume is a consequence of the Mostow rigidity (for $n \geq 3$, and follows from Gauss–Bonnet when $n = 2$), but we have shown this via the invariance of the simplicial volume and Gromov's proportionality, and we will use this to prove the Mostow rigidity.

The proof has four steps, which we will carry out in more detail.

(1) We may assume φ to be smooth up to homotopy. Fix a lift $\tilde{\varphi} : \mathbb{H}^n \cong \widetilde{M} \to \widetilde{N} \cong \mathbb{H}^n$ of φ to the universal covers. Then $\tilde{\varphi}$ is a quasi-isometry and π_1-equivariant.
(2) Using simplicial volume and Gromov's proportionality to show that the boundary map $\partial\tilde{\varphi} : \partial\mathbb{H}^n \to \partial\mathbb{H}^n$ must take the vertex set of a regular ideal n-simplex (which uniquely attains the maximal volume) to another.
(3) Use π_1-equivariance and the above property to show that there is a hyperbolic isometry $F : \mathbb{H}^n \to \mathbb{H}^n$ such that $\partial F = \partial\tilde{\varphi}$.
(4) Conclude that F is π_1-equivariant and thus induces an isometry $f : M \to N$, which is homotopic to φ as desired.

Only step (3) requires $n \geq 3$.

4.3.1. *Step (1).* Up to homotopy, we assume φ to be smooth. Fix a based point $p \in M$ and let $q = \varphi(p)$. Let $\varphi_* : \pi_1(M) \to \pi_1(N)$ be the induced map, where we abbreviate $\pi_1(M, p)$ (resp. $\pi_1(N, q)$) as $\pi_1(M)$ (resp. $\pi_1(N)$) and similarly in the sequel. φ_* is an isomorphism since φ is a homotopy equivalence.

Choose lifts $\tilde{p} \in \widetilde{M}$ and $\tilde{q} \in \widetilde{N}$, which determines the actions of $\pi_1(M, p)$ and $\pi_1(N, q)$ on $\widetilde{M}$ and $\widetilde{N}$ by deck transformations respectively. This also determines a lift $\tilde{\varphi} : \mathbb{H}^n \cong \widetilde{M} \to \widetilde{N} \cong \mathbb{H}^n$ of φ such that $\tilde{\varphi}(\tilde{p}) = \tilde{q}$.

The map $\tilde{\varphi}$ is π_1-equivariant in the sense that the following diagram commutes for all $\gamma \in \pi_1(M)$.

$$\begin{array}{ccc} \mathbb{H}^n & \xrightarrow{\tilde{\varphi}} & \mathbb{H}^n \\ \downarrow{\scriptstyle \gamma} & & \downarrow{\scriptstyle \varphi_*\gamma} \\ \mathbb{H}^n & \xrightarrow{\tilde{\varphi}} & \mathbb{H}^n \end{array}$$

The compactness of M allows us to control the distortion of distance under $\tilde{\varphi}$.

Lemma 4.26. *There is a constant $L > 0$ such that $d(\tilde{\varphi}x, \tilde{\varphi}y) \leq L\varphi d(x, y)$ for all $x, y \in \mathbb{H}^n$.*

Proof. Since φ is C^1, the operator norm $\|D_x\varphi\|$ of the tangent map $D_x\varphi : T_xM \to T_{\varphi x}N$ with respect to the Riemannian metrics depends continuously on $x \in M$. Let $L = \sup_{x\in M} \|D_x\varphi\| > 0$, which is finite since M is compact.

Since $\tilde{\varphi}$ is the lift of φ, we see that the norm of $D_x\tilde{\varphi}$ is also bounded by L for all $x \in \widetilde{M}$. Let $\sigma : [0, d] \to \widetilde{M}$ be a unit speed geodesic connecting x to y, where $d = d(x, y)$. Then

$$\text{length}(\tilde{\varphi}\sigma) = \int_0^d \|\tilde{\varphi}_* \frac{d}{dt}\| \, dt \leq \int_0^d L \, dt = Ld,$$

so $d(\tilde{\varphi}x, \tilde{\varphi}y) \leq \text{length}(\tilde{\varphi}\sigma) \leq Ld(x, y)$. □

Lemma 4.27. *$\tilde{\varphi} : \mathbb{H}^n \to \mathbb{H}^n$ is a quasi-isometry, which is π_1-equivariant.*

Proof. The equivariance is already explained. Let $\psi : N \to M$ be a smooth homotopy inverse of φ and let $\tilde{\psi}$ be its lift to $\mathbb{H}^n$. Then by the previous lemma, it suffices to bound $d(\tilde{\varphi}x, \tilde{\varphi}y)$ from below. Apply the same lemma to ψ, we know

$$Ld(\tilde{\varphi}x, \tilde{\varphi}y) \geq d(\tilde{\psi}\tilde{\varphi}x, \tilde{\psi}\tilde{\varphi}y).$$

So it suffices to show that the right-hand side is uniformly bounded away from $d(x, y)$.

Indeed, there is a compact ball B in $\mathbb{H}^n$ so that $\pi_1(M)B = \mathbb{H}^n$ since M is compact. As both $\tilde{\psi}$ and $\tilde{\varphi}$ are π_1-equivariant, we have

$$d(x, \tilde{\psi}\tilde{\varphi}x) = d(\gamma x, \gamma\tilde{\psi}\tilde{\varphi}x) = d(\gamma x, \tilde{\psi}\tilde{\varphi}\gamma x)$$

for all $\gamma \in \pi_1(M)$. By choosing $\gamma \in \pi_1(M)$ so that $\gamma x \in B$, by compactness and continuity, we know

$$d(x, \tilde{\psi}\tilde{\varphi}x) \leq C,$$

where $C = \sup_{x\in B} d(x, \tilde{\psi}\tilde{\varphi}x)$. Hence $d(\tilde{\psi}\tilde{\varphi}x, \tilde{\psi}\tilde{\varphi}y) \geq d(x, y) - 2C$, and

$$d(\tilde{\varphi}x, \tilde{\varphi}y) \geq \frac{1}{L}(d(x, y) - 2C).$$

This shows that $\tilde{\varphi}$ is a QI embedding, but clearly the above argument also shows that $\tilde{\psi}$ is a QI inverse, and hence $\tilde{\varphi}$ is a quasi-isometry. $\square$

4.3.2. *Step (2).* Since φ is a homotopy equivalence, we know $\|M\|_1 = \|N\|_1$ by Corollary 1.4. Recall the fact that the regular ideal simplex is the unique ideal simplex that achieves that supremum $v_n = \sup_\Delta \operatorname{vol}(\Delta)$ over all ideal n-simplices, and we have $\operatorname{vol}(M) = v_n\|M\|_1$ by Gromov's proportionality 1.36.

Intuitively, we can almost tile $[M]$ by $\|M\|_1$ copies of the regular ideal simplices, and as we map them over by φ and straighten, we get a tiling of $[N]$ by $\|M\|_1 = \|N\|_1$ simplices, and so for volume considerations, each simplex in the tiling of $[N]$ should also have the maximal volume, which requires the simplex to be regular ideal simplex. This is the rough reason why $\partial\tilde{\varphi}$ must take the vertex set of a regular ideal simplex to another such vertex set.

Lemma 4.28. *$\partial\tilde{\varphi}$ takes the vertex set of a regular ideal simplex to another such vertex set.*

Proof. Suppose there is a regular ideal simplex Δ_r in $\mathbb{H}^n$ with vertex set $V \subset \partial\mathbb{H}^n$ such that the ideal simplex Δ_t with vertex set $\partial\tilde{\varphi}(V)$ is not regular. Then there is $\epsilon > 0$ such that $\operatorname{vol}(\Delta_t) < v_n - \epsilon = \operatorname{vol}(\Delta_r) - 2\epsilon$.

Choose a genuine straight simplex Δ approximating Δ_r. Recall that by the smearing construction in Section 1.7, we can represent $[M]$ as a cycle $\sum \lambda_i c_i$ with each $\lambda_i > 0$, $\sum \lambda_i < \|M\|_1(1 + \frac{\epsilon}{v_n}$ and each c_i the projection of some Δ' with each vertex uniformly bounded away from Δ. Thus we may choose Δ close enough to Δ_r so that $\operatorname{vol}(\mathrm{str}\varphi c_i) < v_n - \epsilon$ for all i. Then $[N] = \sum \lambda_i \mathrm{str}(\varphi c_i)$ and

$$\operatorname{vol}(N) = \langle [N], \operatorname{vol}\rangle = \sum \lambda_i \operatorname{vol}(\mathrm{str}\varphi c_i) \leq \sum \lambda_i (v_n - \epsilon) < \|M\|_1(v_n - \epsilon)(1 + \frac{\epsilon v_n}{)} < \|M\|_1 v_n = \|N\|_1 v_n,$$

contradicting Gromov's proportionality for N. $\square$

4.3.3. *Step (3).* Note that any self-homeomorphism of $\partial\mathbb{H}^n$ acts on the set of $(n+1)$-tuples of distinct points in $\partial\mathbb{H}^n$. Consider the subset $\mathcal{V}$ consisting of those $(n+1)$-tuples that appear as the vertex set of some regular ideal simplex. Then the boundary map of any hyperbolic isometry preserves $\mathcal{V}$. When $n \geq 3$, these are the only homeomorphisms of $\partial\mathbb{H}^n$ with this property.

Proposition 4.29. *If a self-homeomorphism h of $\partial\mathbb{H}^n$ preserves $\mathcal{V}$, then there is a hyperbolic isometry $F : \mathbb{H}^n \to \mathbb{H}^n$ such that $\partial F = h$.*

Proof. First note that any two regular ideal simplices differ by a hyperbolic isometry, so $\operatorname{Isom}(\mathbb{H}^n)$ acts transitively on $\mathcal{V}$. This can be seen by noting that any regular ideal simplex with one vertex at ∞ in the upper half space model has the remaining n vertices forming (the vertex set of) a regular Euclidean $(n-1)$-simplex, and the stabilizer of ∞ is the group of Euclidean similarities.

Thus up to composing h with ∂F for a hyperbolic isometry F, we may assume that h fixes each vertex of a regular ideal simplex, and we aim to show that h is the identity. It suffices to show that the fixed point set of h contains dense subset of $\partial\mathbb{H}^n$. The basic observation is that, if h fixes $(v_0, v_1, \cdots, v_n) \in \mathcal{V}$, then it also fixes $(v_0', v_1, \cdots, v_n) \in \mathcal{V}$, where v_0' is the image of v_0 under the hyperbolic reflection across the hyperplane passing through $v_1, \cdots, v_n$. This is true because these are the vertex sets of the only two regular ideal simplices containing $v_1, \cdots, v_n$, which is the only place that we need $n \geq 3$. The density of the fixed point set of h follows from the fact hat the group generated by reflections across the faces of a regular ideal simplex has limit set equal to $\partial\mathbb{H}^n$. We give a direct proof, following [BP92, Proposition C.5.1].

We explain the argument for $n = 3$, but the same works in higher dimensions. Suppose h fixes an ideal regular simplex, which has one vertex v_0 at ∞ in the upper half space model and the other vertices v_1, v_2, v_3 form an equilateral triangle Δ on the Euclidean plane. By the observation above, h fixes all vertices in the tiling of the Euclidean plane by reflections of Δ. In particular, h fixes v_1', the reflection of v_1 across $[v_2, v_3]$. Note that the hyperbolic reflection across the hyperplane through v_1, v_2, v_3 is the inversion with respect to the circumcircle of Δ, which takes v_1' to the midpoint of $[v_2, v_3]$. It follows that all midpoints of the sides of Δ are fixed, which divide Δ into four equilateral triangles whose side length is half of the original one. It follows that h fixes the vertex set of another ideal regular simplex with one vertex being ∞ and the remaining three vertices forming one of those smaller equilateral triangles. This means that we can keep subdividing the equilateral triangles, whose tiling has denser vertex set. Taking the union of vertices in this process shows that h fixes a dense set of the Euclidean plane and thus must be the identity map. ☐

Corollary 4.30. *In the notation from the previous steps of the proof, $\partial\tilde{\varphi} = \partial F$ for some hyperbolic isometry $F : \mathbb{H}^n \to \mathbb{H}^n$. Moreover, F is π_1-equivariant.*

Proof. As we have shown in Lemma 4.28, $\partial\tilde{\varphi}$ has the property as in Proposition 4.29, hence the existence of F. Since $\partial F = \partial\tilde{\varphi}$, which is π_1-equivariant, we know F is also π_1-equivariant, since a hyperbolic isometry is determined by its boundary map. ☐

4.3.4. *Step (4).* Now we are ready to complete the proof.

Proof of Theorem 4.24. As shown above, $\partial\tilde{\varphi} = \partial F$ for some hyperbolic isometry $F : \widetilde{M} \to \widetilde{N}$. The π_1-equivariance of F implies that it induces an isometry $f : M \to N$. Moreover, there is a natural homotopy H between $\tilde{\varphi}$ and F, where for each $x \in \widetilde{M}$, $\{H(x,t) \mid t \in [0,1]\}$ is the geodesic going from $\tilde{\varphi}(x)$ to $F(x)$. As isometry preserves geodesics, the π_1-equivariance of $\tilde{\varphi}$ and F implies that the homotopy H is also π_1-equivariant, and hence it descends to a homotopy between φ and f, which completes the proof. ☐

4.4. **Remarks and consequences of Mostow's rigidity.**

Remark 4.31.

(1) *There are a few other proofs of the Mostow rigidity, with Steps (2) and (3) worked out in different ways. For instance, one can show that $\partial\tilde{\varphi}$ is quasi-conformal, where the distortion is constant due to ergodicity of the action of $\pi_1(M)$ on $\partial\mathbb{H}^n$. A more involved ergodicity (on the double boundary) argument shows that the distortion is* 1*, and thus $\partial\tilde{\varphi}$ is conformal and hence is the boundary map of some isometry.*

(2) *If we view $M = \mathbb{H}^n/\Gamma_1$ and $N = \mathbb{H}^n/\Gamma_2$ as quotients, where $\Gamma_1 \cong \pi_1(M)$ and $\Gamma_2 \cong \pi_1(N)$ are lattices of* $\mathrm{Isom}(\mathbb{H}^n)$*, then Mostow's rigidity states that any abstract isomorphism $\varphi : \Gamma_1 \to \Gamma_2$ is realized as a conjugation by some $F \in \mathrm{Isom}(\mathbb{H}^n)$.*

(3) *Prasad* [Pra73] *showed that the Mostow rigidity also holds for hyperbolic manifolds of finite volume.*

Let $M = \mathbb{H}^n/\Gamma$ be a closed hyperbolic manifold, where we view Γ as a lattice in $\mathrm{Isom}(\mathbb{H}^n)$. Let N_Γ denote the normalizer of Γ, which consists of elements $F \in \mathrm{Isom}(\mathbb{H}^n)$ such that $F\Gamma F^{-1} = \Gamma$. Let

Theorem 4.32. *With the notation as above, if $n \geq 3$, then the following groups are isomorphic and finite:*

(1) *the isometry group* $\mathrm{Isom}(M)$*,*
(2) N_Γ/Γ*,*
(3) *the outer automorphism group* $\mathrm{Out}(\Gamma)$*,*
(4) *the mapping class group* $\mathrm{MCG}(M) := \mathrm{Homeo}(M)/\mathrm{Homeo}_0(M)$*.*

Proof. We first explain the isomorphism $N_\Gamma/\Gamma \cong \mathrm{Isom}(M)$. Each isomorphism $f : M \to M$ lifts to an isometry $F : \mathbb{H}^n \to \mathbb{H}^n$ that is Γ-equivariant in the sense $F\gamma F^{-1} \mapsto \varphi\gamma$ for all $\gamma \in \Gamma$ for an automorphism φ of Γ. This shows that conjugation by F inside $\mathrm{Isom}(\mathbb{H}^n)$ preserves the subgroup Γ, and hence $F \in N_\Gamma$. Conversely, any $F \in N_\Gamma$ gives a Γ-equivariant isometry of $\mathbb{H}^n$ as above and hence induces an isometry $f : M \to M$. Moreover, f is the identity map if and only if F is a deck transformation, i.e. $F \in \Gamma$. This gives an isomorphism $N_\Gamma/\Gamma \to \mathrm{Isom}(M)$.

Now we show that $N_\Gamma/\Gamma \cong \mathrm{Out}(\Gamma)$. Clearly conjugation by any $F \in N_\Gamma$ defines an automorphism on Γ, so we have a homomorphism $h : N_\Gamma \to \mathrm{Out}(\Gamma)$ and clearly $\Gamma \subset \ker h$. The map is surjective by Mostow's rigidity, as any automorphism of Γ is realized by a homotopy equivalence, which is homotopic to an isometry corresponding to some $F \in N_\Gamma$. It remains to see that $\Gamma = \ker h$. Suppose conjugation by $F \in N_\Gamma$ agrees with conjugation by some $\gamma \in \Gamma$, then conjugation by $\gamma^{-1}F$ is the identity map, i.e. $\gamma^{-1}F$ commutes with all elements of Γ. To see this, note that any centralizer z of a loxodromic element γ fixes the endpoints of the axis of γ. As the limit set of Γ is $\partial\mathbb{H}^n$ by compactness of M, if z commutes with all conjugates of γ in Γ then its fixed point contains a dense set of $\partial\mathbb{H}^n$, which implies $z = id$.

To see the isomorphism with the mapping class group, note that there are natural homomorphisms $\mathrm{Isom}(M) \to \mathrm{Homeo}(M) \to \mathrm{MCG}(M)$ and $\mathrm{MCG}(M) \to \mathrm{Out}(\Gamma)$. The composition is an isomorphism as shown above, hence the map $i : \mathrm{Isom}(M) \to \mathrm{MCG}(M)$ is injective. It is also surjective by Mostow's rigidity and hence an isomorphism.

It remains to see that these groups are finite. The action of $\mathrm{Isom}(M)$ on M induces an action on the frame bundle FM, which is compact since M is. The action on FM has no fixed point since any isometry fixing a point on M is purely determined by its tangent map. So it suffices to see that the action on FM has a discrete orbit. If not, then there is a sequence $\{F_n\}$ of lifts of isometries $f_n \in \mathrm{Isom}(M)$ and a point $x \in \mathbb{H}^n$ so that for all $\epsilon, R > 0$, F_n is ϵ close to the identity on $B(x, R)$ for all n large. Then if $\gamma \in \Gamma$ has a fundamental domain of its axis in $B(x, R)$, then $\gamma' = F_n\gamma F_n^{-1}$ has the same translation length and almost the same axis. It follows that for some $p \in B(x, R)$, $d(\gamma'^{-1}\gamma p, p) \le 2\epsilon$. Choosing ϵ less than the injectivity radius of M forces $\gamma'^{-1}\gamma p = p$. Since the deck group $\pi_1(M)$ acts freely, this implies that $\gamma' = \gamma$, i.e. F_n lies in the centralizer of γ. By choosing R large so that this applies to finitely many γ's that generate $\pi_1(M) \cong \Gamma$, we must have F_n in the centralizer of Γ, which we have shown to be trivial. Thus $F_n = id$ for n large, which give a contradiction. Hence $\mathrm{Isom}(M)$ has discrete orbits on FM and thus it is finite. □

Note that when $n = 2$, it is still true that the mapping class group is isomorphic to the outer automorphism of π_1, and both are infinite, while the isometry group stays finite.

5. Actions on the circle and the bounded Euler class

The goal of this section is the introduce the (bounded) Euler class and explain how it characterizes group actions on the circle, generalizing the role of the rotation number in the study of circle homeomorphisms.

As before, let $T = \mathrm{Homeo}^+(S^1)$ and $\widehat{T}$ be group of lifts to $\mathbb{R}$. Then we get a *central extension*

$$1 \to \mathbb{Z} \to \widehat{T} \xrightarrow{\pi} T \to 1, \tag{5.1}$$

where $\mathbb{Z}$ is the group of integer translations on $\mathbb{R}$.

Central extensions of a group G by an abelian group A are in one-to-one correspondence to cohomology classes in $H^2(G; A)$; see [Bro82, Chapter 4] for the general discussion. The associated cohomology class is called the *Euler class*.

We explain the construction in our case. Fix a set theoretic section $s : T \to \widehat{T}$, that is, for each $g \in T$, we choose $s(g) \in \widehat{T}$ such that $\pi(s(g)) = g$. Using inhomogeneous coordinates, define a 2-cochain by setting $e(g, h) := s(gh)^{-1}s(g)s(h)$ for all $g, h \in T$. Note that $\pi e(g, h) = (gh)^{-1}gh = 1$, so $e(g, h) \in \ker \pi = \mathbb{Z}$. Hence e is a $\mathbb{Z}$-valued 2-cochain.

Lemma 5.1. *e is a cocycle.*

Proof. This is to check that $(\delta e)(g, h, k) = 0$ for all $g, h, k \in T$. In the calculation below, note that the addition in $\mathbb{Z}$ is multiplication $\widehat{T}$, and conjugation in $\widehat{T}$ acts trivially on $\mathbb{Z}$ as it is central. Then by definition,

$$
\begin{aligned}
(\delta e)(g, h, k) &= e(h, k) - e(gh, k) + e(g, hk) - e(g, h)\\
&= s(hk)^{-1}s(h)s(k) - s(ghk)^{-1}s(gh)s(k) + s(ghk)^{-1}s(g)s(hk) - s(gh)^{-1}s(g)s(h)\\
&= s(hk)^{-1}s(h)s(k) + s(k)^{-1}s(gh)^{-1}s(ghk) + s(ghk)^{-1}s(g)s(hk) + s(h)^{-1}s(g)^{-1}s(gh)\\
&= s(hk)^{-1}s(h)s(k) + s(k)^{-1}s(gh)^{-1}s(g)s(hk) + s(h)^{-1}s(g)^{-1}s(gh)\\
&= s(hk)^{-1}s(h)s(gh)^{-1}s(g)s(hk) + s(h)^{-1}s(g)^{-1}s(gh)\\
&= s(h)s(gh)^{-1}s(g) + s(g)^{-1}s(gh)s(h)^{-1}\\
&= 0
\end{aligned}
$$

□

Definition 5.2 (Euler class)**.** The Euler class $\mathrm{eu} \in H^2(T; \mathbb{Z})$ associated to the central extension (5.1) is the cohomology class represented by the cocycle e.

Exercise 5.3. eu *does not depend on the choice of the section s.*

Proposition 5.4. *The Euler class is bounded. More precisely, if we choose $s(g)$ so that $s(g)0 \in [0, 1)$ for all $g \in T$, then $e(g, h) \in \{0, 1\}$ for all g, h.*

Proof. For this choice of s, we have $0 \le s(g)0 \le 1$ and $0 \le s(h)0 \le s(h)s(g)0 < s(h)1 < 2$, and thus $-1 < s(gh)^{-1}0 \le e(g, h)0 < s(gh)^{-1}2 < 2$. As $e(g, h) \in \mathbb{Z}$, so we must have $e(g, h) \in \{0, 1\}$. □

This shows that $\mathrm{eu} \in H^2(T; \mathbb{Z})$ is the image of a class $\mathrm{eu}_b^{\mathbb{Z}} \in H_b^2(T; \mathbb{Z})$ (represented by the bounded cocycle e above) under the comparison map.

Definition 5.5 (bounded Euler class)**.** Let $\rho : G \to \mathrm{Homeo}^+(S^1)$ be a homomorphism that defines a G action on S^1. The *bounded Euler class* $\mathrm{eu}_b^{\mathbb{Z}}(\rho) \in H_b^2(G; \mathbb{Z})$ is defined as the pullback $\rho^*\mathrm{eu}_b^{\mathbb{Z}}$, where $\rho^* : H_b^2(T; \mathbb{Z}) \to H_b^2(G; \mathbb{Z})$.

The *(ordinary) Euler class* $\mathrm{eu}(\rho)$ is the image of $\mathrm{eu}_b^{\mathbb{Z}}$ in $H^2(G; \mathbb{Z})$, or equivalently, $\mathrm{eu}(\rho) = \rho^*\mathrm{eu}$.

Denote the image of $\mathrm{eu}_b^{\mathbb{Z}}(\rho)$ in $H_b^2(G; \mathbb{R})$ as $\mathrm{eu}_b^{\mathbb{R}}(\rho)$. When the action ρ or the coefficient is understood, we will omit them in the notation.

Changing the section by a *uniformly bounded* amount does not affect the bounded Euler class.
Here are some basic facts:

Lemma 5.6.

(1) $\mathrm{eu}(\rho) = 0$ *if and only if the action of G on S^1 lifts to the universal cover $\mathbb{R}$.*
(2) There is an exact sequence

$$0 \to \mathrm{Hom}(G; S^1) \to H_b^2(G; \mathbb{Z}) \to H_b^2(G; \mathbb{R})$$

Proof.

(1) If the action lifts to $\mathbb{R}$, then one can choose a section $s : T \to \widehat{T}$ so that $s(\rho(g))$ is the lifted action of $g \in G$ on $\mathbb{R}$. It follows that $e(\rho(g), \rho(h)) = 0$ for all g, h, where e is the cocycle representing eu associated to the section s. Thus $\mathrm{eu}(\rho) = \rho^*\mathrm{eu} = 0$.

Conversely, if $\mathrm{eu}(\rho) = 0$, then the associated central extension of G is the trivial extension $\mathbb{Z} \times G$, and we have a commutative diagram:

$$\begin{array}{ccc} \mathbb{Z} \times G & \longrightarrow & G \\ \downarrow{\scriptstyle \tilde{\rho}} & & \downarrow{\scriptstyle \rho} \\ \widehat{T} & \xrightarrow{\pi} & T \end{array}$$

So the restriction of $\tilde{\rho}$ to the Gfactor is the action of G on $\mathbb{R}$ lifting ρ.

(2) This is part of the long exact sequence associated to the short exact sequence of the coefficients $0 \to \mathbb{Z} \to \mathbb{R} \to S^1 \to 0$. The starting term is zero since $H^1_b(G; \mathbb{R}) = 0$.

□

Lemma 5.7. $\mathrm{eu}_b^{\mathbb{Z}}(\rho) = 0$ *if and only if the action has a global fixed point.*

Proof. If the action has a global fixed point p, then we can choose the lifts such that the preimage $\tilde{p}$ of p in $[0, 1)$ is fixed. Then it is straightforward to check that $e = 0$ for the cocycle associated to this section. Hence $\mathrm{eu}_b^{\mathbb{Z}}(\rho) = 0$.

Conversely, suppose $\mathrm{eu}_b^{\mathbb{Z}}(\rho) = 0$. That is, there is a *bounded* function $f : G \to \mathbb{Z}$ such that $e = \delta f$, where e is the cocycle representing $\mathrm{eu}_b^{\mathbb{Z}}(\rho)$ associated to the section where each lift satisfies $s(\rho(g))0 \in [0, 1)$. This means $s(\rho(gh))^{-1}s(\rho(g))s(\rho(h)) = f(g) + f(h) - f(gh)$ for all $g, h \in G$. Define $s'(\rho(g)) = s(\rho(g)) - f(g)$, then s' is a group-theoretic section, lifting the G-action to $\mathbb{R}$. Since f is bounded, it follows that $|s'(\rho(g))0|$ is uniformly bounded and hence $\widetilde{\mathrm{rot}}(s'(\rho(g))) = 0$ for all g. Thus $s'(\rho(g))$ has fixed points.

The goal is to show the existence of a common fixed point. It suffices to show that every finitely generated subgroup of G has a fixed point: The fixed point set is $\mathbb{Z}$-invariant, hence the fixed points in $[0, 1]$ for each finitely generated subgroup H is a compact set F_H, which form a collection of compact sets satisfying the finite intersection property, so their intersection is nonempty. For finitely generated subgroups, we find fixed points by induction on the number of generators. We have observed that the base case holds. The inductive step follows from the following claim.

Claim 5.8. *If H and K are subgroups of $\widehat{T}$ such that the fixed point sets* $\mathrm{Fix}(H)$*,* $\mathrm{Fix}(K)$ *are nonempty and disjoint. Then there is some g in the subgroup generated by H and K such that* $\widetilde{\mathrm{rot}}(g) > 0$.

Proof. Note that if $g(0) \geq 1$ then $g^n(0) \geq n$ and $\widetilde{\mathrm{rot}}(g) \geq 1 > 0$, so it suffices to find such an element g. Let $C = \sup g(0)$, where the supremum is taken over all g in the subgroup generated by H and K. We show $C = +\infty$. If not, by assumption, we may assume $C \notin \mathrm{Fix}(H)$, i.e. there is some $h \in H$ such that $C \notin \mathrm{Fix}(h)$. Since $\mathrm{Fix}(h)$ is closed, nonempty and $\mathbb{Z}$-invariant, for some $\epsilon > 0$ the interval $J = (C - \epsilon, C]$ lies in some complementary interval I of $\mathrm{Fix}(h)$. By definition, there is some g such that $g(0) \in J \subset I$. Up to replacing h by its inverse, the action of h on I is conjugate to a translation on $\mathbb{R}$ and every forward orbit converges to the right endpoint of I, which is strictly to the right of C. That is, $h^n g(0) > h^n(C - \epsilon) > C$ for some n, but $h^n g$ lies in the subgroup generated by H and K, contracting the definition of C. □

□

More generally, the bounded Euler class $\mathrm{eu}_b^{\mathbb{Z}}(\rho)$ determines the action ρ up to semi-conjugacy, generalizing Poincaré's Theorem 2.62.

Theorem 5.9 (Ghys)**.** *Two actions ρ_1, ρ_2 of G on S^1 have* $\mathrm{eu}_b^{\mathbb{Z}}(\rho_1) = \mathrm{eu}_b^{\mathbb{Z}}(\rho_2)$ *if and only if they are equivalent up to semi-conjugacies.*

Proof. We give a brief sketch. Suppose $\mathrm{eu}_b^{\mathbb{Z}}(\rho_1)=\mathrm{eu}_b^{\mathbb{Z}}(\rho_2)$, then $\mathrm{eu}(\rho_1)=\mathrm{eu}(\rho_2)$. which determine a central extension $1\to\mathbb{Z}\to\widehat{G}\xrightarrow{\pi}G\to 1$ with representations $\tilde{\rho}_i:\widehat{G}\to\widehat{T}$ lifting $\rho_i\pi$, $i=1,2$. Similar to the proof of Poincaré's Theorem 2.62, consider a function $h:\mathbb{R}\to\mathbb{R}$ by $\tilde{h}(p)=\sup_{g\in\widehat{G}}\tilde{\rho}_2(g)^{-1}\tilde{\rho}_1(g)p$. Similar to Lemma 2.65, the map $\tilde{h}$ is monotone, commutes with integral translations and has the property that $\tilde{h}\tilde{\rho}_1=\tilde{\rho}_2\tilde{h}$. Thus it descends to a semi-conjugation h between ρ_1 and ρ_2, except that h is continuous only after collapsing some jumps discontinuities.

Conversely, if two actions are semi-conjugate, one can directly check that the bounded Euler classes agree. □

This generalizes Poincaré's Theorem by restricting to $\mathbb{Z}$ actions.

Example 5.10. *Any homeomorphism $f\in T$, defines an action ρ of $\mathbb{Z}$ on S^1, where $\rho(n)=f^n$ for all $n\in\mathbb{Z}$. Since $\mathbb{Z}$ is amenable, we have $H_b^2(\mathbb{Z};\mathbb{R})=0$. Hence by the exact sequence in Lemma 5.6, we have an isomorphism $\mathrm{Hom}(\mathbb{Z},S^1)\cong H_b^2(\mathbb{Z};\mathbb{Z})$, so the bounded Euler class $\mathrm{eu}_b^{\mathbb{Z}}(\rho)$ corresponds to a homomorphism $\mathbb{Z}\to S^1$, which turns out to be the rotation number $n\mapsto\mathrm{rot}(f^n)=n\mathrm{rot}(f)$, which we explain below in a more general context.*

Of course we can lift f to an element $\tilde{f}\in\widehat{T}$, resulting in a lifted action of $\mathbb{Z}$ on $\mathbb{R}$. This reflects the fact that $\mathrm{eu}(\rho)=0$ as $H^2(\mathbb{Z};\mathbb{Z})=H^2(S^1;\mathbb{Z})=0$.

As $H_b^2(G;\mathbb{R})$ is usually better understood compared to $H_b^2(G;\mathbb{Z})$, we can use the exact sequence in Lemma 5.6 to obtain the following characterization of actions with $\mathrm{eu}_b^{\mathbb{R}}(\rho)=0$.

Proposition 5.11. *An action ρ of G on S^1 has $\mathrm{eu}_b^{\mathbb{R}}(\rho)=0$ if and only if the action is semi-conjugate to an action by rigid rotations on S^1. In this case, the rotation number $\mathrm{rot}_\rho:G\to S^1$ given by $g\mapsto\mathrm{rot}(\rho(g))$ is a homomorphism.*

Proof. By Lemma 5.6, the following sequence is exact:

$$0\to\mathrm{Hom}(G,S^1)\xrightarrow{\delta}H_b^2(G;\mathbb{Z})\to H_b^2(G;\mathbb{R}).$$

Hence the image $\mathrm{eu}_b^{\mathbb{R}}(\rho)$ of $\mathrm{eu}_b^{\mathbb{Z}}(\rho)$ in $H_b^2(G;\mathbb{R})$ vanishes if and only if $\mathrm{eu}_b^{\mathbb{Z}}(\rho)=\delta\varphi$ for some $\varphi\in\mathrm{Hom}(G,S^1)$. As S^1 acts on S^1 by rigid rotations, the homomorphism φ defines an action ρ' of G on S^1 by rigid rotations. By definition of the map $\delta:\mathrm{Hom}(G,S^1)\to H_b^2(G;\mathbb{Z})$, it is straightforward to check that $\delta\varphi=\mathrm{eu}_b^{\mathbb{Z}}(\rho')$. Hence $\mathrm{eu}_b^{\mathbb{Z}}(\rho')=\mathrm{eu}_b^{\mathbb{Z}}(\rho)$, i.e. the two actions are semi-conjugate by Theorem 5.9. In this case, we have $\mathrm{rot}(\rho(g))=\mathrm{rot}(\rho'(g))=\varphi(g)$, so $\mathrm{rot}_\rho=\varphi$ is a homomorphism. □

In the case of amenable groups, this implies a theorem of Hirsch–Thurston.

Corollary 5.12 (Hirsch–Thurston). *If G is amenable, then any action of G on S^1 is semi-conjugate to an action by rigid rotation. In particular, the rotation number is a homomorphism.*

This allows us to classify finite subgroups of $T=\mathrm{Homeo}^+(S^1)$.

Proposition 5.13. *Any finite subgroup of $T=\mathrm{Homeo}^+(S^1)$ is cyclic.*

Proof. Let $G\le T$ be finite. Then by amenability, the rotation number defines a homomorphism $\mathrm{rot}:G\to S^1$. This homomorphism must be faithful: If $\rho(g)=0$, then g has a fixed point on S^1, which implies that g has infinite order unless $g=id$ by Lemma 2.56. Hence G is a finite subgroup of S^1 and thus a cyclic group. □

Corollary 5.14. *A group G cannot act faithfully on S^1 if it has a finite subgroup that is not cyclic.*

Example 5.15. *The mapping class group $\mathrm{Mod}(S)$ of a closed surface S of genus at least one cannot act faithfully on S^1. This because $\mathrm{Mod}(S)$ has a subgroup isomorphic to the Klein group $K=\mathbb{Z}/2\times\mathbb{Z}/2$, which is not cyclic. To visualize this subgroup, note that the rotations by π around* *the x, y, z axis respectively in $\mathbb{R}^3$ together with the identity form a Klein group K, and S can be embedded in a symmetric way around the origin in $\mathbb{R}^3$ so that K acts on S by orientation-preserving homeomorphisms and descends isomorphically to* $\mathrm{Mod}(S)$*.*

In contract, the mapping class group $\mathrm{Mod}(S,p)$ *fixing a base point $p \in S$ acts faithfully on S^1. One way to see this is to note that* $\mathrm{Mod}(S,p)$ *acts by automorphisms on the surface group $\pi_1(S,p)$ and hence acts on its Gromov boundary, which is a circle since $\pi_1(S,p)$ is QI to $\mathbb{H}^2$.*

There is also a cocycle representing $2 \cdot \mathrm{eu}_b^{\mathbb{Z}}$ in terms of circular orders. For the given orientation on S^1, we write $Or(x,y,z) = 1$ if the triple of distinct points (x,y,z) is positively oriented and $Or(x,y,z) = -1$ otherwise; For the degenerate cases we let $Or(x,y,z) = 0$. Fix a base point $p \in S^1$, in homogeneous coordinates, we define a 2-cochain via $c(g,h,k) = Or(\rho(g)p, \rho(h)p, \rho(k)p)$ for all $g,h,k \in G$. It is easy to check that c is a (bounded) cocycle, and thus defines bounded cohomology class in $H^2(G;\mathbb{R})$ associated to the G action ρ on S^1, which turns out to be twice the bounded Euler class.

Proposition 5.16. *We have $[c] = 2\mathrm{eu}_b^{\mathbb{Z}}(\rho)$. In particular, the class does not depend on the choice of p.*

Corollary 5.17. $\|\mathrm{eu}_b^{\mathbb{R}}(\rho)\|_\infty \leq \frac{1}{2}$.

This implies the Milnor–Wood inequality.

Theorem 5.18 (Milnor–Wood)**.** *Let S be a closed oriented surface and let $\rho : \pi_1(S) \to T$ be an action of $\pi_1(S)$ on S^1. Then there is an inequality*

$$|\langle \mathrm{eu}(\rho), [S] \rangle| \leq -\chi^-(S)$$

Proof. We have

$$|\langle \mathrm{eu}(\rho), S \rangle| = |\langle \mathrm{eu}_b^{\mathbb{R}}(\rho), [S] \rangle| \leq \|\mathrm{eu}_b^{\mathbb{R}}(\rho)\|_\infty \cdot \|S\|_1 \leq \frac{1}{2} \cdot -2\chi^-(S) = -\chi^-(S),$$

where $\|\mathrm{eu}_b^{\mathbb{R}}(\rho)\|_\infty = \|\rho^* \mathrm{eu}_b^{\mathbb{R}}\|_\infty \leq \|\mathrm{eu}_b^{\mathbb{R}}\| \leq 1/2$ by functoriality. □

Remark 5.19. *The Euler number $\langle \mathrm{eu}(\rho), [S] \rangle$ varies continuously on the representation variety* $\mathrm{Hom}(\pi_1(S), \mathrm{PSL}_2(\mathbb{R}))$*, where to each representation $\rho : \pi_1(S) \to \mathrm{PSL}_2(\mathbb{R}) = \mathrm{Isom}^+(\mathbb{H}^2)$ we associate the action on the boundary $S^1 = \partial \mathbb{H}^2$. Since $\langle \mathrm{eu}(\rho), [S] \rangle$ takes integer values, it must be constant on connected components. A theorem of Goldman shows that there is exactly one component for each integer value of $\langle \mathrm{eu}(\rho), [S] \rangle$. So by the Milnor–Wood inequality, the representation variety has exactly $4g-3$ components if S has genus $g \geq 2$. Moreover, the components with maximal $|\langle \mathrm{eu}(\rho), [S] \rangle|$ correspond to the Teichmuller space, i.e. the space of discrete faithful representations.*

For a countable group G, there is also a characterization of classes in $H_b^2(G;\mathbb{Z})$ realized as the bounded Euler class for some action on S^1.

Theorem 5.20 (Ghys)**.** *If G is countable, then $\alpha \in H_b^2(G;\mathbb{Z})$ can be represented by cocycle taking values in $\{0,1\}$ if and only if $\alpha = \mathrm{eu}_b^{\mathbb{Z}}(\rho)$ for some action ρ of G on S^1.*

Proof. We give a sketch; More details can be found in [Cal07, Chapter 2]. Clearly the bounded Euler class has such a cocycle representative. Suppose α is represented by such cocycle c taking values in $\{0,1\}$. Then such a cocycle in homogeneous coordinate defines a G-invariant "circular order" by setting $Or(g,h,k) = (-1)^{c(g,h,k)}$. For countable groups, a G-invariant "circular order" can be realized as an action ρ of G on S^1, which is analogous to the better known fact that any left-order of a countable group G can be realized by a G action on $\mathbb{R}$; See [Cal07, Theorem 2.46]. □

References

[Bav91] Christophe Bavard. Longueur stable des commutateurs. *Enseign. Math. (2)*, 37(1-2):109–150, 1991.

[BG88] Jean Barge and Étienne Ghys. Surfaces et cohomologie bornée. *Invent. Math.*, 92(3):509–526, 1988.

[BH99] Martin R. Bridson and André Haefliger. *Metric spaces of non-positive curvature*, volume 319 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[BP92] Riccardo Benedetti and Carlo Petronio. *Lectures on hyperbolic geometry*. Universitext. Springer-Verlag, Berlin, 1992.

[Bro82] Kenneth S. Brown. *Cohomology of groups*, volume 87 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.

[Cal] Danny Calegari. *Kleinian groups*. Lecture notes.

[Cal07] Danny Calegari. *Foliations and the geometry of 3-manifolds*. Oxford Mathematical Monographs. Oxford University Press, Oxford, 2007.

[Cal09] Danny Calegari. *scl*, volume 20 of *MSJ Memoirs*. Mathematical Society of Japan, Tokyo, 2009.

[CW19] Chris Connell and Shi Wang. Positivity of simplicial volume for nonpositively curved manifolds with a Ricci-type curvature condition. *Groups Geom. Dyn.*, 13(3):1007–1034, 2019.

[Fri17] Roberto Frigerio. *Bounded cohomology of discrete groups*, volume 227 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2017.

[Gro82] Michael Gromov. Volume and bounded cohomology. *Inst. Hautes Études Sci. Publ. Math.*, (56):5–99 (1983), 1982.

[Han98] Søren Kold Hansen. Measure homology. *Math. Scand.*, 83(2):205–219, 1998.

[Hat02] Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.

[HL20] Nicolaus Heuer and Clara Löh. The spectrum of simplicial volume. *Inventiones mathematicae*, pages 1–46, 2020.

[HM81] Uffe Haagerup and Hans J. Munkholm. Simplices of maximal volume in hyperbolic n-space. *Acta Math.*, 147(1-2):1–11, 1981.

[Iva17] Nikolai V. Ivanov. Notes on the bounded cohomology theory, 2017.

[LÖ6] Clara Löh. Measure homology and singular homology are isometrically isomorphic. *Math. Z.*, 253(1):197–218, 2006.

[Pra73] Gopal Prasad. Strong rigidity of $\mathbf{Q}$-rank 1 lattices. *Invent. Math.*, 21:255–286, 1973.

[Rol09] Pascal Rolli. Quasi-morphisms on free groups, 2009.

[Tao16] Jing Tao. Effective quasimorphisms on free chains, 2016.

[Thu] William P. Thurston. *Geometry and topology of three-manifolds*. Princeton Lecture Notes.

[Thu97] William P. Thurston. *Three-dimensional geometry and topology. Vol. 1*, volume 35 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1997. Edited by Silvio Levy.

[Zas98] Andreas Zastrow. On the (non)-coincidence of Milnor-Thurston homology theory with singular homology theory. *Pacific J. Math.*, 186(2):369–396, 1998.

Department of Mathematics, The University of Texas at Austin, Austin, TX, USA
Email address, L. Chen: lvzhou.chen@math.utexas.edu