HYPERBOLIC GROUPS

By

M. Gromov
IHES. France

§0. Introduction

0.1 Let us start with three equivalent definitions of hyperbolic groups. First observe that for every finitely presented group Γ there exists a smooth bounded (i.e. bounded by a smooth hypersurface) connected domain $V \subset \mathbb{R}^n$ for every $n \geq 5$, such that the fundamental group $\pi_1(V)$ is isomorphic to Γ. A standard example of such a V is obtained as follows. Fix a finite presentation of Γ and let P be the 2-dimensional cell complex whose 1-cells correspond in the usual way to the generators and the 2-cells to the relations in Γ, such that $\pi_1(P) = \Gamma$. Then embed P into $\mathbb{R}^5$ and take a regular neighborhood of $P \subset \mathbb{R}^5$ for V.

(A) First definition: Call Γ *word hyperbolic* if there exists a constant $C > 0$ (depending on Γ and V), such that every smooth simple closed curve S in V (i.e. an embedded circle) which is *contractible* in V bounds a smooth embedded disk $D \subset V$, satisfying the following *isoperimetric inequality* (compare [Gr_1])

$$(Is_2) \qquad \text{Area } D \leq C \text{ length } S.$$

It is quite easy to see that this hyperbolicity does not depend upon the choice of V (but of course, the implied constant C does depend on V). One could also formulate (Is_2) in a combinatorial fashion (see 2.3) by using (simplically mapped) curves and disks in P.

(B) Second definition: Let Γ be a finitely generated group with a

---

Research supported in part by NSF Grant DMS-8120790.

fixed finite system of generators and let $|\gamma|$ denote the length of a shortest word (in these generators) representing given $\gamma \in \Gamma$. Set (compare 1.1)

$$(\alpha \cdot \beta) = 1/2(|\alpha| + |\beta| - |\alpha^{-1}\beta|)$$

and call Γ word hyperbolic if there exists a constant $\delta \geqslant 0$ (depending on a choice of generators), such that every three elements α, β and γ in Γ satisfy the following *δ-version of the ultrametric triangle inequality*

(*) $$(\alpha \cdot \beta) \geqslant \min((\alpha \cdot \gamma),(\beta \cdot \gamma)) - \delta.$$

(This is equivalent to Rips' definition of hyperbolicity as well as to the definition in [Gr_5].)

It is not apriori obvious why this hyperbolicity is independent of the choice of generators. But after a closer look at (Is_2) and (*) one recognizes both as well known properties of *manifold of negative curvature*. Then it becomes clear that $Is_2 \Leftrightarrow$ (*), though a direct proof of $Is_2 \Rightarrow$ (*) appears cumbersome. We give a rather short proof of that in 6.8 by using the *Riemann mapping theorem* and a version of *Ahlfors-Schwartz lemma*. To see how these analytic tools may help to solve a combinatorial problem we look at the following

(C) <u>Example</u>: Consider two triangulations of $\mathbb{R}^2$, say P_1 and P_2, satisfying the following two conditions.

(a) P_1 and P_2 are *uniformly locally finite*. That is there is an integer N such that every vertex in P_1 has at most N adjacent simplices and the same is true for P_2.

(b) P_1 and P_2 are *hyperbolic*. That is each vertex in P_1 (as well as in P_2) has at least seven adjacent triangles.

Then we claim the existence of simplical subdivisions P_1' and

P_2', such that

(i) The triangulations P_1' and P_2' are isomorphic (as simplical complexes).

(ii) Each simplex in P_1 contains at most M simplices of P_1' for some integer M depending on N, and the same holds for P_2 and P_2'.

__Sketch__ __of__ __the__ __Proof__: There exist (by an easy argument) Riemannian metrics g_1 and g_2 on $\mathbb{R}^2$ with curvatures between -1 and $-1 - K$ for some $K \geq 0$ depending on N, such that the edges of P_1 are g_1-geodesic of unit g_1-length and those in P_2 are g_2-geodesics of unit g_2-length. Then Riemann's theorem provides a *conformal* diffeomorphism $f: (\mathbb{R}^2, g_1) \longrightarrow (\mathbb{R}^2, g_2)$. The norms of the differentials of f and f^{-1} are bounded according to Schwartz lemma,

$$\|D_f\| + \|D_{f^{-1}}\| \leqslant (2 + \sqrt{K})^2,$$

which allows a simplical approximation of f by the required simplical isomorphism $P_1' \longrightarrow P_2'$.

(D) __Third__ __definition__: Let Γ be a group with a fixed finite system of generators. A metric space T is called a *tangent subcone of* Γ *at infinity* if it satisfies the following two properties.

(i) __Geodesic__ __property__: For every two points t_1 and t_2 in T there exists a *geodesic segment* in T between t_1 and t_2. That is, by definition, an *isometric* map $s: [0,\tau] \longrightarrow T$ of the segment $[0,\tau] \subset \mathbb{R}$ with $\tau = \text{dist}(t_1,t_2)$, such that $s(0) = t_1$ and $s(\tau) = t_2$.

(ii) __(Ultra)limit__ __property__: For every finite subset $\{t_1,\dots,t_n\} \subset T$ there exists an infinite sequence of positive numbers $\varepsilon_k \longrightarrow 0$ for $k \longrightarrow \infty$ and n sequences $\gamma_{ik} \in \Gamma$ for $i = 1,\dots,n$, such that

$$\epsilon_k \; |\gamma_{ik}^{-1}\gamma_{jk}| \longrightarrow \mathrm{dist}(t_i,t_j)$$

for $k \longrightarrow \infty$ and $1 \leqslant i,j \leqslant n$.

Now, call Γ word hyperbolic if every tangent subcone T of Γ is a *tree*. That is for every two distinct points t_1 and t_2 in T there exists one and only one *topological segment* $S \subset T$ between t_1 and t_2. (S is homeomorphic to $[0,1]$, such that t_1 and t_2 correspond to the ends of $[0,1]$.)

This definition is most intuitive and, possibly, best suited for generalizations. It says, in effect, that Γ with a word metric looks as a tree when observed from infinity. Notice that inequality (*) immediately implies the tree-likeness of Γ at infinity and the converse (in a more precise quantitative form) is proven in Section 6.

0.2 Examples of hyperbolic groups:

(A) It is obvious (with either of the three definitions) that free groups are word hyperbolic. It is also well known that (small cancellation) 1/6-groups satisfy Is_2 (see 4.7) and, hence, they are word hyperbolic. To get a rough idea of by how much hyperbolicity generalizes small cancellation, we look at these two classes of groups from a probabilistic point of view. Namely, we fix a generating set $\{\gamma_1,\dots,\gamma_p\}$ for $p \geqslant 2$ and then consider the set of presentations given by q relations which are cyclically irreducible words in γ_i, such that the lengths of these words lie in some interval $[\ell_1,\ell_2]$. The number $N = N(p,q,\ell_1,\ell_2)$ of such presentations is approximately $q(2p-1)^{\ell_2}$.

Fix q and let ℓ_1 and ℓ_2 go to infinity, such that the ratio ℓ_2/ℓ_1 is kept bounded, say $\ell_1 \leqslant \ell_2 \leqslant 2\ell_1$. Then, asymptotically, almost all among our presentations are 1/6. Namely, the number $N_{1/6}$ of these satisfies $N_{1/6}/N \longrightarrow 1$ as $\ell_2 \longrightarrow \infty$. However, if we drop the bound $\ell_2/\ell_1 \leqslant$ const, we come up with the opposite conclusion. That is $N_{1/6}/N \longrightarrow 0$ for $\ell_2 \longrightarrow \infty$, provided there are words in our presentation whose lengths, say ℓ_1 and ℓ_2, satisfy $\ell_2 \geq \exp 10p\ell_1$. Yet the number N_h of the word hyperbolic (groups given by these) presentations is still okay, $N_h/N \longrightarrow 1$.

This "genericity" of word hyperbolic groups is important in various constructions of "exotic" finitely generated infinitely presented groups (notice that word hyperbolic groups are finitely presented, see 2.2) such as *infinite torsion groups* and *Olshanski groups* which have no proper non-cyclic subgroups. These are usually obtained by taking an infinite sequence of relations w_i in fixed generators $\gamma_1,\dots,\gamma_p$ with a desired (e.g. torsion) property, and the problem is to show that the resulting group $\{\gamma_1,\dots,\gamma_p \mid w_1,w_2\dots\}$ is infinite. The small cancellation approach becomes increasingly complicated here (for instance, see [Ol]), since intermediate finitely presented groups $\{\gamma_1,\dots,\gamma_p \mid \gamma_1,\dots,\gamma_k\}$ are not small cancellation for large $k \longrightarrow \infty$. Yet they can be chosen hyperbolic and then geometric techniques apply. In fact, once basic geometric features of negatively curved manifolds (such as the above inequalities Is_2 and (*) as well as the equivalence $Is_2 \Leftrightarrow$ (*)) are observed (and proved) for word hyperbolic groups, then there remains nothing mysterious about "exotic" groups. They are seen as clearly as, for example, (the freedom of) subgroups in a free group, once free groups are looked upon as fundamental groups of 1-dimensional polyhedra.

For example, it is geometrically obvious (and a complete proof takes hardly a page, see 4.5.C) that the fundamental group Γ of every closed manifold of negative curvature admits an infinite factor group $\bar{\Gamma}$ all of whose elements are torsion (probably, one can achieve a universal bound on the order of the torsion, but here geometry still lags behind algebra) and such that $\bar{\Gamma}$ isometrically acts on a certain space of non-positive curvature.

In fact, one can build "exotic" groups $\bar{\Gamma}$ by adding (infinitely many) relations to any given hyperbolic group Γ (instead of the free group generated by $\gamma_1,\dots,\gamma_p$). If one starts with a lattice Γ in $Sp(n,1)$ (see (B) below) with $n \geq 2$, then one comes up with a $\bar{\Gamma}$ which among other "pathologies" enjoys *Kazdan's T-property* that is a sharpening of non-amenability (see 5.6).

(B) The hyperbolic groups Γ obtained by adding random relations to a free group are, after all, quite simple *two-dimensional* creatures as they (by an easy argument) act (freely if there is no torsion) on locally finite 2-dimensional polyhedra X with compact quotients $V = X/\Gamma$. One knows in many cases that the quotient polyhedron (or *orbihedron*, if Γ has torsion) V admits a (singular) metric of negative curvature (see Section 4). Similar objects exist in all dimensions $n \geqslant 2$ (see Section 4) and the most studied among them are closed Riemannian *manifolds* (and orbifolds) of negative curvature. The very existence of such manifolds V and of their (hyperbolic!) fundamental groups Γ relies (in known examples for large n, say for $n \geqslant 20$) on (elementary but non-trivial) arithmetic of number fields. Namely, these Γ are cocompact arithmetic subgroups (or simple modifications of arithmetic subgroups) in simple Lie groups with $\mathbb{R}$-rank = 1. These are $O(n,1)$, $U(n,1)$, $Sp(n,1)$ and the Cayley group. The study of discrete (not necessarily cocompact) subgroups Γ in semisimple (of any rank) Lie groups in the last thirty years revealed an astounding geometric (as well as algebraic) beauty and stimulated a search for more general classes of similar groups Γ. It seems word hyperbolic groups provide a good start for this search.

(C) Higher dimensional hyperbolic groups also can be constructed by combinatorial (or geometric) means. For example, *for every finite polyhedron* V_0, *one can find an aspherical polyhedron* V *with word hyperbolic fundamental group* $\Gamma = \pi_1(V)$, *such that* dim V = n = dim V *and* V *admit a continuous map* $f\colon V \longrightarrow V_0$ *which is injective on the cohomology* $H^*(V_0)$. *Furthermore, if* V_0 *is a manifold, then one also can choose* V *a manifold, such that* $f^*\colon H^*(V_0;\mathbb{Q}) \longrightarrow H^*(V,\mathbb{Q})$ *sends the Pontryagin classes of* V_0 *to those of* V. (See 3.4.)

(D) Isometry groups of manifolds (and singular spaces) with non-positive curvature (e.g. lattices in semisimple Lie groups with $\mathbb{R}$-rank $\geqslant 2$) usually are not word hyperbolic. Since we do not know what is the correct definition for such groups (and we do not know if several natural definitions are equivalent) we suggest the following

(E) <u>Non-definition</u>: A group Γ is called *semihyperbolic* if it looks as if it admits a discrete cocompact isometric action on a space of non-positive curvature. (Sometimes we allow *non-cocompact* actions which brings into the picture finitely generated nilpotent groups and related classes of groups.)

Additional motivating examples are 1/5-groups and Cartesian products of hyperbolic (and semihyperbolic) groups. (Notice that $\Gamma_1 \times \Gamma_2$ is not word hyperbolic unless one of the two factors, Γ_1 or Γ_2, is a finite group.)

(F) <u>Relative hyperbolicity</u>: One can handle some weak forms of *semihyperbolicity* in a purely hyperbolic context with a notion of a group Γ which is (word) *hyperbolic relative to a given system of subgroups* in Γ. This notion (defined in 8.5) is similar to that of a small cancellation group over a free product with amalgamations. Basic examples of relatively hyperbolic groups are *non-cocompact* lattices Γ in Lie groups of $\mathbb{R}$-rank = 1. Such a Γ is hyperbolic relative to the *cuspidal* subgroups $\Gamma_i \subset \Gamma$ (which are certain virtually nilpotent groups).

In fact we give in Section 2 a definition of (non-word) hyperbolic groups which automatically includes the relative hyperbolicity (as well as discrete infinite covolume subgroups in Lie groups of $\mathbb{R}$-rank = 1).

0.3 <u>Basic features of hyperbolic groups</u>:

(A) The tree-like behaviour of hyperbolic groups Γ at infinity (see 0.1(D)) shows a strong resemblance between word hyperbolic groups and free groups. A great deal of the small cancellation theory is devoted to generalizing basic properties of free groups to small cancellation groups and a similar generalization carries over to hyperbolic groups. In fact, standard small cancellation arguments become significantly shorter when translated to the hyperbolic language. For example, the solvability of the word and conjugacy problems become totally obvious in this language.

(B) The next stage of a study of Γ starts with a construction of an *ideal boundary* (see 1.8), also called *the hyperbolic boundary* ∂Γ. This is a functorially constructed compact space where Γ acts by homeomorphisms. (This construction is essentially due to Mostow and Margulis.) For example, if Γ is a free group with $p \geqslant 2$ generators then ∂Γ is the Cantor set formed by infinite irreducible words in these generators. This ∂Γ is 1-dimensional for small cancellation groups (e.g., ∂Γ is the usual boundary circle for surface groups Γ). If Γ is the fundamental group of a closed n-dimensional manifold of negative curvature, then ∂Γ is homeomorphic to S^{n-1}, though it does not carry, in general, any natural smooth structure. The topology of ∂Γ for a typical Γ looks quite complicated, both locally and globally, but no systematic study has been conducted so far.

(B') Using the action of Γ on ∂Γ we construct *a compact space (of geodesics)* with an action of ℝ on this space called *the geodesic flow*. This flow is not unique. Yet the one-dimensional foliation into orbits is uniquely (though not functorially) determined by Γ. This rigidity (or stability) of geodesics in manifolds of negative curvature is a famous theorem by M. Morse (see [Mor]) which was brought into the modern context of *hyperbolic dynamics* by Smale and Anosov. In fact, the geodesic flow of Γ is (Anosov-Bowen) hyperbolic. This allows us to apply to Γ the major geometric tool of hyperbolic geometry. That is Thurston's method of *geodesic (hyperbolic) simplices*. A standard application of this method shows, for example, that every *word hyperbolic group Γ contains at most finitely many pairwise non-conjugate subgroups isomorphic to a given finitely presented group Γ_0, provided Γ_0 has only one end.* (Compare 5.3.)

(B") An important feature of the hyperbolic dynamics (also discovered by Morse) is *Markov coding* (or presentation) of hyperbolic systems which bring in combinatorial techniques of *symbolic dynamics* (see [Bow]). The efficiency of the combinatorial methods is seen, for example, in Manning's proof (see [Man]) of Smale's conjecture on *the rationality* of the *ζ-function* (which counts periodic orbits) of every Bowen-hyperbolic $\mathbb{Z}$-action and Cannon's proof (see [Can]) of the rationality of the *counting function* $Z(t) = \sum_k \#\{\gamma \in \Gamma \mid |\gamma| = k\} t^k$ for fundamental groups Γ of manifolds of negative curvature. We shall see in 5.2 and 8.4 that the Markov coding applies to all hyperbolic groups and yields rationality theorems similar to those of Manning and Cannon.

(C) Every lattice Γ in a simple Lie group of rank one naturally acts on $S^{n-1} = \partial\Gamma$ and preserves certain conformal (and, hence, quasiconformal) structure (see [P]) determined by the underlying Lie group. (If $\Gamma \subset O(n,1)$, this is the usual conformal structure.) Margulis (see [Marg]) pointed out that (the first version of) Mostow's proof of rigidity for these Γ produces this conformal structure intrinsically in terms of ("algebraic" properties of) Γ. This remark by Margulis suggests the existence of a natural geometric structure for "most" word hyperbolic groups Γ. (Exceptional Γ, probably, are those where the boundary ∂Γ is disconnected or contains a connected open subset $U \subset \partial\Gamma$, such that $U \setminus A$ is disconnected for some zero-dimensional $A \subset U$.) One wishes to extend Γ to the full automorphism group $L \supset \Gamma$ of this structure, such that L/Γ is compact, and such that L is the unique (in an obvious sense ruling out extensions like $\Gamma \times G$ for compact G) maximal cocompact extension of Γ. For majority of Γ this L is, probably, discrete, for the next class of Γ the group L is totally disconnected, and for the remaining Γ the group L is Lie. Here is a specific conjecture where such structure looks indispensable: Every torsionless word hyperbolic group Γ admits at most finitely many torsionless extensions $\Gamma' \supset \Gamma$ with finite quotients Γ'/Γ.

0.4 The idea of hyperbolicity has been lingering in combinatorial group theory since the basic work by Dehn. An extensive study of a class of word hyperbolic groups Γ with dim $\partial\Gamma = 1$ (in the combinatorial disguise) was conducted by Olshanski (see [Ol]). Deep algebraic results on general hyperbolic groups are contained in the as yet unpublished work by I. Rips who calls them *groups with negative curvature*.

Our approach is motivated by hyperbolic phenomena in the geometry of manifolds and spaces of non-positive, in particular negative, curvature and in topological dynamics (see [Gr1]). One still does not know how to embrace curvature $K \leqslant 0$ (in particular, 1/5-groups and alike) into a consistent (semi)hyperbolic theory. The obvious difficulty is a non-stability of the condition $K \leqslant 0$ in contrast to the stability of $K < 0$. For example, one does not know how to see semihyperbolicity of a 1/5-group if it is given by a non-1/5-presentation. (Notice that 1/6-groups are caught into our word hyperbolic maze.) But the case $K < 0$ is understood (at least on the foundational level) and an exposition is long overdue. There is a "linguistic" difficulty in discussing hyperbolic groups as one translates clear-cut geometric notions into their "quasi" and "approximate" counterparts. (This difficulty would disappear if one could realize every hyperbolic group by isometries on an appropriate space of negative curvature.) The terminology has undergone a slow selection process in several oral presentations of the subject and is now bearable, I hope. The major clarification was achieved at my stay with Tata Institute in 1984-85 thanks to the receptive audience and many stimulating discussions with G. Prasad. The decision of writing everything up was made under the friendly pressure by S. Gersten at MSRI in Berkeley.

Finally, I wish to thank Jim Cannon who generously spent his time on reading this manuscript, found a multitude of errors and suggested many improvements. He also told me of his work in collaboration with D. Epstein and W. Thurston on an automata theoretic approach to group theory that stems from his paper [Ca]. This approach seems deeper in many respects than the "hyperbolic philosophy" and it applies to a wider class of groups.

<u>**The structure of the paper**</u>: We start in Sections 1-5 with an

exposition of examples and basic properties of hyperbolic groups. We give few detailed proofs at this stage but rather explain underlying geometric ideas. These ideas become rigorous as we develop in Sections 6–8 geometry of *hyperbolic metric spaces* (see 1.1). In Section 6 we prove the equivalence of different notions of hyperbolicity. This allows us to translate basic facts on manifolds with $K < 0$ to a "curvature free" language which is also "stable" under small perturbations of metrics (compare [Kl]). This stability is achieved at the cost of "quasi"-fication of terminology. There is also a large amount of numerical constants (e.g. constant δ in 0.1) which are needed for definitions of various "quasi"-notions and for establishing relations between them. We tried, whenever possible, to use specific values of the constants and write $A \leqslant 100B$ rather than: "there exists $C \leqslant 100$, such that $A \leqslant CB$."

The discussion in Sections 6 and 7 is purely geometrical. That is no group appears there. Groups Γ came back into play in Section 8, as isometry groups of hyperbolic spaces X; and algebraic properties of Γ are reflected in geometry of X and X/Γ. Our major tool (which does not come free unlike the case $K < 0$) is (a kind of) a *geodesic flow* over X/Γ which completes our dictionary and justifies the treatment of word hyperbolic groups as of fundamental groups of compact manifolds (and spaces) with curvature $K < 0$.

Our exposition is essentially self-contained. Yet a reader may benefit by looking through Chapter V in [Ly–Sc] on small cancellation theory and Chapter I in [B–G–S] on manifolds with curvature $K \leqslant 0$. Further references can be found in these books. Unfortunately, there is a fair amount of "well known" and "easy to prove" little geometric theorems which are hard to locate in the literature. We bring them forth whenever they help to motivate our abstract discussion but we do not provide the proofs.

HYPERBOLIC GROUPS

Table of Contents

Table of Contents (Cont'd)

Table of Contents (Cont'd)

§1. Examples and basic properties of hyperbolic groups

1.1 Hyperbolic metric spaces: If X is a metric space then the distance between two points x and y in X is denoted by $|x-y|$. If X is a *pointed space*, that is if some $x_0 \in X$ is chosen as a *reference point*, then we set $|x| = |x|_{x_0} = |x-x_0|$ and

$$(x.y) = (x.y)_{x_0} = 1/2(|x| + |y| - |x-y|).$$

Call X *hyperbolic* with respect to x_0 if it satisfies the following *δ-inequality* (compare 2.4.C.)

$$(*) \qquad (x.y) \geq \min((x.z),(y.z)) - \delta$$

for a fixed $\delta \geq 0$ and all x, y and z in X. If we want to specify δ we say X is *δ-hyperbolic* with respect to x_0.

Observe the following obvious

1.1.A Lemma: *If X is δ-hyperbolic then*

$$(**) \qquad (t.y) + (z.x) - \min((t.z) + (y.x)\ ,\ (x.t) + (y.z)) \geq -2\delta$$

for all x, y, z *and* t *in* X.

1.1.B Corollary: *If X is δ-hyperbolic with respect to* $x_0 \in X$ *then it is 2δ-hyperbolic with respect to every point* $x \in X$.

Proof: By an obvious computation the left-hand side of (**) equals $(t.y)_x - \min((t.z)_x,(y.z)_x)$. Q.E.D.

1.1.C Definition: A metric space X is called *δ-hyperbolic* if it is δ-hyperbolic with respect to each point $x \in X$. We call X *hyperbolic* if it is δ-hyperbolic for some $\delta \geq 0$.

1.1.D **Important remark.** Basic examples of hyperbolic spaces are *simplicial trees* (see 1.4.), where the "product" $(x.y)$ equals the distance from the reference point to the edge joining x and y (see 1.4.B.). This shows trees are δ-hyperbolic for $\delta = 0$ (compare 1.4.A.). In fact, every 0-hyperbolic space isometrically embeds into a tree and an arbitrary δ-hyperbolic space can be approximated by trees (see §6).

1.2 First examples of hyperbolic spaces:

(a) The Euclidean space $\mathbb{R}^n$ for $n \geqslant 2$ (obviously) is not hyperbolic but the real line $\mathbb{R}^1$ is δ-hyperbolic for $\delta = 0$.

(b) If X is a bounded space, that is if

$$\operatorname{Diam} X \overset{\text{def}}{=} \sup_{x,y} |x-y| \leqslant d < \infty,$$

then (obviously) X is δ-hyperbolic for $\delta = d$.

(c) Take an arbitrary metric space $(X, |\ |)$ and define a new metric on X by

$$|x-y|' = \log(1 + |x-y|).$$

Then by the *triangle inequality*,

$$|x-y| \leqslant |x-z| + |y-z|,$$

and hence, the new metric satisfies

$$|x-y|' \leqslant \max(|x-z|', |y-z|') + \log 2.$$

It follows (compare 1.1.A) that

$$-|x-y|' - |z-t|' + \max(|x-z|' + |y-t|', |x-t|' + |y-z|')$$

$\geqslant -2 \log 2.$

A straightforward computation identifies the left-hand side of this inequality with

$$(t.y)'_x - \min((t.z)'_x,(y.z)'_x),$$

which shows that $(X,|\ |')$ is δ-hyperbolic for $\delta = 2 \log 2$.

1.3 **Maximal metric**: Consider a non-empty class of metrics on X satisfying a certain condition $\mathcal{C}$. Then the supremum of the metrics in this class is again a metric, unless it becomes infinite at some pair of points in X. This supremum is called the *maximal metric satisfying* $\mathcal{C}$.

1.3.A **Example**: Take a locally finite covering of X by subsets $X_i \subset X$, $i \in I$, where each X_i is endowed with a metric $|\ |_i$. Then we have the maximal metric $|\ |$ on X satisfying $|\ | \leqslant |\ |_i$ on every subset X_i. This metric is nowhere infinite (i.e. an actual metric on X) if and only if the nerve of the covering is connected.

1.4 **Simplical metrics and trees**: A metric $|\ |$ on a simplical polyhedron P is called *simplical* if

(a) the restriction $|\ |_i$ of $|\ |$ on each simplex $\Delta_i \subset P$, $i \in I$, is *Euclidean*. That is $(\Delta_i,|\ |_i)$ admits an affine isometric map into $\mathbb{R}^k$ for $k = \dim \Delta_i$;

(b) the metric $|\ |$ maximal for the condition $|\ | = |\ |_i$ on each Δ_i.

For example, for each $d > 0$ there exists a unique simplical metric on P, where every edge (i.e. a 1-simplex in P) is isometric to $[0,d]$.

1.4.A Recall that P is called *a tree* if it is contractible and dim $P = 1$. Observe that every two distinct points x and y in a tree P can be joined in P by a unique topological segment, denoted $[x,y] \subset P$. Furthermore, for every finite subset $Y = \{x,y,z,t...\} \subset P$ there is a unique minimal subtree (for some refinement of P) $T(Y) \subset P$ containing Y. Namely, $T(Y)$ is the union of the segments between the points in Y. See Figure 1 below for typical trees $T(Y)$, where Y contains 2, 3 and 4 points.

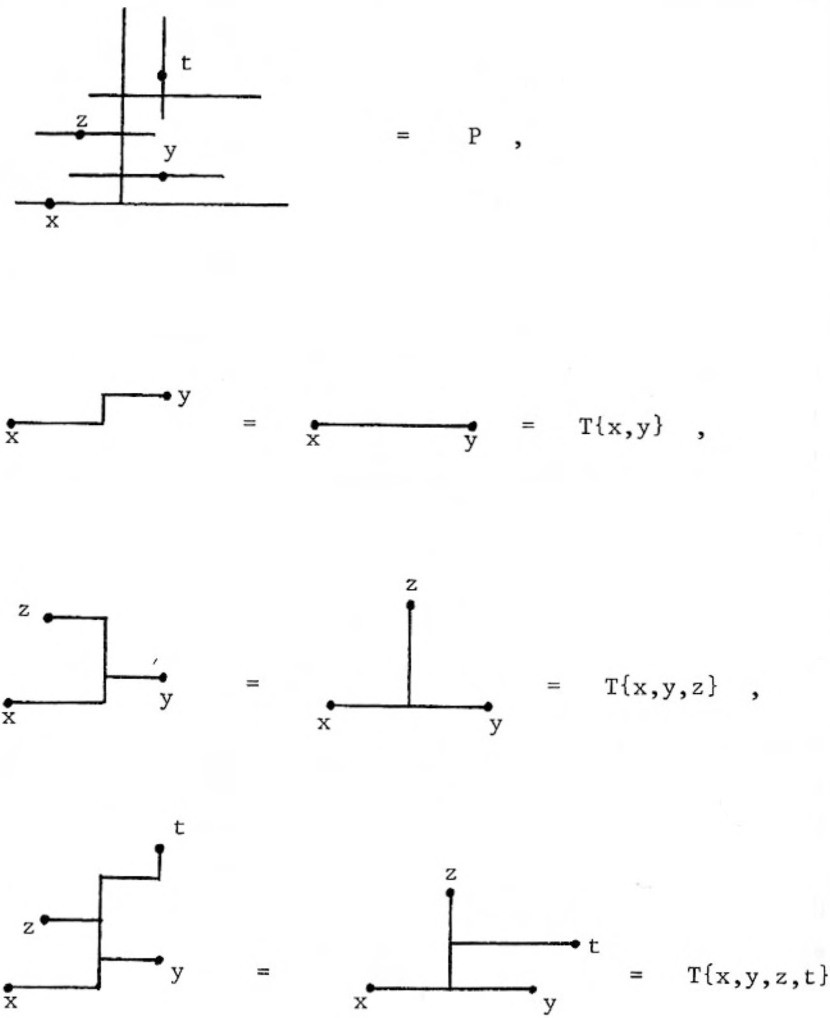

Figure 1

<u>**An important example:**</u> *Every tree P with a simplical*

metric $|\ |$ *is* δ*-hyperbolic for* $\delta = 0$.

<u>Proof</u>: Every segment $[x,y] \subset P$ is divided by the vertices of P meeting $[x,y]$ into finitely many subsegments, say $[p_i,p_j] \subset [x,y]$ isometric to the standard segments $[0,|p_i-p_j|] \subset \mathbb{R}$. Since the metric $|\ |$ is maximal, the total length of these segments equals $|x-y|$. Now, to show that $(t.y)_x \geqslant \min((t.z)_x,(y.z)_x)$, we look at the subtree $T = T\{x,y,z,t\} \subset P$. We ignore the "irrelevant" vertices, where T (and hence P) is locally homeomorphic to $\mathbb{R}$ and thus reduce the story to a tree with at most 6 vertices and 5 edges. We conclude the proof by staring at $T\{x,y,z,t\}$ in Figure 1.

1.4.B <u>Remark</u>: In the case of trees the "product" $(y.t)_x$ equals $\operatorname{dist}(x,[y,t])$, where the distance to a subset is measured by

$$\operatorname{dist}(x,Y) \overset{\text{def}}{=} \inf_{y\in Y} |x-y|.$$

1.5 <u>Geometric examples of hyperbolic spaces</u>: The following statements (1) and (2) are well known (and quite easy to prove).

(1) Every complete simply connected Riemannian manifold with *strictly negative sectional curvature*,

$$K \leqslant -\varkappa^2 < 0,$$

is δ-hyperbolic for $\delta \leqslant \mathcal{C}\varkappa^{-1}$, where $\mathcal{C}$ is a universal constant in the interval $0 < \mathcal{C} < 10$ (compare 2.4.E.).

(2) Lie groups $O(n,1)$, $U(n,1)$, $Sp(n,1)$ with the left invariant Riemannian metrics and the isometry group of the hyperbolic Cayley plane are hyperbolic. In fact these are the only simple non-compact Lie groups which are hyperbolic for left-invariant *Riemannian* metrics. (One has with example 1.2.(c) an invariant non-Riemannian hyperbolic metric on every Lie group.)

(3) Let $\Omega \subset \mathbb{R}^2$ be a connected and simply connected open subset

which contains no disk of radius $\geqslant R_0$ for some fixed $R_0 > 0$. Define the distance in Ω as the infimum of lengths of curves in Ω between pairs of points. Then Ω with this metric is δ-hyperbolic for some $\delta \leqslant \mathcal{C} R_0^{-1}$ where $0 < \mathcal{C} < 10$. The proof is easy and left to the reader.

Notice that this length metric in Ω can be defined as the maximal metric $|\ |'$ in Ω satisfying $|x-y|' = |x-y|_{\mathbb{R}^2}$ for every pair of points x and y in Ω lying on a straight segment contained in Ω.

1.6 __Geodesics__ __metric__ __spaces__: *A geodesic segment between two points* x and y in a metric space X is (the image of) an isometric map $[0,|x-y|] \longrightarrow X$ sending $0 \longmapsto x$ and $|x-y| \longmapsto y$. Such a segment, (if it exists at all) may be not unique. Yet we use the notation $[x,y] \subset X$ for such segments. We also denote by $(1-s)x + sy \in [x,y] \subset X$ the image in X of the point $s|x-y| \in [0,|x-y|]$ for $0 \leqslant s \leqslant 1$. If X is given a reference point then sx denotes a point in X such that $|sx| = s|x|$ and $|sx-x| = (1-s)|x|$. If x can be joined by a segment with the reference point, then (obviously) some sx exists for $0 \leqslant s \leqslant 1$.

A metric space X is called *geodesic* if every two points in X can be joined by a segment. It is well known (and easy to prove) that

1.6.A **(Hopf-Rinow)** *Every complete Riemannian manifold is geodesic.*

1.6.B *Let* P *be a polyhedron with a simplical metric such that every edge has a fixed length* d_0. *Then* P *is geodesic.* (*The proof of this is an exercise for the reader.*)

1.6.C Let X be a metric space with an arbitrary metric $|\ |$ and denote $|\ |_\epsilon$ for some $\epsilon > 0$ the maximal metric, such that $|x-y|_\epsilon = |x-y|$ for all x and y in X having $|x-y| \leqslant \epsilon$. Note that if $(X, |\ |)$ is a geodesic space then, obviously $|\ |_\epsilon = |\ |$ for all $\epsilon > 0$. Let $|\ |_0 = \sup_{\epsilon>0} |\ |_\epsilon$ and assume $|\ |_0$ is a (nowhere infinite) metric in X. Then by an easy argument the metric completion of $(X, |\ |_0)$ is a geodesic space.

Example: Let X be a smooth connected submanifold in $\mathbb{R}^n$ with the induced metric. Then $|\ |_0$ equals the *induced Riemannian metric* in X, where the distance between two points equals the infimum of lengths of smooth curves in X between these points.

1.7 Polyhedron $P_d(X)$: Let X be a metric space, take $d \geqslant 0$ and let $P_d(X)$ be the simplical polyhedron whose set of vertices (i.e. the 0-skeleton) equals X and where a finite subset $Y \subset X$ spans a simplex in $P_d(X)$ if and only if the distance between every two points in Y is $\leqslant d$. Notice that if X is a geodesic space then the 1-skeleton of $P_d(X)$ is equal to that of the nerve of the covering of X by the balls of radius $\leqslant 1/2d$. Recall that the ball in X of radius R around $x \in X$ is

$$B_x(R) = \{y \in X \mid |x-y| \leqslant R\}.$$

Remark: The space $P_d(X)$ serves as a regularization of X. For example, $P_d(X)$ allows us to state Is_2 (see §0) for an arbitrary X and then to prove Is_2 for hyperbolic spaces X (see 1.7.C.). One also needs $P_d(X)$ in the proof of Rips' theorem on the finiteness of the virtual homological dimension of word hyperbolic groups (see 1.7.A., 1.7.D. and 2.2.).

Let K be a subpolyhedron in $P_d(X)$ and let $Y \subset X$ denote the set of vertices of K. Observe that every map $f: Y \longrightarrow X$ defines a simplical map, say $f': K \longrightarrow P_{d'}(X)$, provided $d' \geqslant |f(y_1)-f(y_2)|$ for all pairs y_1 and y_2 in Y satisfying $|y_1-y_2| \leqslant d$. Furthermore, if $d' \geqslant d$ and $|y_1-f(y_2)| \leqslant d'$ for $|y_1-y_2| \leqslant d$, then (obviously) there exists a homotopy in $P_{d'}(X)$ between map f' and the original embedding $K \subset P_d(X) \subset P_{d'}(X)$.

The following important observation is due (up to the terminology) to I. Rips.

1.7.A **Lemma**: *Let* X *be a* δ*-hyperbolic space such that every* $x \in X$ *can be joined by a segment with a fixed reference point* x_0 *in* X. *Then the polyhedron* $P_d(X)$ *is contractible for all* $d \geqslant 4\delta$.

Proof: It suffices to contract every finite subpolyhedron K in $P_d(X)$ which is done by moving the vertices $y \in X$ of K one after another toward the reference point x_0 in X. To see how it works take a vertex, say $y_0 \in X$ of K, move it to $y_0' = |y_0|^{-1}(|y_0|-d_0)y_0$ in X for some $d_0 > 0$ and estimate $|x-y_0'|$ for all $x \in X$ as follows

$$\begin{aligned}|x-y_0'| &= |x| + |y_0'| - 2(x.y_0') \leqslant \\ &\leqslant 2\delta + |x| + |y_0'| - 2\min((x.y_0),(y_0.y_0')) = \\ &= 2\delta + \max(|x-y_0|-d_0, |x|-|y_0|+d_0).\end{aligned}$$

If y_0 is the farthest vertex of K from x_0, then this inequality implies for all vertices y of K and for $2\delta \leqslant d_0 \leqslant 1/2d$, $|y-y_0| \leqslant 2\delta + \max(d-d_0,d_0) \leqslant d$. Hence, this move defines a homotopy of K in $P_d(X)$ and finitely many such homotopies obviously contract K in $P_d(X)$.

1.7.B **Remark**: Suppose K is homeomorphic to the circle and let y_1 and y_2 be the vertices in K adjacent to the farthest (from x_0) vertex y_0. Then we see as earlier,

$$\begin{aligned}|y_1-y_2| &\leqslant 2\delta + \max(|y_1-y_0|+|y_2|-|y_0|, |y_2-y_0|+|y_1|-|y_0|) \\ &\leqslant 2\delta + d.\end{aligned}$$

Consider two kinds of homotopy of K to a shorter circle K' in $P_d(X)$.

Case 1: If $|y_1-y_2| \leqslant d$ then the shortening of K is obvious: remove the edges $[y_1,y_0]$ and $[y_2,y_0]$ and join y_1 with y_2 by an edge in

$K_d(X)$.

<u>Case</u> <u>2</u>: Let $|y_1-y_2| \geqslant d$. Then by the above inequality one of the vertices y_1 and y_2, say y_1, satisfies

$$|y_1| \geqslant |y_0| - 2\delta.$$

Now, move $y_1 \longmapsto y_1' = |y_1|^{-1}(|y_1|-d_1)y_1$ and observe that

$$|y-y_1| \leqslant 2\delta + \max(d-d_1,d_1+2\delta)$$

for all vertices $y \in X$ of K and

$$|y_1-y_2| \leqslant 2\delta + \max(d+2\delta-d_1,d_1+2\delta).$$

Therefore, if $4\delta \leqslant d_1 \leqslant d - 4\delta$, then the move $y_1 \longmapsto y_1'$ yields a homotopy of K in $P_d(X)$ bringing the vertices y_1 and y_2 within distance $\leqslant d$ in X, which allows one to apply the Case 1 shortening of K.

Notice that the shortening homotopy in Case 2 "consists" of three triangles: two take care of the move $y_1 \longmapsto y_2'$ and the third comes from the Case 1 homotopy. So the total number of triangles needed to contract K is $3N_1$ where N_1 is the number of edges in K. In fact, what we have just proved can be expressed more precisely in the following

1.7.C <u>**Lemma**</u>: *Let the circle* S^2 *be subdivided into* N_1 *segments and let* $f\colon S^1 \longrightarrow P_d(X)$ *be a simplical map. If* $d \geqslant 8\delta$, *then there exist a triangulation of the disk* D^2 *into* $N_2 \leqslant 3N_1$ *triangles and a simplical map* $D^2 \longrightarrow P_d(X)$ *which extends* f *from the boundary* $\partial D^2 = S^1$.

1.7.D <u>**Remark**</u>: The proof of 1.7.A also yields the following generalization of 1.7.A.

Let X be a δ-hyperbolic space which satisfies the following condition (*) *for some* $\epsilon > 0$ *and* $d > 0$,

(*) *for every* $x \in X$ *with* $|x| \geq (1/2)d$ *there exists a point* $x' \in X$, *such that* $|x'| \leq \min(|x|-\epsilon, |x|-2\delta)$ *and* $|x-x'| \leq d - 2\delta$.

Then the polyhedron $P_d(X)$ *is contractible.*

1.8 __Hyperbolic boundary ∂X of X__: A sequence $x_i \in X$, $i = 1,2,\ldots$, is called *convergent at infinity* if

$$(x_i.x_j) \longrightarrow \infty \text{ for } i,j \longrightarrow \infty.$$

Observe that this convergence is independent of the choice of the implied reference point in X. Furthermore, if X is hyperbolic then the equality

$$\lim_{i,j\to\infty} \inf(x_i.y_j) = \infty$$

is an equivalence relation on the set of sequences in X convergent at infinity.

__Definition__: The (hyperbolic) *boundary* ∂X of a hyperbolic space is the set of the equivalence classes of sequences in X convergent at infinity. If a sequence x_i is contained in the class $a \in \partial X$ we write $x_i \longrightarrow a$ for $i \longrightarrow \infty$. Thus we define a natural topology on $X \cup \partial X$, such that X is dense in $X \cup \partial X$. Observe that this topology is independent of the choice of the reference point in X. It follows that the isometry group of X acts on ∂X by *homeomorphisms* of ∂X. More generally, every isometric map between hyperbolic spaces, say $f: X \longrightarrow Y$, extends to a unique topological embedding $X \cup \partial X \longrightarrow Y \cup \partial Y$ sending ∂X to ∂Y. Recall that a map f is *isometric* if $|f(x_1)-f(x_2)| = |x_1-x_2|$ for all x_1 and x_2 in X.

1.8.A __Examples__:

(a) Let X be a tree with a simplical metric. Then the boundary ∂X is canonically (and obviously) homeomorphic to the space of isometric maps $f: \mathbb{R}_+ \longrightarrow X$ sending zero to the reference point in X, where the space of maps is given the point-wise convergence topology. It easily follows that the topological dimension of this boundary equals zero. If the tree X is locally finite, then the boundary ∂X (obviously) is compact.

(b) Let X_0 be an arbitrary metric space and let f(t) for $t \in (-\infty,+\infty)$ be a positive monotone increasing function such that $f(t+1) \geq \lambda f(t)$ for all $t \in \mathbb{R}$ and some fixed $\lambda > 1$. Consider the maximal metric $|\ |$ on $X = X_0 \times \mathbb{R}$ for which the embedding $\mathbb{R} \longrightarrow X$ given by $t \longmapsto (x_0,t)$ is isometric for all $x_0 \in X_0$ and

$$|(x_1,t)-(x_2,t)| \leq f(t)|x_1-x_2|_{X_0}$$

for all $t \in \mathbb{R}$ and $x_1,x_2 \in X_0$.

One can easily see that the space $(X,|\ |)$ is hyperbolic and the correspondence $x_0 \longmapsto (x_0,t_i)$ for $x_0 \in X_0$ and $t_i \longrightarrow +\infty$ defines an embedding of X_0 into an open subset in ∂X whose complement is a single point, say $a_0 \in \partial X$, such that $(x_0,t_j) \longrightarrow a_0$ for every fixed $x_0 \in X_0$ and for $t_j \longrightarrow -\infty$. Notice that for $X_0 = \mathbb{R}^n$ and $f(t) = e^t$ the space X is the ordinary hyperbolic space H^{n+1} and $\partial H^{n+1} = S^n = \mathbb{R}^n \cup \{\infty\}$.

(b') Let us replace $\mathbb{R}$ by $\mathbb{R}_+$ in the above construction. The resulting space $X \approx X_0 \times \mathbb{R}_+$ is not necessarily hyperbolic. In fact, it is hyperbolic if and only if X_0 is hyperbolic, as a simple argument shows. The boundary of such a hyperbolic X is canonically homeomorphic to X_0.

1.8.B Let us show that *the boundary* ∂X *of an arbitrary hyperbolic space* X *is metrizable*. In fact, let us define $|a-b| = |a-b|_{x_0}$ for $a,b \in \partial X$ to be the maximal metric satisfying the following condition:

if $x_i \longrightarrow a$ and $x_j \longrightarrow b$, and $\liminf_{i,j\to\infty} (x_i.x_j) \geq 2^k$ for some $k = 1,2,\dots$ then $|a-b| \leq k^{-1}$.

A simple argument (left to the reader) shows that $|a-b|$ is indeed a metric giving the right topology to ∂X and that a change of the reference point $x_0 \in X$ leads to an *equivalent* metric on ∂X. That is the identity map

$$(\partial X, |\ \ |_{x_0}) \leftrightarrow (\partial X, |\ \ |_{x_1})$$

is a *uniform* (i.e. uniformly continuous in both directions) homeomorphism for all x_0 and x_1 in X. Also notice that our metric on ∂X is *complete*. The proof of this is left to the reader.

<u>Remark</u>: If X is a complete Riemannian manifold or, more generally a geodesic space where all closed balls are compact, then the boundary ∂X is compact. Furthermore, if such an X is unbounded then ∂X is non-empty.

<u>Counterexample</u>: Let X be the union of the segments $[0,i]$, for $i = 1,2,\dots$, joined at zero. This X is unbounded geodesic hyperbolic but ∂X is empty.

§2. Hyperbolic metric groups

An abstract group Γ with a left invariant metric $|\ |$ on Γ is called a *metric group.* We always consider Γ as a pointed space where the preferred point is the identity element in Γ.

2.1 Word metrics: Let a subset $G \subset \Gamma$ generate Γ. The corresponding *word metric* $|\ |_G$ is the maximal metric on Γ satisfying $|g| = |g^{-1}| = 1$ for all $g \in G$. If Γ is finitely generated, then the expression "a word metric on Γ" always refers to a word metric for some *finite* generating subset G in Γ.

Observe that our definition agrees with the ordinary definition as $|\gamma_1 - \gamma_2|_G$ equals the length of the shortest word in g and g^{-1} for $g \in G$ representing $\gamma_2^{-1}\gamma_1 \in \Gamma$.

2.1.A Definition: A metric group Γ is called *hyperbolic* (δ-hyperbolic) if it is hyperbolic (δ-hyperbolic) as a metric space.

A finitely generated group Γ is called *word hyperbolic* if some word metric on Γ is hyperbolic.

We shall see later that the word hyperbolicity is independent of the choice of a finite generating set $G \subset \Gamma$ but the implied constant δ may depend on G.

2.2 Contractability of $P_d(\Gamma)$.

Theorem (I. Rips): *Let Γ be a word hyperbolic group. Then Γ admits a faithful discrete simplical action on some finite dimensional locally compact and contractible polyhedron P, such that P/Γ is compact.*

Proof: Take $P = P_d(\Gamma)$ (see 1.7) with the obvious action of Γ on P. Since Γ is finitely generated, P obviously satisfies all requirements with the possible exception of the contractibility.

Now, if Γ is δ-hyperbolic, then it satisfies (*) in 1.7.D for $\epsilon = 1$ and all $d \geqslant 4\delta + 1$ since the metric in question is a *word* metric. Hence, $P_d(\Gamma)$ is contractible for $d \geqslant 4\delta + 1$. Q.E.D.

Corollaries:

2.2.A *The group* Γ *is finitely presented.*

2.2.B *The group* Γ *contains at most finitely many conjugacy classes of torsion elements.*

2.2.C *The rational cohomology* $H^*(\Gamma,\mathbb{Q})$ *is finite dimensional. Furthermore, the* L_2*-cohomology (see* [C-G]*) of* Γ *is* Γ*-finite dimensional.*

2.3 **Isoperimetric inequalities for hyperbolic groups:** Let Γ be *presented* by $\{G \mid W\}$, where $G \subset \Gamma$ is a generating set and W is a subset in the free group $\mathfrak{F}(G)$ freely generated by G, such that W lies in the kernel $\mathfrak{N} \subset \mathfrak{F}(G)$ of the obvious epimorphism $\mathfrak{F}(G) \longrightarrow \Gamma$ and W *normally generates* $\mathfrak{N}$, that is $\mathfrak{N} = \mathfrak{N}(W)$ is the *minimal* normal subgroup in $\mathfrak{F}(G)$ containing W.

Recall that every $f \in \mathfrak{F}(G)$ is an irreducible word $g_1^{a_1} g_2^{a_2} \dots g_k^{a_k}$ for $g_i \in G$ and $a_i \in \mathbb{Z}$ and define the *length* L(f) of f by

$$L(f) = |a_1| + |a_2| + \dots + |a_k|.$$

Next define the *area* A(f) for all $f \in \mathfrak{N}(W)$ as the minimal number A, such that f can be written in the form

$$f = \rho_1 w_1^{b_1} \rho_1^{-1} \rho_2 w_2^{b_2} \rho_2^{-1} \dots \rho_m w_m^{b_m} \rho_m^{-1}$$

for some $\rho_i \in \mathfrak{F}(G)$ and $w_i \in W$, $i = 1,\dots,m$, where

$$L(\rho_1) + L(\rho_2) + \dots + L(\rho_m) \leqslant A$$

and

$$|b_1| L^2(w_1) + |b_2| L^2(w_2) + \dots + |b_m| L^2(w_m) = A.$$

2.3.A __Theorem:__ *Let* Γ *be* δ*-hyperbolic for (the word metric associated with) a generating subset* $G \subset \Gamma$ *and let* $W = W_d \subset \mathcal{F}(G)$, *for some* $d \geq 0$, *be the set of those irreducible words (in letters* $g \in G$*) of length* $\leq 3d$ *which are equal to the identity element in* Γ. *If* $d \geq 8(\delta+2)$, *then* $\Gamma = \mathcal{F}(G)/\mathcal{N}(W)$ *and*

$$A(f) \leq 27\, d^2 L(f)$$

for all $f \in \mathcal{N}(W)$.

__Proof:__ If Γ has no torsion of order ≤ 3, then the (simply connected!) complex $P_d^2(\Gamma)$ is acted on *freely* by Γ and the theorem follows from 1.7.A and 1.7.C since every triangle in $P_d^2(\Gamma)$ defines a word in W_d. Notice that the inequality $d \geq 8\delta$ suffices in this case.

Now, in the general case we recall the Cayley complex (graph) X of (Γ, G) (if $G \cap G^{-1} = \emptyset$ then $X = P_1^1(\Gamma)$) and then apply 1.7.C to some simply connected subcomplex P in $P_d^2(X)$ in which the action of Γ is free. We leave to the reader the (obvious) details of the argument.

2.3.B __Corollary:__ *If* Γ *is word hyperbolic (for some finite generating subset in* Γ*) then for every finite presentation* $\{G | W\}$ *of* Γ *there exists a constant* $C \geq 0$, *such that*

$$A(f) \leq CL(f)$$

for all $f \in \mathcal{N}(W)$. *In particular, the word problem is solvable in* Γ.

2.3.C __Remark:__ We shall see in 7.4 that the conjugacy problem also is solvable in Γ.

2.3.D Let us state a converse to 2.3.B (see 6.8 for the proof).

Theorem: *Let* Γ *be presented by* $\{G|W\}$, *such that all* $f \in \mathcal{R}(W)$ *satisfy*

$$A(f) \leqslant CL(f)$$

for some constant $C \geqslant 0$ *(depending on* $\{G|W\}$ *but not on* f). *Then* Γ *is hyperbolic for the word metric associated to* G.

This theorem immediately implies with 2.3.B the following

2.3.E Corollary: *If* Γ *is hyperbolic for some finite generating subset* $G_0 \subset \Gamma$ *then* Γ *is hyperbolic for every finite generating subset* $G \subset \Gamma$.

2.3.F Let us give a quantitative version of 2.3.D for *triangular* presentations $\{G|W\}$ where every word $w \in W$ has length $\leqslant 3$. Observe that every presentation can be "triangulated" in an obvious way, which allows us, in principle, to apply the following theorem to non-triangular presentations.

Let $\{G|W\}$ *be triangular and suppose there exists a number* $N \geqslant N_0 = 10000$, *such that every* $f \in \mathcal{R}(W)$ *with the area in the interval*

$$N \leqslant A(f) \leqslant 100N$$

satisfies

$$L(f) \geqslant 1000\sqrt{A(f)}.$$

Then the group Γ *with the word metric associated to* G *is* δ*-hyperbolic for* $\delta = 1000\sqrt{N}$. *Furthermore all*

$f \in \mathcal{R}(W)$ *satisfy*

$$A(f) \leqslant CL(f)$$

for $C = N_0\delta^2$. (See 6.8 for the proof.)

2.3.G **Remark:** Given a finite presentation $\{G \mid W\}$ there is an obvious algorithm for writing down all $f \in \mathcal{R}(W)$ with $A_1 \leqslant A(f) \leqslant A_2$ and $L_1 \leqslant L(f) \leqslant L_2$ for given integers A_1, A_2, L_1 and L_2. Thus one obtains an effective (but not very practical) criterion for the word hyperbolicity of finitely presented groups Γ.

2.4 **Convex manifolds:** Consider a Riemannian manifold X with a convex (possibly empty) boundary. For example, X may be a *convex* subset in a complete manifold $X' \supset X$ without boundary, where a subset X in a Riemannian manifold is called convex if the intersection of X with every geodesic segment is connected. Recall that the distance between two points x and y in X is defined as the infimum of lengths of curves in X between x and y.

Let X be simply connected and complete as a metric space and recall the following well known (and easy to prove) geometric fact.

2.4.A **Theorem:** *The sectional curvature* K(X) *satisfies* $K(X) \leqslant 0$ *if and only if every smooth closed curve* F *in* X *bounds a disk* D *in* X *satisfying*

$$\text{area } D \leqslant (4\pi)^{-1}(\text{length } F)^2 \tag{4}$$

Furthermore, if the curvature K(X) *is strictly negative,* $K(X) \leqslant -\epsilon^2 < 0$, *then every curve* F *bounds a disk* D, *such that*

$$\text{area } D \leqslant \epsilon^{-1} \text{ length } F. \tag{5}$$

Remark: Notice that the inequality (4) is sharp as round disks D in $\mathbb{R}^2$ have area $D = (4\pi)^{-1}(\text{length } \partial D)^2$.

2.4.B **Definition**: Call X *convex* (see [Gr1] for a more general definition) if it is complete simply connected with a convex boundary and $K(X) \leqslant 0$.

It is well known (and easy to prove) that the strict inequality $K(X) \leqslant -\epsilon^2 < 0$ implies the hyperbolicity of X. In fact, a convex space X is (easily seen to be) hyperbolic if every smooth closed curve F in X of length $\ell \geqslant \ell_0$ for some constant $\ell_0 = \ell_0(X) > 0$ satisfies either of the following two conditions.

(1) F bounds a disk D, such that area $D \leqslant C_0$ length F for some $C_0 = C_0(X) > 0$.

(2) F bounds a disk D, such that

$$\text{area } D \leqslant [(4\pi)^{-1} - \epsilon_0](\text{length } F)^2$$

for some $\epsilon_0 = \epsilon_0(X) > 0$.

2.4.C **Remark**: The δ-inequality (see 1.1) for $K < -\epsilon^2 < 0$ is a trivial consequence of the CAT-*inequality* (comparison inequality of Alexandrov–Toponogov). This inequality applies to a pointed metric space (X, x_0) and triples of points x, y and z in X where z lies *between* x and y. That is

$$|z-x| + |z-y| = |x-y|.$$

CAT relates distances between the points in X to the distances in the *model* space (X', x_0'), which is a complete simply connected surface with a *constant* curvature $\varkappa$ for a given $\varkappa \in \mathbb{R}$. In other words, this X' is a sphere S^2, the plane $\mathbb{R}^2$ or the hyperbolic plane H^2. Points x', y' and z' in X' are called *comparison* points (for x, y, z) if

$$|x'| = |x|,\ |y'| = |y|,\ |x'-y'| = |x-y|$$

and

$$|x'-z'| = |x-z|, \ |y'-z'| = |y-z|.$$

The CAT($\varkappa$)-inequality claims, by definition, that the comparison points exist (for given x, y and z) and $|z'| \geqslant |z|$.

One of the basic results in Riemannian geometry is the following

2.4.D **Theorem** **(Riemann-Cartan-Alexandrov-Toponogov):** *A Riemannian manifold X with convex boundary has* $K(X) \leqslant \varkappa$ *if and only if each point* $x_0 \in X$ *admits a neighborhood* $U_0 \subset X$ *of* x_0 *where every three points satisfy* CAT($\varkappa$). *Furthermore, every three points in a convex manifold with* $K(X) \leqslant \varkappa \leqslant 0$ *satisfy* CAT($\varkappa$).

2.4.E **Remark:** The CAT($\varkappa$)-inequality for $\varkappa < 0$ easily implies the following sharpening of the δ-inequality.

Let X be a geodesic space satisfying CAT($\varkappa$) *for some* $\varkappa < 0$. (That is every three points satisfy CAT for every reference point in X.) *Take three points* x, y *and* z, *such that*

$$(x.y) \leqslant (x.z) \leqslant (y.z),$$

and let $d = (x.y) - (y.z)$. *Then*

$$(**) \qquad (x.y) \geqslant (x.z) - \delta$$

for some function $\delta = \delta(\varkappa, d) \geqslant 0$ *satisfying* $\delta \leqslant d \exp((-d/4)\sqrt{|\varkappa|})$.

2.5 **Isometry group:** Let Γ be a subgroup of the isometry group Is X. Then for every bounded open subset $U \subset X$ there obviously exists a unique left invariant metric on Γ, such that $|\gamma| = \sup_{x \in U} |x-\gamma(x)|_X$ for all $\gamma \in \Gamma$. If X is hyperbolic, then an obvious consideration shows this metric on Γ also is hyperbolic. In such a case the *orbit map* $\Gamma \longrightarrow X$ defined by $x_0 \longmapsto \gamma x_0$ for a given point $x_0 \in X$ continuously extends to an embedding between the hyperbolic boundaries,

$$\partial\Gamma \longrightarrow \partial X,$$

and the image $\partial\Gamma \subset \partial X$ is called the *limit set of* Γ.

If X is convex then the hyperbolic boundary ∂X of X is naturally homeomorphic to the space of *geodesic rays* in X issuing from a fixed point $x_0 \in X$, where a ray in X by definition is an isometric embedding $\mathbb{R}_+ \longrightarrow X$. This description of ∂X provides an embedding of ∂X into the tangent unit sphere $S^{n-1}_{x_0}$, $n = \dim X$, where each ray r is mapped to the tangent vector to r at x_0. The map $\partial X \longrightarrow S^{n-1}_{x_0}$ is *onto* if and only if the ordinary boundary of X is empty as an obvious consideration shows.

2.6 **Cocompact groups:** A group Γ acting on X is called *cocompact* if the quotient space X/Γ is compact. Every discrete cocompact isometry group obviously is finitely presented and (by an easy argument) Γ *is word hyperbolic if and only if* X *is hyperbolic*. For example, the fundamental group Γ of a *compact* Riemannian manifold V is word hyperbolic if and only if the universal covering X of V is hyperbolic for the *induced Riemannian metric* in X. Observe that this metric is, in fact, the maximal metric on X for which the covering map $X \longrightarrow V$ is locally isometric.

Let us indicate two unsolved problems.

(A) Let Γ be word hyperbolic. Does there exist a convex manifold X which admits a discrete cocompact isometry group isomorphic to Γ?

(B) Let Γ be a discrete cocompact isometry group of a convex manifold X, such that every Abelian subgroup in Γ is virtually cyclic (i.e. a finite extension of a cyclic group). Is Γ hyperbolic? (One can easily show that if Γ is not hyperbolic then there exists an isometric map $\mathbb{R}^2 \longrightarrow X$. But the question is if there is an isometric map $\mathbb{R}^2 \longrightarrow X$ for which the image of $\mathbb{R}^2$ in X/Γ is compact, compare 4.7.)

2.7 __Symmetric spaces__: Let X be a symmetric space of non-compact type without Euclidean factors (i.e. X is not isometric to $X_0 \times \mathbb{R}^k$ for $k > 0$). Then the isometry group Is X is a semisimple Lie group without compact factors. One knows that Is X contains a torsion free cocompact discrete subgroup Γ and every such Γ is residually finite. If rank X = 1 (i.e. there is no isometric map $\mathbb{R}^2 \longrightarrow X$) then Γ is word hyperbolic. Thus, by taking $X = O(n,1)/O(n)$, one obtains for every $n = 2,3,\ldots$, a compact aspherical manifold V, namely $V = X/\Gamma$, whose fundamental group Γ is hyperbolic and residually finite. Notice that the symmetric spaces $U(n,1)/U(n)$ and $Sp(n,1)/Sp(n)$ have dimensions 2n and 4n respectively. The group $Sp(n,1)$ has an additional remarkable feature, namely *Kazdan's T-property* (see 5.6) which is inherited by all lattices Γ in $Sp(n,1)$ and has a profound effect on algebraic properties of Γ. Notice that no *known* purely algebraic construction delivers infinite groups Γ with the T-property.

If X has rank $\geqslant 2$ then no cocompact Γ is hyperbolic. However, there are many word hyperbolic groups Γ operating on X, such as hyperbolic *reflection groups* Γ (see 4.6) which isometrically act on $X = O(p,q)/O(p) \times O(q)$. One usually proves the hyperbolicity of a group Γ operating on X by constructing a hyperbolic (e.g. convex) Γ-invariant submanifold $X_0 \subset X$, such that X_0/Γ is compact.

§3. The action of Γ on $\partial\Gamma$

Consider a hyperbolic (metric) group Γ and look at the hyperbolic boundary $\partial\Gamma$ of Γ. The simplest example is that of a free group G_k on k generators which isometrically acts on the regular infinite tree X with 2k edges of unit length at every vertex. It is obvious that $\partial F_k = \partial X$ and that ∂X is (homeomorphic to) a Cantor set provided $k \geqslant 2$. If $k = 1$, then the boundary consists of two isolated points.

3.1 **Elementary and non-elementary groups:** Let us classify hyperbolic groups according to the cardinality of the boundary.

Case 1: The boundary $\partial\Gamma$ is empty. In this case Γ is bounded as a metric space (see 8.1, 8.2).

Case 2: The boundary $\partial\Gamma$ consists of a single point. Such groups Γ are called *parabolic*. Observe that every infinite group Γ admits a parabolic metric. Namely, start with any unbounded metric $|\ |$ and then take the (hyperbolic!) metric $\log(1 + |\ |)$ (see 1.2). It is easy to see that $(\Gamma, \log(1 + |\ |))$ is parabolic.

Case 3: The boundary of Γ consists of two points. Then (see 8.1) Γ contains an infinite cyclic subgroup, say $\Gamma_0 \subset \Gamma$ which is *cobounded* in Γ. In particular, if Γ is word hyperbolic then Γ is a finite extension of Γ_0.

Recall that a subset $X_0 \subset X$ is cobounded if there is constant $d = d(X_0)$ such that $\mathrm{dist}(x, X_0) \leqslant d$ for all $x \in X$.

The groups in these three cases are called *elementary* hyperbolic groups.

Case 4: The boundary $\partial\Gamma$ contains at least three points. Then (see 8.2) the boundary is infinite uncountable and Γ contains an isomorphic copy of the free semigroup F_2^+ with two generators. In fact, there are two elements α_0 and α_1 in Γ, such that the map of the set of finite diadic sequences into Γ given by

$01101\ldots \longmapsto \alpha_0\alpha_1\alpha_1\alpha_0\alpha_1\ldots$ is injective and extends to an injective continuous map of the Cantor set $\{0,1\}^\infty$ to ∂X. (See 8.2.)

__Case 4'__: Call Γ *quasiparabolic* if there exists a point $p \in \partial X$ which is fixed under Γ. If Γ is *not* quasiparabolic as well as non-elementary, then the above α_0 and α_1 can be chosen freely independent in Γ and such that the monomorphism of the free group $F_2 = F(\alpha_0,\alpha_1)$ into Γ extends to a topological embedding $F_2 \cup \partial F_2 \longrightarrow \Gamma \cup \partial\Gamma$, where F_2 is given the usual word metric. (See 8.2.)

3.1.A __Remarks__:

(a) If Γ is *word hyperbolic* then the boundary $\partial\Gamma$ is a *compact finite dimensional* space (see 7.6.) and isomorphisms $\Gamma_1 \longrightarrow \Gamma_2$ between such groups continuously extend to homeomorphisms between their boundaries (see 7.2.H.). Thus every topological invariant of $\partial\Gamma$ is, in fact, an invariant of Γ. On the other hand a non-bijective homomorphism $\alpha\colon \Gamma_1 \longrightarrow \Gamma_2$ does not necessarily extend to the boundary even if we assume α injective. However, there always exists a natural discontinuous map, called the *Furstenberg map*, sending a "large portion" of $\partial\Gamma_1$ to $\partial\Gamma_2$ (see [Fu]).

(b) If Γ is a subgroup in a word hyperbolic group, then it can not be parabolic or quasiparabolic. (See 8.1.) It follows that every such Γ contains F_2 unless Γ is a finite extension of a cyclic group.

3.2 __Group__ Γ __with__ dim $\partial\Gamma = 0$: Take two metric groups Γ_1 and Γ_2, let $\Gamma_i' \subset \Gamma_i$, $i = 1,2$, be subgroups and let $\Gamma_1' \leftrightarrow \Gamma_2'$ be an isometric isomorphism. Then we consider the amalgamated product $\Gamma = \Gamma_1 \underset{\Gamma'}{*} \Gamma_2$ for $\Gamma' = \Gamma_1' = \Gamma_2'$ with the maximal metric in Γ for which the inclusion homomorphisms $\Gamma_i \longrightarrow \Gamma$, $i = 1,2$, are *short*, where a map $f\colon X \longrightarrow Y$ is called *short* if $|f(x_1)-f(x_2)| \leq |x_1-x_2|$ for all x_1 and x_2 in X. If the group Γ_1 and Γ_2 are hyperbolic and Γ' is bounded, then it is easy to see that Γ is hyperbolic and the topological dimension of $\partial\Gamma$

satisfies

$$\dim \partial\Gamma = \max(\dim \partial\Gamma_1, \dim \partial\Gamma_2).$$

Furthermore, the boundary $\partial\Gamma$ is infinite (and hence $\partial\Gamma$ is non-elementary), provided $\#(\Gamma_1/\Gamma_1') \geq 2$ and $\#(\Gamma_2/\Gamma_2') \geq 3$. For example, if the groups Γ_1 and Γ_2 are elementary (e.g. bounded) then Γ has $\dim \partial\Gamma = 0$ and Γ is non-elementary if $\#(\Gamma_i/\Gamma_i') > 2$ for $i = 1,2$.

__Remark:__ The above discussion (obviously) extends to amalgamated products of finite families of groups and to HNN-extensions.

3.2.A __Proposition:__ *A word hyperbolic group* Γ *has* $\dim \partial\Gamma = 0$ *if and only if it decomposes into amalgamated products and* HNN-*extensions of virtually cyclic (i.e. finite extensions of cyclic) groups over finite subgroups.*

This is an immediate corollary of Stalling's theorem on groups with infinitely many ends.

__Remark:__ There is a close link between the space of ends $E\Gamma$ and the hyperbolic boundary $\partial\Gamma$. Namely every metric $|\ |$ on Γ can be transformed in a natural way to another metric, say $|\ |'$ on Γ which makes $(\Gamma, |\ |')$ hyperbolic and such that $(\partial\Gamma, |\ |') = E\Gamma$ in the case where $|\ |$ is a word metric for a finite generating subset in Γ.

3.3 __Groups with__ $\dim \partial\Gamma = 1$: The simplest example is the fundamental group Γ of a closed surface of genus g. This group is word hyperbolic for $g \geq 2$ and $\partial\Gamma$ is homeomorphic to S^1. Conversely, if a word hyperbolic group Γ has $\partial\Gamma \approx S^1$ then Γ is a finite extension of a surface group as a simple argument shows (compare [GM]).

Most small cancellation groups Γ are word hyperbolic and these have $\dim \partial\Gamma \leq 1$. For example, 1/6-groups are word hyperbolic (see 4.7). But 1/5-groups only are semihyperbolic and instead of the (linear) inequality of 2.3.A they satisfy the *quadratic* isoperimetric inequality $A(f) \leq C(L(f))^2$, where the constant $C \geq 0$ depends on the presentation.

Let Γ_1 and Γ_2 be word hyperbolic without torsion and let $\Gamma_i' \subset \Gamma_i$, $i = 1,2$ be infinite *maximal* cyclic subgroups (i.e. $\gamma^p \in \Gamma_i \Rightarrow \gamma \in \Gamma_i$ for all $p = 1,2,\ldots$, and $i = 1,2$). Then (by an easy argument) the amalgamated product $\Gamma = \Gamma_1 \ast_{\Gamma'} \Gamma_2$ is word hyperbolic and $\dim \partial\Gamma \leq m = \max(1, \dim \partial\Gamma_1, \dim \partial\Gamma_2)$. In fact, $\dim \partial\Gamma = m$, unless Γ_1 and Γ_2 are free products of the form $\Gamma_i = \Gamma_i' \ast \Gamma_i''$, $i = 1,2$.

Start with an arbitrary finite presentation $\{g_1,\ldots,g_p \mid w_1,\ldots,w_q\}$ (where w_j, $j = 1,\ldots,q$ are some cyclically irreducible words in g_i, $i = 1,\ldots,p$ and the implied relations are $w_j = 1$, $j = 1,\ldots,q$). Then introduce new generators g_j' and g_j'' for $j = 1,\ldots,q$ (which correspond to w_j) and define Γ' by the presentation

$$(*) \qquad \{g_i, g_j', g_j'' \mid w_j = [g_j', g_j'']\}$$

for $i = 1,\ldots,p$ and $j = 1,\ldots,q$, where $[g',g'']$ denotes the commutator $g'g''(g')^{-1}(g'')^{-1}$. It is not hard to see that every such Γ' is word hyperbolic and $\dim \partial\Gamma' \leq 1$. The simplest such presentation is $\{g, g', g'' \mid g^k = [g',g'']\}$ for some $k \geq 1$ and the corresponding Γ' has $\dim \partial\Gamma' = 1$ for $k \geq 2$.

A geometric explanation: Let P be a compact connected 2-dimensional cell complex with p 2-cells. Each 2-cell is attached to the 1-skeleton of P by a continuous map of the boundary of the disk, say by $\sigma_j\colon S^1 = \partial D^2 \longrightarrow P^1$, $j = 1,\ldots,p$. Let us replace the j-th 2-cell by a compact connected surface S_j with $\partial S_j = S^1$, attached to P^1 by σ_j, for all $j = 1,\ldots,p$. If the maps σ_j are not contractible (e.g. P is a *simplical* complex) and the surfaces S_j have negative Euler characteristic, then the resulting space, say P', is aspherical and the fundamental group $\Gamma' = \pi_1(P')$ is word hyperbolic with $\dim \partial\Gamma' \leq 1$.

(One recaptures the previous Γ' with punctured tori for S_j.) The proof is left to the reader (compare 4.2.D.).

3.4 **Hyperbolization of polyhedra:** The above geometric construction can be generalized in several ways to n-dimensional polyhedra P for all $n \geq 2$. Here is the simplest version of such a "hyperbolization" of an n-dimensional polyhedron P. First we slightly modify the previous construction for $n = 2$. Start with the 2-cube $\square^2 = [-1,+1] \times [-1,+1]$ and let

$$\square_h^2 = (\partial\square^2 \times [-1,+1])/\mathbb{Z}_2,$$

where the involution (i.e. the non-trivial element in $\mathbb{Z}_2$) acts by

$$(t_1,t_2,t_3) \longmapsto (-t_1,-t_2,-t_3).$$

Observe that $\square_h^2$ is homeomorphic to the Möbius band and that the boundary $\partial\square_h^2$ is *canonically* homeomorphic to $\partial\square^2$.

Now let P be a 2-dimensional *cubical* polyhedron. That is P is a union of 2-cubes, where every two cubes meet (if they meet at all) across a common face (of dimension zero or one). Then we remove from P the interior of every cube $\square^2$ and attach instead the "hyperbolized" cube $\square_h^2$ using the canonical homeomorphisms $\partial\square_h^2 = \partial\square^2$. Notice that the fundamental group $\Gamma = \pi_1(P_h)$ of the resulting (hyperbolized) polyhedron P_h can be presented by

$$(**) \qquad \{g_i, g_i' \mid (w_j)^2 = g_j'\}.$$

Not every presentation (**) gives a word hyperbolic group but our groups Γ (related to cubical polyhedra P) are easily seen to be hyperbolic with $\dim \partial\Gamma \leq 1$.

Next take the 3-cube $\square^3 = \square^2 \times [-1,+1]$ and observe that the boundary $\partial\square^3$ is a cubical 2-polyhedron (homeomorphic to S^2) and that the involution $(t_1,t_2,t_3) \longmapsto (-t_1,-t_2,-t_3)$ on $\partial\square^3$ induces in an obvious way an involution without fixed points on the "hyperbolized" boundary $(\partial\square^3)_h$ (which is homeomorphic to the connected sum of six projective planes). This involution and $t \longmapsto -t$ on $[-1,+1]$ induce an involution on $(\partial\square^3)_h \times [-1,+1]$ which is used to define

$$\square_h^3 = ((\partial\square^3)_h \times [-1,+1])/\mathbf{Z}_2.$$

Notice that $\square_h^3$ is homeomorphic to the total space of a (non-trivial) bundle with the fiber $[-1,+1]$ over the connected sum of four projective plane and the boundary $\partial\square_h^3$ is *canonically* homeomorphic to $(\partial\square^3)_h$.

Now, let P be an arbitrary cubical polyhedron of dimension 3. First we replace the two skeleton P^2 of P by the hyperbolization P_h^2 and then we attach $\square_h^3$ to the hyperbolized boundary $(\partial\square^3)_h \subset P_h^2$ of every 3-cube $\square^3$ in P using the above canonical homeomorphism. Thus we obtain a certain 3-dimensional polyhedron P_h.

This construction obviously generalizes by induction on $n = \dim P$ to all cubical polyhedra P. Namely each $\square^k$-cube in P for $k = 2,\dots,n$ is replaced as earlier by

$$\square_h^k \overset{\mathrm{def}}{=} ((\partial\square^k)_h \times [-1,+1])/\mathbf{Z}_2.$$

This gives us our hyperbolization $P \to P_h$ for all cubical polyhedra.

It is not hard to see that the polyhedron P_h is aspherical and that it admits a (natural) continuous map $P_h \longrightarrow P$, such that the induced cohomology homomorphism

$$H^*(P;\mathbf{Z}_2) \longrightarrow H^*(P_h;\mathbf{Z}_2)$$

is injective. In particular, if P is a closed manifold then P_h also is a closed manifold which is aspherical and which comes with a map $P_h \longrightarrow P$ of degree $= 1 (\mathrm{mod}\ 2)$.

We shall see in 4.3 that *the fundamental group* $\Gamma = \pi_1(P_h)$ *is word hyperbolic for every cubical polyhedron* P and that

$$\dim \partial\Gamma \leqslant \dim P - 1.$$

In fact,

$$\dim \Gamma = \dim P - 1,$$

provided $H^n(P;\mathbb{Z}_2) \neq 0$ for $n = \dim P$. It will become clear later that $\partial\Gamma$ is homeomorphic to S^{n-1} if P is an n-dimensional *manifold*.

3.4.A An obvious drawback in the above construction is the non-orientability of $\square_h^k$ which is due to the fact that the involution $x \leftrightarrow -x$ on the odd dimensional (hyperbolized) sphere is orientation preserving. Let us describe a (slightly less canonical) orientable hyperbolization. Start with a closed oriented manifold S divided into two submanifolds $S = S_+ \cup S_-$ and let $I: S \longrightarrow S$ be an involution permuting S_+ and S_- and keeping fixed $\partial S_+ = \partial S_- = S_+ \cap S_-$. Denote by $\Delta_h(S)$ the cylinder $S \times [-1,+1]$ with the points $s_+ \times (-1)$ and $s_+ \times (+1)$ identified for all $s_+ \in S_+$. Observe that $\Delta_h = \Delta_h(S)$ in a manifold whose boundary is identical to S.

Now, assume by induction, we have a hyperbolization procedure for all (n-1)-dimensional simplical polyhedra. Take an n-dimensional polyhedron P and let P_h^{n-1} be the hyperbolization of the first baricentric subdivision of the (n-1)-skeleton of P. Then the hyperbolized boundary S_h of every n-simplex Δ^n in P has an involution I as above and we hyperbolize P by attaching $\Delta_h(S_h)$ to S_h for all $\Delta^n \subset P$.

It is easy to see that this hyperbolization P_h is aspherical and that the obvious map $P_h \longrightarrow P$ is injective on $H^*(P)$. But the fundamental group $\pi_1(P_h)$ is not hyperbolic for $n \geqslant 3$, as it contains isomorphic copies of $\mathbb{Z} \oplus \mathbb{Z}$. Yet this group is *semihyperbolic* as P_h admits a (singular) metric of non-positive curvature (see Section 4).

If P is a closed (oriented) manifold then, clearly, P_h also is a closed (oriented) manifold. Moreover, a simple application of (the existence of) local combinatorial formula for Pontryagin classes shows that the monomorphism $H^*(P;\mathbb{Q}) \longrightarrow H^*(P_h;\mathbb{Q})$ sends Pontryagin classes of P to those of P_h. In particular, P_h has the same Pontryagin numbers as P. (This also is clear with 3.4.C.)

3.4.B Now let us construct a hyperbolization which is orientable, preserves Pontryagin classes and has word hyperbolic fundamental group. The inductive step $n-1 \Rightarrow n$ is as follows. First we take a sufficiently fine subdivision P' of P, hyperbolize the (n-1)-skeleton of the subdivision and only then fill in the hyperbolized boundaries of *not* subdivided n-simplices Δ in P. We assume as earlier that the hyperbolization of the (n-1)-skeleton of the subdivided Δ agrees with symmetries of Δ. Then we have an even number of n-simplices of P' in Δ, say Δ_i' and Δ_i'' which (for fine enough subdivision) can be assumed pairwise disjoint and oppositely oriented in Δ. Then we attach to each such pair the cylinder $(\partial\Delta_i')_h \times [-1,1]$ where $(\partial\Delta_i')_h \times (-1) = (\partial\Delta_i')_h$ and $(\partial\Delta_i')_h \times 1$ is identified with $(\partial\Delta_i'')_h$.

3.4.C <u>Relative</u> <u>hyperbolization</u>: Let P_0 be a subpolyhedron in P, take the first baricentric subdivision P' of P and apply a hyperbolization to every simplex Δ^k in P' for $k \geqslant 2$ which is not contained in P_0. Call the resulting polyhedron $H = H(P,P_0)$ and observe that

(1) The inclusion homomorphism $\pi_1(P_0) \longrightarrow \pi_1(H)$ is injective.

(2) If P_0 is aspherical then H also is aspherical. In particular, the fundamental group of every aspherical polyhedron embeds into that of some closed aspherical manifold.

(3) If the group $\pi_1(P_0)$ is semihyperbolic then so is $\pi_1(H)$. Moreover if we use a strict hyperbolization (i.e. the one in 3.4 or 3.4.B) then $\pi_1(H)$ is hyperbolic, provided $\pi_1(P_0)$ is such.

(4) If P_0 is a closed manifold and $P = P_0 \times [0,1]$, then $H(P,P_0 \times 0)$ is a cobordism between P_0 and the hyperbolization P_{0h} of

$P_0 = P_0 \times 1$. In particular the hyperbolizations in 3.4, 3.4.A and 3.4.B preserve Stiefel-Whitney numbers of manifolds, as well as (in cases 3.4.A and B) Pontryagin numbers.

(5) If P_0 is a closed aspherical manifold which is the boundary of a manifold $P \supset P_0$ then H is an aspherical manifold with the same boundary $\partial H = \partial P = P_0$.

3.4.D <u>Reflection groups</u>: A group Γ is called a *reflection group* if it admits a presentation of the form

$$\{g_1,\ldots,g_p \mid g_i^2 = 1,\ (g_i g_j)^{d_{ij}} = 1\}$$

for $i,j = 1,\ldots,p$, where d_{ij} are given numbers $\in \{0,1,2,\ldots\}$ and where the relation $g^0 = 1$ by definition means no relation at all. Every reflection group Γ can be realized by a discrete subgroup in the orthogonal group $O(m,n)$ for some m and n depending on Γ (see [Bour]). It follows that every reflection group is residually finite and virtually torsion free (i.e. it contains a subgroup of finite index without torsion). The geometric and topological significance of reflection groups was pointed out by W. Thurston and then clarified by Davis (see [Dav]) who found, for example, an aspherical manifold V of a given dimension $n \geqslant 4$ whose fundamental group Γ is a subgroup of finite index in a reflection group and such that Γ is not simply connected at infinity.

Davis' groups are semihyperbolic as they act on spaces with $K \leqslant 0$ (see 4.6) and some reflection groups are hyperbolic. It is claimed in Section 4.E of [Gr5] that a hyperbolization similar to the one described in the previous section can be achieved with *reflection* groups. Then L. Siebenmann pointed out to me a difficulty in the construction of hyperbolic reflection groups. Now 4.E of [Gr5] should be regarded as a conjecture. This conjecture is justified by some interesting examples of hyperbolic reflection groups found (but yet unpublished) by Siebenmann and Ancel.

§4. Singular spaces and orbispaces with $K \leq 0$

For a geodesic space X the inequality $K(X) \leq \varkappa$ is *defined* (according to Alexandrov, see [Al]) via the CAT-inequality: Every $x_0 \in X$ admits a neighborhood where every three points satisfy CAT($\varkappa$) (compare 2.4). If $\varkappa \leq 0$ then the (local!) condition $K \leq \varkappa$ yields the following (well known and easy to prove) global

Theorem (Cartan-Hadamard-Alexandrov): *If* $K(X) \leq \varkappa \leq 0$ *and* X *is simply connected, then* X *is contractible and satisfies* CAT($\varkappa$) *(for all quadruples of points* x, y, z *and* x_0 *in* X).

Corollary: *Let* f_i: $[0,a_i] \rightarrow X$ *for* $i = 1,2$ *be locally isometric maps. Then the (distance) function* $|f_1(t_1)-f_2(t_2)|$ *is convex on* $[0,a_1] \times [0,a_2]$. *In particular the maps* f_i *are globally isometric.*

From this one easily derives the following

4.1 Hyperbolicity criteria:

(A) *If* X *is simply connected and* $K(X) \leq \varkappa < 0$ *then* X *is hyperbolic. Moreover* X *satisfies the sharpened* δ*-inequality* (**) *of* 2.4.

(B) *Let* X *be simply connected with* $K(X) \leq 0$ *and let the (full) isometry group be cocompact on* X. *Then* X *is hyperbolic unless it receives an isometric map* $\mathbb{R}^2 \rightarrow X$ (see [B-G-S]).

4.2 Polyhedra with $K \leq \varkappa$: Let X be a finite dimensional simlical complex, such that every k-simplex Δ in X is isometric to a simplex Δ^k in the *model space* $(M^k, \varkappa_0)$ which is the complete simply connected manifold of constant curvature $\varkappa_0$. A simplex

$\Delta^k \subset M^k$ by definition is a compact subset which is the intersection of (k+1) half-spaces in M^k in general position. Call such an X an (M,x_0)-*simplical* space and observe that the *link* $L_\Delta \subset X$ of every k-simplex $\Delta \subset X$ has a natural (M,1)-structure. Recall that L_Δ is a (k+1)-codimensional (sub)complex whose simplices correspond to (solid angles along Δ of) the simplices $\Delta' \supset \Delta$ in X. The following (well known and easy to prove) criterion reveals the combinatorial meaning of the curvature.

4.2.A *An* (M,x_0)-*simplical space* X *has* $K(X) \leqslant x$ *if and only if* $x \geqslant x_0$ *and* L_Δ *satisfies* CAT(1) *for all* $\Delta \subset X$ (*compare* [Al]).

This criterion is especially useful in conjunction with another simple (and well known) fact.

4.2.B *An* (M,x_0)-*simplical space* L *satisfies* CAT(1) *if and only if* $K(L) \leqslant 1$ *and one of the following conditions is satisfied.*

(1) *Every two points* ℓ_1 *and* ℓ_2 *in* L *with* $|\ell_1 - \ell_2| < \pi$ *can be joined by at most one segment in* L.

(2) *If the circle* S_ℓ *with the geodesic metric of length* ℓ *admits an isometric map* $S_\ell \longrightarrow L$, *then* $\ell \geqslant 2\pi$.

4.2.C Remark: It is often convenient to use a decomposition of X not into simplices but into larger subspaces, called *blocks* $\square \subset X$ which are subpolyhedra for some (M,x_0)-simplical structure on X (e.g. the building cubes of a cubical polyhedron) and which satisfy the following five conditions

(1) Each $\square$ is assigned a subpolyhedron $\partial\square \subset \square$ called the boundary of $\square$ which is the union of some blocks $\square' \subset \square$ with $\dim \square' < \dim \square$.

(2) The complement $\Box \setminus \partial\Box$ is locally isometric to (M^k, x_0) for $k = \dim \Box$.

(3) The intersection of every two blocks is a union of blocks.

(4) If $\Box_1 \subset \Box_2$ and $\Box_1 \neq \Box_2$ then $\Box_1 \subset \partial\Box_2$ for every two blocks $\Box_1$ and $\Box_2$ in X.

(5) The links L_Δ and $L_{\Delta'}$ of every two k-dimensional simplices Δ and Δ' in $\Box$ are isometric for each k-dimensional block $\Box$. The isometry class of L_Δ then is denoted by $L_\Box$.

Given such a decomposition of X into blocks one can forget the original simplical structure and apply 4.2.A and 4.2.B directly to the (M,k_0)-*block space* X.

4.2.D __Two-dimensional polyhedra__: The condition $K(X) \leqslant \chi$ is easy to verify if $\dim X = 2$. In this case the link of every vertex $x_0 \in X$ is a 1-polyhedron (graph) whose every edge is given length equal the angle of the corresponding triangle in X. Criterion 4.2.B now requires every simple closed curve in the link L_{x_0} to have length $\geqslant 2\pi$.

__Examples__:

(a) Let every simple closed curve in L_{x_0} contain at least $\ell \geqslant 6$ edges. Then the standard simplical metric in X, where all 2-simplices are equilateral Euclidean triangles (with the angles $2\pi/6$), has $K \leqslant 0$. Furthermore, if we replace six by seven and use sufficiently small equilateral hyperbolic triangles with curvature -1 we get a metric on X with $K \leqslant -1$.

(b) Let X be built of regular flat k-gons for $k \geqslant 6$. Then again $K(X) \leqslant 0$ and for $k \geqslant 7$ one gets $K \leqslant -1$.

(c) Let us allow k-gones for $k \geq 4$ and add the inequality $\ell \geq 4$ for the links of all vertices. Then again we get $K(X) \leq 0$. If we additionally assume $\max(k,\ell) \geq 5$, then $K(X) \leq -1$ for some metric on X.

4.2.C __Cubical polyhedra__: Let X be an n-dimensional polyhedron built of unit Euclidean cubes and let us give a (necessary and sufficient) condition for $K(X) \leq 0$. Take the link $L_\square$ of some cube $\square$ in X and say that $L_\square$ satisfies *no-Δ-condition* if for every three vertices in $L_\square$ where every two are joined by an edge in $L_\square$ there exists a 2-simplex in $L_\square$ with these three vertices.

__**Proposition**__: *If* $L_\square$ *satisfies no-Δ-condition for all* $\square$ *in* X *then* $K(X) \leq 0$.

__**Proof**__: Criterion 4.2.B and an obvious induction on dimension reduce the problem to the following

__**Lemma**__: *Let* L *be a simplical polyhedron built of regular spherical simplices of curvature* +1 *and of diameter* $\pi/2$. *If* $K(L) \leq 1$ *and* L *satisfies no-Δ-condition, then every closed geodesic in* L *has length* $\geq 2\pi$.

__**Proof**__: Take the ball $B_v \subset L$ of radius $\pi/2$ around some vertex $v \in L$. This B looks very much like the ordinary hemisphere. Namely, if two points x and y in the boundary ∂B are joined by a geodesic segment s in B of length $< \pi$, then this segment is necessarily contained in ∂B. This is seen by looking at the surface (of curvature ≤ 1) which is the union of the unit segments starting at v and meeting our segment s.

Now, for every closed geodesic γ in L, there exist two simplices Δ_1 and Δ_2 in L such that γ passes from Δ_1 to Δ_2 across a common face in $\Delta_3 \subset \Delta_1 \cap \Delta_2$. Then γ meets the *open* $\pi/2$-balls around some vertices $v_1 \in \Delta_1$ and $v_2 \in \Delta_2$

opposite to Δ_3 and hence, length $\gamma \geqslant 2\pi$.

<u>Remarks</u>:

(a) This argument is similar to the one used in the theory of Tits and Bruhat-Tits buildings which provide most remarkable examples of polyhedra with $K \leqslant 1$ and with $K \leqslant 0$ (see [Ti_1]).

(b) The lemma (and the proof) holds true for non-regular simplices of *size* $\geqslant \pi/2$. This means that the distance from each vertex to the opposite face is $\geqslant \pi/2$. For example, every 3-dimensional polyhedron P built of dodecahedra and satisfying no-Δ-condition has $K \leqslant 0$. In fact, by using hyperbolic dodecahedra with 90-degree dihedral angles one can give P a metric with $K \leqslant -1$.

(c) <u>Siebenmann's</u> <u>no-□-condition</u>: This applies to a simplical complex L and claims in addition to no-Δ that every quadrilateral of edges in L bounds a union of two triangles in L.

One easily sees with the proof of the lemma that every closed geodesic of length 2π in L can be deformed to a geodesic consisting of four edges that is a quadrilateral of edges which bounds no pair of triangles. It follows that no-□ ensures a lower bound $\ell \geqslant 2\pi + \epsilon$ for the length ℓ of all closed geodesics in L where $\epsilon > 0$ only depends on dim L.

<u>Corollary</u>: *Let every* $L_\square$ *in* X *satisfies no-*Δ*-and-no-*□*-conditions and let us give* X *a* $(M,-\epsilon')$ *geometry, where each cube in* X *is isometric to the unit cube in the hyperbolic space of curvature* $-\epsilon'$. *If* $|\epsilon'| > 0$ *is sufficiently small then* $K(X) \leqslant -\epsilon'$. *In particular, if* X *is compact then the fundamental group* $\pi_1(X)$ *is word hyperbolic.*

(d) Let us indicate without a proof (which is easy and left to the reader) a simple general criterion on deformations of metrics without geodesics that are shorter than 2π.

Proposition: *Let* $|\ |_t$ *be a continuous family of metrics with* $K \leqslant 1$ *on a fixed space* L *such that* $|\ |_{t'} \geqslant |\ |_t$ *for all* t *and* $t' \geqslant t$. *If* L *contains no closed* $|\ |_0$*-geodesics of length* $\leqslant \ell < 2\pi$ *then there is no such* $|\ |_t$*-geodesics in* L *for all* $t \geqslant 0$.

Remark: In many (but not in all) cases the strict monotonicity $|\ |_{t'} > |\ |_t$ ensures "no $|\ |_t$-geodesics of length 2π," provided there is no $|\ |_0$-geodesics of length $\leqslant 2\pi$.

4.3 **Cutting and pasting for $K \leqslant 0$**: Let X have $K \leqslant \chi$ and let $Y \subset X$ be a *locally convex* subspace. That is for every $y \in Y$ there exists an $\epsilon > 0$ such that every (distance minimizing) segment $[x_1,x_2] \subset X$ with $x_i \in Y$ and $|x_i - y| \leqslant \epsilon$ for $i = 1,2$, is contained in Y. Let Γ freely and isometrically act on Y and let $\bar{X} = (X \setminus Y) \cup Y/\Gamma$ with the obvious geodesic metric. Then (by an easy and well known argument) the space $\bar{X}$ also has $K(\bar{X}) \leqslant \chi$.

Example: Let Y consist of two disjoint isometric copies, $Y = Y_1 \cup Y_2$ and $\Gamma \approx \mathbb{Z}_2$ permutes the copies. Then $\bar{X}$ is obtained by gluing Y_1 and Y_2 according to the isometry $\alpha: Y_1 \leftrightarrow Y_2$ for the generator $\alpha \in \mathbb{Z}_2$. Algebraically speaking, this gives us amalgamated and HNN products in the category $K \leqslant 0$.

4.3.A The above criterion immediately shows that the natural (M,0)-metrics in the hyperbolized polyhedra P_h in 3.4, 3.4.A and 3.4.B have $K \leqslant 0$ as they are built of standard blocks attached along their locally convex boundaries. Furthermore one can easily perturb the metrics on P_h of 3.4 and 3.4.B to a one with $K < 0$ which implies the hyperbolicity of $\pi_1(P_h)$ when P is compact.

4.4 **Ramified coverings**: Take again a locally convex $Y \subset X$ and a covering $\tilde{X}^*$ of the complement $X \setminus Y$. If $\tilde{X}^*$ is a trivial covering, that is a disjoint union of several copies of $X \setminus Y$, then we can glue together the corresponding copies of X along Y as in 4.3 and obtain a space $\hat{X}$ with $K(\hat{X}) \leq \chi$, provided $K(X) \leq \chi$. Now we generalize this construction to non-trivial coverings $\tilde{X}^*$ by letting $\hat{X} = \tilde{X}^* \cup Y$. This $\hat{X}$ carries a natural geodesic metric induced by the map $p: \hat{X} \longrightarrow X$ which is the covering map on $\tilde{X}^*$ and the inclusion on Y. A simple application of 4.3 shows this metric has $K \leq \chi$.

Example: Let X be a manifold and $Y \subset X$ a codimension two submanifold. Then $\hat{X}$ can be obtained as the metric completion of $\tilde{X}^*$ with the metric induced from X. The above criterion applies if X is Riemannian manifold and Y is a totally geodesic (i.e. locally convex) submanifold. Such submanifolds $Y \subset X$ are abundant if X is compact of constant negative curvature with an *arithmetic* fundamental group. (This leads to useful examples of ramified coverings $\hat{X}$ in [G-T].) Notice that completions $\hat{X}$ of *infinite* coverings $\tilde{X}^*$ of $X^* = X \setminus Y$ provide interesting examples of not locally compact spaces with $K \leq 0$.

4.4.A **Multiple ramifications**: There are many cases where $Y \subset X$ is *not* locally convex but yet some ramified coverings $\hat{X}$ have $K \leq 0$, provided $K(X) \leq 0$. The standard example is the map $\mathbb{C}^n \longrightarrow \mathbb{C}^n$ given by $(z_1,\dots,z_n) \longmapsto (z_1^{i_1},\dots,z_n^{i_n})$ for some integers $i_1,\dots,i_n$. The (singular) metric on the source induced from the flat metric on the target has $K \leq 0$.

Now take an *immersed* codimension two submanifold $Y \subset X$ with transversal selfintersections (normal crossings). A ramified covering $\hat{X}$ is called *normal* if it is locally modelled by the above example for dim X *even*. If dim X is odd we use for the model the Cartesian product of the above map $\mathbb{C}^n \longrightarrow \mathbb{C}^n$ by the identity $\mathbb{R} \longrightarrow \mathbb{R}$.

Now let X be given a Riemannian metric with $K \leq \chi$ and let at every selfintersection point of Y different branches of Y are mutually orthogonal (like transversal coordinate subspaces of codimension two in $\mathbb{R}^n$). Then (by an easy argument) the induced metric in every normal ramified covering $\hat{X}$ also has $K(\hat{X}) \leq \chi$.

Example (Compare [Gr1]): Let X be a flat torus and Y a union of flat codimension two subtori. If these subtori meet orthogonally we have $K \leqslant 0$ for normal ramified coverings $\hat{X}$. Furthermore, if these subtori generate $H_{n-2}(X)$, then, by 4.1.B the fundamental group $\pi_1(\hat{X})$ is hyperbolic, provided $\hat{X}$ *nontrivially* ramifies over each subtorus. If dim X is odd, then every generic homomorphism $X \longrightarrow S^1$ induces a *fibration* $\hat{X} \longrightarrow S^1$ whose fiber F is an even dimensional torus normally ramified over flat subtori. Yet these subtori do not meet orthogonally (whenever they meet). In fact, if dim $F \geqslant 4$, then the fundamental group $\pi_1(F)$ is *not* hyperbolic by 5.4.

4.4.B **Ramifications associated with finite reflection groups:** Let G be a finite group of orthogonal transformations of $\mathbb{R}^n$ such that the quotient $\mathbb{R}^n/G$ is homeomorphic to $\mathbb{R}^n$. Then the map $\mathbb{R}^n \longrightarrow \mathbb{R}^n/G = \mathbb{R}^n$ can be viewed as a ramified covering whose singular locus is the image of fixed points in $\mathbb{R}^n$ of non-identity elements $g \in G$. The standard examples come from groups G' generated by reflections on S^{n-1} and having a simplex for a fundamental domain. Indeed, the subgroup $G \subset G'$ of the orientation preserving transformations has $\mathbb{R}^n/G \approx \mathbb{R}^n$.

Next, one takes a subset Y in an n-dimensional manifold X which locally looks like the ramification locus in $\mathbb{R}^n \approx \mathbb{R}^n/G$ for a given G and then one has a (possibly empty) class of ramified coverings which ramify like the model map $\mathbb{R}^n \longrightarrow \mathbb{R}^n/G$ (an appropriate formalism for such a description is provided by the language of orbifolds, see 4.5).

Example: Let $G \approx \mathbb{Z}_2 \oplus \mathbb{Z}_2$ act on $\mathbb{R}^3$ by $(x,y,z) \longmapsto (-x,-y,z)$ and $(x,y,z) \longmapsto (x,-y,-z)$. Then the ramification locus in $\mathbb{R}^3 = \mathbb{R}^3/G$ consists of three rays from the origin. Now, we take a graph Y in a 3-manifold X, such that there are exactly tree edges at every vertex of this graph. The linear ramification pattern for the above $\mathbb{R}^3 \longrightarrow \mathbb{R}^3/G$ gives a nice class of ramified covers $\hat{X} \longrightarrow X$. If, for example, the manifold X has $K \leqslant 0$, if the graph Y consists of *geodesics* between the vertices in Y and if these geodesics meet at 120-degree angles at the vertices, then (by an easy argument) $\hat{X}$ also has $K \leqslant 0$. We suggest to the reader to find specific examples of such Y in the flat torus T^3 and then look at the standard series (A_n, B_n etc.) of reflection groups.

4.5 Orbifolds and orbispaces: Let a group Γ act discretely and isometrically on a (Riemannian) manifold X. If the action is free then Γ can be recaptured from X/Γ as $\Gamma = \pi_1(X/\Gamma)$. The same can be done for non-free action with an additional *orbistructure* on X/Γ which takes into account the (finite) isotropy subgroups $\Gamma_x \subset \Gamma$ for all $x \in X$. Namely, if $x \longmapsto v \in V = X/\Gamma$, then a small neighborhood $U_v \subset V$ of v is naturally represented as $\tilde{U}/\Gamma_x$ for some neighborhood $\tilde{U} \subset X$ of x.

Now we *start* with a space V, a collection of groups Γ_v for all $v \in V$ operating on some connected manifolds X_v, and maps of X_v/Γ_v onto some neighborhoods $U_v \subset V$, such that these data agree in the way suggested by the example of $V = X/\Gamma$ (see [Th] for details).

We are interested in a more general case where X_v need not be manifolds but rather general geodesic spaces. The *orbispace* V (called *orbifold* if X_v are manifolds and *orbihedron* if X_v are polyhedra) also comes with a geodesic metric and the maps $X_v/\Gamma_v \longrightarrow U_v$ are assumed isometric. The curvature inequality $K(V) \leqslant \chi$ by definition is: $K(X_v) \leqslant \chi$ for all spaces X_v.

In the orbifold category the orbistructure is usually determined by the metric. For example, the unit segment [0,1] carries a unique (maximal) orbifold structure as $[0,1] = \mathbb{R}/\Gamma$ for Γ generated by integral translations and $t \longleftrightarrow -t$. Yet [0,1] has many orbihedral structures. Namely if we divide any infinite regular tree (by its full isometry group (which is not discrete though totally disconnected)), we also get [0,1] but with another orbistructure.

A *universal covering* of an orbispace V is a simply connected space X with an action of Γ such that X/Γ equals Y in the orbispace sense. The group Γ is called *a fundamental group* of V. Notice that neither existence nor uniqueness of Γ is automatic. Yet in the orbifold case X and Γ are unique, though they do not always exist.

If V is an orbivolds with $K(V) \leqslant 0$ then the universal covering

X (which is a manifold with $K \leqslant 0$ in the usual sense) does exist by a trivial argument, but this is unknown for orbihedra. However, there is an additional *rigidity* condition (ruling out tree-like behavior) which ensures the existence and uniqueness of the universal covering. Namely, V is called a *rigid orbispace* if none of the spaces X_v admits a non-trivial isometry which fixes a non-empty open subset in X_v. In the rigid case the uniqueness of X and Γ is obvious and the inequality $K(V) \leqslant 0$ ensures the existence.

<u>Remark</u>: Non-rigid orbihedra, especially those with $K \leqslant 0$, (for example trees) are very interesting objects with abundance of symmetries. Some examples of them can be found in 4.C″ and C‴ of [Gr5].

<u>Unrestricted</u> <u>orbihedra</u>: A geodesic metric space V is called *unrestricted* if every local geodesic can be extended. That is every locally isometric map $[a,b] \longrightarrow V$ extends to a locally isometric map $\mathbb{R} \longrightarrow V$ for all $[a,b] \subset \mathbb{R}$. Then one defines *unrestricted orbispaces* in terms of local extension of geodesics in the corresponding spaces X_v.

If V is a polyhedron with $K \leqslant 0$, then "unrestricted" is (obviously) equivalent to the following property of the links L_Δ of V: for each point $\ell \in L_\Delta$ there exists an *opposite* point ℓ' in L_Δ, such that $|\ell-\ell'|_{L_\Delta} \geqslant \pi$. (If ℓ and ℓ' lie in different components of L_Δ we assume $|\ell-\ell'|_{L_\Delta} = \infty$.)

Every compact unrestricted (orbi)space V with $K \leqslant 0$ obviously has *infinite* universal covering and hence $\pi_1(V)$ also is infinite. Furthermore, if $K < 0$ then $\pi_1(V)$ is a *non-elementary* hyperbolic group unless V is isometric to S^1 or to $[0,1]$ with the standard orbistructure.

4.5.A <u>Examples</u> <u>of</u> <u>orbihedra</u>:

(a) Start with a two dimensional polyhedron V and give it the orbistructure which is *supported* in the centers of the triangles in V (the support of an orbistructure is the set of those $v \in V$ where

Γ_v is non-trivial) and equals there to the $2\pi/k$-*structure*. That is the disk in $\mathbb{R}^2$ divided by the usual $\mathbb{Z}_k$-action. If $k \geqslant 2$ for all triangles, then the obvious (orbi)metric on V has $K \leqslant 0$ and if $k \geqslant 3$ then $K < -\epsilon^2 < 0$. The fundamental groups Γ of these V provide plenty of hyperbolic groups.

(b) Let V be a manifold and V' an immersed submanifold of codimension two with transversal self-intersections. At every p-multiple intersection point we give V the orbistructure of the standard action of $\mathbb{Z}_{i_1} \times \mathbb{Z}_{i_2} \times \dots \mathbb{Z}_{i_p}$ on $\mathbb{C}^p \times \mathbb{R}^m$, where $2p + m = \dim V$ (compare 4.4.A). In order to define an orbistructure we assume an obvious agreement between p and (p+1)-multiple points. For example if V' is the union of p embedded submanifolds then such a structure is given by assigning a ramification index i_j for $i = 1,\dots,p$ to each submanifold. Now if V has a metric with $K \leqslant 0$ and V' is geodesic with *orthogonal* crossings then same metric viewed as orbimetric also has $K \leqslant 0$ and hence the *orbifold* V admits the universal covering, that is the maximal *ramified* covering of the *manifold* V with given ramification along V'.

4.5.B <u>Cutting and pasting orbispaces</u>: This can be done for $K \leqslant 0$ the same way as for ordinary spaces since the condition $K \leqslant 0$ is local.

<u>Example</u>: Let V be an orbispace with $K \leqslant 0$ and let $S \subset V$ be a closed geodesic of length $2\pi r$ which does not meet the support of the orbistructure. (S is the image of locally isometric map of $2\pi r$-circle into V.) Next take the disk $D^2 \subset \mathbb{R}^2$ of radius kr for some $k = 1,2,\dots$, divided by the standard action of $\mathbb{Z}_k$ and attach $D^2/\mathbb{Z}_k$ to V by an isometry between $\partial D^2/\mathbb{Z}_k$ and S. The resulting orbifold $V \cup (D^2/\mathbb{Z}_k)$ fails to have $K \leqslant 0$ since $\partial D^2/\mathbb{Z}_k$ is not locally convex (geodesic) in $D^2/\mathbb{Z}_k$. Yet it becomes such in the limit for $k \longrightarrow \infty$ as $D^2/\mathbb{Z}_k$ converges to the product of $\mathbb{R}_+$ by the $2\pi r$-circle S and the space $V \cup (\mathbb{R}_+ \times S)$ does have $K \leqslant 0$. This follows from 4.3 if S has no self-intersection points in V and an obvious modification of 4.3 allows such points.

Now, let $K(V) < 0$ and show for large $k > k_0 = k_0(V,S)$ that there is a small perturbation of the metric on $V \cup (D^2/\mathbb{Z}_k)$ which achieves $K < 0$. Assume for simplicity's sake that V is an orbihedron with a polyhedral $(M,-\epsilon^2)$-type metric. We may assume (subdividing the polyhedron if necessary) that S contains a vertex v of the triangulation and two adjacent edges. Denote by v' and v" the corresponding vertices in the link L_v and observe that the (angular) distance between v' and v" satisfies $|v'-v"| \geqslant \pi$. Since $K(V) < 0$ there is a small perturbation of the metric on V, such that the new distance in the link satisfies $|v'-v"|_{new} \geqslant \pi+\delta$ for a small $\delta > 0$. Now, instead of D^2 we take the regular k-gone D_k with edge length $2\pi r$ in the hyperbolic plane with (constant) curvature $-\chi < 0$. If k is large and $|\chi|$ is small then the angles α in D_k between adjacent edges satisfy $|\pi-\alpha| \leqslant \delta$. Therefore, if we attach $D_k/\mathbb{Z}_k$ to V along S, such that the corner point in $\partial(D_k/\mathbb{Z}_k)$ goes to v, then the resulting space will have $K < 0$.

__Remark__: The above perturbation easily generalizes to more general orbispaces where a small amount of positive curvature can be engulfed into surrounding negative curvature. For example, one could take a higher dimensional locally convex (immersed) subspace V' in V rather than S and attach orbicones to appropriate coverings of V'.

4.5.C __Construction of infinite torsion groups__: Start with a compact unrestricted polyhedron V with $K < 0$, for example with the figure ∞ or with a surface of constant negative curvature. Take a closed geodesic in V, attach the $2\pi/k_1$-orbicone as above and obtain an orbihedron V_1 with $K(V_1) < 0$ which is obviously unrestricted. Do the same to V_1 and get V_2. Notice that some geodesic in V_1 may pass through the point supporting the orbistructure. To avoid (not at all serious) trouble we perturb our metrics in order to remove geodesics from orbisupports. Now, we continue the process by keeping attaching orbicones and obtain a sequence of unrestricted orbispaces with $K < 0$, say $V \subset V_1 \subset V_2 \subset \ldots V_i$, such that (if we wish) the length of the shortest closed geodesic in V_i, say ℓ_i, satisfies $\ell_i \longrightarrow \infty$ for $i \longrightarrow \infty$. It follows that the fundamental group Γ_∞ of the orbispace $V_\infty = \bigcup_i V_i$ is pure torsion. Furthermore, since V_∞ has $K(V_\infty) \leq 0$ and is rigid, the universal covering $\tilde{V}_\infty \longrightarrow \tilde{V}$ does exist and has infinitely many ramifications of orbitrarily large order. Thus the group Γ_∞ is infinite. Finally, Γ_∞ is the factor group of $\pi_1(V)$ and hence it is finitely generated. Q.E.D.

__Remark:__ If we start this construction with a lattice in Sp(2,1) we obtain an infinite torsion group with Kazdan's T-property.

__Question:__ Let a group Γ be obtained from a non-elementary word hyperbolic group Γ_0 by adding "on random" infinitely many relations to Γ_0. Does Γ satisfy T-property?

4.6 __Reflection orbihedra with__ $K \leq 0$: To grasp the idea start with the unit cube V obtained by dividing $\mathbb{R}^n$ by the group Γ generated by integral translations of $\mathbb{R}^n$ and reflections in the coordinate hyperplanes. The space $\mathbb{R}^n$ can be reconstructed as the universal (orbi)covering of V which is obtained in this case by repeated reflections of V around codimension one faces. The fundamental group Γ ov V is a reflection group (see 3.5) whose generators γ_i (with $\gamma_i^2 = 1$) $i = 1,\ldots,2n$, correspond to the codimension one faces of V. The relations are $(\gamma_i\gamma_j)^2 = 1$ for those pairs of faces which meet at codimension two faces of V.

Notice that the orbistructure of V is supported on the boundary and it is determined by the combinatorial pattern of the codimension one faces. If we take the *nerve* L of the covering of ∂V by these faces, then each face can be identified with the star of a vertex of L in the first barycentric subdivision L' of L while V becomes the cone over L'.

Now we *start* with an *arbitrary* simplical polyhedron L, cover L by the stars of the vertices of L in the first baricentric subdivision L' of L and consider the cone M over L' with the orbifold structure induced by the cover of L by the stars. Notice that the cone over a baricentrically subdivided simplex can be naturally identified with the standard cube. Thus M has a natural structure of a cubical complex and the discussion in 4.2.C applies to the natural (M,0)-cubical metric on M. This metric has $K \leqslant 0$ (in the orbifold sense) if L and the link of each simplex in L satisfy no-Δ-condition. Moreover, if no-□ is satisfied as well, then $K < 0$ for a small hyperbolic perturbation of the (M,0)-metric.

__Corollary:__ *The reflection group* $\Gamma = \Gamma(L)$ *generated by* γ_v *with* $\gamma_v^2 = 1$ *for all vertices* $v \in L$ *and with relations* $(\gamma_v\gamma_w)^2 = 1$ *for all edges* $[v,w]$ *in* L *operates on some polyhedron* X *(that is the universal covering of the orbifold* M) *with* $K \leqslant 0$, *provided no-Δ is satisfied. Moreover no-□ implies* $K < 0$ *and* Γ *is hyperbolic if* L *is compact.*

__Remark:__ The baricentric subdivision of *every* simplical complex satisfies the no-Δ-condition but the no-□ is harder to get as was pointed out to me by L. Siebenmann.

__Examples:__ Let L be a triangulation of the sphere S^{n-1}. Then, obviously, X is a manifold. More generally, let L be a triangulation of a boundary ∂W of some n-dimensional manifold. The partition into stars gives an orbifold structure to W supported on ∂W. This structure usually has no metric with $K \leqslant 0$ but the corresponding structure on the cone M over ∂W does carry such a structure in the no-Δ-case. It follows that the orbifold W admits a universal covering with the fundamental group $\sqcap = \sqcap(W)$ which comes with a natural epimorphism onto the fundamental group $\Gamma = \sqcap(M)$. Of course, if W is simply connected as a topological space, then $\sqcap(W) = \sqcap(M)$. In general, the (ordinary) fundamental group $\pi_1(W)$ embeds into $\sqcap(W)$ and it is annihilated by the homomorphisms $\sqcap(W) \longrightarrow \sqcap(M)$.

4.7 __Small cancellation polyhedra of dimension two__: There are various combinatorial conditions for 2-dimensional orbihedra (in particular, polyhedra) V which are more general than $K < 0$ (or

$K \leqslant 0$) but yet sufficient for hyperbolicity (or semihyperbolicity). First notice that $K < 0$ (and $K \leqslant 0$) is expressed in terms of the geometry of the links L_v which only depends on homothety classes of the triangles in V. Now, we define a *conformal structure* in a simplical 2-polyhedron V by assigning a real number, called an angle, to every corner of each triangle Δ in V, such that the sum of the three angles of every triangle equals π. These angles define *lengths* (positive or negative) of curves in the links L_v and $K < 0$ ($K \leqslant 0$) means that every closed curve in L_v has length $> 2\pi$ ($\geqslant 2\pi$) for all $v \in V$. Notice, that the existence of a conformal structure with $K < 0$ (or $K = 0$) amounts to solvability of a system of *linear* inequalities in unknown angles.

Similarly one defines 2-orbihedra with conformal $K < 0$ (or $K \leqslant 0$) and easily shows that $K < 0$ ensures hyperbolicity of the fundamental group, while $K \leqslant 0$ provides (some kind of) semihyperbolicity.

4.7.A Now, let us formulate the usual small cancellation conditions for 2-orbihedra. Here, our V is a cell complex, such that

(1) The 1-skeleton V^1 of V has no redundancy: there are at least three branches at every vertex in V^1.

(2) The attaching maps for the 2-cells D in V are topological immersions (i.e. locally injective) $\partial D \longrightarrow V^1$.

(3) The orbistructure (if there is any) is supported at finitely many points in the interiors of 2-cells D, where each D contains at most one $2\pi/k$-point for some $k = k(D)$.

A *one layer plane diagram belonging to* V is given by the following data.

1. A closed topological disk $B \subset \mathbb{R}^2$ subdivided into (n+1) convex k_i-gones for $k_i \geqslant 3$, called $D_0, D_1, \ldots, D_n$, such that

(a) Every two k_i-gones meet (if at all) at a common *vertex* or *edge*, and each D_i for $i \geq 1$ does meet D_0.

(b) The *central* k_0-gone D_0 does not meet the boundary ∂B of B and each D_i for $i \geq 1$ does meet ∂B.

2. A continuous map $f: B \longrightarrow V$ such that

(i) The map f sends the vertices of the k_i-gones to some vertices in V and the edges go to the 1-skeleton V^1 of V. Furthermore, f is an immersion (i.e. locally one-to-one) on some neighborhood $U \subset B$ of the union of the edges (that is $\bigcup_i \partial D_i$).

(ii) The interior Int D_i is mapped onto the interior Int D of some 2-cell D in V. This map between open disks is a ramified covering of order $k = k(D)$. Thus, f is a cyclic k-sheeted covering of Int D_i minus a point onto Int D minus the orbipoint for all $i = 0,\dots,n$.

Assign to each vertex v_j of D_0, $j = 1,\dots,k_0$, the angle $\alpha_j = 2\pi/\ell_j$ where ℓ_j denotes the number of k_i-gones adjacent to v_j (notice that $\ell_j \geq 3$) and define the curvature (excess) of B by

$$K(B) = 2\pi + \sum_{j=1}^{k_0} (\alpha_j - \pi).$$

Notice that $K(B) \leq 0$ for $k_0 > 5$ (the familiar 1/5-condition) and $K(B) < 0$ for $k_0 > 6$.

Now, we summarize various small cancellation conditions (for the orbihedra corresponding to a presentation of a group Γ) in the following

<u>**Definition**</u>: The combinatorial curvature $K_c(V)$ is < 0 if $K(B) < 0$ for all diagrams belonging to V and $K_c(V) \leq 0$ if $K(B) \leq 0$.

The basic theorem of the small cancellation theory claims that $K_c < 0$ implies *the hyperbolicity of* $\Gamma = \pi_1(V)$.

<u>**Sketch of the proof**</u>: Let S be an immersed circle in V^1 and let B'

be a *minimal* (area) disk mapped to V with $\partial B' = S$. This disk can be represented by a many layer diagram. Namely, assuming the lift of S^1 to the universal covering of V is embedded, all disks in the diagram are mapped into V by ramified coverings over some 2-cells D in V with ramifications of order $\geqslant k(D)$. Now we give a metric to B' by claiming each k_i-gone in the diagram is isometric to the regular k_i-gone with unit edges in the plane $\mathbb{R}^2$ for the case $K_c \leqslant 0$ and to such a k_i-gone in the hyperbolic plane of curvature $-\epsilon$ for small $\epsilon > 0$ if $K_c < 0$. This metric has $K \leqslant 0$ (respectively $K < 0$) for $K_c \leqslant 0$ ($K_c < 0$) and by a classical (and easy) theorem (see [B-Z])

$$\text{Area } B' \leqslant (4\pi)^{-1}(\text{length } S)^2 \text{ for } K_c \leqslant 0,$$

and if $K_c < 0$, then additionally

$$\text{Area } B \leqslant \text{const length } S.$$

This proof shows that if V has $K_c \leqslant 0$ and is not hyperbolic, then there exists a partition of $\mathbb{R}^2$ into infinitely many bounded convex k_i-gones and a map $\mathbb{R}^2 \longrightarrow V$ satisfying the one layer diagram condition around every k_i-gone of this partition.

It seems to be unknown if the existence of such a map $\mathbb{R}^2 \longrightarrow V$ implies for $K_c(V) \leqslant 0$ that the fundamental group of V contains an isomorphic copy of $\mathbb{Z} \oplus \mathbb{Z}$. There is little evidence in favor of such $\mathbb{Z} \oplus \mathbb{Z}$. In fact, there probably exist compact semihyperbolic (e.g. with $K \leqslant 0$) spaces receiving $\mathbb{R}^2$ but having no $\mathbb{Z} \oplus \mathbb{Z}$ in π_1. Some geometers believe that *closed manifolds* with $K \leqslant 0$ and with non-hyperbolic π_1 do have $\mathbb{Z} \oplus \mathbb{Z} \subset \pi_1$ but well known examples of non-periodic (and non-periodizable) tilings of $\mathbb{R}^2$ by congruent domains make such belief ill founded. There is a similar problem in complex geometry. Let a compact complex algebraic space (or manifold) V receive a non-constant map $\mathbb{C} \longrightarrow V$. Does V receive a non-constant map of some compact complex torus?

§5. <u>Basic properties of word hyperbolic groups</u>

In the following sections we denote by Γ a finitely generated group with a fixed δ-hyperbolic word metric.

Let us apply the discussion in 3.1 to the infinite cyclic subgroup $C_\gamma \subset \Gamma$ generated by a non-torsion element $\gamma \in \Gamma$. Then we see that there are two fixed points, called the *poles* γ^∞ and $\gamma^{-\infty}$ in $\partial\Gamma$ defined by

$$\gamma^\infty = \lim_{i\to\infty} \gamma^i \text{ and } \gamma^{-\infty} = \lim_{i\to-\infty} \gamma^i.$$

Denote by $M_\gamma \subset \Gamma$ the subgroup of transformations which map the subset $\{\gamma^\infty,\gamma^{-\infty}\} \subset \partial\Gamma$ onto itself. It is clear that M_γ is an elementary group and that M_γ contains the normalizer of $C_\gamma \subset \Gamma$. In fact, M_γ is the (unique) maximal finite extension of C_γ in Γ. It follows, that $\gamma = \gamma_0^p$ for some *exponent* $p = p(\gamma) \in \mathbb{N}$ and some *prime* element $\gamma_0 \in M_\gamma$. If M_γ has no torsion, then γ_0 is unique. Namely γ_0 is the genetator of $M_\gamma \approx \mathbb{Z}$ having the same sign (in $\mathbb{Z}$) as γ.

The action of γ on $\partial\Gamma$ is very simple: there are fundamental systems of neighborhoods $U_i^+ \subset \partial\Gamma$ of γ^∞ and $U_i^- \subset \partial\Gamma$ of $\gamma^{-\infty}$, such that γ^i maps $\partial\Gamma \setminus U_i^-$ to U_i^+ for $i = 1,2,\dots$.

All this is well known in the geometric context and the case of a general word hyperbolic group is equally obvious (see 8.1.).

5.1 <u>Density of poles</u>: Denote by $\partial^2\Gamma \subset \partial\Gamma \times \partial\Gamma$ the set of pairs of *distinct* points in $\partial\Gamma$ and let $P(G) \subset \partial^2\Gamma$ denote the subset of pairs $(\gamma^\infty,\gamma^{-\infty}) \in \partial^2\Gamma$ for all non-torsion elements $\gamma \in G$, where $G \subset \Gamma$ is a given subset.

5.1.A (See 8.2.) *If G contains at most finitely many pairwise non-conjugate elements in* Γ *then* P(G) *is discrete. If* G *is a cobounded subset in* Γ (*e.g.* $G = \Gamma$), *then* P(G) *is dense in* $\partial^2\Gamma$.

5.1.B <u>Corollary</u>: *Every non-elementary word* *hyperbolic group* Γ *contains infinitely many conjugacy classes of prime non-torsion elements.*

5.2 <u>Markov coding in</u> Γ: Consider a finite set G and a set H of finite sequences (words) of elements $g \in G$. Call H *Markov* if there exist a finite set S (of states) with a preferred element $s_0 \in S$, a subset $T \subset S \times S$ (of transitions) and a (labeling) map $\alpha: T \longrightarrow G$, such that H equals the set of sequences of the form

$$\alpha(s_0,s_1),\alpha(s_1,s_2),\ldots,\alpha(s_{k-1},s_k)$$

for all sequences $s_0,s_1,\ldots,s_k \in S$, such that $(s_i,s_{i+1}) \in T$ for $i = 0,\ldots,k-1$ and for all $k = 1,2,\ldots$.

Now, let G be a finite *symmetric* generating subset in a group Γ; that is, $G^{-1} = G$. A word w in $g \in G$ representing some $\gamma \in \Gamma$ is called *shortest* if length $w = |\gamma|$ for the word metric associated to G. Next, fix a linear ordering in G and call a shortest w representing γ *normal* if it is minimal for the lexicographic order on the set of the shortest words representing γ.

5.2.A <u>Theorem</u>: (See 8.5.) *If* Γ *is word hyperbolic then the sets of the shortest and of the normal words are Markov.*

5.2.A' <u>Corollary</u>: (Compare [Can].) *Let* N_k *denote the number of* $\gamma \in \Gamma$ *with* $|\gamma| = k$. *Then the counting function* $Z(t) = \sum_{k=0}^{\infty} N_k t^k$ *is rational. Similarly the counting function for the number of shortest words in* Γ *of length* $= k$ *is rational.*

5.2.B Fix some $\gamma_0 \in \Gamma$ and let $N_k(\gamma_0)$ denote the number of $\gamma \in \Gamma$ conjugate to γ_0 and having $|\gamma| = k$.

<u>Theorem</u>: (See 8.5.) *The function* $\sum_k N_k(\gamma_0)t^k$ *is rational for all* $\gamma_0 \in \Gamma$.

This implies with 2.2.B that the functions counting torsion and non-torsion γ with $|\gamma| = k$ separately are rational.

5.2.C Define the *norm* of the *conjugacy class* $[\gamma]$ of $\gamma \in \Gamma$ by

$$|[\gamma]| = \inf_{\gamma' \in [\gamma]} |\gamma'|$$

and let

$$|[\gamma]|_- = \liminf_{n\to\infty} n^{-1} |[\gamma^n]|.$$

We shall see in 8.5 that there exists an integer m_0 depending on (Γ,G), such that $|[\gamma]|_- = m_0^{-1} |[\gamma^{m_0}]|$ for all $\gamma \in \Gamma$. Now, write each non-torsion $\gamma \in \Gamma$ as γ_0^p for some prime γ_0 and $p > 0$, let $\mu[\gamma] = p^{-1}|[\gamma]|_-$ and denote by $[N]_k$ the sum $\sum_{[\gamma]} \mu[\gamma]$ over the classes $[\gamma]$ with $|[\gamma]|_- = k/m_0$ for all non-torsion $\gamma \in \Gamma$.

5.2.D **Theorem**: (See 8.5.) *The counting function* $\sum_k [N]_k t^k$ *is rational.* Furthermore

$$[N]_k \geq (1+\epsilon)^k \text{ for some } \epsilon > 0$$

unless the group Γ is elementary.

5.2.E **Remark**: A group Γ generated by $\gamma_i \in \Gamma$, $i = 1,\dots,k < \infty$ is called *Markov* if there exists a Markov set H of (not necessarily shortest) words in γ_i, such that each $\gamma \in \Gamma$ is represented by a unique $h = h(\gamma) \in H$. Clearly, "Markov" is invariant under isomorphisms of groups and stable under free and Cartesian products of groups. Moreover, every extension of a Markov group by another Markov group is a Markov group. Apparently, little else is known about Markov groups.

5.3 Free and non-free subgroups in Γ: We have seen in 3.1 how to detect freedom of some subgroups Γ' in Γ by looking at the action of Γ' on $\partial\Gamma$. Now we want exploit another geometric concept, namely the idea of convexity.

Definition: A subset Y in a geodesic space X is called ϵ-convex (or *quasiconvex*) if every geodesic segment in X with the end points in Y lies ϵ-close to (i.e. in the ϵ-neighborhood of) Y. If Γ is a group with a word metric then the notion of quasiconvexity applies to the Cayley graph $X \supset \Gamma$. Namely, $\Gamma' \subset \Gamma$ is quasiconvex in Γ iff it is quasiconvex in X.

If Γ is word hyperbolic and Γ' is a quasiconvex subgroup in Γ, then, by an easy argument, Γ' is finitely generated and word hyperbolic. Furthermore, the embedding $\Gamma' \hookrightarrow \Gamma$ extends to a topological embedding $\partial\Gamma' \longrightarrow \partial\Gamma$. Also notice that the quotient space Γ/Γ' with the metric induced from the word metric in Γ (where Γ' acts from the left and hence, isometrically on Γ) is hyperbolic for each quasiconvex subgroups Γ' in every word hyperbolic group Γ.

An important fact is (see Section 7) the *locality* of quasiconvexity.

Geometric example: Let X be a convex manifold with $K \leqslant -1$ and let $Y \subset X$ be a connected *geodesic graph* in X. That is Y consist of geodesic segment joining some pairs in a discrete set $Y^0 \subset X$ of the vertices in Y. Suppose all (geodesic) edges in Y have length $\geqslant \ell$ for a fixed $\ell > 0$ and every two edges at each vertex in Y^0 have angle $\geqslant \alpha > 0$. A simple application of CAT shows that:

If $\alpha \exp \ell \geqslant 10$, *then* Y_0 *is a quasiconvex tree in* X.

A good way to see why some length-angle inequality prohibits any closed curve $Y' \subset Y$ is to take the minimal ball $B \supset Y'$ in X around a fixed point. If Y' is closed, we get at some point $b \in \partial B$ two edges in B starting from b. Then by CAT, there exists a ball B' in the hyperbolic plane and two edges in B' starting from some $b' \in \partial B'$ and having the same lengths and the angle as the original edges in B. This gives a (sharp for Rad B' $\longrightarrow \infty$) bound from above on the angle in terms of the lengths of the edges.

<u>An</u> <u>algebraic</u> <u>application</u>: Let X be the universal covering of a manifold V_0 with fundamental group Γ_0 and let Γ_0 be generated by a map f of some graph (e.g. a Cayley graph of Γ_0) P into V. That is $f_*: \pi_1(P) \longrightarrow \pi_1(X)$ is surjective. Such an f can be made *geodesic* by replacing the map at every edge of P by a locally isometric (i.e. geodesic) map. The lift of this map to X is a (mapped rather than embedded) geodesic graph in X. The edges in Y and the angles between them are the same as for $f(P) \subset X$. Hence, under a suitable edge-angle inequality Y is a quasiconvex tree and consequently Γ_0 is free. Furthermore, if V_0 is a covering of a compact manifold V with $\pi_1(V) = \Gamma \supset \Gamma_0$, then Γ_0 is quasiconvex in Γ.

In order to achieve a suitable edge-angle inequality it is useful to take (P,f) which minimizes some geometric characteristic of P and f. For example, one can let P be homotopy equivalent to the wedge of k circles where k is the number of generators of a given subgroup $\Gamma_0 \subset \Gamma$ and then take map $f: P \longrightarrow V$ of *minimal length* generating $\Gamma_0 \subset \Gamma$ (i.e. $f_*(\pi_1(P)) = \gamma\Gamma_0\gamma^{-1}$ for some $\gamma \in \Gamma$). Then one sees that Γ^0 free and quasiconvex in Γ, provided the injectivity radius of V_0 satisfies

$$\text{Rad } V_0 \geqslant 100 \log(k+10),$$

where Rad V_0 is defined as the upper bound of those R for which every ball of radius R in X *injects* into V_0 under the covering map $X \longrightarrow V$.

In the special case of k = 2 one can additionally benefit by looking at two types of graphs $P \longrightarrow V_0$

(1) two closed geodesics joined by an edge.

(2) a closed geodesic and an edge with the end points on this geodesic.

Then one takes the shortest graph P among all these and observe that the corresponding Γ_0 is free convex unless the total length ℓ of the shortest P is ≤ 10.

These considerations and the quasiconvex discussion in Section 7 lead to the following freedom criteria for subgroups Γ_0 generated by k elements in a word δ-hyperbolic group Γ.

5.3.A *If* Rad Γ_0 *(defined below) satisfies* Rad $\Gamma_0 \geq \delta 1000 \log(k+100)$ *then* Γ_0 *is free and quasiconvex in* Γ.

5.3.B *If* $k = 2$, *and* Γ *has no torsion then* Γ_0 *is free quasiconvex unless it is generated by* $\gamma_1' = \gamma\gamma_1\gamma^{-1}$ *and* $\gamma_2' = \gamma\gamma_2\gamma^{-1}$ *for some* γ, γ_1 *and* γ_2 *in* Γ, *such that* $|\gamma_1| \leq 100\delta$ *and* $|\gamma_2| \leq 100\delta$. *Thus* Γ *contains at most finitely many conjugacy classes of non-free subgroups* Γ_0 *generated by two elements.*

Now let us define Rad Γ_0. First recall that $|[\gamma]| = \inf_{\gamma'} |\gamma'|$ for all $\gamma' \in \Gamma$ conjugate to γ and then set Rad $\Gamma_0 = 1/2 \inf_{\gamma \in \Gamma_0^*} |[\gamma]|$ for $\Gamma_0^* = \Gamma_0 \setminus \{id\}$.

<u>**Example:**</u> Let $\Gamma^1 \supset \Gamma^2 \supset \ldots$ be a sequence of subgroups of *finite index* in Γ. Then Rad $\Gamma^i \longrightarrow \infty$ for $i \longrightarrow \infty$ if and only if

$\bigcap_i \Gamma^i = \{id\}$. Thus 5.3.A applies to subgroups $\Gamma_0 \subset \Gamma^i$ for large i.

<u>**Remark:**</u> Probably, "generic" word hyperbolic groups admit no such sequences of subgroups of finite index.

5.3.C Let us indicate a simple (partial) generalization of 5.3.B. Start with a quasiconvex (but not necessarily free) subgroup $\Gamma_0 \subset \Gamma$ of infinite index. Notice (compare 8.1, 8.2) that $\partial\Gamma_0 \subset \partial\Gamma$ is nowhere dense in $\partial\Gamma$ and that Γ contains infinitely many Γ_0-*equivalence* classes, where γ_1 and γ_2 in Γ are called

Γ_0-equivalent if $\gamma_1 = \gamma_0\gamma_2\gamma_0'$ for some γ_0 and γ_0' in Γ_0. In fact one can define $|\gamma| \bmod \Gamma_0 = \inf_{\gamma'} |\gamma|$ over all γ' which are Γ_0-equivalent to γ and observe that the function $\gamma \longmapsto |\gamma| \bmod \Gamma_0$ is unbounded on Γ. Now, let $\Gamma_1 \subset \Gamma$ be generated by Γ_0 and some $\gamma \in \Gamma$ and notice that Γ_1 only depends on the Γ_0-equivalence class of γ. A useful choice is a *minimal* $\gamma \in \Gamma$, where $|\gamma| = |\gamma| \bmod \Gamma_0$. If $|\gamma|$ is large then adding corresponding geodesic edge to (some graph generating) Γ_0 does not hurt quasiconvexity (see Section 7) and therefore

5.3.C$_1$ *For all but finitely many Γ_0-equivalence classes of γ the subgroup $\Gamma_1 \subset \Gamma$ is quasiconvex and isomorphic to the free product $\Gamma_0 * C_\gamma$ for the cyclic group generated by γ.*

<u>Corollary</u>: Take a random sequence of elements γ_i, $i = 1,2,\ldots$ with large and fast growing length $|\gamma_i|$. Then the subgroups $\Gamma_i \subset \Gamma$ generated by $\gamma_1,\ldots,\gamma_i$ are free quasiconvex, unless Γ is elementary. Thus the subgroup $\Gamma_\infty = \bigcup_i \Gamma_i \subset \Gamma$ also is free (but not quasiconvex as it is infinitely generated).

5.3.D <u>Remark</u>: It seems that subgroups Γ_0 as in 5.3.A having Rad Γ_0 large compared to the number of generators of Γ_0 have *free normal closures* $\mathcal{N}(\Gamma_0) \subset \Gamma$. If Γ is non-elementary one also suspects there exists an epimorphism of Γ onto a non-elementary word hyperbolic group such that the kernel is a free subgroup in Γ containing Γ_0. This is motivated on one hand by the theory of one relator groups and on the other by some results for $SL_2\mathbb{C}$ in [Lub]. Additional evidence comes from the following

5.3.E <u>Theorem</u>: *There exists a constant* $m = m(k,\delta)$, *such that for every k hyperbolic elements* $\gamma_1,\gamma_2,\ldots,\gamma_k$ *in a word δ-hyperbolic group Γ the normal subgroup generated by* $\gamma_i^{m_i}$, $i = 1,\ldots,k$, *is free for all* $m_i \geq m$.

The idea of the proof is clear from the following

5.3.E_1 Example: *Let* V *be a complete manifold with* $K(V) < 0$ *and let* $\gamma \subset V$ *be a closed geodesic. Then there is no non-trivial relation in* $\pi_1(V)$ *between the conjugates of the elements represented by* γ^{m_i} *for* $m_i \geq m$, *where* m *is some constant depending on* V *and* γ.

Proof: Every normal relation between various powers γ^{m_i} is given by a map f of a surface F of genus zero into V such that each component of the boundary, say $S_i \subset \partial\Gamma$, for $i = 1,\ldots,p$, goes onto γ with non-zero degree divisible by m_i. If the relation associated to the map $f: (F,\partial F) \longrightarrow (V,\gamma)$ is formally irreducible, then (without any assumption on K(V)) the area of F in V (obviously) is bounded from below by $\epsilon \sum_{i=1}^{p} m_i$, where $\epsilon > 0$ depends on (V,γ). Now, we take a *minimal* surface F' homotopic to F (as a map $(F,\partial F) \longrightarrow (V,\gamma)$) and observe that the induced curvature in F' is $\leq K(V)$. If $K(V) \leq -\epsilon' < 0$ for $\epsilon' > 0$ we have by Gauss-Bonnet.

$$p - 1 = -x(F) = -(2\pi)^{-1} \int_F K(F') \geq \epsilon'' \text{ Area } F \geq$$

$$\geq \epsilon \text{ const} \sum_1^p m_i \geq \epsilon \text{ const } pm.$$

Hence, no such F' exists for $p \geq 2$ and $m > (p-1)(p\epsilon\epsilon')^{-1} \geq (2\epsilon\epsilon')^{-1}$. Q.E.D.

Remarks:

(a) The above argument works for $K \leq 0$ if $K < 0$ near γ. However it completely fails for locally symmetric spaces V with finite volume and rank ≥ 2 (i.e. $K \leq 0$ and nowhere < 0). Indeed by a famous theorem of Margulis every non-trivial normal subgroup in $\pi_1(V)$ has finite index.

(b) There are two possible generalizations of this geometric argument. First, one can allow surfaces F of positive genus and thus give a lower bound on the genus. Algebraically speaking, surfaces of genus q represent relations between (conjugates of) γ and q commutators in $\Gamma = \pi_1(V)$. The second generalization concerns relation between parabolic elements γ in (non-word) hyperbolic groups, such as fundamental groups of complete non-compact manifolds V with $K < 0$ and with finite volume. Here one uses non-compact minimal surfaces in V properly mapped into V. The area of such a surface can be rather sharply bounded from below by the size of cusps in V which shows in many cases parabolic elements to be normally independent.

(c) Instead of minimal surfaces one can equally use surfaces subdivided into Thurston's hyperbolic geodesic triangles. Since such triangles exist for every word hyperbolic group (see 8.3) our geometric argument proves 5.3.E as well as $5.3.E_1$.

(d) Instead of powers of fixed elements γ_i one could use arbitrary sufficiently long closed geodesics (or conjugacy classes in the algebraic context) satisfying some *geometric small cancellation* condition claiming that no geodesic among γ_i follows too closely another geodesic for a considerable amount of time. (See 8.3 for precise statements of this kind.)

5.3.F **<u>Non-free</u> <u>subgroups</u> <u>in</u> Γ:** If Γ is an arithmetic subgroup in an algebraic group A, say defined over $\mathbb{Q}$, then one finds plenty on non-free subgroups in Γ corresponding to $\mathbb{Q}$-subgroups in A. Apart from such Γ there is no systematic way to locate non-free subgroups, though many can be seen in the examples in Section 4. Yet one may suspect there exist word hyperbolic groups Γ with (arbitrarily) large dim $\partial\Gamma$ (here, large is $\geqslant 1$) where every proper subgroup is free.

Here is a general finiteness result for non-free subgroups which we state, for simplicity's sake, for *torsion free* word hyperbolic groups Γ.

5.3.C' Theorem: *Fix a non-abelian finitely generated group Γ_0 which is not a non-trivial free product. Then Γ contains at most finitely many conjugacy classes of subgroups isomorphic to Γ_0.*

To see the idea we assume as earlier $\Gamma = \pi_1(V)$ for $K(V) \leq -\epsilon < 0$, let Γ_0 be the fundamental group of a compact (which makes Γ_0 finitely presented) simplical polyhedron P of dimension two and consider a continuous map $f_0: P \longrightarrow \Gamma$ representing a conjugacy class (in Γ) of a homomorphism $\Gamma_0 \longrightarrow \Gamma$. Then we use a simple but very powerful

Lemma of Thurston: *The* map f_0 *is homotopic to a* map f, *such that the area of (the induced metric in)* P *is bounded by the topology of* P,

$$\text{Area}_f\ P \leq \pi n_2 \epsilon^{-2},$$

where n_2 *is the number of 2-simplices in* P.

Proof: "Straighten" every two simplex of P in V by first homotoping the edges to (locally) geodesic segments in V and then fill in the boundaries of 2-simplices by minimal disks. We see as earlier with the Gauss-Bonnet theorem, that the new 2-simplices have area $\leq \pi\epsilon^{-2}$ and the lemma is proven.

Now, with the bound on area P one easily finds a "short" systems of generators in $\Gamma_0 = \pi_1(P)$ unless Γ_0 decomposes into a free product (see 6.7 in [Gr3] for details) and with our geodesic simplices in 8.3 the geometric argument applies to all word hyperbolic groups.

Example: Let Γ be presented by

$$\{g_1, g_2, g_3 \mid g_3 g_1 g_3^{-1} = w_1\ ,\ g_3 g_2 g_3^{-1} = w_2\}$$

where w_1 and w_2 are "sufficiently long" words in g_1 and g_2, such that the 1/6-condition is satisfied and Γ is hyperbolic. Then we have an infinite sequence of subgroups $\Gamma_i \subset \Gamma$ generated by g_1, g_2 and g_3^i. It is easy to see that these Γ_i are all word hyperbolic with connected boundaries $\partial\Gamma_i$ of $\dim \partial\Gamma_i = 1$. These groups are not pairwise conjugate as they have different images in $\Gamma/[\Gamma,\Gamma]$ and hence, there are among them infinitely many that are pairwise non-isomorphic. Probably, it is not hard to see directly that all Γ_i are pairwise non-isomorphic (with something like the Alexander polynomial).

Remark: In this example every Γ_i is generated by three elements. No such example is possible with two generators according to 5.3.B. On the other hand cyclic coverings of the ramified torus fibered over S^1 (see 4.4) give us geometric examples with rather large number of generators.

5.4 **Monomorphisms into hyperbolic groups**: Thurston's lemma in 5.4.C was generalized and refined by Thurston for maps of polyhedra of dimension $\geqslant 2$ into spaces with $K < 0$. This allows one to sharpen in some cases the finiteness theorem in 5.3.E by replacing conjugation in Γ by conjugation in Γ_0. Here is an instance of such a result which is due to Thurston in the geometric case (compare [Th], [Gr4]).

5.4.A *Let* V *be a closed aspherical (i.e. the universal covering is contractible) manifold with a word hyperbolic fundamental group* Γ. *If* $\dim V \geqslant 3$ *then the exterior automorphism group* $\operatorname{Aut}\Gamma/\operatorname{Int}\operatorname{Aut}\Gamma$ *is finite.*

To grasp the idea let us show that the image of $\operatorname{Aut}\Gamma$ in $\operatorname{Aut}(\Gamma/[\Gamma,\Gamma])$ is finite. Now, $\Gamma/[\Gamma,\Gamma] = H_1(V)$ which is the Poincare dual to $H^{n-1}(V)$ for $n = \dim V$. Recall that the cohomology $H^*(V;\mathbb{R})$ carries a natural pseudo-norm (which may be infinite) corresponding to the L_∞-norm on the space of singular cochains (see [Gr2]). An application of Thurston's lemma shows this pseudo-norm is an actual (i.e. $< \infty$) norm on $H^{k\geqslant 2}(V)$ if $K(V) < 0$ and a similar fact for $H^{k\geqslant 2}(\Gamma;\mathbb{R})$ holds true for all word hyperbolic groups Γ by the properties of the geodesic flow on $\widehat{G}(\Gamma)$ (see 8.3). The norm on

$H^*(\Gamma)$ (obviously) is unvariant under the action of Aut Γ, which implies the finiteness of the image of Aut Γ in $H^{k\geq 2}(\Gamma;\mathbb{R})$. If $n = \dim V \geq 3$ then, by the Poincare duality, the image of Aut Γ in $H_1(\Gamma)$ also is finite.

5.4.B Another standard application of Thurston's simplices (which are constructed for word hyperbolic groups in 8.3) is the following

__Theorem:__ *Let* V *be a compact orientable aspherical pseudomanifold with word hyperbolic fundamental group* Γ *and let* V' *be an arbitrary compact orientable pseudomanifold with* $\dim V' = n = \dim V$. *Then there exists a constant* $d = d(V,V')$, *such that the degree of every continuous map* $f: V' \longrightarrow V$ *is bounded by* $|\deg f| \leq d$. (*Compare* [Gr2] [Gr4].)

By applying this to iterates of a selfmapping $V \longrightarrow V$ one concludes *that* Γ is *not isomorphic to any* of its *proper subgroups*. Probably, the same is true for every word hyperbolic group Γ connected at infinity.

5.4.C It seems very plausible that Thurston's theory of Kleinian groups can be set up in a general hyperbolic context. Namely, we start with an arbitrary finitely generated group Γ and try to describe the Teichmüller space T (or at least some ideal boundary ∂T) of all (non-word) δ-hyperbolic left invariant metrics on Γ whose "curvature" is bounded from below as well as from above (by $-\delta^{-2}$). For example, one may look at complete simply connected manifolds of a fixed dimension with $-\delta^{-2} \geq K(X) \geq -C$ which are discretely and isometrically acted upon by Γ. The pertinent metrics on Γ are induced by orbit maps $\gamma \longmapsto \gamma x$ for some $x \in X$ and metrics associated to different $x \in X$ should correspond (with an appropriate definition) to same point in the Teichmüller space $T(\Gamma)$. In particular, we do not distinguish metrics on Γ which are equivalent under inner automorphisms of Γ. Yet we must keep track of outer automorphisms. Otherwise we get *Riemann moduli space* rather than the Teichmüller space. Thurston's theory applies to the case dim $X = 3$ and $K(X) = -1$.

One of the basic results (see [Th] and [M-S]) is the compactness of the space T under certain conditions on Γ which rule out, for example, surface groups Γ. Notice that Thurston's compactness theorem (which generalizes Mostow rigidity theorem) implies the finiteness of Aut Γ/Int Aut Γ as well as the compactness of the Riemann moduli space. Notice that theorem 5.3.C' also can be regarded as a (rather weak) compactness theorem for the Riemann moduli space of Γ_0.

The first step in the study of $T(\Gamma)$ was made by J. Morgan and the author who found (unpublished) some higher dimensional versions of Thurston compactness theorem. It seems that non-compactness is related to low dimensional (zero and one) pattern in $\partial\Gamma$, in the case Γ is word hyperbolic. (The "zero-dimensional reason" for non-compactness is disconnectedness of $\partial\Gamma$ and the 1-dimensional source of non-compactness probably is due to a presence zero dimensional subsets A in open connected subsets $U \subset \partial\Gamma$, such that $U \setminus A$ is disconnected.) Yet nothing is known about $T(\Gamma)$ in the non-compact case beyond Thurston-Teichmüller theory. One does not even know the structure of the moduli space and (or) of the exterior automorphism groups of fundamental groups of 2-polyhedra with $K < 0$.

5.5 <u>Factor groups of hyperbolic groups</u>: Let Γ be a non-elementary word hyperbolic group and $F \subset \Gamma$ a finite subset. We shall see in this section that there are many hyperbolic factor groups Γ' of Γ for which the quotient map $\Gamma \longrightarrow \Gamma'$ is injective on F.

5.5.A <u>Theorem</u>: *Let $\Gamma_i \subset \Gamma$, $i = 1,\dots,k$, be non-elementary subgroups in Γ. Then there exists a factor group Γ' with the following four properties.*

(1) *Γ' is a non-elementary word hyperbolic group.*

(2) *The canonical epimorphism $\Gamma \longrightarrow \Gamma'$ is injective*

on the subset $F \subset \Gamma$.

(3) *The epimorphism* $\Gamma \longrightarrow \Gamma'$ *sends* Γ_i *onto* Γ' *for all* $i = 1,\dots,k$.

(4) *The epimorphism* $\Gamma \longrightarrow \Gamma'$ *is bijective on the sets of conjugacy classes of non-free (i.e. not infinite cyclic) elementary subgroups in* Γ *and* Γ'. *In particular, if* Γ *is torsion free then also* Γ' *is torsion free.*

An immediate corollary is the following generalization of a theorem by Olshanski (see [Ol]).

5.5.B *Let* Γ *be a torsion free non-elementary word hyperbolic group (e.g.* Γ *is free with two generators). Then there exists a non-Abelian quotient group* $\bar{\Gamma}$ *of* Γ, *such that every proper subgroup in* $\bar{\Gamma}$ *(proper means* $\neq \bar{\Gamma}$ *or* $\{id\}$*) is infinite cyclic.*

5.5.C Consider an infinite elementary subgroup $E_0 \subset \Gamma$ and let $C \subset E_0$ be a maximal infinite cyclic normal subgroup. Then the subgroup mC is normal in E_0 and we define $E_m = E_0/mC$ for all $m = 1,2,\dots$.

5.5.D __Theorem__: *There exists an integer* $m = m(\Gamma,E_0,F)$ *such that for every* $n \geq m$ *there exists a factor group* Γ' *of* Γ *with the following four properties.*

(1) Γ' *is non-elementary word hyperbolic.*

(2) *The homomorphism* $\Gamma \longrightarrow \Gamma'$ *is injective on* F.

(3) *The image* E_0' *of* E_0 *in* Γ' *is isomorphic to* E_n.

(4) *Every torsion element* γ' *in* Γ' *is conjugate to some* $\gamma'' \in E_0'$ *or it is the image of a torsion element* $\gamma \in \Gamma$.

By combining 5.5.A and 5.5.D we immediately obtain the following

5.5.E **Corollary** **(Compare [Ol]):** *Let* Γ *be a torsion free non-elementary word hyperbolic group. Then there exists a sequence of integers* $m_1, m_2, \dots$, *such that for every sequence* $n_i \geq m_i$, $i = 1,2,\dots$ *there exists an infinite factor group* $\bar{\Gamma}$ *of* Γ *containing subgroups* $\bar{\Gamma}_i \subset \bar{\Gamma}$, $i = 1,2,\dots$ *with the following three properties.*

(1) $\bar{\Gamma}_i \approx \mathbb{Z}/n_i\mathbb{Z}$ for $i = 1,2,\dots$.

(2) $\bar{\gamma}\bar{\Gamma}_i\bar{\gamma}^{-1} \cap \bar{\Gamma}_j = \{e\}$ *for all* $\bar{\gamma} \in \bar{\Gamma}$ *and all* $i \neq j$.

(3) *Every proper subgroup* $\bar{\Gamma}'$ *in* $\bar{\Gamma}$ *is conjugate (in* $\bar{\Gamma}$*) to a subgroup in* $\bar{\Gamma}_i$ *for some* $i = i(\bar{\Gamma}')$. *In particular,* $\bar{\Gamma}$ *is a torsion group.*

Remark: We shall see later that the sequence $m_i = m_i(\Gamma)$ can be chosen primitively recursive, but our proof does not allow bounded sequences m_i. One believes however, that every non-elementary word hyperbolic group Γ admits an infinite quotient groups whose elements have finite orders bounded by a constant $m = m(\Gamma)$.

5.5.F **About the proofs**: Everything boils down to showing that adding "sufficiently random" relations to a non-elementary word hyperbolic group gives us a word hyperbolic group again, and short words in the original group are not affected by the relations. To see this consider some closed geodesics in a manifold V and let

$L_1, L_2, \ldots$ be the connected components of the lifts of these geodesics to the universal covering X of V. We assume as earlier that X is convex with $K \leqslant -1$ and then observe that each L_i is an isometrically embedded (geodesic) line in X. (Notice that every geodesic in V usually lifts to infinitely many lines in X.) Fix a number $\epsilon > 0$ and call a segment $[a,b] \subset L_i$ *a piece* if there exists a segment $[a',b']$ in some L_j for $j \neq i$ such that $|a-a'| \leqslant \epsilon$ as well as $|b-b'| \leqslant \epsilon$. Now we can formulate the usual small cancellation conditions,

$SC_1(\epsilon, 1/k)$: None among our closed geodesics in V can be covered by (projections from X) of k pieces.

Less natural but better suited for the moment is

$SC_2(\epsilon, 1/k)$: The length $|a-b|$ of every piece is less than 1/k (length of the underlying closed geodesic).

In what follows, let $\epsilon = 1$ and assume for simplicity's sake that the injectivity radius of V is bounded from below by Rad $V \geqslant 1$. Let Γ be the fundamental group of V and let $\bar{\Gamma}$ be the fundamental group of the space $\bar{V}$ obtained from V by attaching disks to given closed geodesics γ_i in V.

<u>Claim</u>: *If the closed geodesics γ_i in V satisfies* $SC_2(1\ ,\ 00.1)$ *then $\bar{\Gamma}$ is word hyperbolic.*

<u>Proof</u>: We are going to prove that every contractible closed curve f in $\bar{V}$ bounds a disk D with area $\leqslant$ const length f. First, we may assume that f lies in $V \subset \bar{V}$. Next we can (obviously) homotope f to a closed geodesic in V such that the area of this homotopy (that is a map $S^1 \times [0,1] \longrightarrow V$) is bounded by const length f. (Notice, that existence of such homotopy solves the conjugacy problem in Γ if V is compact.) We can assume moreover, that the lift to X of this geodesic contains no pieces [a,b] which are ϵ-close to $[a',b'] \subset L_i$

for some lift L_i of one of the given closed geodesic, such that $|a'-b'| \geqslant 2/3$ (length of the underlying closed geodesic). Otherwise, one could shorten our geodesic in $\bar{V}$ by roughly $|a'-b'|$ and only adding a fixed small amount of area to the spanning disk. Now, with all these assumptions on f we look at a spanning disk $D^2 \longrightarrow \bar{V}$ with $\partial D^2 = f$. The part of this disk inside $V \subset \bar{V}$ is a surface F of genus zero bounded by f and multiples $m_j \gamma_i$ of the given geodesics γ_i. If we chose F *minimal* (compare 5.3.E1) we obtain a surface with $K \leqslant -1$ bounded by geodesics, such that no component $m_j \gamma_i$ of ∂F lies in a small (of order $\epsilon = 1$) neighborhood of a union of n other components for $n \lesssim 200 m_j/3$. It follows that (a small neighborhood of) each $m_j \gamma_i$ contributes at least -10 (where 10 depends on 100) to the total curvature of F. Then, by Gauss-Bonnet, the number N of $m_j \gamma_i$-component satisfies

$$N = -x(F) = -(2\pi)^{-1} \int_F K(F) \geqslant 10N/2\pi,$$

which makes $N = 0$ and shows that our map (curve) $f: S^1 \longrightarrow V$ is constant. Q.E.D.

This argument extends to word hyperbolic groups if "closed geodesics" are replaced by "closed orbit of the geodesic flow" (see 8.3) and instead of minimal F, we use a subdivision of F into ideal triangles bounded by infinite (as they are lifted to X) geodesics which are orbits of the geodesic flow. Now the proof of 5.5.A-D becomes straightforward and details are left to the reader.

Remark: The logical scheme of the results in this section is essentially due to Olshanski (see [Ol]). The simplification we achieve comes, for the most part, from *explicit* use of the notion of hyperbolicity. However, in deeper question (such as construction of various infinite periodic groups with *bounded* torsion exponents) combinatorial techniques by Olshanski can not be matched so far by simple geometrization of ours. Yet, further development of geometric language will, probably, take care of these questions.

5.6 Kazdan's T-groups: Call Γ a *T-group* if for every generating subset $G \subset \Gamma$ there exists an $\epsilon > 0$ such that every isometric action of Γ on the unit sphere S in the Hilbert space satisfies the following ϵ-fixed point property:

if $|gs_0 - s_0| \leqslant \epsilon$ *for some* $s_0 \in S$ *and all* $g \in G$ *then there exists a fixed point of the action of* Γ *on* S.

The existence of infinite T-groups is a non-trivial fact discovered by Kazdan (see [Ka]) who has shown that every lattice in a simple Lie group of $\mathbb{R}$-rank $\geqslant 2$ satisfies property T. Notice, that these lattices are not word hyperbolic and that the (hyperbolic) lattices in O(n,1) and in U(n,1) never have property T. However, the lattices in Sp(n,1) do satisfy T for all $n \geqslant 2$ by a theorem of Kostant (see [Ko]). Thus for every $d = 4n - 1$, $n = 2,3,\ldots$, there exists a word hyperbolic T-group Γ with $\partial\Gamma \approx S^d$.

Observe the following two properties of T-groups (see [Ka]).

5.6.A *If* Γ *satisfies* T *then every factor group* $\bar{\Gamma}$ *of* Γ *satisfies* T.

5.6.B *If an amenable (e.g. abelian) group* Γ *satisfies* T *then* Γ *is finite. In particular, every infinite factor group of a* T*-group is non-amenable.*

Now we apply 5.5.A to the free product $\Gamma * \Gamma_1$ where Γ_1 satisfies T and obtain the following.

5.6.C Theorem: *Every non-elementary word hyperbolic group* Γ *admits a quotient group* $\bar{\Gamma}$ *which is non-elementary word hyperbolic and satisfies the* T*-property. Furthermore if* Γ *is torsion free then* $\bar{\Gamma}$ *can be also assumed torsion free.*

5.6.D Corollary: *The factor group* $\bar{\Gamma}$ *in* 5.5 *can* *be made T and hence, non-amenable.* (Compare [O1].)

Remark: We shall see in 5.5 that the group Γ' has

$$\dim \partial\Gamma' = \max(1, \dim \partial\Gamma).$$

But if want the T-property we only can obtain Γ' with

$$\dim \partial\Gamma' = \max(7, \dim \partial\Gamma)$$

as all known word hyperbolic T-groups Γ have $\dim \partial\Gamma \geqslant 7$. Probably, there exist word hyperbolic T-groups Γ with $\dim \partial\Gamma = 1$ (or at least, with $\dim \partial\Gamma = 2$). Here is another.

Question: Does there exist for a given integer $d \geqslant 2$ an infinite word hyperbolic group Γ (e.g. a lattice in Sp(d,1)) whose every infinite word hyperbolic factor group Γ' has $\dim \partial\Gamma' \geqslant d$? (For $d = 0$ and $d = 1$ the existence of such a Γ is easy.)

5.6.E **Non-hyperbolic quotient group**: According to the above many word hyperbolic groups Γ (e.g. lattices in Sp(2,1)) have no free factor groups. However, this does not prevent factor groups of Γ from the presence of complicated subgroups as the following theorem shows.

Theorem: *Let Γ be a non-elementary word hyperbolic group and Δ an arbitrary finitely generated group. Then there exists a quotient group $\bar{\Gamma}$ of Γ, such that Δ embeds into $\bar{\Gamma}$ and the factor group $\bar{\Gamma}/\mathcal{N}(\Delta)$ is non-elementary word hyperbolic for the normal closure $\mathcal{N}(\Delta)$ of $\Delta \subset \bar{\Gamma}$.*

The proof is obtained by a simple modification of the arguments in 5.5 and is left to the reader.

§6. Trees, triangles and surfaces in hyperbolic spaces

In the following sections 6.1-6.6 we relate the δ-inequality to the geometry of geodesic triangles. Then (see 6.7 and 6.8) we prove that the δ-inequality is equivalent to the isoperimetric inequality Is_2, provided the space in question is geodesic.

Let (X,x_0) be a metric space δ-hyperbolic with respect to x_0. Consider a sequence of points $x_1,x_2,\dots,x_n$ in X for $n = 2^k + 1$ and observe by induction on k that

(*) $$(x_1.x_n) \geqslant \min_{1\leqslant i\leqslant n-1} (x_i.x_{i+1}) - k\delta.$$

Notice that every sequence $x_1,x_2,\dots,x_i,\dots,x_n$ can be obviously reduced to a sequence $x'_1 = x_1,x'_2,\dots,x'_j,\dots,x'_m = x_n$, such that $m \leqslant n$, all x'_j are *distinct* and $\min_j(x'_j.x'_{j+1}) \geqslant \min_i(x_i.x_{i+1})$. Thus we conclude to the following

Lemma: *If X is a finite space containing at most $2^k + 2$ points then every finite sequence of points in X satisfies* (*).

6.1 **Approximating tree TrX:** Let X be δ-hyperbolic for $\delta = 0$ and let us construct a tree TrX as follows. Take the disjoint union of segments $[0,|x|] \subset \mathbb{R}$ for all $x \in X$ and then call two points $t_1 \in [0,|x_1|]$ and $t_2 \in [0,|x_2|]$ equivalent if $|t_1| = |t_2|$ and $(x_1.x_2) \geqslant t_1 = t_2$. Since X is 0-hyperbolic with respect to x_0, this is indeed an equivalence relation and the quotient space is our tree TrX. We equip TrX with the maximal metric (see 1.3) for which the obvious map $[0,|x|] \longrightarrow$ TrX is isometric for all $x \in X$ and observe that the map $X \longrightarrow$ TrX also is isometric. If X is a finite space, then TrX is a finite simplical metric tree and $X \subset$ TrX consists of *the extremal points* in TrX, where a point x in a tree is called extremal if there is no topological embedding of $\mathbb{R}$ into the tree with Image $\mathbb{R} \ni x$.

6.1.A Let us assume that every finite sequence of points $x_1,\ldots,x_n$ in X satisfies

$$(x_1.x_n) \geq \min_i (x_i.x_{i+1}) - \delta'$$

for some fixed $\delta' \geq 0$. Then define a new "scalar product" $(x.y)'$ in X as follows. Consider all sequences $\mathcal{S}$ of points $x_1 = x, x_2, \ldots, x_n = y$ and set

$$(x.y)' = \sup_{\mathcal{S}} \min_i (x_i.x_{i+1}).$$

The corresponding new metric in X,

$$|x-y|' = |x| + |y| - 2(x.y)',$$

obviously is 0-hyperbolic and satisfies

$$|x| = |x|' \text{ for all } x \in X$$

and

$$|x-y| - 2\delta' \leq |x-y|' \leq |x-y|$$

for all x and y in X.

Thus we conclude to the following

6.1.B <u>Proposition</u>: *Let* X *be* δ*-hyperbolic with respect to a reference point* $x_0 \in X$ *and* $\#X \leq 2^k + 2$. *Then there exists a map of* X *into some simplical metric tree, say* $f: X \longrightarrow \mathrm{Tr}X$ *with the following three properties.*

(1) $|f(x)| = |x|$ *for all* $x \in X$.

(2) *The map* f *is short (i.e. distance non-increasing).*

(3) *The codiameter of the map* f *defined by*

$$\text{Codiam } f = \sup_{x,y\in X} (|x-y| - |f(x)-f(y)|)$$

satisfies

$$\text{Codiam } f \leqslant 2k\delta.$$

6.1.C **Remark:** As it is clear from the proof the assumption $\#X \leqslant 2^k + 2$ can be replaced by the following condition:

X admits a decomposition into the union of $2^{k-1} + 1$ subsets $X_i \subset X$, such that each X_i admits an isometric embedding into $\mathbb{R}$ and $x_0 \in X_i$ for all $i = 1,\dots,2^{k+1} + 1$. (In fact, the condition $x_0 \in X_i$ is not truly essential.)

Then there exists a map of X into a tree, $f: X \longrightarrow TrX$, which satisfies the above (1), (2) and (3).

6.2 **Geodesic trees:** Let $Y = \{y_1,\dots,y_n\}$ be a finite subset is a *geodesic* δ-hyperbolic space X and let us sketch a construction of a *geodesic tree* $T \subset X$ with the following three properties.

(1) T is a union of at most n-1 geodesics segments in X.

(2) The set of extremal points of T equals Y.

(3) Every two points y_i and y_j in Y can be joined by a broken geodesic g in T containing $k \leqslant 1 + 2\log_2 n$ segments and satisfying

$$\text{length } g \leqslant |y_i-y_j| + C\delta(\log_2 n)^2,$$

for some constant $C \leqslant 100$. (See Figure 2 for such graphs for n = 3,4,5,6)

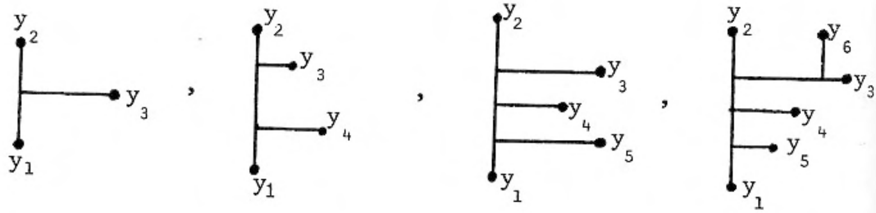

Figure 2

Step 1: Decompose Y into subsets $Y = Y_1 \cup Y_2, \ldots \cup Y_m$ for $m \leqslant \log_2 n$ where Y_1 contains two points and such that for every $i_0 = 1,\ldots,m-1$ and for every two points x and y in Y_{i_0+1} there exist two points x' and y' on the union $U_{i_0} = \bigcup_{i=1}^{i_0} Y_i$, satisfying

$$(x',y')_x + (x',y')_y \leqslant |x-y| + 4(\log_2 n + 1)\delta$$

where $(x',y')_z$ stands for $1/2(|x'-z| + |y'-z| - |x'-y'|)$. The existence of such Y_i is obvious for trees (i.e. for $\delta=0$) and the general case follows via 6.1.B.

Step 2: Let T_1 be a segment between the two points in Y_1 and define (by induction) a geodesic tree $T_{i+1} \subset X$ with the extremal set $Y_1 \cup Y_2 \cup \ldots \cup Y_{i+1}$ as the union of T_i and distance minimizing segments from each point in Y_{i+1} to T_i. One can choose these segments, such that the intersection of any two of them is connected set (usually a single point) containing the T_i-ends of the segments. Then T_{i+1} is inded a tree and a simple induction on i and on n proves our assertion.

6.3 **Thin geodesic triangles**: Consider three pointx x_0, x_1 and x_2 in a metric space X and let T be a (unique up to isometry) simplical metric tree with three extremal points called x_0', x_1', x_2' such that $|x_i'-x_j'| = |x_i-x_j|$ for i,j = 0,1,2. (See Figure 3.)

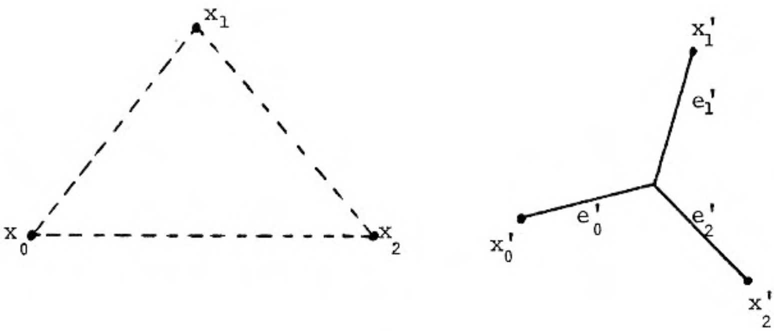

Figure 3

Notice that the edge e_i' in T adjacent to x_i has length $e_i = (x_j.x_k)_{x_i}$.

Let Δ be a geodesic triangle with vertices x_0, x_1, x_2. That is Δ is the union of tree geodesic segments between the pairs of vertices. (Such a Δ always exists, for example, if the space X is geodesic.) The map $\{x_0, x_1, x_2\} \longrightarrow \{x_0', x_1', x_2'\}$ obviously extends to a unique map f_Δ: $\Delta \longrightarrow T$ which is isometric on each segment of Δ. Say that Δ is *δ-thin* if $f_\Delta(x) = f_\Delta(y)$ implies $|x-y| \leqslant \delta$ for all x and y in Δ.

6.3.A <u>**Lemma**</u>: *If X is δ-hyperbolic then Δ is 2δ-thin.*

<u>Proof</u>: We may assume that x lies on the segment $[x_0, x_1]$ and $y \in [x_0, x_2]$. Consider the map f: $\{x_0, x_1, x_2, x, y\} \longrightarrow$ Tr provided by 6.1.B and observe that $f_\Delta(x) = f_\Delta(y)$ implies $f(x) = f(y)$. Hence 6.1.B implies the lemma.

6.3.B Let us prove a converse to 6.3.A. Consider four point x_0, x_1, x_2 and x_3 in X and geodesic segments $[x_i, x_j]$ between these points.

<u>Lemma</u>: *If the geodesic triangles* $[x_0, x_1, x_2]$, $[x_0, x_2, x_3]$ and $[x_0, x_1, x_3]$ *are δ-thin then*

$$(x_1.x_2)_{x_0} \leqslant \min((x_1.x_3)_{x_0},(x_2.x_3)_{x_0}) - 2\delta.$$

Proof: Let $t_0 = \min((x_1.x_3),(x_2.x_3))$ and let $x_i' \in [x_0,x_i]$, $i = 1,2,3$, be some points with $|x_0-x_i'| = t \leqslant t_0$. Since the triangles $[x_0,x_1,x_2]$ and $[x_0,x_2,x_3]$ and δ-thin, we have

$$|x_1'-x_3'| \leqslant \delta \text{ and } |x_2'-x_3'| \leqslant \delta,$$

and then

(*) $$|x_1'-x_2'| \leqslant 2\delta.$$

Since $[x_0,x_1,x_2]$ is δ-thin,

(**) $$|x_1'-x_2'| \geqslant 2(t - (x_1.x_2)_{x_0} - \delta),$$

for all $t = |x_0-x_1'| = |x_0-x_2'|$. Now the lemma follows from (*) and (**) with $t = t_0$.

Let us summarize 6.3.A and B in the following

6.3.C **Proposition**: *Let X be a geodesic metric space. If X is δ-hyperbolic then all geodesic triangles in X are 2δ-thin. Conversely, if all geodesic triangles in X are δ-thin then X is 2δ-hyperbolic.*

6.4 **Hyperbolic geodesic hulls**: Start with a three point metric space $X = \{x_1,x_2,x_3\}$ join every pair of points (x_i,x_j) in X by an edge $[x_i,x_j]$ isometric to $[0,|x_i-x_j|] \subset \mathbb{R}$ and equip the resulting triangle Δ with the maximal metric $|\ |_{2\delta}$ which agrees with the $[0,|x_i-x_j|]$-metric on every edge and for which the triangle Δ is 2δ-thin.

Next, for an arbitrary metric space X we take

$$P^1(X) = \bigcup_{d \geqslant 0} P^1_d(X),$$

where $P^1_d(X)$ is the 1-skeleton of the polyhedron $P_d(X)$ defined in 1.7 and we consider the maximal metric $|\ |_{2\delta}$ in $P^1_d(X)$ for which the canonical map $(\Delta, |\ |_{2\delta}) \longrightarrow (P^d_1(X), |\ |_{2\delta})$ is short (i.e. distance non-increasing) for the above $\Delta = \Delta(x_1, x_2, x_3)$ and for all triples of points x_1, x_2 and x_3 in X.

6.4.A __Lemma__ **(Compare [Dr]):** *If* X *is* δ-*hyperbolic then the above map* $\Delta \longrightarrow P^d_1(X)$ *is isometric for all* x_1, x_2 *and* x_3 *in* X.

__Proof:__ It is enough to consider the case of a four point space X where the proof is obvious.

6.4.B __Corollary:__ *If* X δ-*hyperbolic then the canonical embedding* $X \longrightarrow (P^1(X), |\ |_{2\delta})$ *is isometric and the space* $(P^1(X), |\ |_{2\delta})$ *is* 4δ-*hyperbolic.*

6.4.C Let $\bar{\Delta}$ be the 2-simplex bounded by Δ and let us take a specific simplical metric in $\bar{\Delta}$ for which Δ is thin. Namely, start with the product $\Delta \times [0,2\delta]$ and then collapse $\Delta \times 2\delta$ to the tripod T as in Figure 3. The resulting space is homeomorphic to $\bar{\Delta}$ and we take the maximal metric as Δ for which the (quotient) map $\Delta \times [0,2\delta] \longrightarrow \bar{\Delta}$ is short. Clearly, every edge of $\Delta = \Delta \times 0 = \partial\bar{\Delta} \subset \bar{\Delta}$ is geodesic for this metric and Δ is 4δ-thin. It is also clear that the induced metric on $\Delta \subset \bar{\Delta}$ is $\geqslant |\ |_{2\delta}$.

Now, let

$$P^2(X) = \bigcup_{d \geqslant 0} P^2_d(X)$$

and let $|\ |'_{2\delta}$ be the maximal metric on $P^2(X)$ for which the (canonical) embedding $\bar{\Delta} \longrightarrow P^2(X)$ is short for all $\bar{\Delta} = \bar{\Delta}(x_1, x_2, x_3)$. One sees with the above discussion:

6.4.D **Proposition:** *Let X be δ-hyperbolic. Then* $(P^2(X), |\ |'_{2\delta})$ *is a* δ'*-hyperbolic metric space for* $\delta' = 12\delta$ *and the canonical embedding* $X \longrightarrow P^2(X)$ *is isometric. Furthermore the space* $P^2(X)$ *is geodesic. In fact, the metric in* $P^2(X)$ *is simplical for some subdivision of* $P^2(X)$.

6.4.E **Remark:** The map between the hyperbolic boundaries $\partial X \longrightarrow \partial P^2(X)$ induced by the embedding $X \longrightarrow P^2(X)$ is a *homeomorphism* as it easily follows from the construction of $P^2(X)$.

6.5 **The inscribed triangle** Δ^{in}: Consider a geodesic triangle Δ in X with edges e_i, $i = 1,2,3$ and recall the map f_Δ of Δ onto the tree T which is a tripod whose three extremal points correspond to the vertices of Δ (see Figure 3). Then there is a unique triple of points, called $\Delta^{in} = \{y_1,y_2,y_3\}$ for $y_i \in e_i$, such that all three points go under f_Δ to the central vertex of the tripod T (see Figure 3 and Figure 4 below).

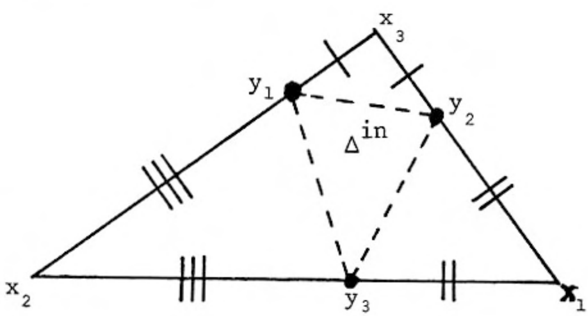

Figure 4

Define

$$\text{insize } \Delta = \text{Diam } \Delta^{in}$$

and observe that every δ-thin triangle obviously has insize $\leqslant \delta$.

6.5.A Suppose that every two points in Δ, say

$y'_1 \in e_1$ and $y'_2 \in e_2$, satisfying $f_\Delta(y'_1) = f'_\Delta(y'_2)$ are joined by a geodesic segment $[y'_1, y'_2]$ in Δ. Denote by Δ' the resulting "subtriangle" $[y'_1, y'_2, x_3]$.

<u>Lemma</u>: *If Δ and all "subtriangles" Δ' have* insize $\leqslant \delta$ *then the triangle Δ is 2δ-thin.*

<u>Proof</u>: We must show for all $y_1^t \in e_1$ and $y_2^t \in e_2$ with

$$|x_3 - y_1^t| = |x_3 - y_2^t| = t \leqslant t_0 = (x_1 . x_2)_{x_1}$$

that

(*) $$|y_1^t - y_2^t| \leqslant 2\delta.$$

If $t = t_0$, then

(**) $$|y_1^t - y_2^t| \leqslant \delta,$$

because insize $\Delta \leqslant \delta$. Then we define $t_i = t_{i-1} - \delta_{i-1}$ where $\delta_{i-1} = \min[y_1^{t_{i-1}}, y_2^{t_{i-1}}, x_3]$ for $i = 1,2,\ldots,$ and obtain by induction on i the inequality (**) for $t = t_0, t, \ldots, t_i, \ldots$. Now (*) follows for all t by the triangle inequality.

6.5.B <u>Corollary</u>: *If all geodesic triangles in a geodesic space X have* insize $\leqslant \delta$ *then X in 4δ-hyperbolic.*

6.6 <u>The minimal size of Δ</u>: Take an arbitrary set Δ in X covered by tree subset $e_i \subset \Delta$, $i = 1,2,3$, and consider triples of points $\Delta' = \{y_1, y_2, y_3\}$ in Δ with $x_i \in e_i$ for $i = 1,2,3$. Define

$$\text{minsize } \Delta = \inf_{\Delta'} \text{Diam } \Delta'.$$

Observe that minsize does not increase under short maps $\Delta \longrightarrow \Delta'$

sending e_i to e'_i, that is

$$\text{minsize } \Delta' \leqslant \text{minsize } \Delta.$$

Also notice that every geodesic triangle Δ covered by the three edges has

$$\text{minsize } \Delta \leqslant \text{insize } \Delta.$$

6.6.A **<u>Lemma</u>:** *Every geodesic triangle Δ satisfies*

$$\text{insize } \Delta \leqslant 2 \text{ minsize } \Delta.$$

<u>Proof</u>: The position of tree points $y_i \in e_i$ are determined by the six distances d_i^j, for $i \neq j$, to the vertices x_j of Δ satisfying three linear equations

$$(*) \qquad \begin{cases} d_1^2 + d_2^3 = |x_2 - x_3| \\ d_2^1 + d_2^3 = |x_1 - x_3| \\ d_3^1 + d_3^2 = |x_1 - x_2|, \end{cases}$$

see Figure 5.

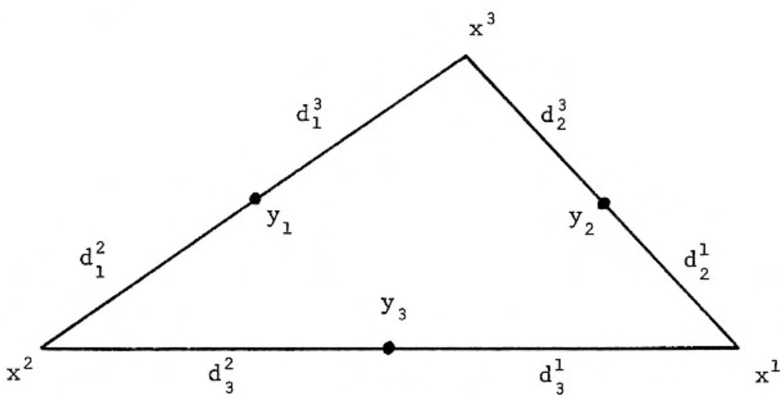

Figure 5

Furthermore, the vertices of Δ^{in} are determined by the numbers d_i^j satisfying (*) and the following three equations

(**) $$d_i^j - d_k^j = 0, \; i \neq k \neq j.$$

If three points $\bar{y}_i \in e_i$ have $|\bar{y}_i - \bar{y}_j| \leqslant \delta$, then, by the triangle inequality, the corresponding distances $\bar{d}_i^j$ satisfy

(***) $$|\bar{d}_i^j - \bar{d}_k^j| \leqslant \delta \text{ for } i \neq k \neq j.$$

Since $\bar{d}_i^j$ also satisfy (*) the solution $\{d_i^j\}$ of (*) + (**) satisfies (by an easy computation)

$$|\bar{d}_i^j - d_i^j| \leqslant \delta/2.$$

Hence, by the triangle inequality

$$\text{insize } \Delta \leqslant 2\delta$$

and the lemma follows.

6.6.B __Corollary__: *If all geodesic triangles* Δ *in a geodesic metric space* X *have* minsize $\Delta \leqslant \delta$, *then* X *is* 8δ*-hyperbolic.*

6.6.C Let us relate minsize Δ to the area of the minimal disk D in X with $\partial D = \Delta$.

6.6.D __Lemma__: *Let* $\Delta_0 \subset \mathbb{R}^2$ *be a regular triangle with the edge length* d_0. *If* minsize $\Delta \geqslant \sqrt{3/2}\, d_0$ *then there exists a short map* $f\colon X \longrightarrow \mathbb{R}^2$ *such that*

(a) *The vertices* x^j *of* Δ *go to the vertices* x_0^j *of* Δ_0 *and the edges of* Δ *go to the corresponding edges of*

Δ_0.

(b) *The image of* f *lies inside* Δ_0.

Proof: Start with the map F: X $\longrightarrow \mathbb{R}^3$ given by

$$F(x) = (1/\sqrt{3})(\text{dist}(x,e_1),\text{dist}(x,e_2),\text{dist}(x,e_3))$$

for the tree edges e_i of Δ. Clearly the map F is short, the image of F lies in the region

$$\{z_1,z_2,z_3 \mid z_i \geqslant 0,\ z_1+z_2+z_3 \geqslant \bar{\delta}/\sqrt{3}\} \subset \mathbb{R}^3 \text{ for } \bar{\delta} = \text{minsize } \Delta,$$

and the edge e_i goes to the plane $z_i = 0$ in $\mathbb{R}^3$ for i = 1,2,3. We compose F with the radial projection from the origin of $\mathbb{R}^3$ to the simplex

$$\bar{\Delta}_0 = \{z_1,z_2,z_3 \mid z_i \geqslant 0,\ z_1+z_2+z_3 = d_0/\sqrt{3}\}$$

thus obtaining the required map f.

6.6.E **Corollary**: *Let* D *be a compact surface in* X *with* $\partial D = \Delta$. *Then*

$$\text{area } D \geqslant (1/2\sqrt{3})(\text{minsize } \Delta)^2$$

Proof: Since Δ is homeomorphic to S^1 and $H_1(D,\Delta;\mathbb{Z}_2) \neq 0$, the map f sends D *onto* Δ_0. Hence,

$$\text{area } D \geqslant \text{area } \Delta_0 = \sqrt{3}/4\ d_0^2$$

(See [Gr1] and [Gr3] for sharper results of this kind.)

Remark: If a triangle Δ in X bounds a surface D, such that

$$\text{area } D \leqslant \epsilon_0^2 (\text{Diam } \Delta)^2,$$

then

(*) $$\text{minsize } \Delta \leqslant \epsilon_0' \text{ Diam } \Delta$$

for $\epsilon_0' = \epsilon_0 \sqrt{2/3}$, as 6.6.E shows. One can prove by an elementary (and rather tedious) argument the following

6.6.F __Proposition:__ *There exists a positive ϵ_0' (one can take $\epsilon_0' = 10^{-3}$) with the following property. If all geodesic triangles Δ in X with Diam $\Delta \geqslant R_0$ for a given $R_0 > 0$ satisfy (*) then X is hyperbolic.*

We shall prove this (as well as Theorem 2.3.D) in Section 6.8 by using the Riemann mapping theorem and elementary properties of conformal metrics presented in the next section.

6.7 __Analytic lemmas:__ Let $D(x,r) \subset \mathbb{R}^2$ be the disk $\{y \in \mathbb{R}^2 \mid \|x-y\| \leqslant r\}$ and $S(x,r)$ be the boundary circle $\partial D(x,r)$. Consider a positive function μ in the disk $D(x,R_0)$ for some $R_0 \geqslant 0$ and let

$$A(r) = A_\mu(x,r) = \int_{D(x,r)} \mu^2 dy$$

and

$$L(r) = L_\mu(x,r) = \int_{S(x,r)} \mu ds,$$

where $r \leqslant R_0$ and ds is the length parameter on the circle $S(x,r)$.

6.7.A *If*

(1) $$L^2(r) \geq 4\pi c A(r)$$

for $r_0 \leq r \leq R_0$ *and some* r_0 *and* $c \geq 0$, *then*

(1') $$A(r) \geq A(r_0)(r/r_0)^{2c}$$

for $r_0 \leq r \leq R_0$.

__Proof__: By Schwartz inequality the derivative A' of A satisfies

(2) $$A'(r) = \int_{D(r)} \mu^2 \geq (2\pi r)^{-1}\Big[\int_{S(r)} \mu\Big]^2 = (2\pi r)^{-1}L^2(r).$$

Then, by (1)

(3) $$A'(r) \geq 2cr^{-1}A(r)$$

and (1') follows.

6.7.B Let $\Omega \subset \mathbb{R}^2$ be a bounded domain and set $d(x) = \mathrm{dist}(x,\partial\Omega)$ for $x \in \Omega$. Next for a continuous positive function μ on Ω define $r(x,A)$ for $A \geq 0$ by the equation

$$A_\mu(x,r(x,A)) = A,$$

provided

$$A_\mu(x,d(x)) \geq A.$$

If $A_\mu(x,d(x)) < A$ we set $r(x,A) = d(x)$.

6.7.C *Let some constants* $a > 1$, $c > 1$ *and* $\alpha > 0$ *satisfy*

(4) $$\frac{a^{\frac{1}{c}}}{a-1} = \alpha < 1.$$

Let $A_0 > 0$ *be another constant, such that*

$$\left[\frac{r(x,a^2A_0)}{r(x,A_0)}\right]^{2c} \leqslant a^2, \tag{5}$$

for all those $x \in \Omega$, *where* $r(x,a^2A_0) < d(x)$. *Then*

$$r(x,A_0) \geqslant (1-\alpha)a^{-\frac{1}{c}}d(x), \tag{6}$$

for all $x \in \Omega$.

<u>Proof</u>: Suppose (6) is violated by some $x_0 \in \Omega$ where

$$r_0 = r(x_0,A_0) < (1-\alpha)a^{-\frac{1}{c}}d(x_0). \tag{$\bar{6}$}$$

To bring ($\bar{6}$) to the contradiction we first observe that the average Av of $A(x,\bar{r}_1)$ over $x \in D(x,R_0)$ satisfies for $\bar{r}_1 = \alpha r_0$ and $R_0 = r(x_0,a^2A_0)$,

$$\mathrm{Av} \geqslant \frac{\bar{r}_1^2 A(x_0,R_0)}{(R_0+\bar{r}_1)^2} \geqslant \frac{\bar{r}_1^2 a^2 A_0}{(r_0 a^{\frac{1}{c}}+\bar{r}_1)^2} = A_0, \tag{7}$$

as it follows from (5). Therefore, there exists a point $x_1 \in D(x_0,R_0) \subset \Omega$, such that

$$r_1 = r(x_1,A_0) \leqslant \bar{r}_1 = \alpha r_0. \tag{8}$$

Since

$$\|x_1-x_0\| \leqslant R_0 \leqslant a^{\frac{1}{c}}r_0.$$

we obtain with $(\bar{6})$,

(9) $$d(x_1) \geqslant d(x_0) - a^{\frac{1}{c}} r_0 > \alpha d(x_0).$$

Hence x_1 also violates (6). Then we pass from x_1 to x_2, x_3 and so on and obtain a sequence $x_i \in \Omega$, such that

(10) $$r_i = r(x_i, A_0) < \alpha^i r_0$$

and

$$\|x_i - x_{i+1}\| < a^{\frac{1}{c}} r_i \leqslant a^{\frac{1}{c}} \alpha^i r_0.$$

Hence, the points x_i lie in the disk $D(x_0, \rho)$ for some ρ, such that

$$\rho < a^{\frac{1}{c}} r_0 \,(1 + \alpha + \alpha^2 + \ldots) = \frac{a^{\frac{1}{c}} r_0}{1 - \alpha} < d(x_0).$$

As the function μ is bounded on $D(x_0, \rho)$ the numbers r_i are bounded away from zero. But this contradicts to (10) as $\alpha < 1$ and the proof is concluded.

6.7.D Consider an equilateral triangle $\bar{\Delta}$ in $\mathbb{R}^2$ and a function $\bar{\mu}$ on $\mathbb{R}^2$ which is positive continuous on $\bar{\Delta}$ and zero outside $\bar{\Delta}$. Let the inequality

(11) $$L_{\bar{\mu}}^2(x,r) \geqslant 4\pi c A_{\bar{\mu}}(x,r)$$

be satisfied by those $x \in \mathbb{R}^2$ and $r \geqslant 0$, for which

(a) the disk $D(x,r) \subset \mathbb{R}^2$ meets at most one edge of $\bar{\Delta}$;

(b) $A_0 \leqslant A_{\bar\mu}(x,r) \leqslant a^2 A_0,$

where $c > 1$, $a > 1$ and $A_0 > 0$ are given constants.

Denote by $\bar T \subset \bar\Delta$ the union of the three segments in $\bar\Delta$ joining the center of $\bar\Delta$ with the centers of the edges of $\bar\Delta$. The length of these segments (that is the radius of the inscribed circle in $\bar\Delta$) is denoted by $\bar d$.

6.7.E *Let* $\alpha = \frac{a^{\frac{1}{c}}}{a-1}$ *be* <1. *Then for every point* $x \in \bar T$ *there exists a number* $\bar r = \bar r(x)$ *in the interval*

(12) $$\bar d \geqslant \bar r \geqslant (1-\alpha) a^{-\frac{1}{c}} \bar d,$$

such that the integral

$$A_{\bar\mu}(x,\bar r) = \int_{D(x,\bar r)} \mu^2 dy$$

satisfies

(13) $$A_{\bar\mu}(x,\bar r) \leqslant A_0.$$

<u>Proof</u>: Let $\bar e$ be the edge of $\bar\Delta$ nearest to a given point $x \in \bar T$. Consider the domain Ω which is the union of $\bar\Delta$ and the reflection of $\bar\Delta$ in $\bar e$. Now 6.7.E follows from 6.7.A and 6.7.C applied to an arbitrary small perturbation of $\bar\mu$ which makes $\bar\mu$ positive and continuous in Ω.

6.7.F <u>**Corollary**</u>: *For every* $x \in \bar T$ *there exists a number* r *in the interval*

(14) $$\bar d \geqslant r \geqslant 1/2(1-\alpha)\, a^{-\frac{1}{c}} \bar d,$$

such that the integral $L_{\bar{\mu}}(x,r)$ *of* μ *over the circle* $S(x,r)$ *satisfies*

$$L_{\bar{\mu}}(x,r) \leq 2\sqrt{\pi A_0} \tag{15}$$

Proof: Apply Schwartz inequality.

6.7.G Let $\bar{\Delta}_{\bar{\mu}}$ be the triangle $\bar{\Delta}$ with the Riemannian metric $\bar{\mu}^2(dx^2+dy^2)$.

6.7.G1 *The minimal size of* $\bar{\Delta}_{\bar{\mu}}$ *satisfies*

$$\text{minsize } \bar{\Delta}_{\bar{\mu}} \leq m_0 = 2(3+(1-\alpha)^{-1}a^{\frac{1}{c}})\sqrt{\pi A_0}. \tag{16}$$

Proof: For every piecewise smooth ark γ in $\bar{\Delta}$ between two points x_1 and x_2 in $\bar{\Delta}$ the $\bar{\mu}$-distance $|x_1-x_2|_{\bar{\mu}}$ and the $\bar{\mu}$-length $L_{\bar{\mu}}(\gamma) = \int_\gamma \bar{\mu}(x)d\gamma$ satisfy

$$|x_1-x_2|_{\bar{\mu}} \leq L_{\bar{\mu}}(\gamma).$$

Therefore, in order to prove 6.7.G1 one must find three arks γ_i between three points on the edges of $\bar{\Delta}$, such that

$$L_{\bar{\mu}}(\gamma_i) \leq m_0 \text{ for } i = 1,2,3.$$

This is done by first covering the tripod $\bar{T} \subset \bar{\Delta}$ by the minimal number of disks with centers in $\bar{T}$ whose radii r satisfy (14) and $\bar{\mu}$-lengths of the boundary circles satisfy (15). The total number N of the disks obviously satisfy

$$N \leq N_0 = 3(1+(1-\alpha)^{-1}a^{\frac{1}{c}}).$$

Now, one constructs an inscribed triangle using arks of boundary circles as in Figure 6 below

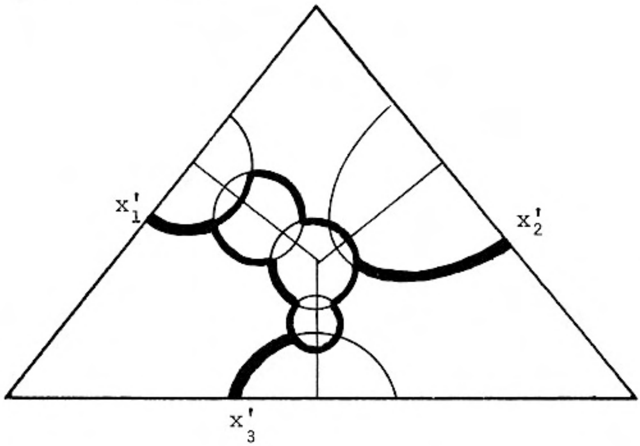

Figure 6

Notice that $\bar{\Delta}'$ degenerates at the circles adjacent to the edges of $\bar{\Delta}$, but in any case, this $\bar{\Delta}'$ obviously has the sides γ_i with

$$L_\mu(\gamma_i) \leqslant m_0' = 2N_0\sqrt{\pi A_0}.$$

The number m_0' is about three times greater than m_0 in (16). The (obvious) improvement of the argument giving us m_0 rather than m_0' (which is not essentially important) is left to the reader.

6.8 Minimal surfaces, conformal maps and a length-area criterion for hyperbolicity: Consider a pair (V,f) where V is a smooth n-dimensional manifold and f is a continuous map of V into a metric space X. We call such a pair a (singular) *manifold* V *in* X and we define the *volume* of V by considering all those Riemannian metrics g on V for which the map f: (V,g) ⟶ X is short (i.e. distance non-increasing) and then by taking the infimum of total Riemannian volumes of the manifolds (V,g), that is

$$\mathrm{Vol}_X V = \inf_g \mathrm{Vol}(V,g).$$

We are especially interested in the cases of *curves* in X that are

1-dimensional manifolds and of surfaces where $n = \dim V = 2$. The volume of a curve is, as usual, called the length and for surfaces it is the area.

<u>**Example**</u>: By projecting the unit circle $S^1 \subset \mathbb{R}^2$ to a coordinate axes one gets a circle S^1 in $\mathbb{R}$ of length = 2.

Next we define *the (filling) area* of a circle S^1 in X as the lower bound of areas of disks D in X bounded by S^1(that are maps $D \longrightarrow X$ extending the structure map $S^1 \longrightarrow X$ from $S^1 = \partial D$).

6.8.A Let P be a polyhedron with a simplical metric. A manifold V in P is called *simplical* if the implied map $V \longrightarrow P$ is simplical for some triangulation of V. The volume of a simplical V in X (obviously) equals the sum of volumes of n-simplices in P covered by the map $V \longrightarrow P$, where each simplex is counted as many times as it covered by $V \longrightarrow P$.

6.8.B In what follows P is a 2-dimensional polyhedron where all 1-simplices have the same length denoted d. Every simplical curve S in P has length L(S) which is an integer multiple of d, that is $L(S) = N_1 d$ and the area of every simplical surface D in P is $A(D) = N_2 \frac{\sqrt{3}}{4} d^2$ for some $N_2 = 0,1,2,\dots$.

Observe that every simple closed curve, that is a circle $S \subset P$, can be homotoped to a simplical circle S' in P whose length L(S') and area A(S') are related to those of S by the inequalities

(17) $$L(S') \leqslant 2L(S)$$

and

(17') $$|A(S')-A(S)| \leqslant (1/2)dL(S).$$

This is done by deforming the intersections of S with the 2-simplices in P to the boundaries of these simplies as it is shown in Figure 7.

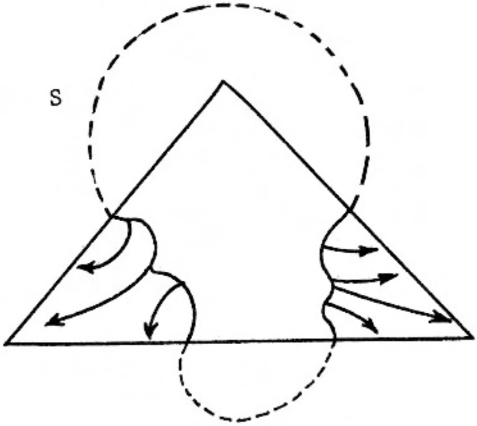

Figure 7

6.8.C **Lemma:** *Let* $A_0, A_1 \geq 4A_0$ *and* $C \leq A_0 d^{-2}$ *be positive numbers, such that every simplical circle* S' *in* P *with the area in the interval*

$$2A_0 \leq A(S') \leq (1/2)A_1$$

satisfies

(18) $$(L(S'))^2 \geq 8CA(S').$$

Then every (continuous) circle S in P with

$$A_0 \leq A(S) \leq A_1$$

satisfies

(19) $$(L(S))^2 \geq CA(S).$$

Proof: Homotope S to S' which satisfies (17) and (17'). Since the inequalities $dL(S) \geq A(S)$ and $C \leq A_0 d^{-2} \leq d^{-2}A(S)$ imply (19), we may assume that $dL(S) \leq A(S)$. Then the inequality (17') implies

$$2A_0 \leq A(S') \leq 1/2A_1$$

and (19) follows from (17), (17') and (18). Q.E.D.

6.8.D __Minimal surfaces__: A disk D in X is called *minimal* if

$$\operatorname{area}_X D = \operatorname{area}_X \partial D < \infty.$$

If X is a "reasonable" space then every contractible circle S in X of finite area bounds a minimal disk. For example, if X is a two-dimensional polyhedron with a simplical metric where all 1-simplices have the same length d, then, obviously, every simplical curve S bounds a minimal disk D in X. In fact, one can find a *simplical* minimal disk D with $\partial D = S$. On the other hand if the 1-simplices have different lengths, then this need not be true as simple examples (of 2-dimensional not locally finite polyhedra) show.

To avoid the existence problem we consider ϵ-*minimal* disks D, where area $D \leqslant$ area $\partial D + \epsilon$ for a small positive ϵ which eventually goes to zero. This ϵ-minimality is good enough for our application given below.

6.8.E __Conformal maps__: A disk D in X is called *conformal* with respect to a fixed Riemannian metric g_0 on D if there exists a non-negative Borel function μ on D, such that the X-length of every smooth curve $S \subset D$ equals the μ-length of S that is the integral $\int_S \mu ds$ for the Riemannian length ds on S, and the X-area of every open $\Omega \subset D$ equals the μ-area $= \int_\Omega \mu^2$.

Consider a curvilinear triangle Δ in X that is a circle subdivided into three arks or edges. Take a plane regular triangle $\bar{\Delta}$ (in place of D) and then look at *triangles* $\bar{\Delta}$ *in* X *bounded by* Δ. These by definition are continuous maps $f: \bar{\Delta} \longrightarrow X$ sending the i-th edge of $\bar{\Delta}$ onto the i-th edge of Δ for some labeling of the edges by $i = 1,2,3$. In most "reasonable" cases the Riemann mapping theorem ensures the existence of a *conformal* minimal (or ϵ-minimal) $\bar{\Delta}$ in X bounded by a given contractible Δ. This is so, for example, if Δ is a simple closed curve in the 1-skeleton of the polyhedron P in 2.6.A. In any case, Rieman's theorem provides, by a trivial approximation argument, an ϵ-minimal map $f\colon \bar{\Delta} \longrightarrow X$ bounded by Δ, such that f also is ϵ-conformal, for some positive continuous function $\bar{\mu}$ on $\bar{\Delta}$. Namely,

$$|\,\mathrm{length}_{\bar{\mu}}\, S - \mathrm{length}_X\, S\,| \leqslant \epsilon$$

for all circle arks $S \subset \bar{\Delta}$ and

$$|\,\mathrm{area}_{\bar{\mu}}\, \Omega - \mathrm{area}_X\, \Omega\,| \leqslant \epsilon$$

for all open $\Omega \subset \bar{\Delta}$, where one can take $\epsilon > 0$ arbitrarily small. This, for $\epsilon \longrightarrow 0$, allows us to use the analytic lemmas in 6.7 (applied to the extension of μ by $\equiv 0$ on $\mathbb{R}^2 \setminus \bar{\Delta}$) in so far as we can control the X-lengthof curves and areas in $\bar{\Delta}$. The pertinent curves are topological circles S in $\bar{\Delta}$ formed by two segments S_1 and S_2, where S_1 lies in an edge of $\bar{\Delta}$ and S_2 is a circle ark in Δ (see Figure 8).

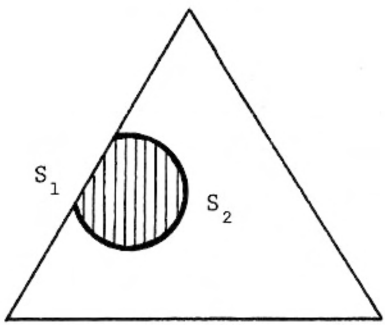

Figure 8

If, for example, the triangle Δ is geodesic, then

$$\mathrm{length}_X\, S_1 \leqslant \mathrm{length}_X\, S_2$$

and thus

(20) $$\text{length}_X S \leq 2 \text{ length}_X S_2.$$

This ensures a bound on the X-area of the disk $D \subset \bar{\Delta}$ bounded by S in term of $\text{length}_X S_2$, provided a similar bound for length S is available. More generally, call Δ a λ-*geodesic* triangle for some $\lambda \geq 1$ if every two points x_1 and x_2 lying in an edge of Δ bound a segment γ in this edge, such that

$$\text{length}_\gamma \leq \lambda | x_1 - x_2 | .$$

If Δ is λ-geodesic then instead of (20) we have

(20) $$\text{length}_X S \leq (1+\lambda)\text{length}_X S_2,$$

which serves the same purpose as (20).

<u>**Example**</u>: Let every edge of Δ be a minimal geodesic in the 1-skeleton of the above polyhedron P. Then, obviously, Δ is 2-geodesic.

6.8.F The above discussion allows us to apply the results in 6.7, in particular, Lemma 6.7.6, to an arbitrary metric space X as follows. Let every circle S in X with the area in the interval

$$A_0 \leq A(S) \leq A_1,$$

for given positive numbers A_0 and $A_1 \geq A_0$ satisfy

$$L^2(S) \geq CA(S),$$

where L(S) denotes the length of S. Let a number $\lambda \geq 1$ be given and suppose there exist numbers $a > 1$, $c > 1$ and $\alpha < 1$ with the following three properties

(i) $A_1 \geq a^2 A_0$,

(ii) $C \geq 4c\pi(1+\lambda)^2$,

(iii) $\alpha = \dfrac{a^{\frac{1}{c}}}{a-1}$.

Now, Lemma 6.7.G_1 leads to the following

6.8.G <u>Proposition</u>: *Let Δ be a λ-geodesic triangle in X of finite area. Then the minimal size of Δ is bounded by*

$$\text{minsize } \Delta \leq m_0 = 2(3+a^{\frac{1}{c}}(1-\alpha)^{-1})\sqrt{\pi A_0}.$$

6.8.H <u>Corollary</u>: *If* $A_1 \geq 16A_0$ *and* $C \geq 48(1+\lambda)^2$, *then*

$$\text{minsize } \Delta \leq 24\sqrt{A_0}.$$

<u>Proof</u>: Take $a = 4$ and $c = \dfrac{C}{4\pi(1+\lambda)^2}$.

6.8.I Let us return to the 2-dimensional polyhedron P with the simplices of size d. Let every simplical circle S' in P with the area in the interval

$$A_0' \leq A(S') \leq 64A_0'$$

for some $A_0' > 0$ satisfies

$$(L(S'))^2 \geq C'A(S')$$

for $C' = 4000$. Then consider a curvilinear triangle Δ in P, where each edge is either a geodesic segment in P or in the 1-skeleton P^1 of P. Notice that such a Δ is λ-geodesic in P for $\lambda = 2$.

By applying the above we obtain the following

6.8.J <u>**Theorem**</u>: *If* Δ *is contractible and* $A_0' \geqslant 500d^2$, *then*

$$\text{minsize } \Delta \leqslant 24\sqrt{A_0'}.$$

This yields with 6.6,

6.8.K <u>**Corollary**</u>: *If* P *is simply connected (and, hence, every* Δ *is contractible) then* P *is* δ*-hyperbolic for* $\delta = 200\sqrt{A_0'}$ *and* P^1 *is* 2δ*-hyperbolic for the simplical metric in* P^1 *with the simplices of size* d.

6.8.L <u>**Remark**</u>: Suppose the above polyhedron P equals the 2-skeleton of $P_d(X)$ for some geodesic metric space X (see 1.7). Then every geodesic triangle Δ in X gives rise to a geodesic triangle Δ' in the 1-skeleton $P^1 = P_d^1(X)$, such that

$$\text{minsize } \Delta' \geqslant \text{minsize } \Delta.$$

this leads to the following

6.8.M <u>**Length-area criterion for hyperbolicity**</u>: *Let for some* $d \geqslant 0$ *and* $A_0' \geqslant 500d^2$ *every simplical circle* S' *in* $P_d^1(X)$ *with the area in the interval*

$$A_0' \leqslant A(S') \leqslant 64A_0'$$

satisfies

$$L^2(S') \geqslant 4000A(S),$$

where the areas are measured in $P_d^2(X) \supset P_d^1(X)$. *If* $P_d(X)$ *is simply connected, then* X *is* δ*-hyperbolic for*

$\delta = 400\sqrt{A_0'}$.

Observe that we have already proven (see 2.3) the converse which can be expressed in the present language as follows.

6.8.N *If* X *is* δ*-hyperbolic and* $d \geq 8\delta$ *then every* $S' \subset P_d^1(X)$ *satisfies*

$$L(S') \geq (d/4\sqrt{3})A(S').$$

6.8.P __The proof of 2.3.D__: If we apply the above results to the 2-dimensional polyhedron associated to the (triangular!) presentation of Γ and take into account the discrepancy factor $9\frac{4}{\sqrt{3}}$ between the two areas, we immediately obtain 2.3.D with exception of the inequality $A(f) \leq CL(f)$. This as well as the (now obvious) proof of 2.3.F is left to the reader.

6.8.Q __The proof of 6.6.F__: The inequality (*) implies the δ-hyperbolicity of every ball B in X of radius $R \geq (\epsilon_0')^{-1}R_0$ for $\delta \leq 8\epsilon_0'R$. This implies an isoperimetric inequality for disks in this ball, which, in turn, yields the hyperbolicity of X.

6.8.R Let us characterize the hyperbolicity by the properties of conformal minimal maps. Assume X is either a Riemannian manifold or a locally compact finite dimensional polyhedron with a simplical metric and let the isometry group Is X be cocompact on X. A simple compactness argument shows that

6.8.S *The universal covering of* X *is hyperboic if and only if every conformal minimal map* $\mathbb{R}^2 \to X$ *is constant (where a map* $\mathbb{R}^2 \to X$ *is called minimal if its restriction to every disk* $D \subset \mathbb{R}^2$ *is minimal).*

6.8.T This proposition suggests the following stronger form of hyperbolicity. Call a map f of a surface S into X *locally minimal* if each point $s_0 \in S$ is contained in a disk $D_0 \subset S$ such that the restriction of f to D_0 is minimal. Notice that for 2-polyhedra X the local minimality is a (local) *topological* invariant of f.

Let X be a space as above. Then a simple argument yields

6.8.U *The following three conditions are equivalent.*

(1) *Every conformal locally minimal map* $\mathbb{R}^2 \longrightarrow X$ *is constant.*

(2) *There exists a constant* $C \geqslant 0$ *such that every locally minimal disk* D *in* X *satisfies*

$$\text{Area } D \leqslant C \text{ length } \partial D.$$

(3) *There exists a constant* $\lambda > 0$ *such that every conformal locally minimal map* $f: D(R) \longrightarrow X$ *for* $D(R) = \{x \in \mathbb{R}^2 \mid \|x\| \leqslant R\}$ *satisfies*

$$\text{Diam}_X f(D(1/2R)) \leqslant \lambda.$$

Remark: A space X satisfying (2) is called *locally hyperbolic* (which is stronger than hyperbolic). Notice that locally hyperbolic 2-polyhedra has been extensively studied (under various names) *in the small cancellation theory*. In particular, compact 2-dimensional polyhedra with $K < 0$ as well as hyperbolic small cancellation polyhedra in 4.7 are locally hyperbolic.

§7. Geodesics, quasigeodesics and quasiconvexity

Let us return to the study of δ-thin triangles in a *geodesic* metric space X (compare 6.3). Observe, that if $\Delta = [x_1,x_2,x_3]$ is a δ-thin geodesic triangle, then every edge of Δ, say $[x_1,x_3]$ is δ-*close* to the union of the remaining two edges. That is

$$(*) \qquad \operatorname{dist}(x,Y) \leqslant \delta$$

for $Y = [x_1,x_2] \cup [x_2,x_3]$ and all $x \in [x_1,x_3]$.

Now, let Y be a broken geodesic consisting of $n \leqslant 2^k$ edges, say $Y = [x_1,x_2] \cup [x_2,x_3] \cup \ldots \cup [x_n,x_{n+1}]$. Then by applying (*) k times we obtain the following

Lemma: *If every geodesic triangle with the vertices in the subset* $\{x_1,\ldots,x_{n+1}\} \subset X$ *is* δ*-thin, then every point* x *in* X *between* x_1 *and* x_{n+1} *(i.e.* x *lies in some geodesic segment* $[x_1,x_{n+1}]$*) satisfies*

$$\operatorname{dist}(x,Y) \leqslant k\delta.$$

7.1 **Exponential growth of balls:** Take a ball $B \subset X$ around $x_0 \in X$ of radius d,

$$B = \{x \in X \mid |x_0-x| \leqslant d\}$$

and let a geodesic segment $[y,z] \subset X$ contain x_0 and have the end points y and z outside B. Let $f\colon [a,b] \longrightarrow X$ for $a,b \in \mathbb{R}$ be a continuous path joining y and z in the complement $X \setminus B$. That is

$$f(a) = y, \; f(b) = z,$$

and

$$|f(t)-x_0| \geqslant d \text{ for all } t \in [a,b].$$

7.1.A *If all triangles with the vertices in the set* $f([a,b]) \subset X$ *are* δ*-thin then the length of the path is bounded from below by*

$$\text{length } f \geqslant \delta(2^{d/\delta}-2).$$

Proof: For every $n = 1,2,\ldots$ there exist, by the very definition of length, $n+1$ points in $[a,b]$, say $a = t_0 \leqslant t_1 \leqslant \ldots \leqslant t_n = b$, such that

$$\delta_i \overset{\text{def}}{=} |f(t_i)-f(t_{i+1})| \leqslant L/n,$$

for $L = \text{length } f$ and $i = 0,1,\ldots,n-1$. Now we make $\delta_i \leqslant 2\delta$ for some $n = 2^k$, where $n-1 \leqslant L/2\delta$, and apply 2.1.E" to some broken geodesic V with the vertices $y = f(t_0),f(t_1),\ldots,f(t_n) = z$. This yields

$$d \leqslant \delta + \text{dist}(x_0,Y) \leqslant (k+1)\delta,$$

and

$$L \geqslant 2n\delta - 2\delta = (2^{k+1}-2)\delta \geqslant (2^{d/\delta}-2)\delta.$$

7.1.B <u>Corollary</u>: *Consider a segment* $[x,y] \subset X$ *and a path* $f\colon [a,b] \longrightarrow X$ *of finite length between* x *and* y, *such that*

$$\text{dist}(f(t_0), [x,y]) \geqslant d_0$$

for some $d_0 \geqslant 20\delta$ *and* $t_0 \in [a,b]$. *If all triangles with the vertices in the union* $[x,y] \cup f[a,b]$ *are* δ*-thin, then for every* $\ell_0 \leqslant d_0$ *there exists a subinterval* $[a_0,b_0] \subset [a,b]$, *such that the length* L_0 *of the path* $f \mid [a_0,b_0]\colon [a_0,b_0] \longrightarrow X$ *satisfies*

$$(*) \qquad \ell_0 \leqslant L_0 \leqslant 2\ell_0$$

and

(**) $L_0 \geqslant \delta(2^{\mu_0/\delta}-2)$ *for* $\mu_0 = 1/2|f(a)-f(b_0)| - 4\delta$

Proof: Since the function $t \longmapsto \text{dist}(f(t),[a,b])$ is continuous, there exists a subinterval $[a_0,b_0]$ satisfying (*) and such that

$$d_1 \overset{\text{def}}{=} \text{dist}(f(a_0),[x,y]) = \text{dist}(f(b_0),[x,y]) \geqslant d_{0/2},$$
$$\text{dist}(f(t),[x,y]) \geqslant d_1 \text{ for all } t \in [a_0,b_0].$$

Take some segment $[x_0,y_0]$ for $x_0 = f(a_0)$ and $f(b_0)$, and show that the center z_0 of $[x_0,y_0]$ satisfies

$$\text{dist}(z_0,f[a_0b_0]) \geqslant \mu = \min(1/2|x_0-y_0| - 4\delta,\ 1/2\ d_0 - 4\delta).$$

To see this take some (nearest) points x_1 and y_1 in $[x,y]$, such that

$$|x_1-x_0| = \text{dist}(x_0,[x,y]) \text{ and } |y_1-y_0| = \text{dist}(y_0,[x,y])$$

and use the fact that z_0 is 2δ-close to the broken geodesic $[x_0,x_1,y_1,y_0]$. Then

$$L_0 \geqslant (2^{\mu/\delta}-2)\delta.$$

Since $L_0 \leqslant 2d_0$ and $d_0 \geqslant 20\delta$ we have

$$L_0 \leqslant (2^{\mu_1/\delta}-2)\delta \text{ for } \mu_1 = 1/2\ d_0 - 4\delta,$$

and (**) follows.

7.1.C <u>A discrete version of 7.1.B</u>: Consider a sequence of points $x_1,x_2,\dots,x_n$ in a δ-hyperbolic space X, such that $x_1 = x$ and $x_n = y$ for a given segment $[x,y]$ in X. Let $d_0 = \max_i \text{dist}(x_i,[x,y])$ and let $\text{dist}(x_{i_0},[x,y]) \geqslant d_0 - \epsilon$ for some $i_0 = 2,3,\dots,n-1$ and some

$\epsilon > 0$. Let

$$\mu_0 = \max(|x_{i_0-1}-x_{i_0}|, |x_{i_0}-x_{i_0+1}|).$$

If X is a tree (i.e. $\delta = 0$), and $\mu_0 \leqslant d_0/2 - \epsilon$, then, obviously,

$$|x_{i_0-1}-x_{i_0+1}| \leqslant \mu_0 + 4\epsilon.$$

Now, let $\delta > 0$. Then the subset $[x,y] \cup \{x_{i_0-1}, x_{i_0}, x_{i_0+1}\}$ in X can be 8δ-approximated by a tree (see 6.1) which leads to the inequality

$$|x_{i_0-1}-x_{i_0+1}| \leqslant \mu_0 + 4\epsilon + 40\delta.$$

Notice, that only the case $\epsilon = 0$ is needed for further applications but $\epsilon > 0$ is necessary when we pass to the approximating tree. In fact, we shall need later the following more technical fact which is obtained by a similar approximation argument.

7.1.D <u>Lemma</u>: *Let* $2 \leqslant i_0 \leqslant n-2$ *and*

$$\operatorname{dist}(x_{i_0}, [x,y]) = d_0 = \max_i \operatorname{dist}(x_i, [x,y]).$$

If

$$\max(|x_{i_0-2}-x_{i_0-1}|+|x_{i_0-1}+x_{i_0}|, |x_{i_0}-x_{i_0+1}|+|x_{i_0+1}-x_{i_0+2}|) \leqslant d_0/2$$

then one of the following three inequalities holds

(1) $|x_{i_0-1}-x_{i_0+1}| \leqslant \max(|x_{i_0-1}-x_{i_0}|, |x_{i_0}-x_{i_0+1}|)$,

(2) $|x_{i_0-2}-x_{i_0+1}|-100\delta \leqslant \max(|x_{i_0-2}-x_{i_0-1}|, |x_{i_0-1}-x_{i_0}|, |x_{i_0}-x_{i_0+1}|)$.

(3) $|x_{i_0-1}-x_{i_0+2}|-100\delta \leqslant \max(|x_{i_0-1}-x_{i_0}|, |x_{i_0}-x_{i_0+1}|, |x_{i_0+1}-x_{i_0+2}|)$.

7.2 <u>Stability of quasigeodesics</u>: A path $f: [a,b] \longrightarrow X$ is called λ-*geodesic* (or *quasigeodesic* if the specific value of the constant λ is irrelevant) if

$$\text{length } f([a',b']) \leqslant \lambda |f(a')-f(b')|,$$

for all $[a',b'] \subset [a,b]$. One immediately obtains with the previous discussion.

7.2.A **Proposition**: *Every* λ-*geodesic path between two points* x *and* y *in a* δ-*hyperbolic geodesic metric space* X *stays* ϵ-*close to every geodesic segment* $[x,y] \subset X$ *for* $\epsilon \leqslant C\delta(1+\log \lambda)$, *where* $C \leqslant 100$ *is some universal constant.*

7.2.B **Remark**: The above obviously generalizes to *locally* λ-geodesic paths f which are, by definition, λ-geodesic on all segments $[a',b'] \subset [a,b]$ with

$$\text{length } f([a',b']) \leqslant \text{const} = \text{const}(\delta,\lambda) \leqslant 1000\delta\lambda.$$

It follows that every local λ-geodesic is, in fact, a (global!) 2λ-geodesic.

Let us state a similar result for "discrete quasigeodesics" which easily follows from the discussion in 7.1.

7.2.C *Let* $x_1,x_2,\dots,x_n$ *be some points in a* δ-*hyperbolic space* X (*which is not assumed geodesic any more*), *such that*

$$|x_i-x_{i-2}|-\epsilon \geqslant \max(|x_i-x_{i+1}|,|x_{i+1}-x_{i+2}|)$$

for some $\epsilon > 100\delta$ *and* $i = 1,2,\dots,n-1$. *Then*

$$\lambda|x_1-x_n| \geqslant \sum_{i=1}^{n-1} |x_i-x_{i+1}|,$$

for

$$\lambda = (0.01\epsilon - \delta)^{-1} \max_i |x_i - x_{i+1}|.$$

7.2.D Let us show that the implication local ⇒ global for quasigeodesics characterizes hyperbolic spaces. In fact we have the following stronger

7.2.E <u>Proposition</u>: *Let X be a non-hyperbolic geodesic metric space and let* $\epsilon_0 > 0$, $L_0 > \epsilon_0$, $L_1 > L_0$ *and* $\lambda > 1$ *be given numbers. Then there exists a circle* S^1 *in X of finite length* $L \geqslant L_1$, *such that every ark* $[a',b']$ *in* S^1 *of length* L' *in the inverval* $\epsilon_0 \leqslant L' \leqslant L_0$ *satisfies*

(*) $$|a'-b'|_X \geqslant \lambda^{-1} L'$$

<u>Proof</u>: If every circle S^1 in X contains an interval which violates (*) then one can shorten S^1 by joining the ends of this interval by a geodesic segment in X. By repeating this process one "subdivides" S^1 into $k \leqslant$ const length circles of length $\leqslant L_0$, which yields the isoperimetric inequality in $P_d^2(X)$ for $d = L_0$ which implies hyperbolicity. Q.E.D.

7.2.F <u>Remarks</u>: One can easily sharpen 7.2.E by requiring S^1 to be an *embedded* circle in X. Furthermore, if X is complete one can allow $\epsilon_0 = 0$.

7.2.G <u>Quasiisometric</u> <u>maps</u>: A map between metric spaces, say f: X ⟶ Y, is called

(a) An ϵ + *isometric map* for some $\epsilon \geqslant 0$ if

$$|x_1 - x_2| - \epsilon \leqslant |f(x_1) - f(x_2)| \leqslant |x_1 - x_2| + \epsilon$$

for all x_1 and x_2 in X. Notice that ϵ + isometry does *not* imply the continuity of f.

(b) An *isometric* $\times$ λ *map* for some $\lambda \geqslant 1$ if

$$\lambda^{-1}|x_1-x_2| \leqslant |f(x_1)+f(x_2)| \leqslant \lambda|x_1-x_2|.$$

(c) An ϵ + *isometric* $\times$ λ *map* if

$$\lambda^{-1}|x_1-x_2|-\epsilon \leqslant |f(x_1)+f(x_2)| \leqslant \lambda|x_1-x_2|+\epsilon.$$

Sometimes we supress ϵ and λ and call such maps + *isometric* $\times$ or *quasiisometric*. In a similar way we speak of + *isometric* and isometric $\times$ maps.

<u>**Examples**</u>:

(a) If X is a geodesic space, then the canonical map $X \longrightarrow P^1_d(X)$ is + isometric for every $d \geqslant 0$, where the polyhedron $P^1_d(X)$ is given the simplical metric with the edges of length d. In fact the map $X \longrightarrow P^1_d(X)$ is d + isometric.

(b) If $X \longrightarrow Y$ is an *injective* + isometric $\times$ map then the induced map $P^1_d(X) \longrightarrow P^1_d(Y)$ is isometric $\times$ for every $d \geqslant 0$. It follows that every + isometric $\times$ map between geodesic spaces, $f: X \longrightarrow Y$, admits a decomposition

$$X \xrightarrow[f_1]{} X_1 \xrightarrow[\bar{f}_2]{} Y_1 \xrightarrow[f_2]{} Y, \qquad f: X \longrightarrow Y$$

where f_1 is an injective + isometric map, $\bar{f}$ is isometric $\times$ and f_2 is subjective + isometric.

7.2.H <u>**Theorem**</u>: *Let* X *and* Y *be hyperbolic spaces where* X *is geodesic. Then every + isometric* $\times$ *map* $f: X \longrightarrow Y$ *(canonically) induces a topological embedding between the hyperbolic boundaries, say*

$\partial f: \partial X \longrightarrow \partial Y$. *Furthermore, if the image* $f(X) \subset Y$ *is cobounded* (*i.e.* $\mathrm{dist}(y, f(X)) \leqslant \mathrm{const} < \infty$ *for all* $y \in Y$) *then* ∂f *is a homeomorphism.*

Proof: Observe that the theorem is obvious for + isometric maps between arbitrary (i.e. non-geodesic) metric spaces. Then using 6.4 and the above discussion we reduce the problem to isometric $\times$ maps $\bar{f}$ between hyperbolic geodesic spaces. Such maps $\bar{f}$ (obviously) send geodesic to quasigeodesics and these lie within finite (and uniformly bounded from above) distance from geodesics. This allows us to control the behaviour of the "product" (x.y) which is equal, up to a bounded error, to the distance from the reference point to some geodesic segment between x and y. Now it is clear that "convergent" sequences $x_i \longrightarrow a \in \partial X$ a sent by f to "convergent" sequences $y_i \longrightarrow b \in \partial Y$ and the proof follows from the definition of the hyperbolic boundary (see 1.8).

7.2.J The conclusion of 7.2.H holds true for *locally* ϵ + isometric $\times \lambda$ maps f. These are, by definition, ϵ + isometric $\times \lambda$ on every ball B in X of radius $R \leqslant C(\delta+\epsilon)\lambda$ where δ is the implied hyperbolicity constant (serving X as well as Y) and $C \leqslant 10{,}000$ is some universal constant. This "localization" of 7.2.H follows from 7.2.B, which also shows that locally + isometric $\times$ maps are, in fact, (globally) + isometric $\times$.

7.2.K *A conformal view on the boundary* ∂X (*compare* 1.8 *and* [Fl]). Let μ be a positive function on a geodesic metric space $(X, |\ |)$. Define the μ-*length* of curves S in X as $\int_S \mu$. Next define the metric $|\ |_\mu$ on X as the infinum of the μ-lengths of paths between given points in X. Let $\bar{X}$ denote the metric completion of X and $\bar{X}_\mu$ be the completion of $(X, |\ |_\mu)$. Define the μ-*boundary* of X by $\partial_\mu X = \bar{X}_\mu \setminus \bar{X}$.

To relate $\partial_\mu X$ and ∂X (for some special μ) we need the following property of $|\ |_\mu$-geodesics and quasigeodesics which immediately follows from the local characterization of quasigeodesics (see 7.2.B).

7.2.L __Lemma:__ Let X be δ-hyperbolic and let

(*) $$|\log \mu(x_1) - \log \mu(x_2)| \leqslant \epsilon_0$$

for all x_1 and x_2 in X satisfying $|x_1 - x_2| \leqslant \delta$ and for some (universal constant) $\epsilon_0 > 0.001$. Then every $|\ |_\mu$-quasigeodesic also is $|\ |$-quasigeodesic. Furthermore $|\ |$- and $|\ |_\mu$-geodesic segments between some points in X are $\epsilon^{-1}\delta$-close for the original metric in X and the distance $|x-y|_\mu$ is related to the μ-length of each $|\ |$-geodesic segment between x and y by

$$\mu\text{-length} \leqslant \epsilon_0^{-1} |x-y|\mu$$

for all x and y in X.

From this we immediately derive the following

7.2.M __Corollary:__ *Let*

$$\mu(x) = \mu_{x_0}(x) = \exp - \epsilon |x - x_0|$$

for some ϵ *in the interval* $0 < \epsilon < 1/2\, \epsilon_0 \delta^{-1}$. Then the identity map $(X, |\ |) \leftrightarrow (X, |\ |_\mu)$ continuously extends to a homeomorphism $X \cup \partial X \leftrightarrow X \cup \partial_\mu X$. Furthermore, if some points a and b in $\partial X = \partial_\mu X$ are limits of sequences x_i and y_i respectively for which the sequence $(x_i.y_i)$ converges to some limit denoted (a.b), then the $|\ |_\mu$-distance between a and b satisfies

$$|\epsilon(a.b) + \log |a-b|_\mu| \leqslant \epsilon_0^{-1} + \delta.$$

It follows that every isometry of X induces a $|\ |_\mu$-Lipschitz map on $(\partial X, |\ |_\mu)$ and every quasi-isometry is Hölder on ∂X.

7.3 __Quasiconvex__ __subsets:__ A subset Y in a geodesic space X is called ϵ-*convex* (or *quasiconvex*) for some $\epsilon \geqslant 0$ if every geodesic segment $[y_1,y_2]$ with the end points in Y lies ϵ-close to Y. In other words, $[y_1,y_2]$ is contained in the *ϵ-neighborhood* of Y that is

$$Y + \epsilon = \{x \in X \mid \text{dist}(x,Y) \leqslant \epsilon\}.$$

<u>**Example**</u>: If X is δ-hyperbolic then every ball $B \subset X$ around some $x_0 \in X$ is 2δ-convex since $[y_1,y_2]$ is 2δ-close to the union of $[x_0,y_1]$ and $[x_0,y_2]$. More generally, we have

7.3.A <u>**Lemma**</u>: *If X is δ-hyperbolic and $Y \subset X$ is ϵ-convex then the ρ-neighborhood $Y + \rho$ is 4δ-convex for all $\rho \geqslant \epsilon$.*

<u>**Proof**</u>: Recall that every segment $[x_1,x_2] \subset X$ stays 4δ-close to the broken geodesic $[x_1,y_1,y_2,x_2]$, where y_1 and y_2 are *normal projections* (see 7.3.B below) of x_1 and x_2 to Y respectively. If x_1 and x_2 lie in $Y + \rho$ for $\rho \geqslant \epsilon$ then $[x_1,y_1,y_2,x_2]$ is contained in $Y + \rho$ and the proof follows.

7.3.B <u>**Definition**</u>: If for every $x \in X$ the function $y \longmapsto |x-y|$ assumes the minimum on Y we assign to each x such a minimum point and thus define a *normal projection* (or *retraction*) called $P_Y: X \longrightarrow Y$. In other words $P_Y(x) \in Y$ is a nearest point to x in Y. If there is no nearest point $y \in Y$ to a given $x \in X$ we use points $y' \in Y$ with $|x-y'| \leqslant \text{dist}(x,Y) + \epsilon'$ where eventually $\epsilon' \longrightarrow 0$. These approximately nearest points y' serve us as well as the actual nearest points. To avoid mess we always use the notation P_Y (rather than $P_Y^{\epsilon'}$ for $\epsilon' \longrightarrow 0$). A similar convention will be applied to other functions (such as the length of a curve) which might not assume their minimal values on the spaces in question.

7.3.C <u>**Corollary to the Proof of 7.3.A**</u>: *The ρ-neighborhood of every geodesic segment in X is 4δ-convex for all $\rho \geqslant 0$.*

7.3.D **Lemma**: *If* X *is* δ*-hyperbolic and* $Y \subset X$ *is* ϵ*-convex then*

$$|P_Y(x_1)-P_Y(x_2)| \leqslant \max(2\epsilon+6\delta, |x_1-x_2|-|x_1-P_Y(x_1)|-|x_2-P_Y(x_2)|)$$

for all x_1 *and* x_2 *in* X.

The proof is similar to that of 6.3.A and is left to the reader.

7.3.E **Remark**: Let us prove a converse to 7.3.D. Namely, let $Y \subset X$ be a subset which admits a *retraction* $P: X \longrightarrow Y$ (i.e. $P|Y = Id$), such that the inequalities

$$|x_1-x_2| \leqslant 2\delta \text{ and } \max_{j=1,2} |x_j-P(x_j)| \geqslant \beta,$$

for given positive numbers α and β, imply that $|P(x_1)-P(x_2)| \leqslant \alpha$ for all points x_1 and x_2 in X. (The map P in 7.3.D satisfies this condition for all $\alpha \geqslant 2\epsilon + 6\delta$ and $\beta \geqslant 2\alpha$.) *Then the subset* Y *is* ϵ*-convex for* $\epsilon = 3\beta + \alpha$.

Proof: To estimate dist(x,Y) for $x \in [y_1,y_2]$, where $y_j \in Y$, $j = 1,2$, we take a subsequent $[x_1,x_2] \subset [y_1,y_2]$ containing x, such that $|x_i-P(x_i)| \leqslant \beta$ for $i = 1,2$ and $|x'-P(x')| \geqslant \beta$ for all $x' \in [x_1+\alpha, x_2-\alpha]$. Observe that $|x_1-x_2| \geqslant 2(\text{dist}(x,Y)-\beta)$. Subdivide $[x_1,x_2]$ into k subintervals of length $\leqslant 2\alpha$ for the minimal possible k and apply the inequality $|P(x_1')-P(x_2')| \leqslant \alpha$ to the end points of these subintervals. This yields

$$|P(x_1)-P(x_2)| \leqslant |x_1-x_2| - \text{dist}(x,Y) + \alpha + \beta$$

and

$$|y_1-y_2| \leqslant |y_1-y_2| - \text{dist}(x,Y) + \alpha + 3\beta$$

by the triangle inequality. Q.E.D.

Notice that the above argument does not use any hyperbolicity. In fact, the existence of certain "contracting" maps $X \longrightarrow X$ ensures the hyperbolicity of X.

7.3.F **Example**: Let a metric space X admits a one-parametric family of maps $P_t: X \longrightarrow X$ for $t \in [0,1]$, such that

(a) For every curve $S \subset X$ the area of the surface $S \times [0,1] \longrightarrow X$ defined by $(s,t) \longmapsto P_t(s)$ satisfies

$$\text{area}(S \times [0,1]) \leqslant C \text{ length } S$$

for some constant C (depending on X but not on S in X).

(b) There exists a constant $C' < 1$ such that

$$\text{area } P_1(D) \leqslant C' \text{ area } (D)$$

for all surfaces D in X.

Then, obviously, every closed curve S of length L(S) and area $A(S) < \infty$ satisfies

$$A(L) \leqslant (1-C')^{-1} CL(A),$$

which implies (see 6.8) the hyperbolicity of X, provided

$$L(S) < \infty \Rightarrow A(S) < \infty$$

for all closed curves S in X.

Notice that every convex manifold X admits transformations P_t satisfying (a) with $C = 1$. There are transformations moving every $x \in X$ toward $x_0 \in X$ with unit speed. Furthermore, if X has (strictly) negative curvature $\leqslant -\epsilon < 0$, then these P_t also satisfy (b), which ensures the hyperbolicity in this case.

7.3.G Let a metric space X be covered by subsets $X_i \subset X$ for $i \in \mathbb{Z}$, such that

$$X_i \subset X_{i+1} \subset X_i + \beta_1$$

for all $i \in \mathbb{Z}$ and some fixed $\beta_1 > 0$. Let $P_i\colon X_i + \beta_2 \longrightarrow X_i$ be retractions for $i \in \mathbb{Z}$ and some $\beta_2 > 0$, such that the inequalities

$$|x_1 - x_2| \leqslant 2\alpha \text{ and } \max_{j=1,2} |x_j - P_i(x_j)| \geqslant \beta$$

for some positive α and β imply that $|P_i(x_1) - P_i(x_2)| \leqslant \alpha$ for all x_1 and x_2 in $X_i + \beta_2$ and all $i \in \mathbb{Z}$.

Proposition: *If X is geodesic, if the intersection $\bigcap_i X_i$ is hyperbolic (for the metric induced from X) and if*

$$10\beta_1 \leqslant \beta \leqslant \frac{\alpha\beta_2}{10\beta},$$

then X is hyperbolic.

Proof: It suffices to show, according to 7.2.E that X contains no long locally quasigeodesic circle. Now let S^1 be such a circle in X and i_0 be the minimal integer such that $X_{i_0} + \beta_2$ contains S^1. Using the retraction P_{i_0} one can shorten some ark in S^1 of a controlled length (compare 1.7.B) and the proof follows by contradiction. We leave to the reader filling in the details and figuring out the value of the hyperbolicity constant.

Example: If $X = \mathbb{R}^2$ then the pertinent S^1 is the round circle of a sufficiently large radius.

7.3.H Let us give a "local" criterion for ϵ-convexity of a *connected* subset Y in a δ-hyperbolic space X.

Proposition: *Let for a given* $\alpha \geqslant 50\delta$ *the center* z *of every segment* $[y_1,y_2]$ *with* $y_1,y_2 \in Y$ *and* $|y_1-y_2| \leqslant 2\alpha$ *satisfies* $\mathrm{dist}(z,Y) \leqslant \alpha - C\delta$ *for* $C = 50$. *Then* Y *is* ϵ*-convex for* $\epsilon = 20\alpha$.

Proof: Let n be the minimal integer, such that there exists a sequence of points $x_1 = y_1,x_2,x_3,\ldots,x_n = y_2$ in Y with $|x_i-x_{i+1}| \leqslant 2\alpha$ for $i = 1,2,\ldots,n-1$. Since n is minimal no point among x_i satisfies the inequality (1) in 7.1.D. Furthermore, if $|x_i-x_{i+3}| \leqslant 2\alpha + 100\delta$ for some $i = 1,\ldots,n-3$ then the center z on $[x_i,x_{i+3}]$ satisfies

$$\max(|x_i-P_Y(z)|,|P_Y(z)-x_{i+3}|) \leqslant 2\alpha,$$

which also is incompatible with the minimality of n. Hence, the inequalities (2) and (3) cannot hold true for any point among x_i. Then, 7.1.B shows that

$$d_0 = \max_i \mathrm{dist}(x_i,[y_1,y_2]) \leqslant 8\alpha$$

and the proof follows by the triangle inequality.

7.3.I **Remarks:**

(a) Instead of the connectivity of Y one only could require the ϵ-neighborhood $Y + \epsilon \subset X$ to be connected for some $\epsilon < \alpha$.

(b) The inequality $\mathrm{dist}(z,Y) \leqslant \alpha - C\delta$ can be replaced by the following more natural condition

(*) there exists a point $z' \in Y$, such that

$$\max(|y_1-z'|,|z'-y_2|) \leqslant 2\alpha - 2C\delta.$$

(c) One can replace (*) by yet another condition for a given integer $n \geqslant 3$.

$(*)_n$ there exists points $z'_1 = y_1, z'_2, \ldots, z'_n = y_2$ such that

$$\max_i |z'_i - z'_{i+1}| \leqslant 2\alpha - 2C_n\delta$$

for $C_n \leqslant 50(n-2)$.

The proof of these remarks is left to the reader.

7.3.J Call a metric space X *quasigeodesic* if it admits an isometric embedding into a geodesic space, say $X \hookrightarrow \bar{X}$, such that

$$\text{dist}(\bar{x}, X) \leqslant \epsilon < \infty$$

for all $\bar{x} \in \bar{X}$ and some $\epsilon > 0$. (If one wants to specify ϵ one says X is ϵ-geodesic.)

Using 6.4 and 7.3.H one obtains the following

Proposition: *Let X be a hyperbolic metric space and fix some numbers $\mu > 1$ and an integer $n \geqslant 3$. If for every two points x and y in X there exists points $x_1 = x, x_2, \ldots, x_n = y$ in X, such that*

$$\max_i |x_i - x_{i+1}| \leqslant \mu^{-1}|x-y| + n,$$

then X is quasigeodesic.

7.3.K Convex hulls: Take a subset Y in a geodesic space X and let $\bar{Y} = \text{Conv } Y \subset X$ be the union of geodesic segments in X.

Lemma: *If X is δ-hyperbolic then $\bar{Y}$ is ϵ-convex for* $\epsilon = 4\delta$.

Proof: One obviously reduces the lemma to the case of a four point subset Y in X where the proof follows from 6.3.

7.4 <u>Convexity of the distance functions</u>: If X is a tree then the distance function $(x,y) \longrightarrow |x-y|$ is convex. Namely, for any two geodesic segments $[x_1,x_2]$ and $[y_1,y_2]$ the distance function is convex on $[x_1,x_2] \times [y_1,y_2]$. Since the definition of convexity uses at most six points in X at a time (namely, x_1, x_2, y_1, y_2, $z_1 = \alpha x_1 + (1-\alpha)x_2$ and $z_2 = \alpha y_1 + (1-\alpha)y_2$ for $0 \leqslant \alpha \leqslant 1$), the approximation lemma in 6.1 shows the distance function to be ϵ-*convex* that is $|z_1-z_2| - \epsilon \leqslant \alpha|x_1-y_1| + (1-\alpha)|x_2-y_2|$ in every δ-hyperbolic space X for $\epsilon \leqslant 6\delta$.

<u>Remark</u>: Notice that the function $x \longmapsto |x-x_0|$ for a fixed $x_0 \in X$ is more than just ϵ-convex. In fact, the restriction of this function to every segment $[x_1,x_2] \subset X$ is ϵ-close to the function $\mu + |x-x_\mu|_{\mathbb{R}}$ where $\mu = \inf_{x \in [x_1,x_2]} |x-x_0|_X$ and $x_\mu \in [x_1,x_2]$ is a point where $|x_\mu-x_0|_X = \mu$. Similarly, the distance to an ϵ-convex subset $X_0 \subset X$ is shaped the same way in every segment $[x_1,x_2]$ which does not meet the ϵ-neighborhood of X_0 for $\epsilon = 6\delta$.

7.4.A <u>The displacement of an isometry</u>: Consider a map $\gamma: X \longrightarrow X$ and let $d_\gamma(x) = |x-\gamma(x)|$. If γ is an isometry and X is geodesic δ-hyperbolic then the (displacement) function d_γ is ϵ-convex for $\epsilon = 6\delta$ by 7.3.A. Let for an isometry $\gamma: X \longrightarrow X$,

$$\bar{d}_\gamma(\leqslant t) = \mathrm{Conv}\ d_\gamma^{-1}[0,t] \subset X$$

and observe that

$$\bar{d}_\gamma(\leqslant t) \subset d_\gamma^{-1}[0,t+\epsilon]$$

for all $t \in \mathbb{R}_+$, by the ϵ-convexity of d_γ. Also observe that the subsets $\bar{d}_\gamma(\leqslant t)$ are γ-invariant and 4δ-convex.

Another useful construction of invariant ϵ-convex subsets in

X is as follows. Let Γ be an arbitrary group isometrically acting on X. Then the convex hull of every orbit of Γ clearly is Γ-invariant and 4δ-convex.

Consider a geodesic segment $S = [x_1,x_2]$ in X and apply the above to the function $d^S(y) = \mathrm{dist}(\gamma(y), [x_1,x_2]$ where $y \in [x_1,x_2]$ and $\gamma: X \longrightarrow X$ is an isometry. This yields

$$(*) \qquad d^S(y) \leqslant 6\delta + \max(0, d_\gamma(x_1) - |x_1 - y|, d_\gamma(x_2) - |x_2 - y|).$$

Next, consider a geodesic triangle Δ with vertices $x_i \in X$, $i = 1,2,3$ and edges $[x_i,x_j]$. Since Δ is 2δ-thin, there exists a point $z \in X$ which is 2δ-close to each edge of Δ. Apply (*) to an isometry γ of X and bound the displacement of z by

$$(**) \qquad d_\gamma(z) \leqslant 12\delta + \max_{i=1,2,3} (0, d_\gamma(x_i) - |x_i - z|).$$

It follows that an arbitrary point $z' \in X$ for which $(x_i,x_j)_{z'} \leqslant \epsilon$ for some $\epsilon \geqslant 0$ satisfies

$$(***) \qquad d_\gamma(z') \leqslant 4\epsilon + 16\delta + \max_i (0, d_\gamma(x_i) - |x_i - z'|).$$

7.4.B <u>Solution of the conjugacy problem</u>: Consider a group Γ with a word metric and suppose the word problem is solvable in Γ. Then the following condition obviously implies the solution of the conjugacy problem.

There exists a constant $C > 0$ such that for every two conjugate elements γ and γ' there exists $\gamma_1 = \gamma, \gamma_2, \gamma_3, \dots, \gamma_n = \gamma'$ in Γ such that $|\gamma_i| \leqslant C(|\gamma| + |\gamma'|)$ and $\gamma_{i+1} = \alpha_i \gamma_i \alpha_i^{-1}$ for some $\alpha_i \in \Gamma$ with $|\alpha_i| \leqslant C$ for all $i = 1,\dots,n-1$. This condition says in the geometric language, that every two homotopic closed curves γ and γ' in a compact manifold V with $\pi_1(V) = \Gamma$ can be joined by a homotopy of closed curves γ_t with length $\gamma_t \leqslant C(\text{length } \gamma + \text{length } \gamma')$. If V has $K \leqslant 0$ such a homotopy is constructed by shortening curves to a closed geodesic. This is essentially equivalent to minimizing the displacement functions of γ and γ' on the universal covering X of V, and then, obviously, the quasiconvexity of the displacements solves the conjugacy problem. In particular we obtain the solution for word hyperbolic groups Γ as well as for discrete cocompact isometry groups of convex spaces. Moreover, in the word hyperbolic case we also have the linear isoperimetric inequality for the conjugation. Algebraically, this means the bound

$$n \leqslant \text{const}(|\gamma| + |\gamma'|).$$

Geometrically, it provides a homotopy $A: S^1 \times [0,1] \longrightarrow V$ between closed curves γ and γ', such that Area $A \leqslant$ const(length γ + length γ'), provided γ and γ' are homotopic in V to start with. Thus linear isoperimetric inequality for disks (which implies hyperbolicity by 6.8) yields such inequality for annuli. As we shall see in 8.3 there is a similar inequality for surfaces in the word hyperbolic case. A similar result (including the solution of the conjugacy problem) is known (and easy to prove) for many classes of semihyperbolic groups, such as small cancellation groups and cocompact isometry groups of (convex) manifolds and spaces with $K \leqslant 0$.

7.5 __Rays, lines, distance-like functions and horofunctions:__ A *ray* in a metric space X is (the image of) an isometric map $r: \mathbb{R}_+ \longrightarrow X$. A *line* is (the image of) an isometric map $\mathbb{R} \longrightarrow X$. If X is hyperbolic then there obviously exists a limit

$$r(\infty) \overset{\text{def}}{=} \lim_{t\to\infty} r(t) \in \partial X,$$

called the *end* of the ray r. Similarly, every line ℓ has two ends $\ell(-\infty)$ and $\ell(+\infty)$ in ∂X. We sometimes say that the ray r *joins* r(0) and $r(\infty)$ and that the line ℓ joins $\ell(-\infty)$ and $\ell(\infty)$.

Usually rays and lines appear as limits of geodesic segments in X. For example let $s_i: [0,t_i] \longrightarrow X$ be a sequence of isometric maps (segments) with $s_i(0) = x_0$ and $s_i(t) = x_i \longrightarrow a \in \partial X$ for $i \longrightarrow \infty$. If the space X is *proper*, that is if every closed ball in X is compact (e.g. X is a complete Riemannian manifold) then some subsequence of maps s_i converges for every fixed $t_0 \in [0,\infty)$ and the limit map $[0,\infty) \longrightarrow \mathbb{R}$ is a ray in X joining x_0 and a. Thus, in a proper geodesic hyperbolic space X the boundary ∂X equals the set of *asymptotic* classes of rays, where two rays r_1 and r_2 are called *asymptotic* if $|r_1(t)-r_2(t)| \leq const < \infty$ for all $t \in [0,\infty)$.

Now, to construct a line joining a and $b \neq a$ in ∂X we take sequences of points in X say $x_i \longrightarrow a$ and $y_i \longrightarrow b$ such that $\big| |x_i| - |y_i| \big| \leq const < \infty$ for $i \longrightarrow \infty$. We parametrize geodesic segments $[x_i,y_i]$ by isometric maps $s_i\colon [-t_i,t_i] \longrightarrow X$ (for $2t_i = |x_i-y_i|$) and observe that in the hyperbolic case the centers $s_i(0) \in X$ lie in a fixed (bounded!) ball as $i \longrightarrow \infty$. Therefore some subsequence converges to a line between a and b, provided X is proper.

If X is non-proper, then the above (limit) rays and lines do not always exist. One may use instead quasigeodesic lines and rays which work equally well in many cases. Another possibility is to slightly perturb (or complete) X in order to have actual lines and rays. Namely call a geodesic hyperbolic space Y *ultra-complete* if any two points in $Y \cup \partial Y$ can be joined by a segment, ray or a line in Y. Then, by a trivial argument, every geodesic hyperbolic X isometrically embeds into some ultra-complete space Y, such that $dist_Y(y,X) \leq const < \infty$ for all $y \in Y$. In particular, $\partial Y = \partial X$. Furthermore, one can choose a Y, such that every isometry of X extends to a unique isometry of Y.

Thus, in most cases, one can work with X as if it were ultra-complete to start with.

7.5.A __Quasiconvex hulls of subsets in the boundary__: If X is a geodesic hyperbolic space then there is an important correspondence between closed subsets $A \subset \partial X$ containing at least two points and certain quasiconvex subsets in X. To go from subsets in X to ∂X one assigns to each $X_0 \subset X$ its boundary $\partial X_0 \subset \partial X$. A more interesting operation is assigning to A the union of all lines in X between the pairs of points in A which is denoted $L(A) \subset X$. If X is ultracomplete and δ-hyperbolic, the, by the arguments in 7.3, this "hull" L(A) is 6δ-convex and $\partial L(A) = A$. If X is not ultracomplete, one first takes L(A) in an ultracompletion $Y \supset X$ (with $\partial Y = \partial X$) and then defines $L'(A) \subset X$ as the intersection of $X \subset Y$ with some ϵ-neighborhood (for possibly large ϵ) of L(A) in Y. An alternative is to define L(A) with quasigeodesics in X.

7.5.B **Distance-like functions and forms on X**: Our objective now is to describe the boundary ∂X using functions $f: X \longrightarrow \mathbb{R}$ rather than rays $\mathbb{R}_+ \longrightarrow X$. The idea comes from the following

7.5.C **Example**: Start with a ray $r: \mathbb{R}_+ \longrightarrow X$ and associate to it the following *ray* (or *Busemann*) *function*

$$f(x) = \lim_{t\to\infty}(|x-r(t)|-t).$$

Observe several obvious properties of this function $f: X \longrightarrow \mathbb{R}$.

(1) Every level set $f^{-1}(-\infty,t) \subset X$ is the union of an increasing family of balls in X.

(2) The function f is "distance-like":

$$f(x) = t + \mathrm{dist}(x,f^{-1}(-\infty,t))$$

for all $t \in \mathbb{R}$ and those $x \in X$, where $f(x) \geqslant t$.

(3) If X is hyperbolic and $f(x_i) \longrightarrow -\infty$ for some $x_i \in X$, then $x_i \longrightarrow r(\infty) \in \partial X$ for $i \longrightarrow \infty$.

(4) Take a point $x_1 \in X$ and a sequence of numbers $t_1 > t_2 > \dots > t_i > \dots$ such that $t_1 < f(x_1)$ and $t_i \longrightarrow -\infty$ for $i \longrightarrow \infty$. Define by induction a sequence of points $x_i \in X$ where x_{i+1} is the normal projection of x_i to the level $f^{-1}(-\infty,t_i] \subset X$. Such an x_{i+1} does exist if X is proper. Otherwise we take an ϵ_{i+1}-approximation to such a projection for some sequence $\epsilon_i \longrightarrow 0$

for $i \longrightarrow \infty$. Then we join each x_i with x_{i+1} by a segment $[x_i, x_{i+1}]$ and thus obtain a map, called *a gradient ray* $r_1: \mathbb{R}_+ \longrightarrow X$ issuing from x_1. If the implied constants ϵ_i are all $= 0$, then this is indeed a ray. Otherwise this is quasiray. If X is hyperbolic, then, obviously, this quasiray is asymptotic to the original ray r issuing from x_0.

7.5.D **Horofunctions:** The above construction of a ray function can be generalized as follows. Fix a reference point $x_0 \in X$ and let a sequence $x_i \in X$, $i = 1, \dots$, have $|x_i| \longrightarrow \infty$ for $i \longrightarrow \infty$. Consider the sequence of functions on X

$$f_i(x) = |x - x_i| - |x_0 - x_i|.$$

If X is proper, then there is a subsequence which converges on every bounded subset in X. Limits f of this kind are called *horofunctions* on X. If X is geodesic, then horofunctions are very similar to (yet more general than) ray functions.

Now let us see what happens to f_i if X is δ-hyperbolic and $x_i \longrightarrow a \in \partial X$. It is clear that

$$\limsup_{i,j\to\infty} |f_i(x) - f_j(x)| \leqslant 4\delta$$

for all $x \in X$. Therefore, even if (no subsequence of) f_i converges, one may take an approximate limit called *quasihorofunction* $f: X \longrightarrow \mathbb{R}$, such that $\limsup_{i\to\infty} |f(x) - f_i(x)| \leqslant 4\delta$ for all $x \in X$. These functions satisfy obvious "quasi-fications" of the properties (1)-(4) of ray functions.

Now, let f and f' be two such quasihorofunctions with reference points x_0 and x_0' in X and limit points a and a' in ∂X correspondingly. If $a = a'$, the difference $f - f'$ obviously is 8δ-constant on X,

$$|(f-f')(x) - (f-f')(y)| \leqslant 8\delta$$

for all x and y in X. On the contrary, if $a \neq a'$, then $f - f'$ is very far from a constant. Namely, if X is geodesic, then for every $d > 0$ there exist points x and y in X with $|x-y| = d$, such that

$$|(f-f')(x)-(f-f')(y)| \geq 2d - 16\delta.$$

This is seen by looking at f and f' on some (geodesic or quasigeodesic) line $\ell(t)$ in X between a and a'. This line is (quasi)-gradient for both f and f', such that f roughly equals t on $\ell(t)$ and f' is $\approx -t$. Thus f–f' looks like 2t on this line.

7.5.E Let us give a *local* description of horofunctions on a geodesic δ-hyperbolic space X by looking at "differentials" $\varphi(x,y) = f(x)-f(y)$. Namely, consider functions φ such that

(a) $\varphi = \varphi(x,y)$ is defined for all x and y in X satisfying $|x-y| \leq 4d$ for some fixed $d \geq 100(\delta+1)$.

(b) *Chain (or 1-form) property.*

$$\varphi(x,y) = -\varphi(y,x).$$

(c) *Local exactness.*

$$\varphi(x,y) = \varphi(x,z) + \varphi(z,y),$$

for all x, y and z in X where $\varphi(x,y)$, $\varphi(x,z)$ and $\varphi(z,y)$ are defined.

(d) *Distance-like property.* Fix an $x \in X$ and consider the function $\bar{\varphi}(y) = -\varphi(x,y)$ on the 4d-ball $B_x \subset X$ around x. Then

$$\bar{\varphi}(y) = t + \mathrm{dist}(y, \bar{\rho}^{-1}(-\infty,t))$$

for all t in the interval $-d \leq t \leq d$ and for all $y \in B_x$, such that $|x-y| \leq d$, and $\bar{\varphi}(y) \geq t$ and for all $x \in X$.

(e) *Quasiconvexity.* The function $\bar{\varphi}$ is quasiconvex for all

$x \in X$. Namely, if $y = (1/2)(y_1+y_2)$ where $|x-y_1| \leqslant 2d$ and $|x-y_2| \leqslant 2d$ then $\bar{\varphi}(y) \leqslant (1/2)(\bar{\varphi}(y_1)+\bar{\varphi}(y_2)) + 10\delta$.

<u>Remarks</u>: Every φ "integrates" to a function $f(x)$ such that $\varphi(x,y) = f(x) - f(y)$, where f is unique up to an additive constant. This follows from (a), (b) and the simply connectedness of the 2-polyhedron $P_d^2(X)$ (see 1.7).

To avoid too much "quasi"-terminology let us assume that X is proper. Then, starting from any point $x_1 \in X$ we construct a (gradient) sequence $x_1,x_2,\ldots,x_i,\ldots$, such that $|x_i-x_{i+1}| = d$ for all $i = 1,2,\ldots$, and

$$\varphi(x_i,x_{i+1}) = d, \qquad i = 1,2,\ldots$$

as follows. For every $i = 1,2,\ldots$, take $\bar{\varphi}_i(y) = -\varphi(x_i,y)$ and take for x_{i+1} the normal projection of x_i to the level $\bar{\varphi}_i^{-1}(-\infty,d)$. It follows from the previous discussion that x_i converge to some point $a = a(\varphi) \in \partial X$ independent of the starting point $x_1 \in X$. Next we observe as earlier the following implications for every two φ_1 and φ_2 satisfying (a)–(e).

$$|\varphi_1-\varphi_2|_d \leqslant 20\delta \Rightarrow a(\varphi_1) = a(\varphi_2) \Rightarrow |\varphi_1-\varphi_2|_d \leqslant 10\delta,$$

where

$$|\varphi_1-\varphi_2|_d = \sup_{|x-y|\leqslant d} |\varphi_1(x,y)-\varphi_2(x,y)|.$$

Therefore, the inequality $|\varphi_1-\varphi_2|_d \leqslant 20\delta$ is an equivalence relation on our functions and ∂X is identified with the set of equivalence classes. Notice that both the definition of the set Φ of functions φ (by (a)–(e)) and the equivalence relation on this set are local: in order to decide whether some φ lies in our set and whether $|\varphi_1-\varphi_2|_d \leqslant 20\delta$ one only should look at φ (and φ_1 and φ_2) in the 4d-ball around each point in X. Also observe that $(\Phi, |\ |_d \leqslant 20\delta)$ is invariant under all isometries of X. Finally assume the balls $B_x \subset X$ of radius 4d around all points $x \in X$ to be *uniformly compact*. That is each B_x can be covered by N balls of radius ϵ for all $\epsilon > 0$ where N depends only on ϵ.

7.5.F __Proposition__: *The quotient map*

$$\pi\colon \Phi \longrightarrow \partial X = \Phi/\,|\ \ |_d \leqslant 20\delta$$

is compact-to-one for the norm $|\ \ |_d$ *in* Φ.

__Proof__: Take a point $a \in \partial X$ and a sequence of points $x_i \longrightarrow \partial X$ on some ray in X converging to a. If two functions φ and φ' in Φ with $a(\varphi) = a(\varphi') = a$ are equal on all balls B_{x_i} then $\varphi \equiv \varphi'$. This follows from the contracting property of the normal projection to quasi-convex sets (see 7.3) and the distance-like property of the (integrated) functions φ. Moreover, if $|\varphi-\varphi'|_d$ is bounded by some ϵ on infinitely many balls B_{x_i} then $|\varphi-\varphi'|_d \leqslant \epsilon$. Hence the pull-back $\pi^{-1}(a) \subset \Phi$ is compact for all $a \in \partial X$. Q.E.D.

7.5.G Let us specialize the above discussion to a locally finite 1-dimensional complex X with a simplical metric having the edges of length one. Denote by $\Phi_0 \subset \Phi$ the subset of function $\varphi(x,y)$ taking *integral* values on the vertices x and y of the complex X. Notice that Φ_0 naturally is the projective limit of finite sets and thus Φ_0 is homeomorphic to a Cantor set. The map $\Phi \longrightarrow \partial X$ restricts to a *surjective* map $\pi_0\colon \Phi_0 \longrightarrow \partial X$ which by the above argument is *finite-two-one*. In fact it is clear that the number $\#\pi_0^{-1}(a)$ does not exceed $(8d)^N$ where N is an upper bound to the number of vertices in X lying in a ball $B_x \subset X$ of radius 4d for all $x \in X$.

7.5.H The above "coding" of ∂X by Φ (and by Φ_0) is invariant under all isometries of X. Let us give a quasi-isometry invariant "coding" of ∂X by "quasigradient fields" on X rather than by "forms" on X.

A *field* ψ on X by definition is a subset $\psi \subset X \times X$ whose two projection to X are onto and such that

$$\text{length } \psi \underset{\text{def}}{=} \sup_{(x,y)\in\psi} |x-y| < \infty.$$

For example, with every function f on X one may associate

$$\psi = \text{grad}_{r,d} f = \{x,y \mid |x-y| \leqslant d,\ f(x)-f(y) \geqslant \lambda\}.$$

An *orbit* of a field ψ is a finite or infinite sequence of points x_i in X, such that $(x_i,x_{i+1}) \in \psi$ for all i. A field ψ of finite length is called *quasigeodesic* if every orbit $x_1,x_2,\ldots,x_k$ has $|x_1-x_k| \geqslant \lambda k$ for some number $\lambda = \lambda(\psi) > 0$. Call ψ *non-expanding* if every two orbits $x_1,\ldots,x_k$ and $y_1,\ldots,y_\ell$ with $|x_1-y_1| \leqslant \epsilon$ and $\big||x_1-x_k|-|y_1-y_\ell|\big| \leqslant 2$ length ψ satisfy

$$|x_k-y_\ell| \leqslant C + \mu(|x_1-x_k|+|y_1-y_\ell|),$$

where $\epsilon > 0$, $C \geqslant 0$ and $\mu < 1$ are some constant depending on ψ.

Now, if ψ is quasigeodesic then every infinite orbit of ψ converges to some point $a \in \partial X$ and if ψ is non-expanding this points a only depends on ψ (but not on a choice of an orbit of ψ). This gives a surjective map $\beta\colon \Psi \longrightarrow \partial X$, where Ψ is the space of the quasigeodesic non-expanding fields ψ on X. Furthermore, observe that $\beta(\psi_1) = \beta(\psi_2) \Leftrightarrow \psi_1 \cup \psi_2 \in \Psi$. Thus we obtained a quasi-isometry invariant "coding" of ∂X by fields $\psi \in \Psi$. This "coding: is *local* in the following sense. Given numbers ℓ, λ, ϵ, C and μ. Then for a field ψ of length $\psi \leqslant \ell$ the above conditions with these constants only need checking within each ball of radius R in X for some R depending on ℓ, λ, ϵ, C and μ. Furthermore, the equivalence relation $\beta(\psi_1) = \beta(\psi_2)$ between fields also is local. It is also clear that the set Ψ and the equivalence relation are quasi-isometry invariant.

7.6 <u>Coding ∂X with trees</u>: Fix a point $x_0 \in X$ and consider the spheres $S_i \subset X$ for $i = 0,1,2,\ldots,$

$$S_i = \{x \in X \mid |x_i| = i\}.$$

Take a subset $V \subset S_i$ for $i \geq 1$ and let $P(V) \subset S_{i-1}$ be the normal projection of V to S_{i-1} defined by

$$P(V) = \{x \in S_{i-1} \mid \mathrm{dist}(x,V) = 1\}.$$

Define for a subset $V \subset S_i$

$$\mathrm{Diam}_k V = \sup_j \mathrm{Diam}\, P^j(V),$$

where j runs over $0,1,2,\ldots,\max(i,k)$ and where $P^j = \underbrace{P \circ P \circ \ldots \circ P}_{j}$.

Now let X be geodesic δ-hyperbolic and consider the graph(1-complex) $T(X)$ whose vertices correspond to the subsets $V \subset S_i$ with $\mathrm{Diam}_{100} V \leq 10\delta$ for all $i = 0,1,\ldots,$ and where V_1 and V_2 are joined by an edge if and only if $V_2 = P(V_1)$. Clearly, this $T = T(X)$ is a tree growing from x_0. Denote by ∂T the hyperbolic boundary of T for the simplical metric with unit edges and observe that every geodesic ray in T issuing from x_0 is represented by a family of asymptotic rays in X. Thus we obtain a continuous map of ∂T onto ∂X.

Let us specialize to the case where X is itself a graph with a simplical metric having unit edges, such that each vertex in X has at most $N < \infty$ adjacent edges. Then the tree T also is locally finite with at most $\exp\exp(\delta+1)N$ edges at every vertex and the map $\partial T \longrightarrow \partial X$ is finite-to-one where "finite" $\leq \exp\exp\exp(\delta+1)N$. Checking all this is left to the reader.

§8. Isometry groups of hyperbolic spaces

According to 6.4 every hyperbolic metric group isometrically acts on some *geodesic* hyperbolic space - and by 7.5 this X can be assumed *ultracomplete*. In this section we fix such a ultracomplete δ-hyperbolic space X and study groups Γ isometrically acting on X.

8.1 Classification of individual isometries: An isometry $\gamma: X \longrightarrow X$ is called *elliptic* if the orbit $\{\gamma^i x\} \subset X$, $i \in \mathbb{Z}$, is bounded for some (and hence, for every) $x \in X$. Call γ *parabolic* if the orbit $\{\gamma^i x\}$ has a unique *limit point* $a \in \partial X$. That is the orbit is unbounded and $|\gamma^{i_j} x| \longrightarrow \infty$ implies $\gamma^{i_j} x \longrightarrow a$ for every subsequence $\{i_j\} \subset \mathbb{Z}$. A parabolic isometry is called *proper* if $|\gamma^i x| \longrightarrow \infty$ for $|i| \longrightarrow \infty$. Finally, an isometry γ is called *hyperbolic* if the orbit map $\mathbb{Z} \longrightarrow X$ for $i \longmapsto \gamma^i x$ is a quasi-isometry for the ordinary metric in $\mathbb{Z}$. If γ is hyperbolic, then, obviously, the limit set of the orbit $\{\gamma^i x\}$ consists of exactly two points:

$$\gamma^{\infty} = \lim_{i \to +\infty} \gamma^i x \text{ and } \gamma^{-\infty} = \lim_{i \to -\infty} \gamma^i x \text{ in } \partial X.$$

Notice that γ^{∞} and $\gamma^{-\infty}$ are independent of x. Also observe that the above classification of isometries is stable under replacement of γ by γ^{i_0} for any $i_0 \neq 0$.

8.1.A Lemma: *Let γ_1 and γ_2 be two non-hyperbolic isometries of X such that the points $x_j = \gamma_j(x)$ for $j = 1,2$, satisfy*

$$(*) \qquad |x-x_j| \geqslant 2(x_1.x_2)_x + 400\delta.$$

Then the isometries $\gamma_1\gamma_2$ and $\gamma_2\gamma_1$ are hyperbolic.

Proof: Since γ_1 and γ_2 are non-hyperbolic, Lemma 7.2.C (applied to the sequences $\gamma_1^i x$ and $\gamma_2^i x)$ shows that $|\gamma_j^2(x)| \leqslant |\gamma_j(x)| + 100\delta$ for $j = 1,2$. Then, by (*) and the δ-hyperbolicity of X, the isometries $\gamma_1\gamma_2$ and $\gamma_2\gamma_1$ satisfy the opposite inequality which yields via 7.2.C the hyperbolicity of these isometries.

8.1.B **Corollary**: *Every isometry γ is elliptic, parabolic or hyperbolic.*

Proof: If γ is non-elliptic and non-parabolic, then there exists two integers i_1 and i_2, such that the Lemma applies to γ^{i_1} and γ^{i_2}.

8.1.C Let us look at the displacement function $d_\gamma(x)$ of γ (see 7.4). If γ is hyperbolic then $d_\gamma(x)$ is bounded on the geodesic $\ell \subset X$ joining the two limit points and $d_\gamma(x) \longrightarrow \infty$ for $\mathrm{dist}(x,\ell) \longrightarrow \infty$ as is seen by looking at the normal projection of x to ℓ. If γ is parabolic, then the $\delta_\gamma(x)$ is bounded on the ray r in X between x and the limit $a \in \partial X$. The translates $\gamma^i(r)$ are 4δ-close to r near infinity for all $i \in \mathbb{Z}$. It follows that for every $i_0 = 1,2,$ there exists a point $x_0 \in X$ such that $d_{\gamma^i}(x_0) \leqslant 8\delta$ for $|i| \leqslant i_0$.

8.1.D **Corollary**: *If Γ is a word hyperbolic group, then every non-torsion element γ in Γ is hyperbolic. That is the homomorphism $\mathbb{Z} \longrightarrow \Gamma$ sending $i \mapsto \gamma^i$ is quasi-isometric.*

This corollary shows that if two elements γ_1 and γ_2 in Γ are conjugate, then there exists an $\alpha \in \Gamma$, such that $\alpha\gamma_1\alpha^{-1} = \gamma_2$ and $|\alpha| \leqslant 2(|\gamma_1| + |\gamma_2| + 8\delta)$. Since the norm $\gamma \longmapsto |\gamma|$ is a primitive recursive function (by 2.3), we see that *the conjugacy problem in Γ is solvable* (compare 7.4).

8.1.E The geometric properties of isometries (see 7.4) and quasi-convex sets in X (see 7.3) lead to the following dynamical corollaries as closed subsets A in ∂X generate (or span) quasi-convex subsets L(A) in X (see 7.5).

8.1.F *Let γ be proper parabolic and let $U \subset X \cup \partial X$ be an arbitrary neighborhood of the limit point $a \in \partial X$. Then there exists an integer i_0 such that the complement $(X \cup \partial X) \setminus U$ is sent to U by the isometry γ^i for all $i \geq i_0$.*

8.1.G *Let γ be hyperbolic and let U_+ and U_- in $X \cup \partial X$ be some neighborhoods of the limit points $\gamma^{+\infty}$ and $\gamma^{-\infty}$ in ∂X correspondingly. Then there is an i_0, such that $(X \cup \partial X) \setminus U_-$ is sent by γ^i to U_+ for all $i \geq i_0$ and $(X \cup \partial X) \setminus U_+$ is sent by γ^{-i} to U_-. Furthermore if i_0 is sufficiently large then the maps*

$$\gamma^i: \partial X \setminus U_- \longrightarrow \partial X \cap U_+$$

and

$$\gamma^{-i}: \partial X \setminus U_+ \longrightarrow \partial X \cap U_-$$

are 2-contracting for the metric in ∂X introduced in 7.2.

(A map $f: A \longrightarrow B$ is called 2-contracting if $|f(a_1)-f(a_2)| \leq 1/2 |a_1-a_2|$ for all a_1 and a_2 in A.)

8.2 **<u>Non-elementary isometry groups Γ and their action on $\partial\Gamma$, $\partial^2\Gamma$ and $\partial^3\Gamma$</u>:** Let Γ be a non-elliptic isometry group of Γ, that is some (and hence every) orbit $\{\Gamma x\} \subset X$ for $x \in X$ is unbounded. Then we see as in 8.1 that either Γ is *parabolic*, that is there is a unique limit point $a \in \partial X$ or Γ contains a hyperbolic element γ. Recall that the limit set $\partial\Gamma \subset \partial X$ by definition is the hyperbolic boundary of an orbit $\{\Gamma x\} \subset X$ which clearly is independent of $x \in X$.

8.2.A **Lemma**: *Let* $A \subset \partial X$ *be a closed* Γ*-invariant subset containing at least two points. Then* $A \supset \partial\Gamma$.

Proof: Take $L(A) \subset X$ (see 7.5) and $x \in L(A)$. Then $\{\Gamma x\} \subset L(A)$ and the lemma follows.

8.2.B If $\partial\Gamma = \partial\{\gamma^i\}_{i\in\mathbb{Z}}$ for some hyperbolic element $\gamma \in \Gamma$, then by the proof of 8.2.A *each orbit* $\{\Gamma x\} \subset X$ *is quasi-isometric to* $\mathbb{Z}$.

8.2.C **Corollary**: *Every non-torsion element* γ *in a word hyperbolic group* Γ *is contained in a unique maximal elementary subgroup* $E_\gamma \subset \Gamma$ *which is a finite extension of the infinite cyclic group generated by* γ. *Hence* $\gamma = \gamma_0^p$ *for some prime element* $\gamma_0 \in \Gamma$ *and* $p \in \mathbb{N}$. *The exponent* $p = p(\gamma)$ *is unique. Moreover if* Γ *has no torsion then* $\gamma_0 = \gamma_0(\gamma)$ *also is unique.*

8.2.D Now, let $\partial\Gamma$ contain at least three points. Then there exists at most one point $a \in \partial\Gamma$ fixed under Γ. If such a exists then Γ is called *quasiparabolic*. It follows from 3.2.A that Γ acts *minimally* away from a. That is the orbit of every non-fixed point $b \in \partial\Gamma$ is dense in $\partial\Gamma$. Hence, the points $\gamma^\infty \in \partial X$ for the hyperbolic $\gamma \in \Gamma$ are dense in ∂X and therefore $\partial\Gamma$ has no isolated points. Since $\partial\Gamma$ is complete it is uncountable. In this case Γ is called *non-elementary*.

8.2.E If Γ is non-elementary then, by the above, there are two hyperbolic elements γ_1 and γ_2 in Γ, such that $\gamma_2^{+\infty} \neq \gamma_1^{\pm\infty}$. Take small neighborhoods $U_j \subset \partial\Gamma$ of $\gamma_j^{+\infty}$ for $j = 1,2$ and observe with that

$$\gamma_j^i(U_1 \cup U_2) \subset U_j$$

for $j = 1,2$, for all $i \geq i_0$ and for some $i_0 > 0$. It follows that the semigroup F_2^+ generated by $\gamma_1^{i_1}$ and $\gamma_2^{i_2}$ in Γ is *free* provided i_1 and i_2 are $\geq i_0$. Furthermore there is a unique closed subset in $U_1 \cup U_2$ univariant under F_2^+ which is homeomorphic to the Cantor set.

8.2.F Now let Γ be non-quasiparabolic as well as non-elementary. Then, by an obvious argument, for any two non-empty open subsets U_1 and U_2 in $\partial\Gamma$ there exists hyperbolic elements γ_1 and γ_2, such that $\gamma_1^{+\infty} \in U_1$, $\gamma_2^{+\infty} \in U_2$ and $\gamma_1^{-\infty} \neq \gamma_2^{-\infty}$. Take disjoint neighborhoods U_1^- and U_2^- in $\partial\Gamma$ of $\gamma_1^{-\infty}$ and $\gamma_2^{-\infty}$ correspondingly and assume that U_1 and U_2 also are disjoint. Take a sufficiently large i_0, such that γ_1^i sends $U_1 \cup U_2 \cup U_2^-$ to U_1 for $i \geq i_0$ (see 8.1) and γ_1^{-i} sends $U_1^- \cup U_2 \cup U_2^-$ to U_1^-, while the transformations γ_2 send $U_2 \cup U_1 \cup U_1^-$ and $U_2^- \cup U_1 \cup U_1^-$ to the subsets U_2 and U_2^- correspondingly. It immediately follows that the subgroup $F \subset \Gamma$ generated by $\gamma_1^{i_1}$ and $\gamma_2^{i_2}$ is free if i_1 and i_2 are $\geq i_0$, and that all isometries in this subgroup are hyperbolic. Furthermore the orbit map $f \longmapsto fx$ is a quasi-isometry of F into X for the word metric in F. The extension of this quasi-isometry to ∂F (which is homeomorphic to the Cantor set) sends ∂F into the union $U_1 \cup U_2 \cup U_1^- \cup U_2^-$. (For large i_0 this ∂F equals the maximal F-invariant subset in this union.) Finally we observe that the product $\alpha = \gamma_1^{i_1}\gamma_2^{-i_2}$ has $\alpha^{+\infty} \in U_1$ and $\alpha^{-\infty} \in U_2$.

8.2.G __Corollary__: *The set of pairs* $(\gamma^{+\infty},\gamma^{-\infty}) \in \partial\Gamma \times \partial\Gamma$ *for all hyperbolic* $\gamma \in \Gamma$ *is dense in* $\partial\Gamma \times \partial\Gamma$. *In particular, every non-elementary word hyperbolic group* Γ *has infinitely many conjugacy classes of prime non-torsion elements.*

By a similar argument one obtains

8.2.H *The action of Γ on $\partial\Gamma \times \partial\Gamma$ is topologically transitive. That is every open non-empty Γ-invariant subset is dense.*

Since $\partial\Gamma$ is complete one concludes,

8.2.I *There is a point in $\partial\Gamma \times \partial\Gamma$ whose Γ-orbit is dense.*

8.2.J **Expansion points:** A point $a \in \partial X$ is called an *E-point* for Γ if for every $\lambda \geqslant 1$ there exists a neighborhood $U \subset \partial X$ of a and an element $\gamma \in \Gamma$, such that the action of γ on U is λ-*expanding* for the metric on ∂X defined in 7.2. That is

$$|\gamma(a_1)-\gamma(a_2)| \geqslant \lambda |a_1-a_2|$$

for all a_1 and a_2 in U. It is easy to see that all E-points lie in the limit set $\partial\Gamma \subset \partial X$. In fact, by the definition of the metric in ∂X, a point $a \in \partial X$ is an E-point if and only if the intersection of the Γ-orbit of some ball $B \subset X$ with some ray r in X from a fixed point $x_0 \in X$ to a is unbounded,

$$\mathrm{Diam}(\Gamma B) \cap r = \infty.$$

For example, if the quotient X/Γ is bounded then every $a \in \partial X$ is an E-point. In particular, *if Γ is word hyperbolic then every $a \in \partial\Gamma$ is an E-point.* More generally let Γ' be a subgroup in the word hyperbolic group Γ and let $\partial\Gamma' \subset \partial\Gamma$ be the limit set. *Then the following two conditions are equivalent.*

(1) *All points $a' \in \partial\Gamma'$ are E-points for Γ';*

(2) *Γ' is word hyperbolic (and, hence finitely*

presented) and the inclusion $\Gamma' \hookrightarrow \Gamma$ *is quasi-isometric for the word metrics in* Γ' *and* Γ.

8.2.K <u>Action of Γ on $\partial^3 X$</u>: Let $\partial^3 X \subset \partial X \times \partial X \times \partial X$ be the set of triples (a_1,a_2,a_3), where $a_1 \neq a_2 \neq a_3 \neq a_1$. To each point $(a_1,a_2,a_3) \in \partial^3 X$ one assigns a *bounded* subset $C \subset X$ as follows

$$C = C_\epsilon = \{x \in X \mid \mathrm{dist}(x,\ell_{ij}) \leqslant \epsilon\},$$

for i,j = 1,2,3 and for all lines ℓ_{ij} in X between a_i and a_j. If $\epsilon \geqslant 10\delta$, then, clearly, the set C_ϵ is non-empty. This construction establishes a correspondence between bounded subsets in X and *codiagonal* subsets $A \subset \partial^3 X$, where "codiagonal" is defined by

$$\inf \min(|a_1-a_2|,|a_2-a_3|,|a_1-a_3|) > 0,$$

where inf is taken over all $(a_1,a_2,a_3) \in A$ and the distance is measured in ∂X. For example, if ∂X is compact then "codiagonal" amounts to "relatively compact".

The above "bounded $\Leftrightarrow$ codiagonal" correspondence leads, in particular, to the following

8.2.L <u>Proposition-Definition</u>: *A set* Δ *of isometries of* X *is called bounded if one of the following four equivalent conditions is satisfied.*

(1) $\sup_{\gamma \in \Delta} |x-\gamma(x)| < \infty$ *for a fixed* $x \in X$.

(2) *The set of homeomorphisms* $\gamma: \partial X \longrightarrow \partial X$ *for* $\gamma \in \Delta$ *is uniformly continuous.*

(3) *There exists a codiagonal subset* $A \subset \partial^3 X$, *such that* $A \cap \gamma(A)$ *is non-empty for all* $\gamma \in \Delta$.

(4) *There exists an* $\epsilon = \epsilon(\Delta) > 0$ *such that for every*

ϵ-ball $B \subset \partial X$ *and each* $\gamma \in \Delta$ *the complement of* $\gamma(B)$ *satisfies*

$$\mathrm{Diam}(\partial X \setminus \gamma(B)) \geq \epsilon.$$

Notice that (4) implies that Δ is bounded $\Leftrightarrow$ Δ^{-1} is bounded. A related notion is that of B-*finiteness* of an isometry group Γ of X which is defined by the conditions: every bounded subset Δ in Γ is finite. If X is a Riemannian manifold this amounts to the usual discreteness of Γ. Here is a similar generalization of cocompactness.

8.2.M __Definition:__ An action of Γ is called *cobounded* if one of the following two equivalent (by the previous discussion) conditions is satisfied.

(1) There is a bounded subset $Y \subset X$, such that $\Gamma(Y) = X$.

(2) There is a codiagonal subset $A \subset \partial^3 X$, such that $\Gamma(A) = \partial^3 X$.

Notice, that for word hyperbolic groups Γ the boundary $\partial\Gamma$ is compact and the action of Γ on $\partial^3 X$ is discrete cocompact (hence it is B-finite and cobounded).

8.2.N __Example:__ If X is a complete convex manifold with $K < -\epsilon < 0$, then the space $\partial^3 X$ can be identified with the total space S_2X of the orthonormal 2-bundle of X by the following well-known construction of Ahlfors-Cheeger. Take a point $(a_1,a_2,a_3) \in \partial^3 X$ join a_1 and a_2 by the (unique for $K < -\epsilon < 0$) line $(a_1,a_2) \subset X$ and observe that the normal projection $P: X \longrightarrow (a_1,a_2)$ continuously extends to $\partial X \setminus \{a_1,a_2\}$. Now we define a homeomorphism $h: \partial^3 X \longrightarrow S_2X$ by sending each (a_1,a_2,a_3) to the frame of vectors at $P(a_3) \in (a_1,a_2) \subset X$ tangent to the line (a_1,a_2) and to the ray $[P(a_3),a_3) \subset X$. Notice that h commutes with isometries and thus "descends" to quotients X/Γ.

8.2.P __Corollary__ (Cheeger): *Let* V_1 *and* V_2 *be* *compact manifolds with* $K < 0$ *and let* Γ_1 *and* Γ_2 *be their fundamental groups. Then every isomorphism* $\Gamma_1 \to \Gamma_2$ *induces a homeomorphism of (the total space of) the 2-frame bundle of* V_1 *to that of* V_2*, such that compositions of isomorphisms go to compositions of homeomorphisms.*

Example: The group of exterior automorphisms of the fundamental group of every closed surface V acts by homeomorphisms on the unit tangent bundle of V (which is essentially the same as the 2-frame bundle for dim V = 2).

8.2.Q **Remark**: Many properties of hyperbolic groups can be seen by looking at the action of Γ on ∂X, without ever mentioning X. The key property to use is the equivalence

(*) $$(2) \Leftrightarrow (3) \Leftrightarrow (4)$$

for all $\Delta \subset \Gamma$ (see 8.2.L). It is conceivable that the existence of a complete metric space D and of an action of a group Γ on D by uniform homeomorphisms satisfying (*) makes Γ hyperbolic for a suitable metric on Γ. For example, let "bounded" stand for "codiagonal" in $D \times D \times D$ and consider a B-finite (now B is for codiagonal rather than for bounded) and cobounded action of a group Γ on D. Such a Γ looks very much like a word hyperbolic one. Notice that B-finite cobounded actions are "rather rigid": for two such actions of Γ on D_1 and on D_2 every Γ-equivariant map $D_1 \to D_2$ is injective.

As another example let us generalize 8.1 (compare [GM]).

8.2.R *Let* $\gamma: D \to D$ *be a uniformly continuous map such that for some* $k \geq 2$ *one of the following three conditions is satisfied.*

(a_k) *For every codiagonal subset A in*

$D^k = \underbrace{D \times D \times \ldots \times D}_{k}$ *there is at most finitely many*

$i \in \mathbb{Z}_+$, *such that* $A \cap \gamma^i(A) \neq \emptyset$, *where "codiagonal" means* $\mathrm{dist}(A,\Sigma) > 0$ *for the union* Σ *of the diagonals (defined by* $d_i = d_j$ *for* $(d_1,\ldots,d_k) \in D)$ *in* D.

(b_k) *For every codiagonal* A *there is an integer* m *such that every orbit of* γ *(acting on* D^k*) meets* A *at most* m*-times. (Here "an orbit" means a map* $\mathbb{Z}_+ \longrightarrow D^k$ *and the meeting points with* A *are the pull-backs of* A*; thus every fixed point in* A *meets* A *infinitely many times).*

(c_k) *For the above* A *there is an* m, *such that* A *contains no segment of any orbit of length* m, *that is* $x, \gamma(x), \gamma^2(x), \ldots, \gamma^m(x)$, *for* $x \in D^k$.

Then there is a non-empty finite γ*-invariant subset* $D_0 \subset D$, *such that every orbit asymptotically approaches* D_0.

The proof is straightforward and we leave it to the reader. Here we only notice that $(a_k) \Rightarrow (b_k) \Rightarrow (c_k)$ and that (a_3) is our earlier B-finiteness condition.

8.2.S The above result extends to many groups and semigroups acting on D. Furthermore, if D is compact and Γ is an *amenable* group acting on D, such that no point in $D^k \setminus \Sigma$ is recurrent (this generalizes (c_k)), then there exists a non-empty finite Γ-invariant subset $D_0 \subset D$. In fact, the non-recurrency of $D^k \setminus \Sigma$ condition shows that the support of every Γ-invariant measure on D is finite.

<u>Remark</u>: Many interesting actions of Γ on D have invariant *non-diagonal* subsets $\Sigma \subset D^k$ to which the above discussion applies. For example, if D is a projective space then a useful Σ consists of k-tuples of points lying in a (k-2)-subspace. A

fundamental result here is

Tits' free group theorem: *If a finitely generated group* Γ *of projective transformations of* $\mathbb{C}P^n$ *contains no free subgroups then there is a nonempty finite union of* k-*dimensional subspaces in* $\mathbb{C}P^n$ *invariant under* Γ *for all* $k = 0,1,\dots,n-1$.

(See [Ti] and compare [Fu].)

8.3 Geodesic flows and hyperbolic *simplices*: Start with two geodesic rays or segments r_1 and r_2 in a convex manifold X with $K \leqslant -4\epsilon^2 < 0$ which converge to some point $a \in X \cup \partial X$. Then one knows that r_1 and r_2 converge *exponentially* fast. Namely, one can parametrize r_1 and r_2 by length,

$$r_i\colon [0,t_i] \longrightarrow X \quad i = 1,2 \text{ and } r_i \in [0,\infty],$$

such that

$$(+) \qquad |r_1(t+\theta)-r_2(t+\theta)| \leqslant 2(\exp-\epsilon\theta)|r_1(t)-r_2(t)|,$$

for all $t \in [0,t_1) \cap [0,t_2)$ and for those $\theta \geqslant 0$ which satisfy

$$t + \theta \in [0,t_1) \cap [0,t_2)$$

and

$$\theta \geqslant 2|r_1(t)-r_2(t)|.$$

In particular every two distinct points a and b in $X \cup \partial X$ can be joined by a *unique* geodesic in X. It follows that the space of double-infinite geodesics in X is canonically homeomorphic to $\partial^2 X = \{x,y \in \partial X \mid x \neq y\}$.

If X is a general hyperbolic space then one might have several geodesics joining two points in X or in ∂X. Our goal is to find a quasi-isometric perturbation of X to another space X' which would satisfy some version of (+). This is quite easy to achieve if one does not care about the isometries. But what one needs is an X' isometrically acted upon by the isometry group Γ of X. Here one can meet a non-trivial cohomological obstruction even for an X which is quasi-isometric to $\mathbb{R}$.

8.3.A __Example:__ Let Γ be an infinite group, such that $H^1(\Gamma;\mathbb{R}) = 0$ and which admits an unbounded *quasi-homomorphism* $h\colon \Gamma \longrightarrow \mathbb{R}$. That is $h(\gamma) = h(\gamma^{-1})$ for all $\gamma \in \Gamma$ and

$$|h(\gamma_1\gamma_2)-h(\gamma_1)-h(\gamma_2)| \leq d$$

for some $d > 0$ and all γ_1 and γ_2 in Γ. It follows from [Gr2] that every non-elementary word hyperbolic group Γ (e.g. $\Gamma = \{\alpha,\beta \mid \alpha^7 = 1,\ \beta^7 = 1\}$ which has $H^1(\Gamma;\mathbb{R}) = 0$) admits such an h.

Now, take a constant c, such that $h^{-1}[0,c] \subset \Gamma$ generates Γ and consider the word metric $|\ |$ in Γ corresponding to the (infinite!) generating subset $h^{-1}[0,c'] \subset \Gamma$ for $c' = 10(c+d)$. Clearly the map $h\colon (\Gamma,|\ |) \longrightarrow \mathbb{R}$ is quasi-isometric. On the other hand since $H^1(\Gamma;\mathbb{R}) = 0$, there is no (non-trivial) isometric action of Γ on $\mathbb{R}$. Notice that $\mathbb{R}$ is the only candidate for X' in this case and so $(\Gamma,|\ |)$ can not be *equivariantly* "straightened out".

This example may seem non-conclusive, as the action of Γ on $\partial\Gamma = \partial\mathbb{R} = \{\pm\infty\}$ is trivial. One gets a true example by taking a free product of two copies of $(\Gamma,|\ |)$.

Notice that amenable (in particular finite) groups have trivial bounded cohomology (see [Gr2]) and so the above phenomenon can not appear for B-finite (or amenable) Γ-action. In fact we shall see below that every word hyperbolic group does admit a well behaved "geodesic flow".

8.3.B First we define *the geodesic flow* for an arbitrary metric space X by considering the space GX of all isometric maps $g\colon \mathbb{R} \longrightarrow X$ with the metric

$$|g_1-g_2|_{GX} = \int_{-\infty}^{+\infty} \left(|g_1(t)-g_2(t)|_X\right) 2^{-|t|}\, dt.$$

The translations in $\mathbb{R}$ define a (non-isometric) action of $\mathbb{R}$ on GX called the *geodesic flow*. Observe that the projection $GX \longrightarrow X$ given by $g \longmapsto g(0)$ is a quasi-isometry. Also observe that the isometry group Is = IsX isometrically acts on GX and the actions of Is and of $\mathbb{R}$ agree in an obvious sense. Furthermore, the group $\mathbb{Z}_2$ acts on GX by the involution $g(t) \longleftrightarrow g(-t)$ and this action of $\mathbb{Z}_2$ agrees with those of Is and $\mathbb{R}$.

Now, let X be hyperbolic and ultracomplete. Then there is a surjective projection

$$D: GX \longrightarrow \partial^2 X = \{a,b \in \partial X \mid a \neq b\},$$

given by $g \longmapsto (g(-\infty), g(+\infty))$. The pull-back $D^{-1}(a,b)$ consists, in general, of several $\mathbb{R}$-orbits. Our main objective is to construct a quotient space of GX by identifying these orbits to a single one for all $(a,b) \in \partial^2 X$, such that a given action on an isometry group Γ on X (and hence on GX) goes to an action on the quotient space. Our example indicates a cohomological obstruction to such an identification of orbits. In particular, one can not expect any *canonical* construction of the desired quotient space. However we have the following

8.3.C <u>**Theorem:**</u> *Let Γ be a word hyperbolic group. Then there exists a locally compact finite dimensional hyperbolic metric space $\hat{G}$ acted upon by the groups Γ, $\mathbb{R}$ and $\mathbb{Z}_2$ such that the following six conditions are satisfied.*

(1) *The three actions agree in the same way as those on GX.*

(2) *The actions of Γ and of $\mathbb{Z}_2$ are isometric.*

Furthermore, the action of Γ *is B-finite (i.e. discrete) and the orbit map* $\Gamma \longrightarrow \hat{G}$ *for* $\gamma \longmapsto \gamma g_0$ *is a quasi-isometry for every* $g_0 \in \hat{G}$. *This quasi-isometry continuously extends to a canonical (in particular, independent of* g_0*)* Γ*-equivariant homeomorphism* $\partial\Gamma \leftrightarrow \partial\hat{G}$.

(3) *The action of* $\mathbb{R}$ *on* $\hat{G}$ *is free and every orbit* $\mathbb{R} \longrightarrow \hat{G}$, *denoted by* $t \longmapsto g(t)$, *is a quasi-isometric embedding. The map* $g \longmapsto (g(-\infty), g(+\infty)) \in \partial^2\hat{G} = \partial^2\Gamma$ *defines a canonical homeomorphism* $\hat{G}/\mathbb{R} \leftrightarrow \partial^2\Gamma$, *where the projection* $D\colon \hat{G} \longrightarrow \hat{G}/\mathbb{R} = \partial^2\Gamma$ *is a locally trivial fibration.* (Since $\partial^2\Gamma$ is paracompact and $\mathbb{R}$ is contractible this fibration is trivial, that is $\hat{G} \approx \partial^2\Gamma \times \mathbb{R}$. Yet there is no *canonical* homeomorphism between $\hat{G}$ and $\partial^2\Gamma$.) As it follows from the construction, the map $D\colon \hat{G} \longrightarrow \partial^2\Gamma$ is $\Gamma \times \mathbb{Z}_2$-equivariant for the canonical Γ action on $\partial^2\Gamma$ (coming from that on $\partial\Gamma$) and the involution $(a,b) \leftrightarrow (b,a)$ on $\partial^2\Gamma$.

(4) *If two* $\mathbb{R}$*-orbits* $g_1(t)$ *and* $g_2(t)$ *converge for* $t \longrightarrow \infty$, *that is if* $g_1(\infty) = g_2(\infty)$, *then the distance from* $g_1(t)$ *to* $\bar{g}_2 = \bigcup_t g_2(t) \subset \hat{G}$ *exponentially decays for* $t \longrightarrow \infty$. *That is* $d(t) = \mathrm{dist}(g_1(t), \bar{g}_2) \leqslant \mathrm{const}\ \exp{-\epsilon t}$, *for some* const > 0 *and* $\epsilon > 0$. *Moreover, there exists an* $\epsilon = \epsilon(\hat{G}) > 0$, *such that the distance* $d(t)$ *for every two (convergent or not) orbits satisfies*

$$(++) \qquad d\left(\frac{t_1+t_2}{2}\right) \leqslant \exp(-\epsilon\,|t_1-t_2| + d(t_1) + d(t_2)),$$

for all t_1 *and* t_2 *in* $\mathbb{R}$.

(5) *Let X be an ultracomplete hyperbolic metric space which is isometrically acted upon by* Γ *such the action is B-finite and cobounded. Then there exists a (non-canonical) continuous quasi-isometric*

$\Gamma \times \mathbb{Z}_2$-equivarient map $GX \longrightarrow \hat{G}$ *which homeomorphically maps every geodesic in* GX *onto an* $\mathbb{R}$*-orbit in* $\hat{G}$.

(6) *The space* $\hat{G}$ *satisfying* (1), (2) *and* (3) *is unique in the following sense. For every two such, say* $\hat{G}_1$ *and* $\hat{G}_2$, *there exists a (non-canonical)* $\Gamma \times \mathbb{Z}_2$*-equivariant homeomorphism* $\hat{G}_1 \leftrightarrow \hat{G}_2$ *which maps* $\mathbb{R}$*-orbits in* $\hat{G}_1$ *onto those in* $\hat{G}_2$.

__Proof__: Start with some X as in (5), take two points a and b on ∂X, consider two geodesics g_1 and g_2 in X between a and b and let us produce an equivalence relation between points $x_1 = g_1(t_1) \in g_1$ and $x_2 = g_2(t_2) \in g_2$ on these geodesics, such that the resulting quotient space will be our $\hat{G}$. First we fix a point $x_0 \in X$ and take a non-negative continuous function $\psi(x) = \psi(x,x_0)$ for $x \in X$ of the form $\psi(x,x_0) = \varphi(|x-x_0|)$, where $\varphi(\rho) = 1$ for $\rho \leqslant \rho_0$ and $\varphi(\rho) = 0$ for $\rho < \rho_1$ for two numbers $0 < \rho_0 < \rho_1 < \infty$. Next we take a non-negative continuous function $\eta = \eta(a,b)$ on $\partial^2 X$ of the form $\eta(a,b) = \zeta(|a-b|)$ for some metric on ∂X such that

(i) if every geodesic in X between a and b meets the ball in X of radius ρ'_0 around x_0, then $\eta(a,b) = 1$;

(ii) if some geodesic in X between a and b misses the ball of radius ρ'_1 around x_0, then $\eta(a,b) = 0$.

Here ρ'_0 and ρ'_1 are some constants in the interval $0 < \rho'_0 < \rho'_1 < \rho_0$.

Now take a geodesic $g = g(t)$ in X between two points a and b in ∂X such that $\eta(a,b) \neq 0$ and let

$$L_g(t) = \int_{t_0}^{t} \psi(g(t))dt,$$

where t_0 is the "center of gravity" of the function $\psi(g(t))$. That is

t_0 is determined by the equality

$$\int_{-\infty}^{+\infty} (t_0-t)\psi(g(t))dt = 0.$$

Then for points $x_1 = g(t_1)$ and $x_2 = g_2(t_2)$ on two different geodesics between a and b we set

$$L_0(a,b;x_1,x_2) = \eta(a,b)(L_{g_1}(t_1)-L_{g_2}(t_2))$$

where L_0 is extended by zero to those (a,b) where L_g were not previously defined.

Since the action of Γ is B-finite on X (and $\partial^3 X$) we can define

$$L(a,b;x_1,x_2) = \sum_{\gamma\in\Gamma} L_0(\gamma a,\gamma b,\gamma x_1,\gamma x_2).$$

8.3.D Let us enumerate basic properties of the function L.

(1) L is Γ-invariant.

(2) L is a cocycle in (x_1,x_2). That is $L(a,b;x_1,x_2) = -L(a,b;x_2,x_1)$ and $L(a,b;x_1,x_2) + L(a,b;x_2,x_3) + L(a,b;x_3,x_1) = 0$.

(3) $L(a,b;x_1,x_2) = -L(b,a;x_1,x_2)$.

(4) If ρ_0' is sufficiently large, then $|L(a,b;x_1,x_2)| > 0$ for all $a \neq b$ and all points $x_1 \neq x_2$ which lie on *same* geodesic between a and b. This follows from the coboundness of the action of Γ on X.

Now, assuming ρ_0' large, we identify any two geodesics between a and b, say g_1 and g_2 by the equivalence relation $x_1 \sim x_2 \Leftrightarrow L(a,b;x_1,x_2) = 0$ for $x_i \in g_i$, $i = 1,2$, and define $\hat{G}$ to be the resulting quotient space. If we start with a locally compact X, for example with the polyhedron $P_d^1(\Gamma)$, then $\hat{G}$ with the quotient topology is (obviously) locally compact and finite dimensional and (1).

(2) and (3) hold for any $\Gamma \times \mathbb{Z}_2$-invariant metric on $\hat{G}$. (Such a metric exists since the action of $\Gamma \times \mathbb{Z}_2$ is discrete.) The metric satisfying (4) is constructed in 8.3.F below. Now, let us prove (5). First we observe that there is a $\Gamma \times \mathbb{Z}_2$-invariant discontinuous quasi-isometry $GX \longrightarrow \hat{G}$ sending $\mathbb{R}$-orbits to $\mathbb{R}$-orbits. This map can be made continuous with a $\Gamma \times \mathbb{Z}_2$-invariant partition of unity in GX. What remains is to make such a map $GX \longrightarrow \hat{G}$ homeomorphic on every $\mathbb{R}$-orbit. Let us observe that for every quasi-isometry $f: \mathbb{R} \longrightarrow \mathbb{R}$ and for every continuous non-negative function $\varphi: \mathbb{R} \longrightarrow \mathbb{R}$ with compact support the φ_2-*diffusion* of f defined by

$$f_\lambda(t) = \int_{\mathbb{R}} f(t+\tau)\varphi(\lambda^{-1}\tau)d\tau \Big/ \int_{\mathbb{R}} \varphi(\lambda^{-1}\tau)d\tau$$

is a *homeomorphism* for every sufficiently large $\lambda > 0$, which only depends on the implied quasi-isometry constant of f. This diffusion applies to the map $GX \longrightarrow \hat{G}$ on every orbit thus making this map homeomorphic on the orbits. The same construction yields maps $\hat{G}_1 \longrightarrow \hat{G}_2$ required by (6).

8.3.E Corollary: *The geodesic flow on a compact manifold* V *of negative curvature is uniquely determined up to a time reparametrization by* $\Gamma = \pi_1(V)$. *(In other words the* 1*-dimensional foliation into orbits is determined by* Γ.*) Furthermore let* A *be a finite group of exterior* (*i.e. in* Aut Γ/Int Γ) *automorphisms of* Γ. *Then there is a homeomorphic action of* A *on the unit tangent bundle* GV *of* V *sending geodesic orbits to orbits and inducing (in an obvious sense) the given action of* A *on* Γ. (Compare 8.2.)

Proof: Apply the theorem to the (hyperbolic!) semi-direct product $\Gamma \ltimes A$.

8.3.F **Metrics in $\hat{G}$ satisfying (4):** Consider a locally compact space H and a partition $\mathcal{L}$ of H into 1-dimensional manifolds called *leaves* (or *orbits*). A *split neighborhood* in H by definition is an open subset in H split into Carterian product of some metric space by the real line, say $B \times \mathbb{R} \subset H$, such that every fiber $b \times \mathbb{R} \subset \mathbb{R}$, $b \in B$ is contained in a single leaf (depending on b) of $\mathcal{L}$. Call $\mathcal{L}$ a *foliation* if each point in H admits a split neighborhood.

Basic example: The quotient space $\hat{G}/\Gamma \times \mathbb{Z}_2$ partitioned into the images of $\mathbb{R}$-orbits in $\hat{G}$ is a foliation.

Definition: A *transversal Lipschitz structure* in a foliation $\mathcal{L}$ is given by a metric $|\ |_B$ in B for each split neighborhood $B \times \mathbb{R} \subset H$, such that for every two neighborhoods $B_1 \times \mathbb{R} \subset B_2 \times \mathbb{R} \subset H$ the obvious map $B_1 \longrightarrow B_2$ is a bi-Lipschitz homeomorphism of B_1 onto its image in B_2.

Basic example: Every metric in $\partial^2\hat{G}$ for which the action of $\Gamma \times \mathbb{Z}_2$ is Lipschitz induces in an obvious way a transversal Lipschitz structure in the $\mathbb{R}$-orbits foliation on $H = \hat{G}/\Gamma \times \mathbb{Z}_2$.

Now, let $\mathcal{L}$ be a foliation with a fixed transversal Lipschitz structure. A metric $|\ |_H$ on H is called *Lipschitz* if for every split neighborhood the identity map

$$(B \times \mathbb{R}, |\ |_H) \longleftrightarrow (B \times \mathbb{R}, |\ |_B \times |\ |_{\mathbb{R}})$$

is bi-Lipschitz for the standard metric $|\ |_{\mathbb{R}}$ in $\mathbb{R}$ and where $|\ |_B \times |\ |_{\mathbb{R}}(b,t) = |b|_B + |t|_{\mathbb{R}}$.

8.3.G **Lemma:** *Every $\mathcal{L}$ admits a Lipschitz metric. Furthermore such a metric is unique in the following sense. If $|\ |_1$ and $|\ |_2$ are two Lipschitz metrics on* H *then there exists a bi-Lipschitz homeomorphism* $(H, |\ |_1) \longleftrightarrow (H, |\ |_2)$ *which sends every leaf of $\mathcal{L}$ into itself and which is arbitrarily close* *to the identity map.*

<u>Proof</u>: First we prove the uniqueness. The identity map $(H,|\ |_1) \longrightarrow (H,|\ |_2)$ can be approximated by a leaves preserving Lipschitz map $f: (H,|\ |_1) \longrightarrow (H,|\ |_2)$ with a $|\ |_1$-Lipschitz partition of unity in H. Then we apply a φ_λ-diffusion to f with a small (depending on the precision of the preceding approximation) $\lambda > 0$, in order to make f bi-Lipschitz. Routine details are left to the reader.

Now, to prove the existence, decompose H into countable union of split neighborhoods, $H = \bigcup_i B_i \times \mathbb{R}$.

The desired metric on H is obtained by extending such from $H_j = \bigcup_{i \leq j} B_i \times \mathbb{R}$ to H_{j+1}, where the uniqueness of the metric on $(B_{j+1} \times \mathbb{R}) \cap H_j$ ensures the possibility of such an extension. Again, the details are left to the reader.

8.3.H <u>Corollary</u>: *Let $|\ |$ be a metric on $\partial^2\hat{G}$ for which the action of $\Gamma \times \mathbb{Z}_2$ is Lipschitz. Then there exists a $\Gamma \times \mathbb{Z}_2$-invariant metric in $\hat{G} = \partial^2\hat{G} \times \mathbb{R}$ for which the identity map*

$$(\hat{G},|\ |_{\hat{G}}) \leftrightarrow (\partial^2\hat{G} \times \mathbb{R},|\ | \times |\ |_{\mathbb{R}})$$

is bi-Lipschitz.

8.3.I <u>The proof of (4)</u>: Now we involve the metric $|\ |_\mu$ on $\partial\hat{G}$ defined in 7.2 and observe that the exponential convergence property is "transversally" satisfied for the corresponding "transversally defined metric" on $\hat{G} = \partial^2\hat{G} \times \mathbb{R}$. Hence, this property is satisfied for the (true) metric provided by 8.3. Q.E.D.

8.3.J <u>Remark</u>: The construction of $\hat{G} = \hat{G}(\Gamma)$ can be performed in a more canonical way without using any auxiliary space X. Namely $\hat{G}$ can be obtained as a quotient of the space $\partial^3\Gamma$ by using functions $L(a,b,x_1,x_2)$ where now x_1 and x_2 are points in $\partial\Gamma$ different from a and b, where L satisfies (1), (2) and (3) of 8.3.D and where (4) is replaced by

$$L(a,b;x_1,x_2) \longrightarrow \infty$$

for $x_1 \longrightarrow a$ and $x_2 \longrightarrow b$.

8.3.K <u>Example</u>: If $\partial\Gamma$ consists of two points, then our theorem reduces to the following well known (and nearly obvious) fact.

Let Γ *be a finite extension of* $\mathbb{Z}$. *Then there exists a homomorphism* h *of* Γ *into the (only) non-trivial semidirect product* $\mathbb{Z} \ltimes \mathbb{Z}_2$, *such that the image* $h(\Gamma) \subset \mathbb{Z} \ltimes \mathbb{Z}_2$ *is infinite.*

8.3.L <u>Ends of non-hyperbolic groups</u>: Theorem 8.3.C and its proof resemble Stallings theorem on groups with infinitely many ends. In fact Stallings theorem admits an extremely short geometric proof (essentially due to Matthew Brin) based on the regularity of solutions of Plateau's problem (see [Gr5]). Here is another version of this proof where we only need the regularity of harmonic functions on Riemannian manifolds.

<u>Step 1</u>: To simplify the matter we assume Γ to be finitely presented (the general case can be treated along the lines indicated in [Gr5]) and then take a closed Riemannian manifold V with $\pi_1(V) = \Gamma$ such that the space of ends of Γ, denoted $\partial_0\Gamma$, is identical to that of the universal covering X of V.

<u>Step 2</u>: We assume Γ to have more than one end. Then, either Γ is a finite extension of $\mathbb{Z}$ and $\partial_0\Gamma = \partial\Gamma = \{\pm\infty\}$, or $\partial_0\Gamma$ is a Cantor set. Cover $\partial_0\Gamma$ by two disjoint non-empty open (and hence closed subsets) A and B and consider all smooth functions f on X which continuously extend to $X \cup \partial_0 X$ and have $f|A = 0$ and $f|B = 1$. Then we consider the Dirichlet functional $\|f\|_p^{\frac{1}{p}} = \int_X \|\mathrm{grad}\, f\|^p$ on these functions and define

$$|A\text{-}B|_p = \inf_f \|f\|_p^{\frac{1}{p}}.$$

Clearly $|A\text{-}B|_p < \infty$ for all $p > 0$. If $p = 1$, then $|A\text{-}B|_p$ equals the (n-1)-dimensional (for $n = \dim X$) volume of a minimal hypersurface $H \subset X$ separating A from B. If $\partial_0\Gamma = \{\pm\infty\}$ and $p > 1$, then $|A\text{-}B|_p = 0$. However, if $\partial_0 X = \partial_0\Gamma$ is infinite, then $|A\text{-}B|_p > 0$. Moreover $\|\partial_0 X\|_p \overset{def}{=} \inf_{A,B} |A\text{-}B|_p > 0$. This follows from the (obvious) exponential lower bound on the number of ends of X detected in the ball of radius $R \longrightarrow \infty$ around a fixed point in X. Notice that the group Γ plays little role here: the same conclusion would hold for every X' quasi-isometric to X.

<u>Step</u> <u>3</u>: If $p \geqslant 1$ and $|A\text{-}B|_p > 0$, then there exists a function f for which $\|f\|_p^{\frac{1}{p}} = |A\text{-}B|_p$, though f may be non-smooth for $p \neq 2$. For example if $p = 1$ this extremal f equals zero on the A-side of the hypersurface H and one on the B-side. However, for $p = 2$ this function (obviously) is harmonic and hence smooth. Notice that Γ plays no role here either.

<u>Step</u> <u>4</u>: There exists a function $\bar{f}$ on $X \cup \partial_0 X$ which is continuous on $\partial_0 X = \partial_0\Gamma$ with values 0 and 1 and such that $\|\bar{f}\|_p^{\frac{1}{p}} = \|\partial_0 X\|_p$, where $p \geqslant 1$ and provided $\|\partial_0 X\|_p \neq 0$ (e.g. $\partial_0\Gamma$ is infinite and $p = 2$). This is obvious since the isometry group of X (which contains Γ) is cocompact. This is the only moment where Γ is seriously used.

<u>Step</u> <u>5</u>: Notice that for every f and γ the functions

$$\gamma_+ f = \max(f, \gamma f)$$

and

$$\gamma_- f = \min(f, \gamma f)$$

satisfy

$$\|\gamma_+ f\|_p^1 + \|\gamma_- f\|_p^1 = 2\|f\|_p^1.$$

Hence, if $\bar{f}$ is our absolutely minimizing function, then

$$\|\gamma_+ \bar{f}\|_p^1 = \|\gamma_- \bar{f}\|_p^1 = \|\bar{f}\|_p^1.$$

Now, if *every* extremal $\bar{f}$ is known to be sufficiently regular (C^1-smooth is enough), we conclude that the function $\bar{f}$ has the following property: for every $\gamma \in \Gamma$ one of the two inequalities holds

(1) $\bar{f}(x) \geqslant \bar{f}(\gamma(x))$ for all $x \in X$;

(2) $\bar{f}(x) \leqslant \bar{f}(\gamma(x))$ for all $x \in X$.

In particular this is satisfied for $p = 2$ if $\partial_0\Gamma$ is infinite. Now we recall that $\bar{f}$ is continuous on $X \cup \partial_0 X$ and that $\bar{f}\,|\,\partial_0 X$ is non-constant. The existence of such an $\bar{f}$ immediately implies that Γ is a non-trivial free or HNN-product amalgamated over a finite subgroup.

8.3.M It is not hard to show that every finitely generated group Γ admits a left invariant metric $|\ |$ such that $(\Gamma, |\ |)$ is hyperbolic $\partial\Gamma = \partial_0\Gamma$ and the action of Γ on $\partial^3\Gamma$ (and hence the canonical action of Γ on $\partial_0^3\Gamma$) is B-finite (but not necessarily cobounded). For example, if Γ_1 and Γ_2 are infinite groups connected at infinity, we start with some word metric $|\ |_i$ in Γ_i for $i = 1,2$, then take $|\ |_i' = \log(1 + |\ |_i)$ and finally take the *metric* free product $\Gamma = (\Gamma_1 * \Gamma_2, |\ |_1' * |\ |_2')$. A similar construction applies to amalgamated and HNN-products, but it would be more interesting to find a direct (not using Stallings' theorem) construction and then to derive Stallings theorem from a suitable generalization of 8.3.

Notice that the action of each $\Gamma_i \subset \Gamma$, $i = 1,2$, is *parabolic* on $\partial_0\Gamma$. Namely, Γ_i keeps fixed a unique point

$a_i \in \partial_0\Gamma$, $i = 1,2$, and $\gamma(a) \longrightarrow a_i$ for every $a \in \partial_0\Gamma$ as $a \in \Gamma_i$ goes to infinity.

The compactification of *any* Γ by $\partial_0\Gamma$ suggests a general notion of (partially) hyperbolic boundaries of a non (word) hyperbolic group Γ. In particular one may seek a *maximal* hyperbolic boundary similar to the Furstenberg boundary (which is $\partial\Gamma$ if Γ is word hyperbolic). An approach to this problem is indicated in [Fl].

8.3.N Remarks on groups of rank ≥ 1: Recall that the space $\hat{G}$ is acted by three groups, Γ, $\mathbb{R}$ and $\mathbb{Z}_2$, and the three actions agree in a certain manner. A similar situation arises for groups Γ operating on symmetric spaces of range $p \geqslant 1$, for example for subgroups in $O(p,q)$ for $1 \leqslant p \leqslant q \leqslant \infty$. In this case one takes the flow space of chambers for $\hat{G}$ and the three groups are Γ, $\mathbb{R}^p$ and W, where W is the Weil group and where $\hat{G}|\mathbb{R}^p$ can be identified with the Furstenberg boundary of Γ, provided Γ is a (finite dimensional) lattice. However, one does not know what corresponds to this picture for more general semihyperbolic groups, such as cocompact isometry groups of convex manifolds. This is not even known for cocompact groups acting on convex domains in *symmetric* spaces of $\mathbb{R}$-rank $\geqslant 2$.

8.3.P Some corollaries of 8.3: We see with 8.3.C that word hyperbolic groups Γ are indeed very similar to fundamental groups of convex manifolds. In particular, there is a one to one correspondence between periodic orbits of the $\mathbb{R}$-flow on $\hat{G}/\Gamma$ and conjugacy classes of non-torsion elements γ in Γ.

8.3.Q According to 2.2 every word hyperbolic group Γ admits a discrete cocompact action on a contractible finite dimensional polyhedron P. Then an obvious partition of unity argument delivers a Lipschitz Γ-equivariant map $\hat{G} \longrightarrow P$ which necessarily (and obviously) is a quasi-isometry. The images of $\mathbb{R}$-orbits give a distinguished (depending on the map $\hat{G} \longrightarrow P$) class of quasigeodesics in P, called $\hat{G}$-*lines*, which serve in most respects better than geodesics in P. Any such $\hat{G}$-line joins two points at the boundary $\partial P = \partial\Gamma = \partial\hat{G}$.

If ℓ_1, ℓ_2 and ℓ_3 are three $\hat{G}$-lines between a_1, a_2 and a_3 at ∂P, then the resulting (ideal) $\hat{G}$-*triangle* Δ with vertices a_i and sides ℓ_i exponentially narrows at infinity like an ideal geodesic triangle Δ_ϵ in the hyperbolic plane H_ϵ of curvature $-\epsilon^2 < 0$ for some constant $\epsilon = \epsilon(\Gamma) > 0$. Namely, there is a *distance decreasing map of the (2-dimensional) triangle* Δ_ϵ *into* P *such that the sides of* Δ_ϵ *go to the sides* ℓ_i *of* Δ. In particular, Δ bounds a disk (that is the image of Δ_ϵ) of area $\leqslant \pi\epsilon^{-2}$. This property of $\hat{G}$-triangles immediately implies a similar property of all ideal G-simplices in P, which leads to the following important corollary.

8.3.R Denote by $Q^n = Q^n(\partial\Gamma)$ the n-dimensional polyhedron whose k-simplices for $k = 0,1,\dots,n+1$ corresponds to all (k+1)-tuples of distinct points in $\partial\Gamma$ and let $Q_0^n = Q^n \setminus \partial\Gamma$, where $\partial\Gamma$ is identified with the 0-skeleton $Q^0 \subset Q^n$.

<u>**Theorem**</u>: *There exists a metric* $|\ | = |\ |_\epsilon$ *on* Q_0^n *for some* $\epsilon = \epsilon(\Gamma,n) > 0$ *and a continuous* Γ*-equivarient map* $T: Q^n \longrightarrow P \cup \partial P$ *with the following three properties.*

(1) *Each* k-*simplex* Δ^k *in* Q^n *is isometric to some ideal* k-*simplex* Δ_ϵ^k *in the hyperbolic space* H_ϵ^k *of curvature* $-\epsilon^2$.

(2) *The map* T *is the identity on* $Q^0 = \partial\Gamma = \partial P \subset Q^n$ *and* Q_0^n *is sent into* P.

(3) *The map* T *is distance decreasing on* Q_0^n *(for the metric* H_ϵ *and a given simplical metric* on P). *Furthermore* T *is uniformly quasi-isometric on every simplex* Δ^k *in* Q^n. *That is* T *is* ϵ- + isometric $\times\ \lambda$ *for some constants* ϵ' *and* λ *depending on* Γ *and* n.

The property (1) needs a comment. An ideal k-simplex Δ in

H^k_ϵ by definition is the convex hull of k+1 distinct points (vertices) $a_0,\ldots,a_k$ in ∂H^k_ϵ. That is Δ is the minimal convex subset in H^k_ϵ with $\partial\Delta = \{a_0,\ldots,a_k\} \subset \partial H^k_\epsilon$. For example, every edge $\Delta^1 \subset H^1 = \mathbb{R}$ is isometric to $\mathbb{R}$. Also all ideal triangles in H^2 are isometric. However, simplices Δ^k for $k \geq 3$ are not mutually isometric. In fact, one can assign in a canonical way a *geodesic tree* $S \subset \Delta^k \subset H^k$ with the following properties.

(a) The tree S has $\ell \leq k - 1$ vertices $x_1,\ldots,x_\ell$ inside Δ^k.

(b) S is the union of some geodesic segments between x_j and of k+1 geodesic rays joining each $a_i \in \partial H^k$ with some x_j.

(c) Denote by d(s) the distance from $s \in S$ to the nearest vertex x_j and let $B_s \subset H^k$ be the ball around s of radius 10 exp$-\epsilon d(s)$. Then the union $\bigcup_{s\in S} B_s \subset H^k$ contains the simplex Δ^k. In other words, Δ^k exponentially narrows away from the vertices x_j.

(d) The volume of Δ^k is bounded by $\pi\epsilon^{-\frac{k}{2}}$ exp$-\epsilon d$ for $d = 1/2\ \mathrm{diam}\{x_1,\ldots,x_\ell\}$. In fact, the simplex of the maximal volume is the regular one (see [Th]) for which S is the union of k+1 rays joining the points a_i with a single point $x \in H^k$.

(e) The geometry of Δ^k is essentially determined by the (simplical) geometry of the tree S with the distance defined by lengths of paths in S between pairs of points. Namely the embedding $S \longrightarrow \Delta^k$ is a + isometry where the implied constant only depends on k and ϵ. Furthermore, the geometry of Δ within given ball B of radius R "essentially" depends on $B \cap S$ with an error $\sim$ exp$-\epsilon R$. More generally let B be a bounded convex subset in H^k, let $S_B \subset \Delta^k \cap B$ be the convex hull of the intersection $S \cap \partial B$ and let d(y) for $y \in \Delta^k \cap B$ denote the distance from y to the (topological) boundary ∂B of B. Then

$$\text{dist}(y, S_B) \leqslant 10 \exp - \epsilon d(y)$$

for all $y \in \Delta^k \cap B$ where $d(y) \geqslant 10\epsilon^{-1}$.

Example: If $k = 3$ then the graph (generically, for $\ell = 2$) looks like the one sketched in Figure 9 below.

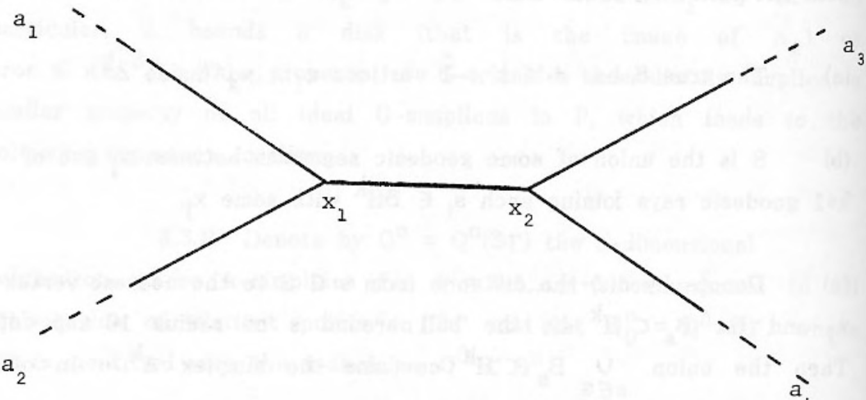

Figure 9

The basic invariant of Δ^3 is the distance $|x_1 - x_2|$. If $|x_1 - x_2|$ keeps bounded then the geometry of Δ^3 is close to that of the regular ideal simplex. If $|x_1 - x_2| \longrightarrow \infty$ then Δ^3 (exponentially fast) converges to the union of two triangles spanned by the triples $(a_1, a_2, 1/2(x_1+x_2))$ and $(a_3, a_4, 1/2(x_1+x_2))$ correspondingly.

These properties of hyperbolic simplices play a crucial role in Thurston's approach to discrete groups isometrically acting on H^3 and Theorem 8.3 allows us to extend his methods to all word hyperbolic groups. Let us see how it works.

8.3.S **Application of geodesic simplices**: Let us start with 2-simplices in Γ and to simplify the matter we assume Γ has no torsion. Thus Γ is the fundamental group of a compact aspherical manifold V with boundary and the orbits of the geodesic flow in $\hat{G}(\Gamma)$, mapped to V, are called here *geodesics in* V. We also use some auxiliary Riemannian metric in V to measure length and area but "geodesics" refer to those coming from $\hat{G}$.

Theorem: *Let* F_0 *be a compact connected surface mapped to* V. *Then there is a homotopy of* F_0 *to a surface* F_1, *such that this homotopy does not move the boundary* $\partial F_0 \longrightarrow V$ *and*

$$\text{Area } F_1 \leq C(\text{length}(\partial F_0 = \partial F_1) + |\chi(F_0)|),$$

where χ *denotes the Euler characteristic and where the constant* $C \geq 0$ *only depend on* V.

Proof: The surface F_1 is built of two pieces. First is a *collar* of $\partial F_0 = \partial F_1$ which is obtained by homotopying each component of ∂F_0 (which may be assumed non-contractible) to the unique closed geodesic in the respective homotopy class. This collar can be chosen, such that

$$\text{Area Collar} \leq C \text{ length } \partial F_0,$$

according to 7.4.

Now we have a surface with geodesic boundary ∂' and we homotope it without moving the boundary to a surface with area $\leq C|\chi(F_0)|$. To do this we subdivide F_0 into ideal triangles (compare [Th]) Δ_i for $i = 1,\dots,k = -2\chi(F_0)$ whose sides, in the universal covering of F_0, are asymptotic to geodesics covering boundary components. Then we take corresponding geodesics in V and thus fill in ∂' by k our triangles. Q.E.D.

Remark: The above construction gives a bound on the total geometry of F_1, not only on area F_1. Namely, there exists a constant $\epsilon = \epsilon(V) > 0$, and a metric g of constant curvature $-\epsilon$ on F_1 for which the map into V is distance decreasing and each component S of $\partial F_0 \hookrightarrow V$ has

$$\text{length}_g S \leq C \text{ length}_V S.$$

Moreover, if some S is a closed geodesic in V then S is geodesic in (F_1,g). Thus we have a perfect replacement of minimal surfaces in manifolds with $K < 0$.

8.3.T Another standard application of the boundness of volumes of our simplices is the boundness of the cohomology $H^{k\geq 2}(\Gamma;\mathbb{R})$ (compare [Gr2]). Since we only have *ideal* simplices the discussion in [Gr2] needs a minor readjustment. (A judicious reader may notice a little problem for k = 2 which disappears on second thoughts.)

<u>**Remark**</u>: The above applications of the geodesic flow can be generalized to most (non-word) hyperbolic groups, by extracting and generalizing the essential information recorded in the flow. Unfortunately, precise statements and proofs become quite cumbersome. (Of course, there is no problem for groups operating on convex spaces with $K \leq -\epsilon < 0$.)

8.4 <u>**Symbolic dynamics for word hyperbolic groups**</u>: Consider a graph X with the vertex set $X^0 \subset X$ and let Γ be a semigroup of simplical embedding $X \longrightarrow X$. Take a finite set S and let Σ be the space of maps $\sigma\colon X^0 \longrightarrow S$ with the following topology: a sequence σ_i converges to σ for $i \longrightarrow \infty$ iff $\sigma_i(x) = \sigma_{i_0}(x)$ for every given $x \in X^0$ and all $i \geq i_0$ for some $i_0 = i_0(x)$. The space Σ with the induced action by Γ is called the *Bernoulli shift* defined by X, Γ and S.

<u>**Examples**</u>: Let Γ be a group with a fixed generating set $G \subset \Gamma$. Then Γ acts on the corresponding Cayley graph and we have the associated Bernoulli shift on $\Sigma = \Sigma(\Gamma,G,S)$. If Γ is the free cyclic group generated by a single element then our definition agrees with the classical notion of *the two sided Bernoulli shift*. Recall that *one sided shifts* correspond to the action of $\mathbb{Z}_+$ on the obvious graph $\approx \mathbb{R}_+ \supset \mathbb{Z}_+$.

8.4.A **Markov shifts:** A Γ-invariant subset $\Phi \subset \Sigma$ is called *Markov* if there exists a number $R < \infty$ and a subset $\Sigma(R,x)$ in the set of map $B(x,R) \longrightarrow S$, for all $x \in X^0$, where B denotes the R-ball around x, such that a function $\sigma \in \Sigma$ is contained in Φ if and only if $\sigma \mid B(x,R) \in \Sigma(R,x)$ for all $x \in X^0$.

In most cases of interest the graph X is locally finite and the action of Γ is cofinite; that is there are only finitely many Γ-orbits of points $x \in X^0$. Such shifts are called of *finite type* as they are described by finite combinatorial data given by (finitely many of finite) subsets $\Sigma(R,\bar{x})$ corresponding to Γ-orbits $\bar{x}$ in X^0.

Example: The diagonal action of Γ on $\Sigma \times \Sigma$ is Markov for the obvious embedding

$$\Sigma \times \Sigma \subset \Sigma(X,\Gamma,S \times S).$$

Also the diagonal $\Delta \subset \Sigma \times \Sigma$ is Markov. Furthermore, if $\Phi \subset \Sigma$ is Markov or of finite type then $\Phi \times \Phi$ and the diagonal in $\Phi \times \Phi$ also are such.

8.4.B **Finitely presented systems:** Let X and Γ be as earlier and let Ω be a topological space acted upon by Γ. Call Ω *finitely presented* over X if there exists a subshift of finite type in some Bernoulli shift Σ and a surjective Γ-equivariant map $\pi: \Phi \longrightarrow \Omega$ such that the equivalence relation $R_\pi \subset \Phi \times \Phi$ defined by

$$(\sigma_1,\sigma_2) \in R_\pi \Leftrightarrow \pi(\sigma_1) = \pi(\sigma_2),$$

is Markov and the topology in Ω equals that in the quotient space Φ/R_π.

3.4.C **Theorem:** *Let X be a hyperbolic locally finite graph and let Γ be a cobounded automorphism (isometry) group of X. Then the induced action of Γ on ∂X is finitely presented. Moreover there is a*

Markov presentation $\pi: \Phi \longrightarrow \partial X$ *which is a finite-to-one map. In particular the action of every word hyperbolic group* Γ *on* $\partial\Gamma$ *admits such a presentation over the Cayley graph of* Γ.

<u>Proof</u>: The "coding" π_0 defined in 7.6 clearly does the job.

8.5 <u>**Markov trees**</u>: Let T be a locally finite simplical tree with our usual metric where the edges have unit length. Fix a reference vertex $x_0 \in T$ and assign to each vertex $x \in T$ the subtree $T_x \subset T$ consisting of those $y \in T$ for which the segment $[x_0,y]$ contains x. Call T *Markov* or of *finite type* if there are at most finitely many isomorphisms classes of pointed trees (T_x,x). Label these classes by indices 0,1,...,i,...,k (where 0 corresponds to the class of (T,x_0)) and define a *Markov chain* by the matrix $M = M(i,j)$ for $0 \leqslant i,\ j \leqslant k$ where M(i,j) is the number of j-vertices following an i-vertex. Here x is called an i-vertex if T_x is labeled by i and y follows x if $|x-y| = 1$ and $|x_0-y| = |x_0-x| + 1$. Now one can easily count the number N(r) of vertices in the sphere of radius r around x_0 for r = 1,2,... as well as the number $N_i(r)$ of i-vertices in this sphere. Namely, $N_i(r) = M^r(0,i)$ and $N(r) = \sum_i N_i(r)$ where M^r is the r-th (matrix) power of M. In particular one sees that the counting (generating) functions $\sum_r N_i(r)t^r$ and $\sum_r N(r)t^r$ are rational.

8.5.A <u>**Markov** Γ**-trees**</u>: Suppose, for some pairs of vertices x and y in T, we are given a set $\Gamma_{x,y}$ of isomorphisms $T_x \longrightarrow T_y$, such that

(1) $\quad Id \in \Gamma_{x,x}$ for all vertices x,

(2) $\quad \Gamma_{yx} = \Gamma_{xy}$,

(3) $\quad \Gamma_{xz} \supset \Gamma_{xy} \circ \Gamma_{yz}$.

Then we call such a Γ-tree *Markov* if there is at most finitely many of Γ-isomorphism classes of trees T_x. Again, we label these Γ-classes by $0,1,\ldots,i,\ldots,k$, assign a Markov chain $M(i,j)$ to (T,Γ) and conclude to the *rationality of the corresponding counting function* $\sum_r N_i(r)t^r$.

8.5.B Let X be a locally finite graph with a reference point x_0 and the usual simplical metric and let T be the set of all minimal segments of integer length issuing from x_0. This T has a natural structure of a tree, such that the sphere of radius r in T around x_0 consists of segments of length r (joining x_0 with the points of the r-sphere in X).

8.5.B_1 **Theorem:** *If X is hyperbolic and the isometry group Γ of X is cofinite on X, then the tree T of minimal segments in X issuing from x_0 is Markov. Furthermore if Γ_0 is an isometry group of X having an invariant quasiconvex subset $X_0 \subset X$ on which the action of Γ_0 is cofinite, then the corresponding tree T for the quotient graph X/Γ_0 also is Markov.*

Proof: Take two vertices x and y in X and suppose there exists an isometry $\gamma: X \longrightarrow X$, such that $\gamma(x) = y$ and such that the function $|x| - |\gamma(x)|$ is constant on the ball of radius d around x. If $d \geqslant 100\delta$ (for the implical hyperbolicity constant δ of X) then for any two segments $[x_0,x]$ and $[x_0,y]$ the corresponding subtree $T_x \subset T$ is sent onto T_y by γ and the proof follows from the (obvious) finiteness of the number of pertinent (distance) functions/const on d-balls in X.

8.5.C Let X be the Cayley graph of a group Γ with fixed generators $\gamma_1,\gamma_2=\gamma_1^{-1},\gamma_3,\gamma_4=\gamma_3^{-1},\ldots,\gamma_{2n}=\gamma_{2n-1}^{-1}$. Then every oriented edge e of X is labeled by some generator γ_i and the change of the orientation switches γ_i to γ_i^{-1}. Now, the ordering of the generators gives an order to the edges issuing from a fixed vertex $x \in X$ which induces a lexicographic ordering on the set of segments issuing from x. In particular, there is an order on the set of segments from x to a fixed point x_0. Let $\overleftarrow{[x_0, x]}$ be the first segment for this order. The set of first segments obviously form a tree, say $\overleftarrow{T}$, which is embedded into X by $\overleftarrow{[x_0, x]} \longmapsto x$. Clearly this embedding is bijective on the set of vertices and preserves the distance to the reference point.

8.5.D *If Γ is word hyperbolic then the tree $\overleftarrow{T}$ is Markov (compare [Can]). In particular, the counting function for the number of points in the sphere of radius r in Γ is rational* (compare 5.2). *Furthermore, if Γ_0 is a quasiconvex subgroup in Γ then the counting function for the number of Γ_0-orbits in Γ (here we consider the left action of Γ_0 which is isometric for the word metric in Γ) within distance r from the identity element in Γ also is rational.* (Notice that every cyclic subgroup in Γ is quasiconvex.)

The proof is identical to that of $8.5.B_1$.

8.5.E Observe that the full isometry group Γ of a locally finite graph X is locally compact and totally disconnected but not necessarily discrete. If X is hyperbolic, the action of Γ is cofinite on X and if Γ admits a discrete cocompact (and hence word hyperbolic) subgroup Γ' then the above theorem implies the rationality for the counting function of the number of concentric spheres in X. In fact, one can do without Γ' by using the tree T(X) constructed in 7.6 instead of $\overleftarrow{T}$. Indeed, this tree is Γ-Markov by the previous argument and the set of vertices in T(X) corresponding to one-point subsets in X clearly is Γ-invariant. Hence, the previous argument applies and leads to the following

8.5.F <u>**Theorem**</u>: *Let X be a locally finite hyperbolic graph whose full isometry (i.e. automorphism) group is cofinite on X (i.e. has at most finitely many orbits on the vertex set of X) and let*

X_0 *be a quasiconvex subgraph (i.e. there exists a constant* $\epsilon > 0$ *such the distance minimizing segment in* X *between any two points in* X_0 *is* ϵ*-close to* X_0*), such that the isometry group* Γ_0 *of the pair* (X,X_0) *is cofinite on* X_0. *Then the counting function for the number of points on concentric spheres* $S(r) \subset X/\Gamma$, $r = 1,2,\dots$, *around a fixed point in the quotient graph* X/Γ *is rational. In particular, the counting function for the spheres in* X *is rational. Moreover the counting function for the intersections of concentric* r*-spheres in* X *with* X_0 *also is rational.*

8.5.G <u>**Remark**</u>: If X is an arbitrary geodesic hyperbolic space, $X_0 \subset X$ is a quasiconvex subspace and Γ is an isometry group of the pair (X,X_0) whose action is cobounded on X_0, then, obviously, the quotient space X/Γ is hyperbolic.

8.5.H <u>**Markov spaces**</u>: Consider a Markov action of $\mathbb{Z}_+$ on some (necessarily compact) space Φ of functions $\varphi(i)$, $i \in \mathbb{Z}_+$, and call a (non-invariant!) subset $\Psi \subset \Phi$ *semi-Markov* if it is the intersection of a (invariant) Markov subset (subshift) and a *basic* subset $\Psi_0 \subset \Phi$. Recall that, by definition, every basic subset is determined by finitely many (finite) subsets $S_i \subset S$, $i = 1,\dots,k$, by

$$\Psi_0 = \{\varphi(i) \in \Phi \mid \varphi(i) \in S_i \text{ for } i \leqslant k\}.$$

A compact topological space Ω is called *semi-Markov* if it admits a presentation $\Omega = \Psi/R$ for some semi-Markov equivalence relation $R \subset \Psi \times \Psi \subset \Phi \times \Phi$, where Ψ is semi-Markov in (the space of) a Markov shift Φ.

Semi-Markov spaces can be equivalently described in terms of Γ-trees T with an additional *neighborhood* structure, that is a subset $\mathcal{N}$ in the set of pairs of vertices x and y in T, such that

(i) $(x,y) \in \mathcal{N} \Rightarrow |x| = |y|$,

(ii) the (induced) subset $\bar{\mathcal{N}} \subset \partial T \times \partial T$ consisting of pairs of rays $r_1, r_2: \mathbb{R}_+ \longrightarrow T$ with $r_1(0) = r_2(0) = x_0$ and having $(r_1(i), r_2(i)) \subset \mathcal{N}$ for $i = 1,2,\ldots$, is an *equivalence* relation on ∂T.

8.5.I **Lemma**: *Let $\mathcal{N}$ be Γ-invariant and let there are at most finitely many Γ-classes of subtrees $T_x \subset T$ for all x. If the following condition (A) is satisfied then the space is semi-Markov.*

(A) *If two different vertices y_1 and y_2 follow the same vertex x (i.e. y_1 and y_2 lie in T_x and have $|y_1| = |y_2| = |x| + 1$) then the trees T_{y_1} and T_{y_2} are not Γ-equivalent.*

Proof: Let S be the set of Γ-equivalence classes of subtrees T_x where we can assume the class of T_{x_0} to consist only of T_{x_0} and let Φ consist of the maps $\varphi: \mathbb{Z}_+ \longrightarrow S$, such that for every $i = 0,1,\ldots$ every vertex x in the Γ-class $\varphi(i)$ is followed by a (necessarily unique by (A)) vertex y in the Γ-class $\varphi(i+1)$. Clearly, ∂T *embeds* onto a semi-Markov subset in Φ and $\bar{\mathcal{N}} \subset \partial S \times \partial S$ becomes semi-Markov in $\Phi \times \Phi$. Q.E.D.

8.5.J Let us return to our hyperbolic graph X with a cofinite isometry group Γ. For any of the three trees T, $\overleftarrow{T}$ and T(X) associated to X one defines a neighborhood relation N as follows. To be specific, we do it here for the tree T of all segments $[x_0, x]$ from x_0 by

$$(x_1, x_2) \in N \Leftrightarrow |x_1| = |x_2| \text{ and } |x_1 - x_2| \leqslant 10\delta.$$

Recall that the pertinent Γ-structure on T (as well as on $\overleftarrow{T}$ and T(X)) is given by isometries in T moving $x \longmapsto y$ and preserving the function $|x| \pm$ const on a ball of radius d around x. Now, if d is sufficiently large then (A) clearly is satisfied in case Γ is discrete (hence, word hyperbolic) and X equals a Cayley graph of Γ. Therefore we have

8.5.K **Proposition**: *The boundary* $\partial\Gamma$ *of every word hyperbolic group is semi-Markov. Moreover the semi-Markov presentation* $\partial\overleftarrow{T} \longrightarrow \partial\Gamma$ *and* $\partial T(X) \longrightarrow \partial\Gamma$ *are finite to one.*

8.5.L **Remark**: It is unclear whether ∂X is semi-Markov if the full isometry group Γ of X is non-discrete. (Of course this follows from 8.5 if Γ admits a cocompact lattice.) Notice that there are only countably many (homeomorphism classes of) semi-Markov spaces and so most compact spaces are not semi-Markov. (But, obviously, all finite polyhedra are semi-Markov.) In fact, one does not know if there are only countably many homeomorphism classes of ∂X for the above spaces X and (or) whether there are only countably many of isomorphism classes of (compact) isotropy subgroups $\Gamma_x \subset \Gamma$ for the full isometry group Γ of X. A similar question also is open for locally finite simply connected non-hyperbolic 2-polyhedra X whose full isomorphism group Γ has finitely many orbits on the vertices of X.

8.5.M Let Γ be a word hyperbolic group with a fixed word metric. Take $\gamma \in \Gamma$ and consider the following four subsets,

(1) The centralizer $C_\gamma \subset \Gamma$. Notice that if γ has infinite order then C_γ is a finite extension of the cyclic group generated by γ.

(2) The fixed point set $F_\gamma \subset \partial\Gamma$.

(3) The level

$$D_\gamma(r) = \{x \in \Gamma \mid d_\gamma(x) \leqslant r\}$$

for $d_\gamma x = |x-\gamma(x)|$ and for some $r \geqslant |\gamma|$.

(4) The quasiconvex hull $L_\gamma = L(F_\gamma) \subset \Gamma$ (see 7.5). To be

precise, we take L_γ in the Cayley graph in Γ and then intersect it with Γ.

Observe the following obvious

Lemma: *The above sets satisfy*

(a) $C_\gamma \subset D_\gamma(r)$

(b) $D_\gamma(r)$ *is invariant under* C_γ *and the action of* C_γ *on* $D_\gamma(r)$ *is cofinite.*

(c) $\partial C_\gamma = \partial D_\gamma(r) = F_\gamma$.

(d) $D_\gamma(r)$ *is contained in some* ϵ*-neighborhood of* L_γ *where* ϵ *depends on* γ *and* r.

Corollaries: It follows from (a) and (b) that the subgroup C_γ is quasiconvex (in the Cayley graph of Γ). Therefore C_γ is finitely generated and word hyperbolic.

According to (c) and 8.1 there are the following four possibilities for $F_\gamma \subset \partial\Gamma$

(i) F_γ is empty,

(ii) F_γ consists of two points,

(iii) F_γ is a nowhere dense, uncountable closed subset in $\partial\Gamma$,

(iv) $F_\gamma = \partial\Gamma$. In this case Γ/C_γ is finite.

8.5.N Observe that the (displacement) function d_γ on Γ is invariant under the (left) action of C_γ and thus defines a function, say $\bar{d}_\gamma$, on Γ/C_γ such that each level

$$\bar{D}_\gamma(r) = \{x \in \Gamma/C_\gamma \mid \bar{d}_\gamma(x) \leqslant r\}$$

is finite.

<u>Theorem</u>: *The counting function* $\sum_r N_\gamma(r)t^r$ *for*

$$N_\gamma = \#(\bar{D}_\gamma(r+1)\setminus\bar{D}_\gamma(r)),\ r = 0,1,\ldots,$$

is rational.

<u>Proof</u>: Let us construct the pertinent Markov tree. First, for every finite subset $V \subset \Gamma$ consider the set all segments $[v,\gamma(w)]$ for v and w in V. Let ℓ be the maximum of the lengths of these segments and let $[v_i,\gamma(w_j)]$ be the segments of length ℓ. Now, move these v_i toward w_j by one. That is take the set V' of the points of the form $\ell^{-1}v_i + \ell^{-1}(\ell-1)\gamma(w_j)$. Then remove from V all points v_i (having $|v_i-\gamma(w_j)| = \ell$ for some $w_j \in V$), add V' to what is left and call the result $P_\gamma(V)$. Finally, apply the construction in 7.6 to P_γ instead of P and to subsets $V \subset \Gamma$ of diameter $\leqslant D$ for $D = 100(|\gamma| + \delta + 1)$. Since P_γ is C_γ-invariant the resulting graph, say T_γ also C_γ-invariant and we take the quotient $\bar{T}_\gamma$. This $\bar{T}_\gamma$ is not always a tree. However, if we remove the (finite!) subgraph spanned by those $V \subset \Gamma$ on which the function d_γ is $\leqslant 200(|\gamma| + \delta + 1)$ then what remains is a union of finitely many trees. One sees as earlier that these trees are Markov and the proof is concluded.

8.5.P One can reformulate Theorem 8.5.N as follows.

The counting function for the intersections of the set $\Gamma\gamma\Gamma^{-1}$ *with concentric spheres in* Γ *is rational for all* $\gamma \in \Gamma$.

Since Γ contains only finitely many conjugacy classes of torsion elements, we see with the above that *the counting function for torsion (as well as for non-torsion) elements in* Γ *is rational (compare* 5.2).

8.5.Q Let X be the Cayley graph of a word hyperbolic group Γ and $\pi: \Phi \longrightarrow \partial X = \partial\Gamma$ be the Markov presentation over Γ provided by 8.4. Recall that points in Φ are represented by certain (distance like) functions $\varphi: \Gamma \longrightarrow \mathbb{Z}$ defined up to additive constant. This presentation of $\partial\Gamma$ is closely related to the Markov chain (over $\mathbb{Z}_+$) associated to the tree $\overleftarrow{T}$. Namely, let $\gamma = \gamma(\varphi) \in \Gamma$ be the first (for a fixed ordering) generator γ of Γ, such that $\varphi(\gamma) = \varphi(\mathrm{id}) - 1$, and let us define $\alpha: \Phi \longrightarrow \Phi$ by $\gamma^{-1}\varphi$ for all $\varphi \in \Phi$ and $\gamma = \gamma(\varphi)$. Clearly, this α is identical to the (one-sided) Markov shift associated to $\overleftarrow{T}$. Then we complete α to a *two-sided Markov shift* $\bar{\alpha}$ on the space $\bar{\Phi}$ of $\mathbb{Z}$-*orbits* of α that are maps $\psi: \mathbb{Z} \longrightarrow \bar{\Phi}$ satisfying $\psi \circ s = \alpha \circ \psi$ for the translation $s: i \longmapsto i + 1$ on $\mathbb{Z}$. Let us give a description of $\bar{\Phi}$ as of a Γ-quotient of the following space $\tilde{\Phi}$. Each $\tilde{\varphi} \in \tilde{\Phi}$ is given by a (distance-like) function $\varphi: X \longrightarrow \mathbb{R}$ and a (gradient) line $\ell: \mathbb{R} \longrightarrow X$, such that

(a) $\ell(\mathbb{Z})$ lies in the zero skeleton of X

(b) for every $i \in \mathbb{Z}$ the vertex $\ell(i+1)$ is the first vertex adjacent to $\ell(i)$ such that $\varphi(\ell(i+1)) = \varphi(\ell(i)) - 1$, where "the first" refers to the order on edges corresponding to that on the generators of Γ.

Recall that function φ is quasiconvex and distance-like:

$$\varphi(x) = \varphi(y) + \mathrm{dist}(x, \varphi^{-1}(-\infty,y])$$

for all x and y in X satisfying $\varphi(x) \geqslant \varphi(y)$. Hence, the values of φ on 20δ-neighborhood of $\ell(\mathbb{R}) \subset X$ determine φ on all of X. Now, Γ obviously acts on $\tilde{\Phi}$ and also $\mathbb{Z}$ acts on $\tilde{\Phi}$ by $\bar{\alpha}: (\varphi,\ell) \longmapsto (\varphi,\ell \circ s)$ for the translation s on Σ. Clearly

$$(\tilde{\Phi},\tilde{\alpha})/\Gamma = (\bar{\Phi},\bar{\alpha}).$$

Let us relate $\tilde{\Phi}$ to the space $\hat{G} = \hat{G}(\Gamma)$ of geodesics in X. Denote by $\bar{\Phi}_{\mathbb{R}}$ the *mapping torus* of $\bar{\alpha}$. That is

$$\bar{\Phi}_{\mathbb{R}} = (\bar{\Phi} \times \mathbb{R})/\mathbb{Z},$$

where $\mathbb{Z}$ acts by $\bar{\alpha}$ on $\bar{\Phi}$ and by the unit shift on $\mathbb{R}$. There is an (obviously defined) action of $\mathbb{R}$ on $\bar{\Phi}_{\mathbb{R}}$, called $\bar{\alpha}_t: \bar{\Phi}_{\mathbb{R}} \hookleftarrow$ for $t \in \mathbb{R}$, such that the map $\bar{\alpha}_1$ sends $\bar{\Phi} = \bar{\Phi} \times 0 \subset \bar{\Phi}_{\mathbb{R}}$ onto itself and $\bar{\alpha}_1 \mid \bar{\Phi} = \bar{\alpha}$. One sees as earlier that $\bar{\Phi}_{\mathbb{R}} = \tilde{\Phi}_{\mathbb{R}}/\Gamma$ where $\tilde{\Phi}_{\mathbb{R}}$ consists of pairs (φ,ℓ), where the above condition (a) on ℓ is not required any more and (b) applies to the pairs $(\varphi(t),\varphi(t+1))$, such that $\varphi(t)$ lies in the zero skeleton of X.

Now the projection $(\varphi,\ell) \longmapsto \ell$ maps $\tilde{\Phi}_{\mathbb{R}}$ into the space G(X) of all lines in X. Then we recall the (non-unique) map of G(X) onto $\hat{G} = \hat{G}(\Gamma)$ and we denote the composed map by $\tilde{\pi}_{\mathbb{R}}: \tilde{\Phi}_{\mathbb{R}} \longrightarrow \hat{G}$. Clearly, $\tilde{\pi}_{\mathbb{R}}$ is surjective and the (identification) relation, called

$$\tilde{R}_{\mathbb{R}} = \{(\tilde{\varphi}_1,\tilde{\varphi}_2) \mid \pi_{\mathbb{R}}(\tilde{\varphi}_1) = \pi_{\mathbb{R}}(\tilde{\varphi}_2)\} \subset \tilde{\Phi}_{\mathbb{R}} \times \tilde{\Phi}_{\mathbb{R}}$$

satisfies for all $\tilde{\varphi}_1 = (\varphi_1,\ell_1)$ and $\tilde{\varphi}_2 = (\varphi_2,\ell_2)$ in $\tilde{\Phi}_{\mathbb{R}}$:

$$(\tilde{\varphi}_1,\tilde{\varphi}_2) \in \mathbb{R} \Rightarrow |\ell_1(t)-\ell_2(t)| \leqslant d,$$

for all $t \in \mathbb{R}$, where $d = d(\Gamma)$ is some constant. Also notice that the map $\tilde{\pi}_{\mathbb{R}}$ is *injective* on every line in $\tilde{\Phi}_{\mathbb{R}}$ and every $\mathbb{R}$-orbit $\tilde{\ell}$ of $\tilde{\Phi}_{\mathbb{R}}$ is homeomorphically mapped onto the unique $\mathbb{R}$-orbit in $\hat{G}$ having some limit points in $\partial\Gamma$ as $\tilde{\ell}$. Thus the relation $\tilde{R}_{\mathbb{R}}$ is *strictly* contained in the (neighborhood) relation $|\ell_1(t)-\ell_2(t)| \leqslant d$, called $\tilde{R}(d) \subset \tilde{\Phi}_{\mathbb{R}} \times \tilde{\Phi}_{\mathbb{R}}$. In fact, one could define (non-uniquely) $\tilde{R}_{\mathbb{R}}$ (and thus $\hat{G}$) as a closed Γ-invarian subset $\tilde{R}$ in $\tilde{R}(d)$ for some (sufficiently large) d, which is an *equivalence* relation, such that

(a) No two distinct points on same $\mathbb{R}$-orbit in $\tilde{\Phi}_{\mathbb{R}}$ are $\tilde{R}$-equivalent.

(b) If two $\mathbb{R}$-orbits have same ends in $\partial\Gamma$, then $\tilde{R}$ establishes a homeomorphism between these orbits.

(The existence of such an $\widetilde{R}$ is proven in 8.3.)

Also notice that since the presentation $\pi: \Phi \longrightarrow \partial\Gamma$ is finite-to-one, say at most N-to-one, the (quotient) map $\widetilde{\pi}_{\mathbb{R}}: \widetilde{\Phi}_{\mathbb{R}} \longrightarrow \widehat{G}$ is at most M-to-one for some M satisfying $M \leqslant Nk^d$, where k is the number of generators in Γ.

Now we observe that the intersection $\widetilde{R}(d) \cap \widetilde{\Phi} \times \widetilde{\Phi}$ is univariant under the $\mathbb{Z}$- and Γ-actions and thus it projects to a certain subset (relation) in $\bar{\Phi} \times \bar{\Phi}$, called $\bar{R}(d) \subset \bar{\Phi} \times \bar{\Phi}$ which is invariant under the (diagonal) action of $\mathbb{Z}$. By looking through the definitions, one easily sees the following

8.5.R <u>Properties</u> <u>of</u> <u>$\bar{R}(d)$</u>: *The subset* $\bar{R}(d)$ *is Markov; it is symmetric under the involution* $(\bar{\varphi}_1,\bar{\varphi}_2) \leftrightarrow (\bar{\varphi}_2,\bar{\varphi}_1)$ *but (in general) it is not an equivalence relation. However, there exists an integer* $i_0 > 0$, *such that if* $(\bar{\varphi}_1,\bar{\varphi}_2) \in \bar{R}(d)$ *and* $(\bar{\varphi}_2,\bar{\varphi}_3) \in \bar{R}(d)$ *then* $(\bar{\varphi}_1,\bar{\alpha}^i\bar{\varphi}_3) \in \bar{R}(d)$ *for some* $i \in \mathbb{Z}$ *satisfying* $|i| \leqslant i_0$.

Finally, the projection of $\bar{R}(d)$ *to* Φ *is at most* M_d*-to-one for some constant* $M_d < \infty$.

<u>Remark</u>: The *weak transitivity* of $\bar{R}(d)$ stated above shows that $\bar{R}(d)$ induces an *equivalence* relation on the set of $\mathbb{Z}$-orbits in $\bar{\Phi}$. In fact this equivalence relation identifies two orbits if and only if they go to same orbit under the map $\bar{\Phi} \longrightarrow \bar{G} \overset{def}{=} \widehat{G}/\Gamma$ obtained by comparing the inclusion $\bar{\Phi} \subset \bar{\Phi}_{\mathbb{R}}$ and the (quotient) map $\bar{\Phi}_{\mathbb{R}} \longrightarrow \bar{G}$. Notice that the map $\bar{\Phi} \times \bar{G}$ induces a surjective map on the sets of orbits (it sends $\mathbb{Z}$-orbits of $\bar{\Phi}$ to $\mathbb{R}$-orbits of $\bar{G}$), such that the set of periodic orbits in $\bar{\Phi}$ bijectively goes onto the set of periodic orbits in $\bar{G}$.

8.5.S Let us explain the algebraic meaning of periodic orbits in $\bar{G}$. Call γ_1 and γ_2 in Γ *stably conjugate* if γ_1^m is conjugate to γ_2^m for some integer m. Denote the stable conjugacy class of γ by $[\gamma]_{st}$, (observe that it contains the conjugacy class $[\gamma]$ of γ) and define the *stable exponent*

$$p[\gamma]_{st} = \liminf_{i\to\infty} i^{-1} p(\gamma^i)$$

where $p(\gamma)$ denote the exponent of γ that is the maximal integer p such that $\gamma = \gamma_0^p$ for some $\gamma_0 \in \Gamma$.

Now we recall the maximal elementary (i.e. isomorphic to a finite extension of $\mathbb{Z}$) subgroup $E_\gamma \ni \gamma$ in Γ attached to every non-torsion $\gamma \in \Gamma$. Every stable conjugation between γ_1 and γ_2 also conjugate E_{γ_1} onto E_{γ_2}. It follows that the stable exponent is an integer. In fact, non-torsion elements γ_1 and γ_2 are stably conjugate if $p[\gamma_1]_{st} = p[\gamma_2]_{st}$ and if some $\gamma \in \Gamma$ acting on $\partial\Gamma$ sends $(\gamma_1^{-\infty}, \gamma_1^{\infty}) \longmapsto (\gamma_2^{-\infty}, \gamma_2^{\infty})$.

Now we see that there is a one-to-one correspondence between stable conjugacy classes of non-torsion $\gamma \in \Gamma$ and periodic orbits in $\bar{G}$, such that the stable exponents of γ equals the multiplicity of the corresponding orbits.

8.5.T Let us turn to periodic points (and orbits) of $(\bar{\Phi}, \bar{\alpha}) = (\tilde{\Phi}, \tilde{\alpha})/\Gamma$. Take a fixed point of some power of $\bar{\alpha}$, say $\bar{\varphi} \in \mathrm{Fix}\ \bar{\alpha}^k \subset \Phi$, and let $\tilde{\varphi} \in \tilde{\Phi}$ over $\bar{\varphi}$ be represented by (φ, ℓ), where ℓ is a line in X and φ is a (distance-like) function in the 20δ-neighborhood of $\ell \subset X$. Since $\bar{\varphi}$ is fixed under π^k there exists a non-torsion $\gamma \in \Gamma$ with $\gamma^{\pm\infty} = \ell(\pm\infty)$, such that the line ℓ is invariant under γ and $\varphi \circ \gamma = \varphi - k$. It is clear that the stable conjugacy class of such a γ only depends on k and the (finite) $\tilde{\alpha}$-orbit of $\bar{\varphi}$ and that $|[\gamma]|_- = |[\gamma]| = k$, where $|[\gamma]|_- = \liminf_{i\to\infty} i^{-1} |\gamma^i|$ and $|[\gamma]| = \inf_{\gamma' \in [\gamma]} |\gamma|$ (compare 5.2). Furthermore, since the correspondence between the orbits in $\tilde{\Phi}$ and $\hat{G}$ is finite-to-one, there exists an integer m_0 depending only on X, such that every non-torsion element in Γ whose stable exponent is divisible by m_0 is represented by some k and $\bar{\varphi}$. This in particular, implies that the number $m_0 |[\gamma]|_-$ is an *integer* for all $\gamma \in \Gamma$.

The above relation between periodic orbits and classes $[\gamma]_{st}$ can be reformulated as follows. Recall the embedding

$\bar{\Phi} = \Phi \times 0 \subset \bar{\Phi}_{\mathbb{R}}$, let

$$\bar{\Phi}' = (\bar{\Phi}\times 0) \cup (\bar{\Phi}\times(\frac{1}{m_0})) \cup (\Phi\times(\frac{2}{m_0})) \cup \ldots \cup (\bar{\Phi}\times 0 = \Phi\times 1)$$

and consider the transformation $\bar{\alpha}' = \bar{\alpha}_t$ for $t = 1/m_0$ on $\bar{\Phi}'$. Notice that $(\bar{\alpha}')^{m_0} \mid \Phi \times 0 = \bar{\alpha}$ and that $\bar{\Phi}' = \tilde{\Phi}'/\Gamma$ where $\tilde{\Phi}'$ consists of the pairs (φ,ℓ) where $\ell(0) \in X$ is of the form,

$$\ell(0) = \frac{1}{m_0} x_0 + \frac{m_0 - i}{m_0} x_1,$$

or some *integer* i and two adjacent *vertices* in X.

Now one can obviously find (by no means uniquely) an $\bar{\alpha}'$-invariant equivalence relation $\bar{R}' \subset \bar{\Phi} \times \bar{\Phi}'$, such that the periodic orbits of the corresponding quotient of $(\tilde{\Phi}',\bar{\alpha}')$, say of $(\hat{\Phi}',\hat{\alpha}')$ are in a one-to-one correspondence with stable conjugacy classes in Γ. (Notice that we make no assumptions on the topology of $\bar{R}'$ and view $\hat{\Phi}'$ as an abstract set.) Namely, the pairs $(k,\hat{\psi})$, where $k = 1,2,\ldots,$ and $\hat{\psi} \subset \hat{\Phi}'$ is a k-periodic orbit in $\bar{\Phi}'$, correspond to the classes $[\gamma]_{st}$, such that $|[\gamma]|_{-} = k/m_0$.

Next we consider the relation $\bar{R}'(d) \subset \bar{\Phi}' \times \bar{\Phi}' \subset \bar{\Phi}_{\mathbb{R}} \times \bar{\Phi}_{\mathbb{R}}$ which is the Γ-quotient of the $|\ell_1 - \ell_2| \leqslant d$ relation on $\tilde{\Phi}'$. Clearly, this $\bar{R}'(d)$ has the same properties as $\bar{R}(d)$ discussed earlier. In particular, $\bar{R}'(d)$ is Markov, the projections of $\bar{R}'(d)$ to $\bar{\Phi}'$ is finite-to-one and $\bar{R}'(d)$ (though not an equivalence relation) satisfies a weak transitivity property. It follows that one can find the above relation $\tilde{R}'$ such that $\bar{R}'(d) \supset \bar{R}'$.

Let us summarize pertinent properties of $\bar{R}'$ in the following

<u>Definition</u>: Let Ψ be a Markov shift over $\mathbb{Z}$ and $R \subset \Psi \times \Psi$ be a $\mathbb{Z}$-invariant equivalence relation. Call R *sub-Markov* if there exists a subshift $R_0 \subset \Psi \times \Psi$ containing R and an integer i_0, such that the set of pairs $(\psi_1, i\psi_2)$ for all $(\psi_1,\psi_2) \in R$ and $i = 1,2,\ldots,i_0$, contains R_0, where $i\psi$ denotes the action of $i \in \mathbb{Z}$ in $\psi \in \Psi$.

Now we can say that the relation $\bar{R}'$, for which the periodic orbits in $\bar{\Phi}'/\bar{R}'$ one-to-one correspond to stable conjugacy classes, *can be chosen sub-Markov.*

8.5.U **Theorem**: *Let* R *be a sub-Markov equivalence relation on* Ψ *and let* $\sqcap(k)$ *denote the number of k-periodic (i.e. fixed for* $k\mathbb{Z} \subset \mathbb{Z}$*) points in* Ψ/R. *If the (quotient) map* $\pi: \Psi \longrightarrow \Psi/R$ *is at most N-to-one for some* $N < \infty$, *then the counting function* $Z(t) = \sum_k \sqcap(k)t^k$ *is rational.*

Remarks:

(a) This theorem for R *Markov* is due to Manning (see [Ma]) who, in fact proves that the corresponding ζ-function $\exp \sum_k k^{-1}\sqcap(k)t^k$ is rational. Our argument below is similar to that of Manning and implies the rationality of the ζ-function.

(b) It is unclear if the "finite-to-one" assumption is essential.*

(c) If Γ has no torsion, then the counting function for the action on $\bar{\Phi}'/\bar{R}'$ is identical to that in 5.2.D. Thus Theorem 5.2.D follows from 8.5 and the preceding discussion.

8.5.V **The proof of 8.5.U**: Start with the following simple

Lemma: *Let* (Φ,α) *be a Markov shift over* $\mathbb{Z}$ *and let* $\beta: \Phi \hookleftarrow$ *be a continuous map of finite order commuting with* α. *Then the counting function* $Z(t) = Z(t;\alpha,\beta)$ *defined by*

$$Z(t) = \sum_k (\#\{\varphi \in \Phi \mid \alpha^k\varphi = \beta\varphi\})t^k$$

---

*The "finite to one" condition is redundant for finitely presented $\mathbb{Z}$-systems, see [Fr].

is rational.

<u>Proof</u>: One can assume without loss of generality that (Φ,α) is represented by some zero-one (Markov) matrix $M = M(i,j)$ for $1 \leq i,j \leq n$, such that the action of β is given by some permutation β' on the index set $\{1,2,\ldots,n\}$. Then $Z(t) = \sum_k t^k \sum_{i=1}^{n} M^k(i,\beta'(i))$ and the proof follows.

Now we prove the theorem by breaking our original Z(t) into a finite linear combination of the function $Z(t,\alpha_\nu,\beta_\nu)$ for some auxiliary Markov shifts $(\Phi_\nu,\alpha_\nu,\beta_\nu)$. These shifts will be subshifts in the Cartesian products Ψ^m for $m = 1,2,\ldots,n$ for some fixed n depending on N and i_0. Namely we consider in Ψ^m all possible intersections of the subshifts of the following two types

$$R = R(i_1,j_1;i_2,j_2) = \{\psi_1,\ldots,\psi_m \mid (i_1\psi_{j_1},i_2\psi_{j_2}) \in R_0\}$$

and

$$\Delta = \Delta(i_1,j_1;i_2,j_2) = \{\psi_1,\ldots,\psi_m \mid (i_1\psi_{j_1}=i_2\psi_{j_2})\}$$

where $1 \leq j_1,j_2 \leq m$ and $-i_m \leq i_1,i_2 \leq i_m$ for some constant i_m depending on N and i_0.

Some among these intersections $\Phi_\nu \subset \Psi^m$ are invariant under certain permutation β_ν (of the index set $\{1,\ldots,m\}$) acting on Φ^m and thus we get all our $(\Phi_\nu,\alpha_\nu,\beta_\nu)$ where α_ν stands for the (diagonal) action of the original shift on $\Phi_\nu \subset \Psi^m$. Besides the Z-functions corresponding to these Φ_ν our Z(t) also will contain finitely many terms of the form $\sum_k t^{ik}$ corresponding to i-periodic points with $i \leq i(i_0,N)$.

The counting process is naturally divided into two steps. The first can be expressed by the following trivial lemma.

Consider an arbitrary action of $\mathbb{Z}$ on a set Ψ with finitely many points of every fixed period and let $R \subset \Psi \times \Psi$ be a $\mathbb{Z}$-invariant equivalence relation, such that the projections $R \longrightarrow \Psi$ is at most N-to-one for $N < \infty$. Denote by $R_m \subset \underbrace{\Psi \times \Psi \times \ldots \times \Psi}_{m}$

the subset,

$$R_m = \{\psi_1,...,\psi_m \mid (\psi_i,\psi_j) \in R,\ 1 \leq i,j \leq m\}.$$

Notice that R_m naturally acted upon by $\mathbb{Z}$ and by the permutation group S_m.

<u>Lemma</u>: *Let (counting) Z-function for* Ψ/R *is a universal rational linear combination of Z-functions for* (R_m,σ) *for finitely many (depending on* N) $n = 1,2,...$ *and* $\sigma \in S_m$.

In order to conclude the proof of 8.5.U we must relate σ-periodic points of R_m^n to those of the (Markov!) relation $(R_0)_m$. Since $R_m \subset (R_0)_m$ every σ-periodic point in R_m corresponds to one or more (but finitely many) σ-periodic points in $(R_0)_m$. To count the multiplicity we project $(R_0)_m$ onto R_m by

$$(\psi_1,...,\psi_m) \longmapsto (i_1\psi_1,...,i_m\psi_m)$$

where $(i_1,...,i_m)$ is the first (for the lexicographic order) n-tuple of integers i_ν with $|i_m| \leq i_m$ which is contained in R_m. Though this projection does not commute with permutations $\sigma \in S_m$, one can obviously express the counting function for σ-periodic points in R_m in terms of similar functions for $R'_{m'} \subset R_m^{m'} \subset \Phi^{mm'}$ where $R' \subset (R_0)_m \times (R_0)_m$ is the (Markov!) equivalence relation for the projection $(R_0)_m \longrightarrow R_m$, provided there is no point in Ψ with period $\leq k_0$ for some constant $k_0 = k_0(N,i_0)$. Since these (infinitely many) short period points contribute a rational summand to our Z-functions, the theorem follows from the lemmas.

8.5.W <u>The proof of 5.2.D</u>: Theorem 8.5.U implies the rationality of the counting function for *stable* conjugacy classes in Γ. Moreover, it implies the following more general fact. Let $\Gamma_0 \subset \Gamma$ be a quasiconvex subgroup and set

$$[N_k^0]_{st} = \sum p^{-1}(\gamma)| [\gamma] |_-,$$

where the sum is taken over the stable conjugacy classes of non-torsion elements in Γ_0 whose stable norm (measured for the given word metric in Γ) satisfies $| [\gamma] |_- = k/m_0$.

A simple generalization of the earlier argument leads to the following

8.5.X **Theorem**: *The (counting) function* $Z(\Gamma_0,t) = \sum_k [N_k^0]_{st} t^k$ *(as well as* $\xi(\Gamma_0,t) = \exp \sum_k k^{-1}[N_k^0]_{st} t^k)$ *is rational.*

Now one easily sees that the Z-function of 5.2.D is a linear combination of $Z(\Gamma_i,t)$, where Γ_i are centralizers of (finitely many) finite subgroups in Γ which are pairwise non-conjugate. Thus the proof of 5.2.D is concluded.

Remarks and open questions: It would be interesting to understand the Z-function with the (unstable) norm $| [\gamma] |$ instead of $| [\gamma] |_-$. Notice that $e_i(\gamma) = i| [\gamma] |_- - | [\gamma^i] |$ is eventually periodic in i for all $\gamma \in \Gamma$. Namely, there exist an integer i_0 depending only on Γ and the generating subset, such that $| e_i(\gamma) | \leqslant i_0$ and $e_i(\gamma) = e_{i+i_0}(\gamma)$ for $i \geqslant i_0$. This suggests a similar periodicity in i for the (rational) functions $\sum_r N_{\gamma^i}(r)t^r$ (see 5.2) and this raises the counting problem for the conjugacy classes $[\gamma]$ according to the "type" of $\sum_r N_\gamma(r)t^r$ rather than $| [\gamma] |_-$.

Finally, one wishes to use these Z-functions to produce some invariants of Γ (independent of the word metric). One possible approach consists in a study of some generalization of Z-function for non-word metrics on Γ and of continuous families of those. Then one may try to control the modification of such a Z-function as one continuously deforms one word metric to another. A specific invariant of Γ (which, in fact, is a quasi-isometry invariant) is the dimension of the L_p-*cohomology* of Γ. This dimension is either zero or $+\infty$, but by varying $p \in [1,\infty)$ one gets (many?) non-trivial invariants of Γ.

8.5.Y __Semistability of the action of Γ on $\partial\Gamma$__: Let a word hyperbolic group Γ with a fixed generating set $G \subset \Gamma$ act on a metric space D. Fix $\epsilon > 0$ and $\lambda > 1$ and call $\gamma \in \Gamma$ *expanding* at a given point $p \in D$ if γ^{-1} sends the ball $B(\gamma p, \lambda\epsilon) \subset D$ into the ball $B(p,\epsilon) \subset D$. Then define the subset $\psi_p \subset \Gamma \times \Gamma$ of the pairs (γ_1, γ_2) such that $\gamma = \gamma_1\gamma_2^{-1}$ is expanding and contained in G. Call the action of Γ on D *regular expanding* if this ψ_p is a quasigeodesic nonexpanding field on Γ in the sense of 7.6 for some G, ϵ and λ and all $p \in D$. (Strictly speaking, the "quasigeodesic non-expanding" applies to the Cayley graph X of (Γ, G).) Notice that the action of Γ on $\partial\Gamma$ is regular expanding for the metric on $\partial\Gamma$ constructed in 7.2. Furthermore, small perturbations of a regular expanding action on a compact space D (obviously) are regular expanding.

Now, according to 7.6, every quasigeodesic non-expanding field on Γ defines a unique point in $\partial\Gamma$. Hence, there is a continuous Γ-equivariant map $D \longrightarrow \partial\Gamma$ for every regularly expanding action of Γ on D. In particular *the action of Γ on $\partial\Gamma$ is semistable*. Namely, for every small perturbation of this action, say $(\partial\Gamma, \Gamma)_{pert}$, there exists a continuous Γ-equivariant map $S: (\partial\Gamma, \Gamma)_{pert} \longrightarrow (\partial\Gamma, \Gamma)$. Clearly such an S is unique and converges to the identity as the perturbation goes to zero. (This semistability result in a geometric context is due to Sullivan, see [Sul].)

The case of groups Γ of rank $\geqslant 2$. Let X be symmetric space of non-positive curvature, let ∂X be the space of asymptotic classes of Weil chambers and let Γ be a discrete cocompact isometry group acting on X and thus on ∂X. By moving (compact pieces of) chambers from X to a fixed Γ-orbit in X one can easily produce "quasiflat non-expanding fields" in Γ and show that the action of Γ on ∂X admits a Markov presentation (over Γ) and that this action is semistable. However, one does not know how to make such a construction quasi-isometry invariant if $\text{rank}_{\mathbb{R}} X \geqslant 2$. One also does not know how to find a Markov presentation of ∂X which is finite-to-one and how to count words (and conjugacy classes of Abelian subgroups) in Γ.

8.6 <u>Relative hyperbolicity</u>: Let us give a hyperbolic version of small cancellation theory over free products by adopting geometric language of manifolds with cusps.

Let X be a complete hyperbolic locally compact geodesic space with a discrete isometric action of a group Γ such that the quotient space $V = X/\Gamma$ is quasi-isometric to the union of k copies of $[0,\infty)$ joined at zero. To simplify the matter we assume the action of Γ on X is free and then lift the k rays in X (corresponding to the k points in $\partial V \approx \{1,2,\dots,k\}$) to k rays r_i: $[0,\infty) \longrightarrow X$, $i = 1,\dots,k$. Denote by h_i the corresponding (ray) horofunctions and by $r_i(\infty) \in \partial X$ the limit points of r_i. Denote by $\Gamma_i \subset \Gamma$ the isotropy subgroups of $r_i(\infty)$ for the action of Γ on ∂X and assume that Γ_i preserves h_i for $i = 1,\dots,k$. Denote by $B_i(\rho)$ the horoballs $h^{-1}(-\infty,\rho) \subset X$ and assume that for a sufficiently small ρ the intersection $\gamma B_i(\rho) \cap B_j(\rho)$ is empty unless $i = j$ and $\gamma \in \Gamma_i$. Denote by $\Gamma B(\rho) \subset X$ the union $\bigcup_{i,\gamma} \gamma B_i(\rho)$ over $i = 1,\dots,k$ and all $\gamma \in \Gamma$, set $X(\rho) = X \setminus \Gamma B(\rho)$ and assume the action of Γ on $X(\rho)$ is cocompact for all $\rho \in (-\infty,+\infty)$.

8.6.A <u>Definition</u>: A group Γ is called word hyperbolic relative to some subgroups $\Gamma_1,\dots,\Gamma_k$ in Γ if Γ admits an action on some X with the above properties, where Γ_i are the isotropy subgroups of h_i.

<u>Examples</u>:

(a) Take the free product of some groups Γ_i and add finitely many relations which are "generic sufficiently long" words in $\gamma_i \in \Gamma_i$. The resulting group Γ is hyperbolic relative to Γ_i.

(b) Take a group Γ with a subgroup $\Gamma_0 \subset \Gamma$ and let $\gamma_1,\dots,\gamma_m$ in Γ_0 generate Γ_0 and $\gamma_1,\dots,\gamma_m,\dots,\gamma_{m+n}$ generate Γ. Let $|\ |$ be the maximal left invariant metric in Γ for which $|\gamma_i| \leqslant 1$

for $i = 1,\dots,m+n$ and $|\gamma| \leqslant \log(1+|\gamma|_0)$ for all $\gamma \in \Gamma_0$, where $|\ |_0$ is the word metric in Γ_0. If $(\Gamma, |\ |)$ is hyperbolic and the embedding $(\Gamma_0, \log(1+|\ |_0)) \longrightarrow (\Gamma, |\ |)$ is quasi-isometric then Γ is word hyperbolic relative to Γ_0.

(c) Let Γ be a finite covolume discrete isometry group of a complete simply connected Riemannian manifold X with *pinched* negative curvature,

$$0 > -a \geqslant K(X) \geqslant -b > -\infty.$$

Then the quotient space $V = X/\Gamma$ is quasi-isometric to the wedge of several copies of $[0,\infty)$ and Γ is hyperbolic relative to the (cuspided) isotropy subgroups of $r_i(\infty) \in \partial X$. For example, every finite covolume lattice in a simple Lie group with $\mathbb{R}$ - rank = 1 is hyperbolic relative to the cuspidal subgroups.

(d) The relative hyperbolization in 3.4.C gives us a group $\pi_1(H)$ which is hyperbolic relative to $\pi_1(P_0)$.

8.6.B All our results on word hyperbolic groups extend with minor modifications to the relative case. Here we only look on the relative version of Markov coding of Γ.

8.6.B_1 <u>Definition</u>: A graph Y is called *geo-Markov* if for every finite subset of vertices $Y_0 \subset Y$ there exists a finite union $\tilde{Y}$ of Markov trees of finite types and a simplicial map $\sqcap: \tilde{Y} \longrightarrow Y$ bijective on the set of vertices, such that every point $\tilde{y} \in \tilde{Y}$ can be joined by a path (in $\tilde{Y}$) with some vertex $\tilde{y}_0 \in \tilde{Y}_0 \underset{def}{=} \sqcap^{-1}(Y_0)$ and

$$\operatorname{dist}(\tilde{y}, \tilde{Y}_0) = \operatorname{dist}(y, Y_0)$$

for all $\tilde{y} \in \tilde{Y}$ and $y = \sqcap(\tilde{y})$. Next, a group Γ is called *geo-Markov* if every locally finite graph Y with the edges of integral length which admits a cocompact isometric action of Γ is geo–Markov. Clearly, every geo–Markov group Γ has a *rational* word counting Z–function.

8.6.C <u>Examples</u>:

(a) We know already that every word hyperbolic group is geo–Markov.

(b) It is likely that *every Abelian group is geo–Markov*. To see this we order the edges issuing from $x \in Y$, for all vertices $x \in Y$, such that this local order is Γ–invariant. This gives us as earlier, a (lexicographic) order on shortest paths from x to Y_0 for all vertices x in Y. The first path for this ordering consists of at most N "pieces" for some N depending on (Y,Γ) and Y_0, where every piece P_i, i = 1,...,N is the composition of Γ–translates of a fixed path P_i,

$$P_i = \gamma_1 P_1 * \gamma_2 P_1 * \ldots * \gamma_m P_1,$$

for some m depending on the original shortest path. Therefore, this path looks like a broken line with at most N break–points, where the "segment" between two breaks is parallel to one of N given "directions" p_i. The "type" of such a path is given by a finite sequence of directions. Now every $x \in Y$ is characterized by *finitely many* possible types of first shortest paths from z to Y_0 via x for all $z \in Y$ and the geo–Markov property becomes feasible.

8.6.D <u>Proposition</u>: *Let Γ be word hyperbolic relative to some subgroups Γ_i, i = 1,...,k, in Γ. If all Γ_i are geo–Markov then Γ also is geo–Markov. In particular, the counting function for* $\#\{\gamma \in \Gamma \mid |\gamma| = k\}$ *is rational for every word metric in* Γ. (This suggests an approach to the rationality question in [Can] for lattices in $O(n,1)$.)

<u>Proof</u>: Let us look at the geometry of $X(\rho) \subset X$ defined at the

beginning of this section. Take a segment $[x_1,x_2]$ in X and let $[x_1,x_2]_\rho$ be the union of $[x_1,x_2]$ with the Γ-translates of the horoballs $B_i(\rho)$ which intersect $[x_1,x_2]$. Then one easily sees with the convexity analysis in 7.3 that the subset $[x_1,x_2]_\rho \subset X$ is ε-convex, where ε only depends on X and ρ. Therefore, if x_1 and x_2 lie in X(ρ) then every shortest path in X(ρ) between x_1 and x_2 (if such path exists) lies ε'-close to $[x_1,x_2]_\rho$ where ε' depends only on X and ρ. Moreover, the same holds true for quasigeodesics in X(ρ) for the geodesic metric defined by the lengths of minimal paths in X(ρ) between pairs of points. Now, we may assume X(ρ) equals our graph Y as the two are quasi-isometric. Then two shortest paths issuing from nearby points in Y and heading for Y_0 stay close unless they follow for some time same Γ_i-orbit or its Γ-translate. Then, as the paths go away from such an orbit, they again become mutually close. Therefore, the geometry of $C(Y_0,x)$ is determined by the pattern of Γ-translates of Γ_i-orbits lying "behind" x (as looked upon from Y_0) and by the geometry of paths in Y_0. The Markov property for the former follows by our "hyperbolic" discussion in 7.6 and the latter is taken care of by "geo-Markov" for Γ_i. Making this geometric discussion rigorous is left to the reader.

<u>Questions</u>: Are finitely generated nilpotent groups geo-Markov? Are Cartesian products of geo-Markov (i.e. word hyperbolic) groups geo-Markov? Are 1/5-groups (and semihyperbolic groups, in general) geo-Markov? What are examples (if there are any) of Markov groups (see 5.2.E) which are not geo-Markov?

<u>Remark</u>: Jim Cannon pointed out to me that the stronger Γ-invariant (see 8.5.) version of the geo-Markov property may fail for Cayley groups of finite extension of Abelian groups (and also of nilpotent groups). His criticism forced me to modify the definition and to leave the above (b) inconclusive.

REFERENCES

[Al] A.D. Alexandrov, A theorem on triangles in a metric space and some of its applications, Trudy Math. Inst. Steks. 38 (1951), 5-23.

[Bour] N. Bourbaki, XXXIV, Groups et algébres de Lie, Part 2, Hermann, Paris, 1968.

[Bow] R. Bowen, Markov partitions for axiom A diffeomorphisms, Am. J. Math. 92 (1970), 725-747.

[B-G-S] W. Ballman, M. Gromov and V. Schroeder, Manifolds of non-positive curvature, Progress in Math., Vol 61, Birkhäuser, 1985.

[B-Z] Yu. Burago and V. Zalgaller, Geometric Inequalities, Egrebn. der Math., Springer-Verlag, to appear.

[Can] J.W. Cannon, The combinatorial structure of cocompact discrete hyperbolic groups, Geometr. Dedicata, to appear.

[C-G] J. Cheeger and M. Gromov, L^2-cohomology and group cohomology, Topology, 25:2, (1986), 189-217.

[Dav] M. Davis, Groups generated by reflections and aspherical manifolds not covered by Euclidean space, Ann. of Math. 117 (2), (1983), 293-325.

[Dr] A. Dress, Trees, extensions of metric spaces and the cohomological dimension of certain groups, Preprint, Bielefeld.

[Fl] W.J. Floyd, Group completions and Kleinian groups, Ph.D. Thesis, Princeton, 1978.

[Fr] D. Fried, Finitely presented dynamical systems, preprint, 1986.

[Fu] H. Furstenberg, Boundary theory and stochastic processes on homogeneous spaces, Proc. Symp. in Pure Math. XXVI (1973), 193-233.

[GM] F. Gehring, G. Martin, Discrete quasiconformal groups, Proc. Lond. Math. Soc., to appear.

[Gr_1] M. Gromov, Hyperbolic manifolds, groups and actions, In. Riemannian surfaces and related topics, Ann. Math. Studies 97 (1981), 183-215.

[Gr_2] ________________, Volume and bounded cohomology, Publ. Math. IHES 56 (1983), 213-307.

[Gr_3] ________________, Filling Riemannian manifolds, Journ. Diff. Geom. 18 (1983), 1-147.

[Gr_4] ________________, Hyperbolic manifolds according to Thurston and Jorgensen, Springer Lect. Notes 842 (1981), 40-53.

[Gr_5] ________________, Infinite groups as geometric objects, Proc. ICM Waszawa, Vol 1 (1984), 385-391.

[G-T] M. Gromov and W. Thurston, Pinching constants for hyperbolic manifolds, to appear in Inv. Math.

[Ka] D. Kazdan, On a connection between the dual space of a group with the structure of it's closed subgroups, Funct. An. and Applications, 1:1 (1967), 71-74.

[Kl] W. Klingenberg, Lectures on closed geodesics, Section 5.3, Springer-Verlag (1978), 189-200.

[Ko] B. Kostant, On the existence and irreducibility of certain series of representations, BAMS 75 (1969), 627-642.

[Lub] A. Lubotzki, Group presentations, p-adic analytic groups and lattices in $SL_2(\mathbb{C})$, Ann. Math. 118 (1983), 115-130.

[L-S] R.C. Lyndon and P.E. Schupp, Combinatorial group theory, Springer-Verlag (1977).

[Man] A. Manning, Axiom A diffeomorphisms have rational zeta functions, Bull. Lond. Math. Soc. 3 (1971), 215-220.

[Marg] G. Margulis, The isometry of closed manifolds of constant negative curvature with the same fundamental group, Dokl. Ak. Nauk. SSSR 192 (1970), 736-737.

[Mor] M. Morse, A fundamental class of geodesics on any closed surface of genus greater than one, Trans. Am. Math. Soc. 26 (1924), 25-60.

[M-S] J. Morgan and P. Shalen, Valuation, trees and degeneration of hyperbolic structures, I. Ann. Math. 120 (1984), 401-476.

[Ol] A.Y. Olshanski, On a geometric method in the combinatorial group theory, Proc. ICM, Warszawa, Vol 1 (1984), 415-423.

[P] P. Pansu, Métriques de Carnot-Carathéodory et quasiisométries des espaces symmetriques, preprint.

[Sul] D. Sullivan, Quasiconformal homeomorphisms and dynamics, II, Preprint IHES, 1982.

[Th] W. Thurston, Geometry and topology of 3-manifolds, Princeton, 1978.

[Ti] J.L. Tits, Free subgroups of linear groups, J. Algebra 20 (1972), 250-270.

[Ti_1] ________________, On buildings and their applications, Proc. I.C.M. 1974, pp. 209-221, Vancouver.