Groups, geometry, and rigidity

Kathryn Mann

Notes from a minicourse given at MIT, March 2017

Abstract

This mini-course is an introduction to some central themes in geometric group theory and their modern offshoots. One of the earliest and most influential results in the area (in fact a precursor to the field of geometric group theory) is Mostow's celebrated strong rigidity theorem. This course begins with an "annotated" proof of Mostow's theorem, using the framework of the proof as a means to introduce foundational ideas in large-scale geometry. The later lectures will touch on recent developments and still-open questions in rigidity theory that can be framed as natural offshoots of Mostow.

Introduction

Most people think of Gromov as the founder of geometric group theory. In a series of highly influential papers in the 80s and early 90s – perhaps most notably *Hyperbolic groups* (1987)[1] – he popularized the idea that abstract groups can be thought of as geometric objects and profitably studied with geometric techniques.

However, there were many precursors to Gromov's ideas. A particularly important one is the notion of *rigidity* of groups (in particular of lattices in semisimple groups) by Mostow, Margulis, and others, in the 60's and 70's. "Rigidity" is a broad term; it means that some structure – usually geometric, topological, or dynamical in nature – is determined by just a small part of it. A good (and early) example is Mostow's strong rigidity theorem:

Theorem 0.1 (Mostow rigidity)**.** Let M and N be closed manifolds of dimension $n \geq 3$, each equipped with a metric of constant curvature -1. If $\pi_1(M) \cong \pi_1(N)$, then M and N are isometric.[2]

In other words, the fundamental group completely determines the geometry of such a manifold: in principle, one should be able to read off all geometric invariants of M (diameter, volume, length of longest closed geodesic, etc.) from a presentation for $\pi_1(M)$.

One can also rephrase Mostow's result purely in terms of Lie groups:

Theorem 0.2 (Mostow rigidity, algebraic version)**.** Let Γ_1 and Γ_2 be discrete subgroups of $\mathrm{SO}(n,1)$, $n \geq 3$, so that $\mathrm{SO}(n,1)/\Gamma_i$ is compact. Then any isomorphism $\Gamma_1 \cong \Gamma_2$ is realized by conjugation in $\mathrm{SO}(n,1)$.

⟨SOn aside⟩ **Aside remark 0.3.** *These are related by the following* $\mathrm{SO}(n,1) \subset \mathrm{SL}(n+1,\mathbb{R})$ *is the group of determinant 1 matrices preserving the symmetric form* $x_1y_1 + ... + x_ny_n - x_{n+1}y_{n+1}$ *on* $\mathbb{R}^{n+1}$*. One model for hyperbolic space* $\mathbb{H}^n$ *is one of the sheets of the hyperboloid* $\{x \in \mathbb{R}^{n+1} : \langle x, x\rangle = -1\}$ *with the metric induced from the symmetric form. Hence,* $\mathrm{SO}(n,1) = \mathrm{isom}(\mathbb{H}^n)$.

[1] note the title! "Hyperbolic" refers to a space of constant negative curvature - in what sense can this be applied to a discrete,algebraic object?

[2] This is not the most general possible statement of this theorem

Given a hyperbolic manifold M, we identify its universal cover with $\mathbb{H}^n$, whence π_1 is a discrete group of isometries.

Both the result and the techniques in Mostow's proof were highly influential. The first phrasing hints at "hyperbolicity" as a source of rigidity, and led to the study of hyperbolic groups (a la Gromov), while the second phrasing is the precursor to the study of rigidity of lattices in Lie groups in a much broader sense and setting.

My goal in these lectures is to introduce you to some fundamental concepts of geometric group theory through the proof of Mostow's theorem. Along the way, I'll comment on modern offshoots and perspectives.

1 GGT basics: Groups as geometric objects

One of the earliest instances of viewing a group as a geometric object is the work of Dehn (circa 1910) on finitely generated groups, and specifically the *word problem*. He makes the following definition.

Definition 1.1. Let G be any group, and S a symmetric generating set. The *word norm* on G is $\|g\|_S := \min\{k : g = s_1 s_2 ... s_k,\ s_i \in S\}$, and *word metric* $d_S(f,g) := \|f^{-1}g\|_S$

One easily checks that $\|\cdot\|_S$ is indeed a norm, and d_S a metric, moreover, d_S is invariant under left-multiplication in G.

⟨bilip prop⟩ **Proposition 1.2** (bilipschitz equivalence)**.** If S_1, S_2 are two finite, symmetric generating sets for a discrete group G, or compact symmetric generating sets for a locally compact topological group G Then $\exists K > 1$ such that

$$1/K d_{S_1}(g,h) \leq d_{s_2}(g,h) \leq K d_{S_1}(g,h)$$

holds for all $g, h \in G$.

Aside remark 1.3. *This holds in an even more general context, where compact is replaced with "coarsely bounded". This perspective, and a number of applications, comes from very recent work of C. Rosendal.*

I gave the (easy) proof of Prop. 1.2 in the discrete case. Note that the locally compact case is not a totally trivial generalization. One argues as follows: We want to find K so that $S_2 \subset S_1^k$. Now $\bigcup_n (S_1)^n = G$, so the Baire cateogry theorem (we're assuming G is Hausdorff) implies that some $(S_1)^n$ has nonempty interior. Thus, $(S_1)^{2n}$ contains a neighborhood of the identity. If $s \in S_2$ has length l in S^1, then $s \in \mathrm{int}(S_1^{2n+l})$. This shows that $\bigcup_{l>0} \mathrm{int}(S_1^{2n+l})$ is an open cover of S_2, so has a finite subcover.

Note that finiteness (or more generally compactness) is a necessary condition. Taking $S = G$ gives the group diameter 1.

Exercise 1.4. The (additive) group $\mathbb{Q}$ with topology induced from $\mathbb{R}$ is not locally compact, but is compactly generated. Can you find two compact generating sets with non bilipschitz-equivalent word norms?

Exercise 1.5. What groups can be distinguished from each other by their geometry? Show that $\mathbb{Z}$ is not quasi-isometric to $\mathbb{Z}^2$. In fact, it is true (but much harder) that $\mathbb{Z}^m$ is QI to $\mathbb{Z}^n$ iff $m = n$. However, the free group on 2 generators, F_2 is quasi-isometric to F_n for any $n \geq 2$. Can you prove this?

The important takeaway from the discussion above is

Corollary 1.6. Metric properties invariant under bilipschitz maps[3] are intrinsic to a (finitely

[3] we'll see some examples shortly

or compactly generated locally compact) group G. In other words, groups have a well-defined "large-scale" geometry.

It turns out that bilipschitz equivalence is not exactly the right notion of large-scale geometry. For example, consider $\mathbb{R}^n$ (as an additive group) with the standard Euclidean unit ball as generating set S. Then $d_S(x, y)$ is always within 1 of the standard Euclidean distance between x and y, so we's like to view this discrete metric d_S as essentially the same as the usual Euclidean one. However, the two metrics are not bilipschitz equivalent since $d_S(x, y) \geq 1$ for all $x \neq y$.

Definition 1.7. Let X, d_X and Y, d_Y be metric spaces. A map $f : X \to Y$ is a (K, C)-*quasi-isometric embedding* if

$$\tfrac{1}{K} d_X(x, z) - C \leq d_Y(f(x), f(z)) \leq K d_X(x, z) + C$$

holds for all $x, z \in X$. It is a (K, C)-*quasi-isometry* if additionally for any $y \in Y$ there exists $x \in X$ with $d_Y(f(x), y) \leq C$. If such a quasi-isometry exists, we say that X is quasi-isometric to Y, or $X \overset{QI}{\sim} Y$.

Exercise 1.8. Check that "quasi-isometric to" is an equivalence relation. *Large-scale geometry* refers to geometric properties that are well defined on QI classes.

This leaves with some basic questions:

1. To what extent does the large-scale geometry of a group determine or reflect its algebraic structure?
2. If G is quasi-Isometric to a metric space X, what geometric properties of (or structures on) X translate to interesting algebraic properties of G?
3. Given G, what spaces can we build that are QI equivalent to G? More generally, we are interested in geometric spaces X on which G acts in a way that preserves some geometry.

The *Cayley graph* gives a straightforward answer to question 3.

Definition 1.9. The *Cayley graph* $\Gamma(G, S)$ of a group G with generating set S is a graph with vertex set G and edges between each g and sg for $s \in S$.

We have $G \overset{QI}{\sim} \Gamma(G, S)$; in fact, more generally:

Lemma 1.10 (Milnor-Schwarz[4] Lemma)**.** Let X be a proper (closed balls are compact) geodesic metric space. Let Γ act cocompactly and properly discontinuously on X.[5] Then Γ is finitely generated and, for any $x_0 \in X$ the map $\Gamma \to X$, given by $\gamma \mapsto \gamma x_0$ is a quasi-isometry.

Exercise 1.11. (Easy corollary of Milnor-Schwarz) If $\Lambda \subseteq \Gamma$ is a finite index subgroup of a finitely generated group then Λ is quasi-isometric to Γ.

Another important corollary is the following – it is our first hint at where Mostow rigidity starts, with a relationship between the geometry of $\pi_1(M)$ and the geometry of the universal cover.

Corollary 1.12. Let M be a compact Riemannian manifold and let $\tilde{M}$ be the universal cover with metric lifted from M. Then $\pi_1(M) \overset{QI}{\sim} \tilde{M}$

[4] also spelled Svarc

[5] Properly discontinuously means that for all compact $K \subseteq X$, the set$\{\gamma \in \Gamma | \gamma K \cap K \neq \emptyset\}$ is finite

Milnor's original motivation for the Lemma above was to study the growth[6] of fundamental groups by comparing this to growth of (volume) of balls in universal covers. In particular, this lemma lets him do some standard Riemannian geometry and show that the fundamental group of a compact Riemannian manifold with negative sectional curvature has *exponential growth*.

Here are some properties invariant under quasi-isometry:

- Finite presentability
- Growth rate
- Number of ends*
- Hyperbolicity – our next topic.

* The number of *ends* of a topological space X is the number of components of $X \setminus K$ for all large enough compact K (if finite, this number eventually stabilizes). The number of ends of a group is the number of ends of its Cayley graph (exercise: why is this well defined?) It is a remarkable theorem of Stallings that every group has either 1, 2, or infinitely many ends.

⟨qi rig rem⟩ **Remark 1.13.** The best one can hope for is that the quasi-isometry type of a group allows one to recover, not just some of its algebraic features, but (essentially) the group itself. Since passing to finite index subgroups and quotients by finite groups do not affect the QI type of a group, the best determination one can hope for is the following:

Say G_1 and G_2 are "virtually isomorphic" if there are finite index subgroups $H_i \subset G_i$ and finite normal subgroups $F_i \subset H_i$ so that $H_1/F_1 \cong H_2/F_2$.

Remarkably, "QI implies virtually isomorphic" is true within several classes of groups, most classes being related to lattices (discrete cocompact or co-finite volume) subgroups of Lie groups.

2 Hyperbolicity

Recall that Mostow's theorem wants to turn an isomorphism $\phi : \pi_1(M) \to \pi_1(N)$ into an isometry. The outline of the proof is:

1. Use ϕ to build a map $f : \mathbb{H}^n = \tilde{M} \to \mathbb{H}^n = \tilde{N}$, equivariant with respect to π_1, and show that f is a quasi-isometry.

2. Use "hyperbolicity" to show that f extends to a homeomorphism of a suitable compactification $\overline{\mathbb{H}^n} \to \overline{\mathbb{H}^n}$ that is topologically a closed ball.

3. Study the properties of the restriction of f to the boundary $\partial\overline{\mathbb{H}^n}$. In particular, show that it is *quasi-conformal*

4. Using some analysis, show that the boundary map is in fact *conformal*, this will imply that f was homotopic to an isometry.

There are many alternative proofs of Mostow, but they all share steps 1 and 2 (and most of 3). In order to do steps 1 and 2, we some basic hyperbolic geometry. A good reference is Thurston's book [11].

[6]meaning number of elements of G that can be written as words of length $< r$ in a fixed generating set S, as a function of r; one looks as the equivalence class of this function under a suitable notion of equivalence

Hyperbolic geometry basics

Hyperbolic space $\mathbb{H}^n$ is the unique simply connected homogeneous space with constant negative (sectional) curvature -1. Here are two models.

Definition 2.1 (Poincaré Disc Model of $\mathbb{H}^n$)**.** Hyperbolic space is the open unit radius disc $\mathbb{D}^n \subset \mathbb{R}^n$ with the metric

$$ds^2(x) = \frac{4}{(1 - \|x\|^2)^2} dx^2$$

where dx is the usual Euclidean metric.

Geodesics are circular arcs that meet the boundary at right angles. Note that the metric is Euclidean at the center, but rescaled drastically towards the outside.

The ball model comes from a stereographic projection of the hyperboloid (from Aside 0.3) onto the hyperplane $x_0 = 0$. For another model, we can apply an inversion in a circle (with centre a boundary point of the unit ball above, and radius 2) to take the ball conformally to the upper half of $\mathbb{R}^n$.

Definition 2.2 (Upper half space model of $\mathbb{H}^n$)**.** Hyperbolic space is the set $\{(x_1, ...x_n) : x_n > 0\} \subset \mathbb{R}^n$ with the metric

$$ds^2(x) = \tfrac{1}{x_n^2} dx^2$$

where dx is the usual Euclidean metric.

Geodeiscs here are semi-circles meeting the boundary and vertical straight lines. Note what geodesic triangles look like.

The following generalization of hyperbolicity to arbitrary (non-Riemannian) metric spaces, including cayley graphs of groups, is due to Gromov.

⟨delta def⟩ **Definition 2.3.** A metric space X is δ-hyperbolic if, for any geodesic triangle, each side of the triangle lies in the union of the δ-neighborhoods of the other two sides.

Exercise 2.4. (An exercise strictly in hyperbolic trigonometry). Check that this indeed holds for geodesic triangles in $\mathbb{H}^2$, and therefore in $\mathbb{H}^n$.

One shows (we will later) that δ-hyperbolicity is a QI invariant (although the constant δ may change), hence can declare a group to be hyperbolic if its Cayley graph is a δ-hyperbolic space for some δ.

Step 1 of Mostow:

Since M and N are $K(\pi, 1)$ spaces, there is a homotopy equivalence $M \underset{g}{\overset{f}{\leftrightarrows}} N$ so that f induces ϕ on π_1. Since every map is homotopic to a C^1 one, we can assume f has continuous derivatives. Since M is compact, f is Lipschitz.

Lift f to a map $\tilde{f} : \tilde{M} \to \tilde{N}$. This will also be Lipschitz with the same Lipschitz constant as f. We can apply the same argument to g and lift $\tilde{g}$.

Now $g \circ f$ is homotopic to identity on M by some homotopy moving points distance $< C$. Provided we have chosen appropriate lifts of f and g, we may take a lift of this homotopy and get a homotopy from $\tilde{g} \circ \tilde{f}$ to identity, moving all points distance $< C$. Let K be the larger of the Lipschitz constants for f and g. Now for $x, y \in \tilde{M}$ we have

$$Kd(x, y) \geq d(\tilde{f}(x), \tilde{f}(y)) \geq 1/Kd(\tilde{g}\tilde{f}(x), \tilde{g}\tilde{f}(y)) \geq 1/K(d(x, y) - 2C)$$

where the first inequalities are Lipschitz, and the last from the homotopy. This shows that $\tilde{f}$ is a $K, 2C/K$ QI.

Step 2 of Mostow: Boundary maps

The Poincaré disc model suggests that we should compactify $\mathbb{H}^n$ with the unit sphere. We do this by declaring boundary points to be equivalence classes of geodesic rays.

Definition 2.5. Let X, d_X be a geodesic metric space and $\gamma_1, \gamma_2 : [0, \infty) \to X$ geodesic rays. We say $\gamma_1 \sim \gamma_2$ if there exists K so that $d_X(\gamma_1(t), \gamma_2(t)) < K$ for all t. (Equivalently, for t sufficiently large).

Exercise 2.6. Describe equivalence classes of geodesic rays in the Euclidean space $\mathbb{R}^n$.

In $\mathbb{H}^n$, one checks that two geodesic rays are equivalent iff their images limit onto the same boundary point in the disc (or, equivalently upper half plane) model. Moreover, rays that are not equivalent (bounded distance apart) diverge exponentially.

Aside remark 2.7. *More generally, in any δ-hyperbolic space, if two geodesic rays eventually separate by more than 2δ and are in the same end, then they diverge at an exponential rate. One can even use this to* characterize *hyperbolicity.*

Definition 2.8. The *boundary* of a δ-hyperbolic space is the set of equivalence classes of geodesic rays.

There is a natural way to topologize this space. In our case (for $\mathbb{H}^n$) the situation is simple: we may take representatives of geodesic rays based at the origin in the disc model, and say two endpoints are close if the rays have small angle at 0. Generally, one uses what is called the *Gromov product* to measure how close two rays are to each other.

The next step of the proof of Mostow is to show that f induces a map on the boundary. However, f does not send geodesics to geodesics. However, it does something almost as good.

Definition 2.9. Let X be a metric space. A *quasi-geodesic* in X is a QI embedded copy of the real line. A (K, C) quasi-geodesic segment is a (K, C) QI embedded copy of a segment of $\mathbb{R}$.

In hyperbolic spaces, quasi-geodeiscs are uniformly close to geodesics. Precisely, we have:

Lemma 2.10 (Mostow–Morse Lemma)**.** For all $\delta \geq 0, K \geq 1, C \geq 0$, there exists $R = R(\delta, K, C)$ with the following property: If X is a δ-hyperbolic metric space, $\gamma : [a, b] \to X$ a (K, C)-quasi-geodesic, and $[\gamma(a), \gamma(b)]$ is any geodesic from $\gamma(a)$ to $\gamma(b)$, then $[\gamma(a), \gamma(b)]$ is Hausdorff distance at most R from γ.

Note that this is very far from true in Euclidean space!

Exercise 2.11. Find a QI from $\mathbb{R}^2$ to $\mathbb{R}^2$ that takes the positive real axis to a logarithmic spiral.

We gave the proof of Mostow–Morse in the case $X = \mathbb{H}^n$. A proof for general δ-hyperbolic metric spaces can be found in Theorem 9.38 of [6] . (Their proof uses convergence of rescaled metric spaces – if you are familiar with *asymptotic cones*, this is that kind of machinery.)

A straightforward, but important consequence of Mostow–Morse is the following.

Corollary 2.12. Hyperbolicity in the sense of Definition 2.3 is a QI invariant: if X is δ-hyperbolic and Y quasi-isometric to X, then Y is δ'-hyperbolic for some δ'.

Corollary 2.13 (Local to global principle)**.** If γ is locally a quasi-geodesic in a δ-hyperbolic space, then γ is globally a quasi-geodesic. Precisely, there exists M (depending on K, C, δ) such that, if every length M segment of γ is a (K, C) quasi-geodesic, then γ is a (K', C') quasi-geodesic.

For our immediate purposes, the most important corollary is the following.

⟨boundary cor⟩ **Corollary 2.14** (Boundary map exists)**.** Let $\tilde{f} : X \to Y$ be a quasi-isometry between δ-hyperbolic spaces (for example $X = Y = \mathbb{H}^n$). For each geodesic ray $\gamma : [0, \infty) \to X$, there is a unique equivalence class of hyperbolic ray $[\alpha]$ in Y that is a bounded distance from $f \circ \gamma$. Moreover, $[\alpha]$ depends only on the equivalence class of γ in X.

This gives a well defined map $F : \partial X \to \partial Y$.

A short digression on hyperbolic groups

Hyperbolic groups are abundant, in fact, there is a good sense in which the "random" finitely presented group is hyperbolic. If one chooses relations "at random", with probability one, one obtains either the trivial group, $\mathbb{Z}/2\mathbb{Z}$, or a hyperbolic group. This is theorem of Gromov.

Hyperbolic groups also have many special properties. For instance, the word problem is solvable (following ideas of Dehn) and they all admit finite presentations. A proof of this, and a great introduction to hyperbolic groups in general can be found in the lecture notes [8] of Gersten.

Here's an advertisement for the richness of the boundary of hyperbolic groups, taken from these notes:

The boundary is one of the reasons that topologists are so interested in hyperbolic groups. N. Benakli showed in her thesis that the Menger and Sierpinski curves occur naturally as boundaries of hyperbolic groups. The characterization of hyperbolic groups which have the circle as boundary, achieved independently by Gabai, Casson-Jungreis, and Tukia, led to the solution of a classical conjecture of Seifert, that closed irreducible 3-manifolds containing normal infinite cyclic subgroups in their fundamental groups are Seifert fibred. One of the outstanding problems of 3-dimensional topology is the conjecture that a closed irreducible 3-manifold with an infinite hyperbolic fundamental group admits a Riemannian metric of constant negative curvature. As a result of the work of Bestvina and Mess, it is known that the boundary of such a group is homeomorphic to the 2-sphere.

3 Boundary maps and quasi-conformal maps

The following are easy consequences of the definition of the boundary map from Corollary 2.14.

Proposition 3.1. Let $f, g : \mathbb{H}^n \to \mathbb{H}^n$ be quasi-isometries, with boundary maps F and $G : S^{n-1} \to S^{n-1}$. Then

1. F and G are injective
2. The boundary map of $f \circ g$ is $F \circ G$.
3. If there exists D such that $d(f(x), g(x)) < D$ holds for all x, then $G = F$.

Using the Mostow–Morse lemma, one also can show the following regularity properties.

⟨boundary prop⟩ **Proposition 3.2.** Let $f : \mathbb{H}^n \to \mathbb{H}^n$ be a (K, C) quasi-isometry and let $F : S^{n-1} \to S^{n-1}$ be the boundary map of f. Then

1. F is continuous (and thus a homeomorphism)
2. F is quasi-conformal, with constant depending only on (K, C).

Recall that a *quasi-conformal* homeomorphism of a domain U in $\mathbb{R}^n$ is one such that

$$\limsup_{r \to 0} \frac{\sup\{\|f(x) - f(y)\| : \|x - y\| \leq r\}}{\inf\{\|f(x) - f(y)\| : \|x - y\| \geq r\}} \leq k$$

or, equivalently

$$\limsup_{x\in U}\limsup_{r\to 0}\left(\sup\left\{\frac{\|f(x)-f(y)\|}{\|f(x)-f(z)\|} : \|x-y\| = \|z-y\| = r\right\}\right) \leq k$$

for some $k \geq 1$. We'll work with the second definition.

The key tool to prove both quasi-conformality and continuity is the following "bounded projection" lemma.

Lemma 3.3. Let $f : \mathbb{H}^n \to \mathbb{H}^n$ be a quasi-isometry, with boundary map F. There exists $R > 0$ such that, for any hyperplane P in $\mathbb{H}^n$ and geodesic γ orthogonal to P, the diameter of the orthogonal projection of $f(P)$ onto a geodesic ray in the class of $F(\gamma)$ is bounded by R.

Here is a surprising immediate corollary of continuity of boundary maps (perhaps the statement is not surprising, but how else would one prove it?)

Corollary 3.4. $\mathbb{H}^n$ and $\mathbb{H}^m$ are QI as metric spaces iff $m = n$

Proposition 3.2 generalizes to δ-hyperbolic groups, giving:

Corollary 3.5. If G and H are quasi-isometric hyperbolic groups, their boundaries are homeomorphic.

3.1 Facts about quasi-conformal maps

This section and the following one introduce some (important and fundamental) tools from analysis and ergodic theory used to finish Mostow's proof. Since our focus is geometric group theory, we will black-box most of these results. However, I encourage the interested reader to consult given references to appreciate how GGT interacts with other fields.

If $L : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear map, then it is easy to see that L is quasiconformal and the constant k in the definition above is equal to the *eccentricity* of the ellipsoid given by the image of a sphere under L. More generally, we have the following theorem:

⟨qc linear⟩ **Proposition 3.6.** A k-quasiconformal map is almost-everywhere differentiable. Where defined, the derivative is k-quasiconformal as a linear map.

The proof is not completely trivial, and originally due to Rademacher-Stepanov. C. Butler has some quite readable (though not written for publication) notes giving a self-contained exposition, at the level of a first course in analysis: `http://math.uchicago.edu/~cbutler/Quasiconformality.pdf`

⟨qc Liouville⟩ **Proposition 3.7** (QC Liouville's theorem)**.** Suppose $F : S^n \to S^n$ is differentiable a.e. with 1-quasiconformal derivative where defined. Then F is *conformal*.

A sketch proof of this appears in [4]. As a historical note, the original "Liouville's theorem" states that a differentiable, conformal map between domains in $\mathbb{R}^n$, $n \geq 3$ is a (conformal) Mobius transformation: a composition of inversions in spheres. This is not true for domains in $\mathbb{R}^2$ (use the Riemann mapping theorem!), but does hold for conformal homeomorphisms $\mathbb{R}^2 \to \mathbb{R}^2$ or $S^2 \to S^2$. The modern improvement is the lower regularity hypotheses on F.

In our context, conformal maps are particularly important because of the following proposition:

Proposition 3.8. The conformal homeomorphisms of S^{n-1} are precisely the boundary maps of isometries of $\mathbb{H}^n$.

As an example, in the $n = 2$ and $n = 3$ cases, these are the (real and complex, respectively) fractional linear transformations of $S^1 = \mathbb{R}P^1$ and $S^2 = \hat{\mathbb{C}}$. More generally, the conformal homeomorphisms of S^{n-1} contain inversions in hypercircles, as well as the *Euclidean similarities* of $\mathbb{R}^{n-1}$ (dilations and isometries) – fixing a point of S^{n-1}, and identifying the complement of this point with $\mathbb{R}^{n-1}$ using stereographic projection. Note that these all naturally extend to isometries of $\mathbb{H}^n$.

The important takeaway from this section is the following:

Corollary 3.9. If $F : S^{n-1} \to S^{n-1}$ has a.e. conformal derivative, then F is the boundary map of an isometry of $\mathbb{H}^n$ (and hence everywhere differentiable and conformal).

3.2 Finishing the proof of Mostow

At this point we have constructed a quasi-isometry $\tilde{f} : \mathbb{H}^n \to \mathbb{H}^n$, equivariant with respect to the action of $\pi_1(M)$ and $\pi_1(N)$ by deck transformations, i.e for our isomorphism $\phi : \pi_1(M) \to \pi_1(N)$, we have $\phi(\alpha)\tilde{f}(x) = \tilde{f}(\alpha x)$ for all $\alpha \in \pi_1(M)$.

We showed that $\tilde{f}$ induces a map on equivalence classes of geodesic rays, hence on the boundary $S^{n-1} \to S^{n-1}$. Using our basic properties of boundary maps, this equvariance passes to equivariance of the boundary action.

Let E be the projectivization of the tangent bundle of S^{n-1}. Geometrically, thinking of S^{n-1} as the unit sphere in $\mathbb{R}^n$, points in E are lines in $\mathbb{R}^n$ tangent to S^{n-1}.

Lemma 3.10. Isometries of $\mathbb{H}^n$ act transitively on E, in other words E is a *homogeneous space* for isom$(\mathbb{H}^n)$.

Thus, we can identify $E = G/H$ where H is the stabilizer of a point in E. In $\mathbb{H}^3$, we have $\mathrm{isom}_+(\mathbb{H}^2) \cong \mathrm{PSL}(2, \mathbb{C})$ and taking $x \in E$ to be any tangent vector at $0 \in \hat{C}$, we have that H is the subgroup of matrices of the form $\begin{pmatrix} \lambda & 0 \\ c & \lambda^{-1} \end{pmatrix}$ where $\lambda, c \in \mathbb{R}$. In general, H is a non-compact subgroup – it always contains an $\mathbb{R}$-subgroup of dilitations.

Since F is almost everywhere differentiable, we have an almost-everywhere defined function $h : E \to \mathbb{R}$ given as follows. Elements of E can be represented by unit tangent vectors to points in S^{n-1}, for a unit tangent vector v at p, define $h(v) := \frac{\|DF_p(v)\|}{\|DF_p\|}$.

Think of h as measuring the failure of F to be conformal, if DF is conformal at p, then $h(v) = 1$ for all v tangent at p. Since $\|DF_p\| = \sup\{\|DF_p(w) : w$ a unit tangent vector at $p\}$, it is easy to check that if g is conformal, then $h(g_*(v)) = h(v)$. Now, elements of $\pi_1(M)$ act on $\mathbb{H}^n$ by isometries, hence the induced boundary maps are conformal. We conclude that h is $\pi_1(M)$–invariant.

The last ingredient in the proof of Mostow is a classical result of Moore, proved about 10 years before Mostow's work.

Theorem 3.11 (Moore Ergodicity)**.** Let G be a non-compact, connected simple Lie group with finite center (isom$(\mathbb{H}^n)$ is an example). Let $H \subset G$ be a closed, non-compact subgroup of G. Let $\Gamma \subset G$ be a discrete subgroup such that G/Γ is compact. Then any measurable Γ-invariant function on G/H is almost everywhere constant[7].

See Chapter 2 of [12] for a detailed discussion and proof.

Using Moore's theorem, we conclude that h is a.e. constant. We need to show that it is a.e. 1. Note that, by construction, $h(v) = 1$ for *some* v tangent to every point; and for any $\epsilon > 0$, we have $h(v) < 1 + \epsilon$ on some positive measure set of tangent vectors. Since h is a.e. constant, it must be a.e. equal to 1. Using Propositions 3.6 and 3.7, we now conclude that

[7]in fancier terminology, the action of Γ on G/H is *ergodic*

F is conformal, so we can identify it with an isometry $\tilde{M} \to \tilde{N}$. Since $F\gamma F^{-1} = \phi(\gamma)$ for all $\gamma \in \pi_1(M)$, this isometry descends to an isometry $M \to N$ that induces ϕ on π_1.

Notice that here is another way boundary maps are helping us: $\tilde{f} : \tilde{M} \to \tilde{N}$ is not an isometry, but is "bounded distance from an isometry". The induced boundary map tells us that a genuine isometry exists.

4 Alternative approaches to the proof

In all approaches, we build f and the induced boundary map F as before, and need to show that F is conformal. The first, due to Tukia (perhaps building on a similar argument by D. Sullivan), avoids using Moore Ergodicity by a clever rescaling argument.

Tukia's "zooming" argument.

?⟨zooming⟩? As before we know that F is differentiable almost everywhere, so has nonzero jacobian almost everywhere. (Here's where we're using that the dimension of S^{n-1} is at least 2 – on S^1 there are "cantor staircase" functions, differentiable a.e. with derivative a.e. zero). The idea of this proof, coming from a Theorem of Tukia [**?**], is to "zoom in" on a point of differentiability. By using a suitable choice of rescaling, one shows that the linear approximation to F at this point sends a group of conformal maps (coming from $\pi_1(M)$) to conformal maps, so must be conformal.

Let x be a point of differentiability of F, and switch to the upper half space model where $x = 0 \in \mathbb{R}^{n-1}$. By composing F with a conformal map, we may assume $F(0) = 0$ and $F(\infty) = \infty$. Let L be the vertical line above 0, and let $y = (0, ..., 0, 1) \in L$. Since $\pi_1(M)$ acts cocompactly, we may find a sequence $\gamma_i \in \pi_1(M)$ so that $\gamma_i(y) \to x$ in the *Euclidean metric* on $\mathbb{R}^{n-1}$, and so that $\gamma_i(y)$ is always hyperbolic distance at most D from L, where D is the diameter of a fundamental domain for $\pi_1(M)$[8]

Let y_i be the (hyperbolic) closest point projection of $\gamma_i(y)$ to L, and let $\lambda_i \in \mathbb{R}$ be such that $\lambda_i y = y_i$. Consider the sequence of quasi-conformal homeomorphisms $\lambda_i^{-1} F \lambda_i$. Since $\lambda_i \to 0$, we have

$$\lim_{i\to\infty} \lambda_i^{-1} F \lambda_i = DF_0 \in \mathrm{GL}(n, \mathbb{R}).$$

Let $g_i = \gamma_i^{-1}\lambda_i$. These send the point y near itself, so belong to a *compact* subset of $\mathrm{isom}(\mathbb{H}^n)$, hence converge along a subsequence to some $g \in \mathrm{isom}(\mathbb{H}^n)$. Now we conjugate:

$$g_i^{-1}\pi_1(M) g_i = \lambda_i^{-1}\pi_1(M)\lambda_i$$

since $\gamma_i \in \pi_1(M)$. Let $\Gamma = g^{-1}\pi_1(M) g \subset \mathrm{isom}(\mathbb{H}^n)$. This is the limit (in the Chabauty topology, after passing to a subsequence) of $\lambda_i^{-1}\pi_1(M)\lambda_i$.

Take any sequence $h_i \in g_i^{-1}\pi_1(M) g_i$ converging to $h \in \Gamma$. Let $A = DF_0$. Since $\lambda_i^{-1} F \lambda_i h_i (\lambda_i^{-1} F \lambda_i)^{-1}$ is a sequence in $\mathrm{isom}(\mathbb{H}^n)$, and converges to AhA^{-1}

We claim now that $A^{-1}\Gamma A \subset \mathrm{isom}(\mathbb{H}^n)$. To see this, think of Γ as the limit of $\lambda_i^{-1}\pi_1(M)\lambda_i$. Since $A = \lim_{i\to\infty} \lambda_i^{-1} F \lambda_i$, one argues (some care is required to argue in what sense this limit is defined) that

$$\Gamma = \lim_{i\to\infty} \lambda_i^{-1} F \pi_1(M) F^{-1} \lambda_i$$

Since $F\pi_1(M)F^{-1} = \pi_1(N)$, we have $\lambda_i^{-1} F \pi_1(M) F^{-1} \lambda_i \subset \mathrm{isom}(\mathbb{H}^n)$, and since the group of isometries is closed, $\Gamma \subset \mathrm{isom}(\mathbb{H}^n)$.

---

[8] as a simplification, imagine the case where $\pi_1(M)$ contains a map of the form $\gamma(x) = \lambda x$ for some $\lambda < 1$. Then we take the sequence $\gamma_i = \gamma^i$. The rest of the proof becomes much simpler in this case too.

Since Γ is a (fairly large) group of isometries, this actually implies that A is a *euclidean similarity* i.e. a composition of a dilatation $x \mapsto \lambda x$ and a rotation about L. (Take a hyperplane P through 0 in $\mathbb{R}^n$, and $\gamma \in \Gamma$ so that $\gamma A(P)$ does not contain ∞, i.e. is a round hemisphere. If A is not a similarity, then $A^{-1}\gamma A$ would be an ellipsoid rather than a round sphere).

We conclude that F is conformal at 0 and hence conformal everywhere the derivative exists. Theorem 3.7 now implies that F is conformal.

The Gromov norm

One consequence of Mostow's theorem is that, for hyperbolic manifolds of dimension ≥ 3, the hyperbolic volume $\mathrm{vol}(M)$ is a topological invariant.

Gromov takes a different perspective on this. He defines a homological invariant of a manifold (the *Gromov norm*, and shows that, for hyperbolic manifolds this is proportional to the volume. With some extra work, one can use this to finish the proof of Mostow.

Definition 4.1. We work in singular homology. For an element $\alpha \in H^n(M;\mathbb{R})$, define

$$\|\alpha\| = \inf\{\Sigma|c_i| : \Sigma c_i\sigma_i \text{ is a cycle representing } \alpha\}$$

Heuristically, this measures the complexity of the class α.

Definition 4.2. For a manifold M, define $\|M\|$ to be the norm of the fundamental class of M.

Note that this immediately gives some information about the topology of M. For example, you can show:

Exercise 4.3. Show from the definition that $\|S^n\| = 0$. Now, more generally, show that if $\|M\| \neq 0$, then any map $f : M \to M$ has degree ± 1 or 0.

This follows from the easy fact that $\|f_*(\alpha)\| \leq \|\alpha\|$ for any continuous map f.

Gromov's theorem is the following rigidity statement for hyperbolic manifolds.

Theorem 4.4 (Gromov)**.** If M is a hyperbolic manifold, then $\|M\| \neq 0$. In fact, $\|M\| = \mathrm{vol}(M)/v_n$, where v_n is the maximal volume of an ideal simplex in $\mathbb{H}^n$.

This has been generalized by Lafont–Schmidt and Connell-Farb to show $\|M\| \neq 0$ for many examples of $M = \Gamma\backslash G/K$ where G is semi-simple, K a maximal compact subgroup.

One half of Gromov's theorem is easy. Let ω be a volume form on M. Then

$$\mathrm{vol}(M) = \langle \omega, [M] \rangle = \sum c_i \langle \omega, \sigma_i \rangle = \sum c_i \,\mathrm{vol}(\sigma_i)$$

for any chain $\sum c_i\sigma_i$ representing $[M]$. Now show that each σ_i can be replaced with a *geodesic* simplex, without changing the representative in homology, and without increasing this sum. Since any geodesic simplex has volume less than v_n, we have $\mathrm{vol}(M) \leq \sum |c_i| v_n$, hence $\|M\| \geq \mathrm{vol}(M)/v_n$. To get the other direction, one need to find a way to "efficiently" represent $[M]$ by a cycle.

To apply this to Mostow rigidity, one shows that, since $\tilde{f}_*[M] = [N]$ (by construction), and volume is proportional to norm, the boundary map F must map maximal volume ideal simplices to maximal volume ideal simplices. In S^{n-1}, for $n-1 \geq 2$, not every ideal simplex is maximal volume, in fact the maximal ones are precisely those that are *regular*, i.e. where the vertices can be permuted transitively by isometries. In the upper half-space model, if one vertex of such a simplex is ∞, then the others form a regular (all side lengths equal)

Euclidean simplex in $\mathbb{R}^{n-1}$. Thus, there exists some isometry ψ such that $\psi \circ F$ fixes the vertices of an ideal regular simplex. By reflecting vertices in opposite faces (iteratively), we conclude that $\psi \circ F$ fixes a dense set; continuity of F now implies that $F = \psi^{-1}$, an isometry.

We have glossed over many details here, a complete exposition is given in [**?**] and [2].

A recent proof of Besson–Courtois–Gallot

Finally, it's worth mentioning a recent proof technique that applies in a much wider context. The general result is:

Theorem 4.5 (Besson–Courtois–Gallot [3])**.** Let (Y, g) be a Riemannian manifold, (X, g_0) a closed, connected, locally symmetric[9] Riemannian manifold with negative curvature, both of dimension $n \geq 3$, and suppose $f : Y \to X$ is a map of nonzero degree. Then

$$(h(g))^n \operatorname{vol}(Y, g) \geq |\deg(f)|(h(g_0))^n \operatorname{vol}(X, g_0)$$

with equality if and only if f is homotopic to a locally isometric covering map.

From this one can derive "Mostow rigidity" for any negatively curved locally symmetric manifolds. Take X and Y to be negatively curved, connected, locally symmetric, compact manifolds with isomorphic fundamental groups, and f a homotopy equivalence between them. Then the general theorem above gives an inequality in both directions, implying that X and Y are isometric after rescaling the metrics so their volumes agree.

Amazingly, the machinery of the proof "reconstructs" the local isometry out of the boundary map.

5 Failure of Mostow in $\mathbb{H}^2$

Much as there are many distinct Euclidean structures on the torus, there are many (in fact, a $6g - 6$ dimensional family) distinct hyperbolic structures on a surface of genus $g \geq 2$. However, the first two steps of Mostow rigidity carry through: given two closed hyperbolic surface of genus g (i.e. with isomorphic fundamental groups), one builds a π_1-equivariant quasi-isometry of $\mathbb{H}^2$ that extends to a continuous map F of S^1.

While the notion of quasi-conformal doesn't make sense in dimension 1, the same argument as in the higher dimensional case proves that the boundary map is *quasi-symmetric*. Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is symmetric (at 0) if $f(x) = -f(-x)$. Hence,

⟨S1qsdef⟩ **Definition 5.1.** A function $F : S^1 \to S^1$ is *k-quasi-symmetric* if, for all $x \in S^1$

$$\lim_{r \to 0} \frac{|F(x - r)|}{|F(x + r)|} < k$$

The standard generalization of this to arbitrary metric spaces is the following:

Definition 5.2. Let ν be an increasing function $[1, \infty) \to [1, \infty)$. A function $f : X, d_X \to Y, d_Y$ is *ν-quasi-symmetric* if, for all $x, y, z \in X$ we have

$$\frac{d_Y(f(x), f(y))}{d_Y(f(x), f(z))} \leq \eta \left(\frac{d_X(x, y)}{d_X(x, z)} \right).$$

[9]a Riemannian manifold is *locally symmetric* if its universal cover has the property that, for every point p, there is an isometry fixing p and reversing geodesics through p. The negatively curved ones all have the form $\Gamma \backslash G/K$ for a semisimple Lie group G, with maximal compact subgroup K and Γ a lattice

If X and Y are domains in $\mathbb{R}^n$, $(n \geq 2)$, then ν-quasi-symmetric implies k-quasi-conformal, for a constant depending only on ν, and vice versa.

Proposition 5.3. A quasi-symmetric map of the circle is differentiable a.e.

However, q.c. maps are not necessarily absolutely continuous, so may have derivative a.e. zero!

Exercise 5.4 (straightforward exercise)**.** Check your understanding of our proof of Mostow by re-doing it for 2 dimensional manifolds, and showing that the boundary map is continuous and quasi-symmetric in the sense of definition 5.1.

In fact, Tukia's "zooming" argument (together with an argument using ergodicity) works well in this case, giving the following:

Proposition 5.5 (Mostow rigidity on the line)**.** Let $F : \mathbb{H}^2 \to \mathbb{H}^2$ be a quasi-isometry, equivariant with respect to the action of $\pi_1(\Sigma_g)$ for two hyperbolic structures on a surface Σ_g. If F is differentiable with nonzero derivative at any point, then F is conformal, and the hyperbolic structures are equivalent.

In fact, one needs much weaker hypotheses than this. There is a delightful little survey called *Mostow Rigidity on the Line* by S. Agard [1] that discusses this family of ideas.

We now discuss what else one can recover in dimension 2.

5.1 QI rigidity of surface groups

A class of groups $\mathcal{C}$ is *QI rigid* if any group quasi-isometric to a group in $\mathcal{C}$ is virtually isomorphic (see Remark 1.13) to a group in $\mathcal{C}$. This is known for many classes of groups, including free groups (due to Stallings) and free abelian groups.

One consequence of the first half of the proof of Mostow in $\mathbb{H}^2$ (with a lot of added work!) is QI rigidity for the class of surface groups, meaning fundamental groups of surfaces of genus $g \geq 2$. Any group QI to one of these is QI to the hyperbolic plane, and the "boundary map" construction gives an action of G on S^1 with finite kernel. The action has the additional property that it is properly discontinuous and cocompact on the *space of distinct triples of points in* S^1. Such actions have been well studied, and are known as *convergence group action.* It was proved in a series of papers by Tukia, Gabai, and Casson–Jungreis that a group with such an action is essentially conjugate to a discrete subgroup of $\mathrm{PSL}(2, \mathbb{R})$; this gives our virtual isomorphism between G and a surface group.

References

Agard [1] S. Agard *Mostow rigidity on the line: a survey* In Holomorphic Functions and Moduli II. Springer (1988), 1-12.

BP [2] R. Benedetti and C. Petronio, *Lectures on Hyperbolic Geometry.* Universitext. Springer-Verlag 1992

BCG [3] G. Besson, G. Courtois, S. Gallot *Minimal entropy and Mostow's rigidity theorems*, Ergodic theory and Dynamical Systems 16 (1996), 623-649.

Bourdon [4] M. Bourdon *Quasi-conformal geometry and Mostow rigidity.* Géométriesa courbure négative ou nulle, groupes discrets et rigidités, 18 (2000): 201-212.

?CH? [5] Y. Cornulier, P. De La Harpe. *Metric geometry of locally compact groups*. Preprint arXiv:1403.3796 (2014).

DK [6] C. Drutu, M. Kapovich *Lectures on Geometric Group Theory*. Book in preparation. Draft available at https://www.math.ucdavis.edu/~kapovich/EPR/kapovich_drutu.pdf

?WordProc? [7] D. B. A. Epstein, M. Paterson, J. W. Cannon, D. Holt, S. Levy, W. Thurston. *Word Processing in Groups*. AK Peters, Ltd., 1992.

Gersten [8] S. Gersten, *Introduction to hyperbolic and automatic groups*. In Summer School in Group Theory in Banff. O. Kharlampovich (ed.) CRM Proc. Lecture Notes 17, Amer. Math. Soc., Providence, RI, 1999. Pp 45-70.

?Harpe? [9] P. De La Harpe *Topics in Geometric Group Theory*. Chicago Lectures in Mathematics, University of Chicago Press, 2000.

?Thurston notes? [10] W. Thurston *Geometry and topology of three-manifolds*. Lecture notes available at http://library.msri.org/books/gt3m/

Thurston book [11] W. Thurston, S. Levy. *Three-dimensional geometry and topology*. Vol. 1. Princeton university press, 1997.

Zimmer [12] R. Zimmer. *Ergodic Theory and Semisimple Groups*. Monographs in Mathematics 81, Springer, 1984.