GROUPS OF POLYNOMIAL GROWTH AND EXPANDING MAPS

by MIKHAEL GROMOV

Introduction

Consider a group Γ generated by $\gamma_1, \ldots, \gamma_k \in \Gamma$. Each element $\gamma \in \Gamma$ can be represented by a word $\gamma_{i_1}^{p_1} \gamma_{i_2}^{p_2} \ldots \gamma_{i_\ell}^{p_\ell}$ and the number $|p_1| + |p_2| + \ldots + |p_\ell|$ is called *the length* of the word. *The norm* $||\gamma||$ (relative to $\gamma_1, \ldots, \gamma_k$) is defined as the minimal length of the words representing γ. Notice, that one can have several shortest words representing the same $\gamma \in \Gamma$.

Examples. — Let Γ be the free Abelian group of rank 2 generated by γ_1, γ_2. Each $\gamma \in \Gamma$ can be represented as $\gamma_1^p \gamma_2^q$, $p, q \in \mathbf{Z}$, and $||\gamma|| = |p| + |q|$. (For the identity element $e \in \Gamma$ we set $||e|| = 0$.)

Let Γ be the free (non-Abelian) group of rank 2 generated by γ_1, γ_2. Each $\gamma \neq e$ can be uniquely written as $\gamma_1^{p_1} \gamma_2^{p_2} \gamma_1^{p_3} \ldots \gamma_{i_k}^{p_k}$, or as $\gamma_2^{p_1} \gamma_1^{p_2} \gamma_2^{p_3} \ldots \gamma_{i_k}^{p_k}$, where $i_j = 1, 2$ and $p_1, \ldots, p_k$ are non-zero integers. The norm of such a γ is equal to $\sum_{i=1}^{k} |p_j|$.

Let Γ be the free cyclic group generated by γ_0 by let us use the generators $\gamma_1 = \gamma_0^2$, $\gamma_2 = \gamma_0^3$ and $\gamma_3 = \gamma_0^4$. Relative to these $\gamma_1, \gamma_2, \gamma_3$ we obviously have $||\gamma_0|| = 2$, $||\gamma_0^2|| = 1$, $||\gamma_0^6|| = 2$ (because $\gamma_0^6 = \gamma_2^2 = \gamma_1 \gamma_3$), $||\gamma_0^{100}|| = 25$ and so on.

Elementary properties of the norm. — For any group one obviously has

$$||\gamma|| = ||\gamma^{-1}||,$$
$$||\gamma\gamma'|| \leq ||\gamma|| + ||\gamma'||.$$

Let $\gamma_1, \ldots, \gamma_k$ and $\delta_1, \ldots, \delta_\ell$ be two systems of generators in Γ. The corresponding norms $||\ ||^{\text{old}}$ and $||\ ||^{\text{new}}$ are not necessarily equal but there obviously exists a positive constant C such that for each $\gamma \in \Gamma$ one has

$$C||\gamma||^{\text{old}} \geq ||\gamma||^{\text{new}} \geq C^{-1}||\gamma||^{\text{old}}.$$

For a group Γ with fixed generators we denote by $B(r) \subset \Gamma$, $r \geq 0$, *the ball of radius r* centered at the identity element e. In other words, $B(r)$ consists of all $\gamma \in \Gamma$ with $||\gamma|| \leq r$. We denote by $\# B(r)$ the number of elements in $B(r)$.

For a free Abelian group of rank two generated by γ_1, γ_2 one has

$$\sharp\, B(r) = 2N^2 + 2N + 1 \quad \text{for} \quad r \in [N, N+1).$$

For the free (non-Abelian) group with two generators one has

$$\sharp\, B(r) = 2 \cdot 3^N - 1 \quad \text{for} \quad r \in [N, N+1).$$

Growth of a group. — One says that a group Γ with generators $\gamma_1, \ldots, \gamma_k$ has *polynomial growth* if there are two positive numbers d and C such that for all balls $B(r)$, $r \geq 1$, one has

$$\sharp\, B(r) \leq Cr^d.$$

One can easily see that this definition does not depend on the particular choice of the generators and so this notion is correctly defined for the finitely generated groups.

Examples. — The finitely generated Abelian groups are easily seen to be of polynomial growth. Also the finitely generated nilpotent groups are of polynomial growth (see [14] and the appendix).

If Γ is a finite extension of a group of polynomial growth, then Γ itself has polynomial growth. So we conclude:

If a finitely generated group Γ has a nilpotent subgroup of finite index then Γ has polynomial growth.

The free groups with $k \geq 2$ generators do not have polynomial growth. They even have *exponential growth*, i.e.

$$\sharp\, B(r) \geq C^r, \quad r \geq 1$$

for some real constant $C > 1$. One can immediately see that this property does not depend on the choice of the generators.

The following theorem settles the growth problem for the solvable groups:

(Milnor-Wolf [8] [14].) — *A finitely generated solvable group Γ has exponential growth unless Γ contains a nilpotent subgroup of finite index.*

This result together with a theorem of Tits (see [13] and § 4) implies:

(Tits.) — *A finitely generated subgroup Γ of a connected Lie group has exponential growth unless Γ contains a nilpotent subgroup of finite index.*

In this paper we prove the following.

Main theorem. — *If a finitely generated group Γ has polynomial growth then Γ contains a nilpotent subgroup of finite index.*

The proof if given in § 8.

One can combine this theorem with Shub's criterion (see § 1) and obtains the

Geometric corollary. — An expanding self-map of an arbitrary compact manifold is topologically conjugate to an infra-nil-endomorphism.

The proof and the definitions are given in § 1.

1. Expanding maps

A map f from a metric space X to a metric space Y is called *globally expanding* if for any two points $x_1, x_2 \in X$, $x_1 \neq x_2$, one has

$$\operatorname{dist}(f(x_1), f(x_2)) > \operatorname{dist}(x_1, x_2).$$

We call f *expanding* if each point $x \in X$ has a neighbourhood $U \subset X$ such that the restriction of f to U is globally expanding.

Suppose that X and Y are connected Riemannian manifolds of the same dimension. If X is a complete manifold without boundary, then each expanding map is a covering. In particular, when Y is simply connected such a map is a globally expanding homeomorphism.

Let X be a compact connected Riemannian manifold and let $f: X \to X$ be an expanding map. One can see that X has no boundary, and hence, the map $\tilde{f}: Y \to Y$ induced on the universal covering $Y \to X$ is a globally expanding homeomorphism. The inverse map $\tilde{f}^{-1}: Y \to Y$ is contracting. Moreover, for each $\delta > 0$ there is a positive ε such that for any two points $y_1, y_2 \in Y$ with $\operatorname{dist}(y_1, y_2) \geq \delta$ one has

$$\operatorname{dist}(\tilde{f}^{-1}(y_1), \tilde{f}^{-1}(y_2)) \leq (1 - \varepsilon) \operatorname{dist}(y_1, y_2).$$

This is obvious. (Notice, that we use in Y the Riemannian metric induced from X by the covering map $Y \to X$.) It follows that $\tilde{f}^{-1}$ has a unique fixed point and that Y is homeomorphic to the Euclidean space $\mathbf{R}^n$, $n = \dim X$. Now, it is clear that $f: X \to X$ also has a fixed point.

All these facts were established by M. Shub (see [11]). (Notice that the definitions used in [11] are slightly different from ours.)

Examples. — Consider the torus $T^n = \mathbf{R}^n/\mathbf{Z}^n$. Each linear map $\mathbf{R}^n \to \mathbf{R}^n$ which sends the lattice $\mathbf{Z}^n \subset \mathbf{R}^n$ into itself induces a map $T^n \to T^n$. This map is expanding if and only if all eigenvalues of the covering linear map $\mathbf{R}^n \to \mathbf{R}^n$ have absolute values greater than one.

Flat manifolds. — Let Γ be a discrete fixed point free group of motions of $\mathbf{R}^n$ with compact quotient $X = \mathbf{R}^n/\Gamma$. A linear map $\mathbf{R}^n \to \mathbf{R}^n$ which respects Γ induces a map $X \to X$ and this map is expanding if and only if the covering linear map $\mathbf{R}^n \to \mathbf{R}^n$

has only eigenvalues of absolute value greater than one. It is known (see [4]) that any flat manifold X has an expanding map of the type we have just described.

Nil-manifolds. — Let L be a simply connected nilpotent Lie group with a left invariant Riemannian metric and let $\Gamma \subset L$ be a discrete subgroup with compact quotient $X = L/\Gamma$. (Such an X is called a nil-manifold.) An automorphism $A: L \to L$ which sends Γ into itself induces a map $X \to X$ and this map is expanding if and only if the linear map $a: \ell \to \ell$ induced by A in the Lie algebra ℓ of L has all its eigenvalues greater than one in absolute value. Observe, that not all nilpotent Lie groups admit an expanding automorphism.

Infra-nil-manifolds. — Let L be as above and denote by Aff(L) the group of transformations of L generated by the left translations and by all automorphisms $L \to L$. Let $\Gamma \subset \mathrm{Aff}(L)$ be a group which acts freely and discretely on L. When the quotient $X = L/\Gamma$ is compact it is called an *infra-nil-manifold.* Each expanding automorphism $L \to L$ which respects Γ induces an expanding map $X \to X$. Such maps are called expanding *infra-nil-endomorphisms.*

Topological conjugacy. — Two maps $f: X \to X$ and $g: Y \to Y$ are called topologically conjugate if there exists a homeomorphism $h: X \to Y$ such that $h \circ f = g \circ h$, i.e. the following diagram commutes

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "h"'] & X \arrow[d, "h"] \\ Y \arrow[r, "g"'] & Y. \end{tikzcd}$$

M. Shub discovered the following remarkable fact (see [11]):

An expanding self-map of a compact manifold X *is uniquely determined, up to topological conjugacy, by its action on the fundamental group* $\pi_1(X)$.

The following two results of Shub and Franks (see [11]) are especially important for our paper.

Shub's criterion. — *An expanding self-map of a compact manifold* X *is topologically conjugate to an expanding infra-nil-endomorphism if and only if the fundamental group* $\pi_1(X)$ *contains a nilpotent subgroup of finite index.*

The polynomial growth property (Franks). — *If a compact manifold* X *admits an expanding self-map then the fundamental group* $\pi_1(X)$ *has polynomial growth.*

We prove this in the next section.

These two facts explain why the geometric corollary is a consequence of the main theorem.

Some partial results on the classification of the expanding maps were obtained earlier by Shub [11] and Hirsch [6]. An idea of Hirsch's paper plays an important role in our approach (see § 5).

2. Geometric growth

Consider a Riemannian manifold Y and denote by $\mathrm{Vol}_y(r)$, $y \in Y$, the volume of the ball of radius r around y. The growth of Y is defined as the asymptotic behavior of $\mathrm{Vol}_y(r)$ as $r \to \infty$.

This concept is due to Efremovič (see [3]) who pointed out that the growth of a manifold Y, which covers a *compact* manifold X, depends only on the fundamental groups $\pi_1(X)$, $\pi_1(Y)$ and the inclusion $\pi_1(Y) \subset \pi_1(X)$.

The corresponding algebraic notion of the growth was introduced by Švarc and by Milnor who, in particular, proved the following theorem.

(Švarc, Milnor). — *Let* $Y \to X$ *be the universal covering and let us fix a set of generators in the fundamental group* $\Gamma = \pi_1(X)$. *Then there is a constant* $C > 0$ *such that for each* $y \in Y$ *and all* $r \geq 1$ *one has*

$$\mathrm{Vol}_y(Cr + C) \geqslant \# B(r) \geqslant \mathrm{Vol}_y(C^{-1} r),$$

where $B(r)$ *denotes the ball in* Γ *(see the introduction).*

Proof. — Let us identify Γ with the orbit of $y \in Y$ under the action of Γ, so that y corresponds to the identity. Denote by $\widetilde{B}_y(r)$ the intersection of Γ with the Riemannian ball of radius r in Y centered at y. It is not hard to show (see [8], [12]) that

$$C_1 \mathrm{Vol}_y(r) \geq \# \widetilde{B}_y(r) \geq C_1^{-1} \mathrm{Vol}_y(r), \quad r \geq 1$$

and

$$\# \widetilde{B}_y(C_2 r) \geqslant \# B(r) \geqslant \# \widetilde{B}_y(C_2^{-1} r), \quad r \geq 0.$$

This implies the theorem.

Corollary. — *The fundamental group of a compact manifold* X *has polynomial growth if and only if the universal covering* Y *of* X *has polynomial growth, i.e. if for some* C *and* d *one has*

$$\mathrm{Vol}_y(r) \leq C r^d, \quad r \in [1, \infty).$$

Observe that most (complete, non-compact) manifolds have exponential growth, i.e. $\mathrm{Vol}_y(r) \geq C^r - 1$, $C > 1$, $r \in [1, \infty)$, but there are some interesting instances of polynomial growth.

Examples of manifolds of polynomial growth:

(a) Complete manifolds of non-negative Ricci curvature;
(b) Real algebraic submanifolds in $\mathbf{R}^q$;
(c) Nilpotent Lie groups with left invariant metrics;
(d) Leaves of Anosov foliations.

Remarks. — Polynomial growth for a manifold of non-negative Ricci curvature follows from Rauch's comparison theorem (see [2]); (b) and (c) are easy exercises; (d) has the same nature as the polynomial growth in the presence of an expanding map: a slight modification (required by a minor discrepancy between the notion of expansion) of the following simple lemma yields both facts.

Let $f: Y \to Y$ be a totally expanding smooth map. Suppose that the Jacobian of this map is bounded by a constant C and that f is uniformly expanding, i.e. for any two points $x, y \in Y$ with $dist(x,y) \geq 1$ one has

$$dist(f(x), f(y)) \geq (1+\varepsilon) dist(x,y), \quad \varepsilon > 0.$$

Then Y *has polynomial growth.*

Proof. — Each ball B of radius $r \geq 1$ is sent by f onto a set containing a ball of radius $(1+\varepsilon)r$ and whose volume is at most C Vol(B). It follows that

$$\mathrm{Vol}_y((1+\varepsilon)r) \leq \mathrm{C}\, \mathrm{Vol}_y(r), \quad r \geq 1,$$

where y is the fixed point of f. This inequality implies polynomial growth.

3. Elementary properties of the growth

Let Γ be a group with a fixed finite system of generators. The norm $\|\ \|$ (see the introduction) provides Γ with a left invariant metric

$$\mathrm{dist}(\alpha, \beta) = \|\alpha^{-1}\beta\|.$$

Consider a subgroup Γ' and the left action of Γ' on Γ. Denote by X the corresponding factor space Γ/Γ' and by $f: \Gamma \to X$ the natural projection. Define $\mathrm{dist}(x,y)$, $x, y \in X$ as $\inf_{\alpha, \beta} \mathrm{dist}(\alpha, \beta)$, $\alpha \in f^{-1}(x)$, $\beta \in f^{-1}(y)$. Since the action of Γ' on Γ is isometric, the function $\mathrm{dist}(x,y)$ is a metric in X.

Connectivity. — The space $X = \Gamma/\Gamma'$ has the following two equivalent properties:

(a) for any two points $x, y \in X$ with $\mathrm{dist}(x,y) = p$, where p is an integer, there exist points $x = x_0, x_1, \ldots, x_p = y$, such that $\mathrm{dist}(x_i, x_{i-1}) = 1$, $i = 1, \ldots, p$;

(b) take a ball $B \subset X$ of radius p, where p is an integer, and take its ε-neighbourhood $U_\varepsilon(B) \subset X$ where ε is also a non-negative integer. Then $U_\varepsilon(B)$ is exactly the ball of radius $p + \varepsilon$ concentric to B.

Both properties (as well as their equivalence) are obvious for Γ itself and they are preserved when we pass to X.

As an immediate application we have:

If X *is infinite then each ball in* X *of radius $r = 0, 1, \ldots,$ contains at least $r+1$ elements.*

This simple fact provides a useful relation between the growth of a group and its subgroups.

We define growth(Γ) as the lower bound of the numbers $d \geq 0$, such that

$$\# \mathrm{B}(r) \leq \text{const.}\, r^d, \qquad r \geq 1.$$

Splitting lemma. — *If* $\Gamma' \subset \Gamma$ *is a finitely generated subgroup of infinite index, then*

$$\operatorname{growth}(\Gamma') \leq (\operatorname{growth}(\Gamma)) - 1$$

Proof. — The connectivity properties of $X = \Gamma/\Gamma'$ imply that each ball $\mathrm{B}(r)$, $r = 0, 1, \ldots$, in Γ contains at least $r+1$ elements $\alpha_0, \ldots, \alpha_r \in \Gamma$ such that $f(\alpha_i) \neq f(\alpha_j)$ for $i \neq j$ (recall that $f: \Gamma \to X$ is the factor-map). Consider the intersection $\mathrm{B}' = \mathrm{B}(r) \cap \Gamma'$ and its translates $\mathrm{B}'\alpha_i$, $i = 0, \ldots, r$. These sets are disjoint and they are contained in the ball $\mathrm{B}(2r)$. It follows that $\# \mathrm{B}(2r) \geq (r+1)\,(\# \mathrm{B}')$. This yields the lemma.

Regular growth. — All balls in Γ of a given radius r have the same number of elements. We denote this number by $b(r) = \# \mathrm{B}(r)$. For a group of growth $d < \infty$ we call a number r *i-regular* $i = 1, 2, \ldots$, if it satisfies the following two conditions:

(a) $\log(b(2^{-j}r)) \geq \log(b(r)) - j(d+1)\log 2, \quad j = 1, 2, \ldots, i,$

(b) $\log(b(2^j r)) \leq \log(b(r)) + \beta_j, \quad j = 1, 2, \ldots, i, \quad$ where $\quad \beta_j = 16^{j+1}(d+1).$

Regularity lemma. — *There is a sequence* (r_i) *tending to* ∞ *such that each* r_i *is i-regular.*

Proof. — Start with the sequence $r'_k = 2^k$. Since $\operatorname{growth}(\Gamma) = d$ we have

$$\log(b(r'_k)) \leq C + k\, d \log 2.$$

This inequality implies that there is an infinite subsequence $r_i = 2^{k_i}$ which satisfies (a), i.e.

$$\log(b(2^{k_i - j})) \geq \log(b(2^{k_i})) - j(d+1)\log 2, \quad j = 1 \ \ldots, i.$$

Let us show that this sequence must also satisfy (b); in other words (a) implies (b) for large r.

We first prove the following general inequality which is valid for all finitely generated groups

$$b(5r) \leq \frac{(b(4r))^2}{b(r)}, \quad r = 1, 2, \ldots. \tag{*}$$

Proof. — Consider a maximal system of points $\gamma_1, \gamma_2, \ldots \in \mathrm{B}(3r)$ such that the distance between any two of them is at least $2r+1$. It is clear that the balls of radius r centered at $\gamma_1, \gamma_2, \ldots$, do not intersect and the concentric balls of radius $2r$ cover $\mathrm{B}(3r)$. Using the connectivity property of Γ we conclude that the concentric balls of radius $4r$ cover $\mathrm{B}(5r)$. This proves (*), because the original balls of radius r were contained in $\mathrm{B}(4r)$ and the total number of their points could not exceed $b(4r)$.

We simplify the notations by setting $\ell(r)=\log(b(r))$. Then the inequality (*) becomes

$$\ell(5r)\leq 2\ell(4r)-\ell(r).$$

When r is divisible by 4 this implies

$$\ell(6r)\leq \ell\left(5r+\frac{5r}{4}\right)\leq 2\ell(5r)-\ell(r),$$

and so

$$\ell(6r)\leq 4\ell(4r)-3\ell(r).$$

In the same way we get

$$\ell(8r)\leq \ell\left(6.\frac{6r}{4}\right)\leq 16\ell(4r)-15\ell(r).$$

It means that for an r divisible by 16 we have

$$\ell(2r)\leq 16\ell(r)-15\ell\left(\frac{r}{4}\right).$$

Applying this inequality j times we get

$$\ell(2^j r)\leq 16^j\left(\ell(r)-\ell\left(\frac{r}{4}\right)\right)+\ell\left(\frac{r}{4}\right).$$

In our case $\ell(r)-\ell\left(\frac{r}{4}\right)\leq 2(d+1)\log 2$ and so

$$\ell(2^j r)\leq 16^{j+1}(d+1)+\ell\left(\frac{r}{4}\right)\leq 16^{j+1}(d+1)+\ell(r),$$

q.e.d.

4. Linear representations

A group Γ is called, for brevity, *almost nilpotent (almost solvable)* if it contains a nilpotent (solvable) subgroup of finite index.

This section is devoted to the proof of the following.

Algebraic lemma. — Let Γ be a finitely generated group of polynomial growth and let L be a Lie group with finitely many connected components. Suppose that for each finitely generated infinite subgroup $\Gamma'\subset\Gamma$ there is a subgroup $\Delta\subset\Gamma'$ of finite index in Γ' with the following property: for every $p=1, 2, \ldots,$ there is a homomorphism $\Delta\to\mathrm{L}$ such that its image contains at least p elements. Then Γ is almost nilpotent.

Our proof is based on the following fundamental facts.

(Jordan). — *For each Lie group* L *with finitely many components there is a number* q *such that every finite subgroup in* L *contains an Abelian subgroup of index at most* q. (See [10].)

(Tits). — *Let* L *be as above and let* $G \subset L$ *be an arbitrary finitely generated subgroup. Then there are only two possibilities*:

(a) G *contains a free group of rank* 2. *In this case* G *has exponential growth.*

(b) G *is almost solvable. In this case* G *has exponential growth unless it is almost nilpotent.*

(See [13].)

We first prove two simple lemmas.

(a) *Let* L *be as above and let* G *be an arbitrary finitely generated group. Suppose that for every number* $p = 1, 2, \ldots,$ *there is a homomorphism* $G \to L$ *such that its image is finite and has at least* p *elements. Then* G *contains a subgroup* $G' \subset G$ *of finite index such that the commutator group* $[G', G'] \subset G'$ *has infinite index and, consequently,* G′ *admits a non-trivial homomorphism in* $\mathbf{Z}$.

Proof. — Let q be as in Jordan's theorem. Take for $G' \subset G$ the intersection of all subgroups in G of index at most q. It is clear that G′ satisfies all the requirements.

(b) *Let* Γ *be a finitely generated group of polynomial growth. Then the commutator subgroup* $[\Gamma, \Gamma]$ *is also finitely generated.*

Proof. — It is sufficient to show that the kernel $\Delta \subset \Gamma$ of any surjective homomorphism $g : \Gamma \to \mathbf{Z}$ is finitely generated.

Take a system of generators $\gamma_0, \gamma_1, \ldots, \gamma_k \in \Gamma$ with the following properties:

$$g(\gamma_0) = z_0 \in \mathbf{Z}, \quad \text{where } z_0 \text{ denotes the generator in } \mathbf{Z},$$

$$\gamma_i \in \Delta, \quad i = 1, \ldots, k.$$

Denote by $\Delta_m \subset \Delta$ the subgroup generated by

$$\{\gamma_0^j \gamma_i \gamma_0^{-j}\}, \quad i = 1, \ldots, k; \quad j = -m, \ldots, 0, 1, \ldots, m.$$

One obviously has

$$\bigcup_0^\infty \Delta_m = \Delta.$$

If for some number m one has $\Delta_m = \Delta_{m+1}$ then $\Delta_m = \Delta$ and the proof is finished.

Otherwise, there is a sequence $\alpha_m \in \Delta$, $m = 0, 1, \ldots,$ such that each α_m is of the form $\alpha_m = \gamma_0^m \gamma_i \gamma_0^{-m}$ or $\alpha_m = \gamma_0^{-m} \gamma_i \gamma_0^m$, for some $i = 1, \ldots, k$ and α_m is not contained in the group generated by $\alpha_0, \alpha_1, \ldots, \alpha_{m-1}$.

Consider all the products $\beta = \beta(\varepsilon_0, \ldots, \varepsilon_m) = \alpha_0^{\varepsilon_0} \alpha_1^{\varepsilon_1} \ldots \alpha_m^{\varepsilon_m}$ where $\varepsilon_i = 0, 1$. It is clear that the equality $\beta(\varepsilon_0, \ldots, \varepsilon_m) = \beta(\varepsilon'_0, \ldots, \varepsilon'_m)$ implies $\varepsilon_0 = \varepsilon'_0$, $\varepsilon_1 = \varepsilon'_1, \ldots,$ $\varepsilon_m = \varepsilon'_m$. So we have 2^{m+1} different β's.

On the other hand $||\beta|| \leq ||\alpha_0|| + ||\alpha_1|| + \ldots + ||\alpha_m|| \leq (m+1)(2m+1)$ and for the ball $B((m+1)(2m+1)) \subset \Gamma$ we have

$$\# B((m+1)(2m+1)) \geq 2^{m+1}, \quad m = 1, 2, \ldots$$

This contradicts the polynomial growth, q.e.d.

Proof of the algebraic lemma. — According to the splitting lemma (see § 3) we can use induction and assume that all finitely generated subgroups in Γ of infinite index are almost nilpotent. Let $\Delta \subset \Gamma$ be a subgroup of finite index which has the required homomorphisms into L. If all these homomorphisms have finite images we use lemma (a) and get a subgroup $\Delta' \subset \Delta$ of finite index such that the commutator subgroup $[\Delta', \Delta'] \subset \Delta'$ has infinite index.

If there is a homomorphism $\Delta \to L$ with infinite image, we apply Tits' theorem to this image and again obtain $\Delta' \subset \Delta$ with the same property.

According to lemma (b) the commutator subgroup $[\Delta', \Delta']$ is finitely generated and by the induction hypothesis it is almost nilpotent. It follows that Γ is almost solvable and, by the theorem of Milnor-Wolf, Γ is almost nilpotent, q.e.d.

Corollary. — *Let Γ and Γ' be as in the algebraic lemma. If each Γ' has a subgroup Δ of finite index such that either Δ satisfies the condition of the lemma or Δ is Abelian, then Γ is almost nilpotent.*

This is a trivial consequence of the lemma.

5. Topological transformation groups

The following deep theorem plays a crucial role in our proof.

(Montgomery-Zippin). — *Let Y be a finite dimensional, locally compact, connected and locally connected metric space. If the group L of the isometries of Y is transitive (on Y) then L is a Lie group with finitely many connected components.*

The proof immediately follows from the first corollary in § 6.3 of the book [9].

We shall also need an obvious corollary of this theorem.

Localization lemma. — *Let Y be as above, let $U \subset Y$ be a non-empty open set and let $p = 1, 2, \ldots$ There exists a positive ε with the following property:*

If $\ell : Y \to Y$ is a non-trivial (i.e. ℓ is not the identity) isometry such that $\mathrm{dist}(u, \ell(u)) \leq \varepsilon$, $u \in U$, *then ℓ generates in L a subgroup of order at least p.*

The idea of applying the theory of Montgomery-Zippin to the classification of expanding maps is due to Hirsch (see [6]). He proceeds as follows.

An expanding map $X \to X$ lifts to a globally expanding homeomorphism $f: Y \to Y$ of the universal covering $Y \to X$ and f respects the action of $\Gamma = \pi_1(X)$ on Y. Hirsch views Γ as a subgroup of the group of all homeomorphisms of Y and he constructs subgroups $\Gamma = \Gamma_0 \subset \Gamma_1 \subset \ldots \subset \Gamma_i \subset \ldots$ by setting $\Gamma_i = f^{-i} \Gamma f^i$.

The closure of the union $\bigcup_i \Gamma_i$ is a topological group acting on Y, and Hirsch shows that, in some cases, this group satisfies the requirements of the theory of Montgomery-Zippin (the same corollary in § 6.3 of their book) and thus, he realizes Γ as a subgroup of a Lie group.

In our approach we do not use the universal covering but construct Y as a limit of discrete spaces.

6. Limits of metric spaces

Consider a space Z with a metric δ and take two sets $X, Y \subset Z$. *The Hausdorff distance* $H^\delta(X, Y)$ is defined as the lower bound of the numbers $\varepsilon > 0$ such that the ε-neighbourhood of X contains Y and the ε-neighbourhood of Y contains X. The Hausdorff distance can be infinite but it has all properties of a metric.

Consider now two arbitrary metric spaces X and Y and denote by Z their disjoint union. A metric δ on Z is called *admissible* if its restrictions to X and Y are equal to the original metrics in X and Y respectively.

We define the *Hausdorff distance* $H(X, Y)$ as the lower bound $\inf_\delta H^\delta(X, Y)$ where δ runs over all admissible metrics on $Z = X \cup Y$.

When X and Y are *compact* spaces the Hausdorff distance enjoys all the properties of a metric. In particular, one has:

$$H(X, Y) = 0 \text{ if and only if } X \text{ and } Y \text{ are isometric.}$$

When the spaces are not compact, it is convenient to have reference points in them and to use the following " modified Hausdorff distance " [1]. For two metric spaces X, Y with distinguished points $x \in X$ and $y \in Y$, we define $\widetilde{H}((X, x), (Y, y))$ as the infimum of all $\varepsilon > 0$ with the following property: there exists an admissible metric δ on the disjoint union $X \cup Y$ such that $\delta(x, y) < \varepsilon$, that the ball $B_x(1/\varepsilon)$ of radius $1/\varepsilon$ centered at x in X is contained in the ε-neighborhood of Y (with respect to δ) and, similarly, that $B_y(1/\varepsilon)$ in Y is contained in the ε-neighborhood of X. For three spaces with distinguished points, the function $\widetilde{H}$ satisfies the triangular inequality provided that at least two of the three " distances " involved are small enough $\left(\text{say, } \leqslant \frac{1}{2}\right)$.

Proper spaces. — A metric space X is called *proper* if for each point $x_0 \in X$, the distance function $x \to \operatorname{dist}(x_0, x)$ is a proper map $X \to \mathbf{R}$, i.e. if each closed ball (of

[1] The definition of the modified distance $\widetilde{H}$ is due to O. Gabber who kindly pointed out that another function $\widetilde{H}$ introduced in an earlier version of this paper was inadequate.

finite radius) in X is compact. Observe that a Riemannian manifold is proper if and only if it is complete. This is the theorem of Hopf-Rinow (see [2]).

Convergence. — We say that a sequence of spaces X_j with distinguished points $x_j \in X_j$ converges to (Y, y), and we write $(X_j, x_j) \underset{j\to\infty}{\longrightarrow} (Y, y)$ if $\lim \widetilde{H}((X_j, x_j), (Y, y)) = 0$. If the spaces X_j are compact *with uniformly bounded diameter*, this implies that $H(X_j, Y) \to 0$. It is not difficult to see that if the spaces X_j are proper and if there exist arbitrarily large real numbers r such that the sequence $(B_{x_j}(r))$ of balls of radius r in the X_j converges for the Hausdorff distance H, then a subsequence of (X_j, x_j) converges (in the above sense) to a proper space with distinguished point.

Uniform compactness. — A family $\{X_j\}$, $j \in J$, of compact metric spaces is called *uniformly compact* if their diameters are uniformly bounded and one of the following two equivalent conditions is satisfied:

(a) for each $\varepsilon > 0$ there is a number $N = N(\varepsilon)$ such that each space X_j, $j \in J$, can be covered by N balls of radius ε;

(b) for each $\varepsilon > 0$ there is a number $M = M(\varepsilon)$ such that in each space $X_j, j \in J$, one can find at most M disjoint balls of radius ε.

Compactness criterion. — *Let* $(X_j, x_j)_{j=1,2,\dots}$, *be a sequence of proper metric spaces with distinguished points. If for each* $r \geq 0$ *the corresponding family of balls* $\{B_j(r)\}_{j=1,2,\dots}$, *is uniformly compact, then there is a subsequence* $(X_{j_k}, x_{j_k})_{k=1,2,\dots}$ with $\lim\limits_{k\to\infty} j_k = \infty$, *which converges to a proper metric space* (Y, y).

Proof. — To prove the criterion it is sufficient to find a convergent subsequence of $(B_j(r))$ for an arbitrary but fixed number r and thus we can assume without loss of generality that all X_j are compact and that the family $\{X_j\}_{j=1,2,\dots}$ is uniformly compact.

Take the sequence $\varepsilon_i = 2^{-i}$ and let N_i be natural numbers such that each X_j can be covered by N_i balls of radius ε_i. Denote by A_i the set of all finite sequences of the form $(n_1, n_2, \dots, n_i)$, $1 \leq n_1 \leq N_1$, $1 \leq n_2 \leq N_2$, $\dots$, $1 \leq n_i \leq N_i$, and denote by $p_i : A_{i+1} \to A_i$ the natural projection.

For each space $X_j, j = 1, \dots$, there are maps $I_j^i : A_i \to X_j$ with the following properties:

(a) the image of I_j^i forms an ε_i-net in X_j, i.e. the ε_i-balls centered at the points of this image cover X_j;

(b) for each $a \in A_{i+1}$, $i = 1, 2, \dots$, the point $I_j^{i+1}(a)$ is contained in the $2\varepsilon_i$-ball centered at $I_j^i(p_i(a))$.

These maps are constructed as follows. First we cover X_j by N_1 balls of radius ε_1 and we take any bijective map from A_1 onto the set of centers of these balls. This is our map I_j^1. After that we cover each ε_1-ball by N_2 balls of radius ε_2 and map A_2 on the set of centers of these ε_2-balls so that (n_1, n_2) goes to the center of a ball which we used to cover the ε_1-ball centered at $I_j^1((n_1))$. This is our I_j^2. Then we cover each ε_2-ball by N_3 balls of radius ε_3 and map A_3 onto the set of centers of these ε_3-balls, so that $(n_1, n_2, n_3) \in A_3$ goes to the center of a ball which was used in covering the ε_2-ball centered at $I_j^2((n_1, n_2))$, and so on.

Denote by A the union $\bigcup_{i=1}^{\infty} A_i$ and by $I_j : A \to X_j$ the map corresponding to all I_j^i, $i = 1, \ldots$ Denote by F′ the space of all bounded functions $f : A \to \mathbf{R}$ with the norm $\|f\| = \sup_{a \in A} |f(a)|$. Denote by $F \subset F'$ the set which consists of all functions satisfying the following inequalities:

$$\text{if} \quad a \in A_1 \subset A, \quad \text{then} \quad 0 \leq f(a) \leq \sup_j \operatorname{Diam} X_j,$$
$$\text{if} \quad a \in A_i,\ i > 1, \quad \text{then} \quad |f(a) - f(p_{i-1}(a))| \leq 2\varepsilon_{i-1}.$$

The set F is compact.

Let us define a map $h_j : X_j \to F'$ as follows

$$(h_j(x))(a) = \operatorname{dist}(x, I_j(a)), \quad x \in X_j, \quad a \in A,$$

and "dist" is taken relative to the metric in X_j.

The property (a) of I_j^i implies that the map h_j is isometric and the property (b) shows that the image of h_j is contained in F. So we have proved:

If the family $\{X_j\}$ *is uniformly compact then there is a compact metric space* F *such that each* X_j *can be isometrically embedded into* F.

To complete the proof of the compactness criterion we invoke the following well-known fact.

Let F *be a compact metric space with metric* δ. *Then the space of all compact subsets of* F *is a compact space relative to the Hausdorff distance* H^δ.

We now identify each X_j with its image $h_j(X_j) \subset F$ and take a subsequence X_{j_k} which converges to a compact set $Y \subset F$, i.e. $\lim_{k \to \infty} H^\delta(X_{j_k}, Y) = 0$, where δ is the metric associated to the norm in $F' \supset F$. It is clear that the distance $H(X_{j_k}, Y)$ also converges to zero as $k \to \infty$, q.e.d.

Example. — Let (X_j) be a system of complete n-dimensional Riemannian manifolds with Ricci curvature bounded from below by a negative constant. Then the sequence (X_j, x_j) has a subsequence which converges to a metric space (Y, y), but this space is not, in general, a manifold. (See [5] for additional examples.)

Definite convergence. — Let (X_j, x_j) be a sequence of spaces which converges to (Y, y). By definition, there exists a system of metrics δ_j in the disjoint unions $X_j \cup Y$ such that, for any given $r \geqslant 0$ and $\varepsilon > 0$, the following properties hold for almost all j: one has

$\delta_j(x_j, y) \leqslant \varepsilon$, the ball $B_{x_j}(r)$ in X_j is contained in the ε-neighborhood of Y (with respect to δ_j) and the ball $B_y(r)$ in Y is contained in the ε-neighborhood of X_j. When such metrics δ_j are chosen and fixed, we say that there is *definite convergence* and we write

$$(X_j, x_j) \underset{j \to \infty}{\Rightarrow} (Y, y).$$

Now it makes sense to speak about convergence of a sequence $x'_j \in X_j$ to a point $y' \in Y$: this just means that $\lim_{j \to \infty} \delta_j(x'_j, y') = 0$. In particular, the reference points x_j converge to $y \in Y$.

Convergence of maps. — Consider a sequence $(X_j, x_j) \underset{j \to \infty}{\Rightarrow} (Y, y)$ and a system of maps $f_j : X_j \to X_j$. We say that the maps f_j *converge to a map* $f : Y \to Y$ if for each $r \geq 0$ and each positive ε there is a number $\mu = \mu(r, \varepsilon) > 0$ and an integer $N = N(r, \varepsilon)$ such that, for all $j \geq N$, one has:

If the points $x' \in B_{x_j}(r) \subset X_j$ and $y' \in B_y(r) \subset Y$ satisfy $\delta_j(x', y') \leq \mu$ then $\delta_j(f_j(x'), f(y')) \leq \varepsilon$.

Isometry Lemma. — Let $(X_j, x_j) \underset{j \to \infty}{\Rightarrow} (Y, y)$ *and let* Y *be a proper space. Let* $f_j : X_j \to X_j$ *be isometries such that* $dist_j(x_j, f_j(x_j)) \leq C$ *(where* C *is a constant which does not depend on* j *and* $dist_j$ *denotes the metric in* X_j*). Then there is a subsequence* (X_{j_k}, x_{j_k}) *such that the maps* f_{j_k} *converge to an isometry* $Y \to Y$.

Proof. — Take a sequence (ε_i), with $\lim_{i \to \infty} \varepsilon_i = 0$ and $\varepsilon_i \leqslant 1/4$, and a sequence (r_i) such that $r_{i+1} \geqslant r_i + C + 1$. Upon passing to a subsequence, we may—and shall—assume that, for all j, $\delta_j(x_j, y) \leqslant \varepsilon_j$, that the ball $B_y(r_j)$ is contained in the ε_j-neighborhood of X_j and that the ball $B_{x_j}\left(r_j + C + \frac{1}{2}\right)$ is contained in the ε_j-neighborhood of Y. For all i, choose a finite ε_i-net in $B_y(r_i)$. Now, construct a system of maps $g_{ij} : E_i \to E_{i+1}$ as follows. For $e \in E_i$, choose a point $x \in X_j$ such that $\delta_j(e, x) \leqslant \varepsilon_j$. Thus, $x \in B_{x_j}\left(r_i + \frac{1}{2}\right)$ hence $f_j(x) \in B_{x_j}\left(r_i + C + \frac{1}{2}\right)$. Now choose $x' \in Y$ such that $\delta_j(x', x) \leqslant \varepsilon_j$. Then, $x' \in B_y(r_{i+1})$. Finally, take for $g_{ij}(e)$ any point e' of E_{i+1} such that $\mathrm{dist}(x', e) \leqslant \varepsilon_{i+1}$.

There exists a sequence $j_1, \ldots, j_k, \ldots$, such that for each $i = 1, 2, \ldots$, the map g_{ij_k} does not depend on k for $k \geqslant i$, i.e. for any two sufficiently large k and ℓ we have $g_{ij_k} = g_{ij_\ell}$. It is clear that the corresponding sequence f_{j_k} converges to an isometry $g : Y \to Y$.

Corollary. — If each space X_j *is homogeneous (i.e. the group of all isometries of* X_j *is transitive) then* Y *is also homogeneous.*

Proof. — Let us construct an isometry $Y \to Y$ which sends y to $y' \in Y$. Take an arbitrary sequence (x'_j) with $x'_j \in X_j$ which converges to y' and take a system of isometries $f_j : X_j \to X_j$ such that $f_j(x_j) = x'_j$. According to the lemma we can assume that (f_j) converges to an isometry $g : Y \to Y$, and $g(y) = y'$.

7. Limits of discrete groups

We start with a general construction. Let X be a metric space with metric " dist ". We denote by λX, $\lambda > 0$, the same X but with a new metric

$$(\mathrm{dist})^{\mathrm{new}} = \lambda(\mathrm{dist}).$$

Examples. — When $X = \mathbf{R}^n$ then all spaces λX, $\lambda \in (0, \infty)$, are isometric.

When X is a manifold of constant curvature K then λX has curvature $\lambda^{-2}K$.

When X is compact and has diameter D then λX has diameter λD.

Let X be an n-dimensional manifold of dimension n and let $x \in X$. If $\lim_{i\to\infty} \lambda_i = \infty$ then the sequence (X_i, x_i), $X_i = \lambda_i X$, $x_i = x$, converges to $\mathbf{R}^n$.

When X is a metric space such that each ε-ball, $0 \leq \varepsilon \leq 1$ centered at $x \in X$ can be covered by p balls of radius $\varepsilon/2$, where p is an arbitrary but fixed number, then there is a sequence $\lambda_i \to \infty$ such that $(\lambda_i X, x)$ converges to a proper space (Y, y) which can be regarded as a tangent cone of X at x. (The proof follows from the compactness criterion of § 6.)

We are now concerned with the limits of $\lambda_i X$ when $\lambda_i \to 0$. The following examples serve only as illustrations and the proofs (quite simple) are left to the reader.

When X is compact and $\lambda_i \to 0$, then the sequence $\lambda_i X$ converges to the one-point space.

When X is a complete manifold of non-negative Ricci curvature then for some sequence $\lambda_i \to 0$ there is a limit (Y, y) of $(\lambda_i X, x)$, but Y is not always a manifold.

Let X be the free Abelian group of rank two with two fixed generators. If we use in X the metric associated to the norm as in § 3 then the sequence $(\lambda_i X, x)$, $\lambda_i \to 0$, converges to the plane $\mathbf{R}^2$ with the following Minkowski metric

$$\mathrm{dist}((a, b), (a', b')) = |a - a'| + |b - b'|.$$

Take a non-Abelian nilpotent simply connected Lie group X of dimension 3 (notice, that there is only one such group) with a left invariant Riemannian metric. When $\lambda_i \to 0$, then the sequence $(\lambda_i X, x)$ converges to a space (Y, y) which is homeomorphic to X, but the limit metric in Y is not Riemannian. This metric can be described as follows. When we divide the Lie group by its center (which is isomorphic to $\mathbf{R}$) we get a Riemannian submersion $X \to \mathbf{R}^2$ with one dimensional fibers. Take two points $x_1, x_2 \in X$ and consider all smooth curves which are normal to the fibers and which join x_1 with x_2. Define $\mathrm{dist}(x_1, x_2)$ as the lower bound of the lengths of these curves. This is exactly the limit metric in Y (which is homeomorphic to X). Notice that for each $\lambda > 0$ the space λY is isometric to Y. Notice also that the Hausdorff dimension of Y is 4 rather than 3. (The definition of the Hausdorff dimension can be found in Ch. VII of [7].)

Let us return to our major topic.

Main construction. — Let Γ be a group of polynomial growth with a fixed system of generators and with the corresponding metric dist. Let $\{r_i\}$, $i=1, 2, \ldots$ be a sequence of i-regular numbers such that $\lim\limits_{i\to\infty} r_i=\infty$ (see the regularity lemma of § 3). We denote by $e\in\Gamma$ the identity element and we consider the sequence (Γ_i, e_i), $\Gamma_i=r_i^{-1}\Gamma$, $e_i=e$. It follows from the definition of the regularity (see inequalities (a) and (b)) that the family of r-balls in $(\Gamma_i)_{i=1,2,\ldots}$ is uniformly compact (each Γ_i is, obviously, a proper space) and, by the compactness criterion (see § 6), there is a convergent subsequence. To avoid double indices we assume that the sequence (Γ_i, e_i) itself converges to a space (Y, y).

Properties of the limit space Y

(1) Y is a locally compact space, because it is proper.

(2) Y is connected and locally connected. Moreover each ball in Y is connected and even path-connected.

Proof. — The connectivity property of Γ (see § 3) implies that for any two points $\alpha, \beta\in\Gamma_i$ there is a point $\gamma\in\Gamma_i$ such that

$$\operatorname{dist}_i(\alpha, \gamma)\leq\frac{1}{2}\operatorname{dist}_i(\alpha, \beta)+r_i^{-1},$$

$$\operatorname{dist}_i(\gamma, \beta)\leq\frac{1}{2}\operatorname{dist}_i(\alpha, \beta)+r_i^{-1}.$$

It follows that for any two points $y_1, y_2\in Y$ there is an $x\in Y$ such that

$$\operatorname{dist}(y_1, x)=\operatorname{dist}(x, y_2)=\frac{1}{2}\operatorname{dist}(y_1, y_2)$$

("dist_i" denotes the metric in Γ_i and "dist" is the limit metric in Y). This property not only implies the required connectivity of Y, but also shows that any two points $y_1, y_2\in Y$ can be joined by a curve with the length equal to $\operatorname{dist}(y_1, y_2)$.

(3) The group L of all isometries of Y is transitive on Y.

This follows from the corollary to the isometry lemma (§ 6).

(4) Y is finite dimensional.

Proof. — The regularity condition (see inequality (a) in § 3) implies that for $j\leq i$ each ball in Γ_i of radius $\frac{1}{2}$ can be covered by N_j balls of radius 2^{-j+1} where $N_j=2^{j(d+1)}$ and d denotes the growth of Γ. It follows that for *each* $j=1, 2, \ldots,$ one can cover every $\frac{1}{2}$-ball in Y by N_j balls of radius 2^{-j+1}. This shows that the Hausdorff dimension of Y is at most $d+1$ and hence (see chapter VII in [7]) Y is finite dimensional.

Main conclusion. — *The group* L *of all isometries of* Y *is a Lie group with finitely many components.*

Proof. — Use the theorem of Montgomery-Zippin (see § 5).

8. Proof of the main theorem

Take an arbitrary group Γ with a fixed finite system of generators and the associated metric. Define

$$D(\gamma, r) = \sup_{\beta} \operatorname{dist}(\gamma\beta, \beta)$$

where $\gamma \in \Gamma$, $r \in [0, \infty)$ and β runs over the r-ball in Γ centered at the identity.

Take now a subgroup $\Gamma' \subset \Gamma$ generated by $\gamma_1, \ldots, \gamma_k$ and set (with an abuse of notations)

$$D(\Gamma', r) = \sup_{1 \leq j \leq k} D(\gamma_j, r).$$

If the function $D(\Gamma', r)$, $r \in [0, \infty)$, *is bounded, then* Γ' *contains an Abelian subgroup of finite index.*

Proof. — If $D(\gamma, r)$ is bounded when $r \to \infty$, then the conjugacy class $\{\beta^{-1}\gamma\beta\}$, $\beta \in \Gamma$, of γ is finite and the centralizer of γ has finite index in Γ, q.e.d.

Suppose that $D(\Gamma', r)$ is unbounded but for a divergent sequence r_i the ratio $r_i^{-1} D(\Gamma', r_i)$ converges to zero.

Displacement lemma. — *For each* $\varepsilon > 0$ *there is a sequence* α_i, $i = 1, 2, \ldots$, *such that*

$$\lim_{i \to \infty} r_i^{-1} D(\alpha_i^{-1} \Gamma' \alpha_i, r_i) = \varepsilon,$$

where $D(\alpha^{-1}\Gamma'\alpha, r) = \sup_{1 \leq j \leq k} D(\alpha^{-1}\gamma_j\alpha, r)$.

Proof. — The connectivity property of Γ (see § 3) implies that for an arbitrary integer m one has

$$D(\Gamma', r+m) \leq D(\Gamma', r) + 2m.$$

On the other hand it is clear that for any $\alpha \in \Gamma$ and $r \geq 0$

$$D(\alpha^{-1}\Gamma'\alpha, r) \leq D(\Gamma', r + \|\alpha\|).$$

So we have

$$D(\alpha^{-1}\Gamma'\alpha, r) \leq D(\Gamma', r) + 2\|\alpha\|. \tag{*}$$

Since $D(\Gamma', r)$ is unbounded as a function of r, the function $D(\alpha^{-1}\Gamma'\alpha, r)$ is unbounded as a function of $\alpha \in \Gamma$, when r is kept fixed.

When r_i is sufficiently large our assumption $r_i^{-1}D(\Gamma', r_i) \to 0$ implies that

$$D(\Gamma', r_i) < \varepsilon r_i.$$

On the other hand for some $\mu \in \Gamma$ we have

$$D(\mu^{-1}\Gamma'\mu, r_i) > \varepsilon r_i.$$

Using (*) and the connectivity again we conclude that there is an $\alpha_i \in \Gamma$ such that

$$|D(\alpha_i^{-1}\Gamma'\alpha_i, r_i) - \varepsilon r_i| \leq 2,$$

q.e.d.

Let Γ_i be as in the main construction (see § 7). The group Γ acts isometrically on each Γ_i and if $\gamma_i \in \Gamma$ satisfy $r_i^{-1}\|\gamma_i\| \leq C < \infty$, $i = 1, 2, \ldots$, then the corresponding isometries $\widetilde{\gamma}_i : \Gamma_i \to \Gamma_i$ satisfy the condition of the isometry lemma (see § 6) and we can assume (using a subsequence when it is necessary) that these isometries converge to an isometry $\ell : Y \to Y$.

Let $\Gamma' \subset \Gamma$ be an arbitrary subgroup and $\gamma \in \Gamma'$. By taking $\gamma_i = \gamma$ we get an isometry $\ell = \ell_\gamma : Y \to Y$ and so we get a map $\Gamma' \to L$, where L is the isometry group of Y. (Because Γ' is countable we can assume that the convergence takes place for all $\gamma \in \Gamma'$.) This map is, obviously, a homomorphism. (To be precise we must fix a definite convergence $\Gamma_i \Rightarrow Y$ as in § 6 and only then our consideration becomes meaningful.)

The kernel of this homomorphism $\gamma \to \ell_\gamma$ consists of all γ in Γ' such that

$$\lim_{i\to\infty} r_i^{-1}D(\gamma, r_i) \to 0.$$

This limit exists because we have the convergence of the isometries $\gamma_i = \gamma : \Gamma_i \to \Gamma_i$ to ℓ.

We are ready to prove the main theorem. Take a subgroup $\Gamma' \subset \Gamma$ generated by $\gamma_1', \ldots, \gamma_j', \ldots, \gamma_k'$. According to the algebraic lemma (§ 4) and the main conclusion of § 7 we must only find a subgroup $\Delta \subset \Gamma'$ of finite index in Γ' such that either Δ is commutative or Δ has as many homomorphisms into L as is required by the algebraic lemma. (See the corollary to the algebraic lemma.) If the homomorphism $\gamma \to \ell_\gamma$ we constructed above has infinite image, the conditions of the algebraic lemma are satisfied and the proof is finished.

Suppose that the kernel $\Gamma'' \subset \Gamma'$ of the homomorphism $\gamma \to \ell_\gamma$ has finite index in Γ'. For the group Γ'' we have

$$\lim_{i\to\infty} r_i^{-1}(D(\gamma, r_i)) = 0, \quad \gamma \in \Gamma'',$$

and we can shorten the notations by assuming $\Gamma' = \Gamma''$.

If the function $D(\Gamma', r)$, $r \in [0, \infty)$, is bounded we have an Abelian $\Delta \subset \Gamma'$ and the proof is finished. Now we come to the last case.

The function $D(\Gamma', r)$ is unbounded but

$$\lim_{i\to\infty} r_i^{-1}D(\Gamma', r_i) = 0.$$

Fix an $\varepsilon > 0$ and construct $\alpha_i = \alpha_i(\varepsilon) \in \Gamma$ as in the displacement lemma. We obtain a new isometric action of Γ' on Γ_i as follows. First we send γ to $\gamma_i = \alpha_i^{-1}\gamma\alpha_i$ and then we act by γ_i on Γ_i as usual by the left translation. For each generator $\gamma'_1, \ldots, \gamma'_j, \ldots, \gamma'_k \in \Gamma'$ the translations $\gamma_{ji} = \alpha_i^{-1}\gamma'_j\alpha_i : \Gamma_i \to \Gamma_i$ satisfy

$$\limsup_{i \to \infty} \operatorname{dist}_i(\gamma_{ji}(e_i), e_i) \leq \varepsilon,$$

where dist_i denotes the distance in Γ_i, $e_i \in \Gamma_i$ denotes the reference point in Γ_i corresponding to $e \in \Gamma$ and $j = 1, \ldots, k$.

The isometry lemma (§ 6) allows us to assume that for each j the sequence γ_{ji}, $i \to \infty$, converges to an isometry $\ell_j : Y \to Y$ and so we obtained a homomorphism $A = A(\varepsilon) : \Gamma' \to L$.

Let us show that when $\varepsilon > 0$ is small then the image of $A(\varepsilon)$ is large. The properties of $\alpha_i = \alpha_i(\varepsilon)$ guarantee (see the displacement lemma) that for some γ'_j, say for γ'_1, one has

$$\lim_{i \to \infty} r_i^{-1} D(\alpha_i^{-1}\gamma'_1\alpha_i, r_i) = \varepsilon.$$

But

$$r_i^{-1} D(\alpha_i^{-1}\gamma'_1\alpha_i, r) = \sup_x \operatorname{dist}_i(\gamma_{1i}(x), x),$$

where x runs over the unit ball $B_{e_i}(1) \subset \Gamma_i$. So for the limit $\ell_1 = A(\gamma'_1)$ we also have

$$\sup_{y'} \operatorname{dist}(\ell_1(y'), y') = \varepsilon,$$

where y' runs over the ball $B_y(1) \subset Y$.

We apply now the localization lemma (§ 5) and finish the proof of the main theorem.

Final remarks. — Let Γ be a finitely generated almost nilpotent group *without torsion.* By a result of Auslander-Schenkman (see [1]) one can easily show that Γ contains a nilpotent subgroup $\Delta \subset \Gamma$ of index q with $\log\log q \leq 2^d$, where $d = \operatorname{growth}(\Gamma)$.

This allows us to apply the main theorem to infinitely generated groups.

If Γ has no torsion and each finitely generated subgroup $\Gamma' \subset \Gamma$ has growth at most d, then Γ contains a nilpotent subgroup of finite index.

Let us now give a more effective version of the Main Theorem.

For any positive integers d and k there exist positive integers R, N and q with the following properties. If a group Γ with a fixed system of generators satisfies the inequality $\# B(r) \leq kr^d$ for $r = 1, 2, \ldots, R$, then Γ contains a nilpotent subgroup Γ' of index at most q and whose degree of nilpotency is at most N.

Proof. — We start with a definition. Let Δ and Γ^i, $i = 1, 2, \ldots,$ be a system of groups endowed with a fixed system of k generators. We say that the sequence (Γ^i) converges to Δ if there is a sequence of balls $B_i = B(r_i) \subset \Delta$ with $r_i \to \infty$ for $i \to \infty$, and a sequence of bijective maps f of each ball B_i onto the r_i-ball in Γ^i centered at the identity such that $f(\delta_1\delta_2^{-1})=f(\delta_1)\,(f(\delta_2))^{-1}$ for any two elements δ_1, δ_2 in Δ satisfying $\|\delta_1\|+\|\delta_2\|\leq r_i$.

An arbitrary sequence of groups (Γ^i), each with k given generators, always has a convergent subsequence. If the limit group Δ has a nilpotent subgroup of index $\leq R$ and of degree of nilpotency $\leq N$, then such subgroups exist in the groups Γ^i for all sufficiently large i's.

Now, if we suppose that our theorem is false, we get a sequence of groups Γ^j, $j=1, \ldots,$ such that the balls in each group Γ^j satisfy the inequality $B(r)\leq kr^d$ for $r=1, \ldots, j$, but none of the groups Γ^j has a nilpotent subgroup of index $\leq j$ and of degree of nilpotency $\leq j$. Taking a convergent subsequence and passing to the limit we get a group Δ such that all balls in Δ satisfy $B(r)\leq kr^d$, $r=1, 2, \ldots,$ and such that Δ contains no nilpotent subgroup of finite index. This contradicts the Main Theorem.

Question. — What is the dependence of the numbers R, N and q on d and k? In particular, does there exist an *effective* estimate of these numbers in terms of d and k?

A geometric application. — Let $V=V^n$ be a complete Riemannian manifold such that the values of the Ricci tensor on all unit tangent vectors of V are bounded from below by $-(n-1)K$, $K\geq 0$. Let Γ be the group generated by some isometries $\gamma_1, \ldots, \gamma_k$ of V. Suppose that for a point $v\in V$ we have the following inequalities :

$$\operatorname{dist}(v, \gamma(v))\geq \varepsilon > 0, \quad \text{for all} \quad \gamma\in\Gamma\backslash e;$$
$$\operatorname{dist}(v, \gamma_i(v))\leq C\varepsilon, \quad \text{for} \quad C\geq 1 \quad \text{and} \quad i=1, \ldots, k.$$

In this case, the geometric growth theorem of Milnor (see [8]) reduces to the inequality

$$\# B(r)\leq 4^n C^n r^n \exp(2nCr\varepsilon\sqrt{K}), \qquad r=1, 2, \ldots,$$

where $B(r)$ denotes the r-ball in Γ relative to the word metric associated to $\{\gamma_i\}$. This gives the following

Geometric theorem. — *If* ε *is sufficiertly small compared to* n, C *and* K, *that is* $\varepsilon\leq\mu=\mu(n, C, K)>0$, *then the group* Γ *is almost nilpotent.*

Acknowledgements. — I want to thank H. Bass and J. Tits for their interest in this paper and their many helpful comments and suggestions. I am particularly thankful to J. Tits who kindly agreed to present his short proof of the theorem of Milnor-Wolf in an appendix.

REFERENCES

[1] L. Auslander, E. Schenkman, Free groups and Hirsch-Plotkin radicals, *Proc. A.M.S.*, **16** (1965), 784-788.

[2] J. Cheeger, D. Ebin, *Comparison theorems in Riemannian geometry*, North-Holland, 1975.

[3] V. Efremovič, The proximity geometry of Riemannian manifolds, *Uspehi Mat. Nauk*, **8** (1953), 189 (Russian).

[4] D. Epstein, M. Shub, Expanding endomorphisms of flat manifolds, *Topology*, **7** (1968), 13-141.
[5] M. Gromov, Synthetic geometry of Riemannian manifolds, *Proc. I.C.M. Helsinki*, 1978, vol. I, 415-419 (1979).
[6] M. Hirsch, Expanding maps and transformation groups, *Proc. Symp. Pure Math.*, XIV (1970), 125-133.
[7] W. Hurewicz, H. Wallman, *Dimension Theory*, Princeton U. Press, 1948.
[8] J. Milnor, A note on curvature and fundamental group, *J. Diff. Geom.*, **2** (1968), 1-7.
[9] D. Montgomery, L. Zippin, *Topological transformation groups*, Interscience, N. Y., 1955.
[10] M. Raghunathan, *Discrete subgroups of Lie groups*, Spiinger, Berlin, 1972.
[11] M. Shub, Expanding maps, *Proc. Symp. Pure Math.* XIV (1970), 273-277.
[12] A. Švarc, A volume invariant of coverings, *Dokl. Akad. Nauk SSSR*, **105** (1955), 32-34.
[13] J. Tits, Free subgroups in linear groups, *J. Algebra*, **20** (1972), 250-270.
[14] J. Wolf, Growth of finitely generated solvable groups and curvature of Riemannian manifolds, *J. Diff. Geom.*, **2** (1968), 421-446.

Institut des Hautes Études scientifiques,
35, route de Chartres,
91440 Bures-sur-Yvette.

Manuscrit reçu le 11 mars 1980.

APPENDIX

by Jacques TITS

This appendix to M. Gromov's paper " Groups of polynomial growth and expanding maps "—hereafter referred to as [G]—arose from an attempt to understand J. Wolf's article [10] (1). Our main purpose is to provide a short and self-contained account of the results on nilpotent and solvable groups which are needed for the proof of Gromov's main theorem and its converse. In A1 we show—in a slightly more general context—that a finitely generated nilpotent group has polynomial growth, and, as a by-product of the proof, we obtain a formula of H. Bass [*Proc. Lond. Math. Soc.* (3), 25 (1972), 603-614] giving the degree of that growth (proposition 2). In A2, we observe that one can, at little cost, make Gromov's proof of his main theorem independent of the special case of that theorem for solvable groups, thus providing a new proof for that special case as well (the statement, part of the result of Milnor-Wolf, that the growth of a finitely generated solvable group is either polynomial or exponential, is not included). The present text is made up of excerpts of a conference at the Séminaire Bourbaki (February 1981, exposé n° 572), slightly expanded and adapted.

A 1. Growth of filtered groups

Let Γ be a nilpotent group and let $(\Gamma_i)_{i \in \mathbf{N}^*}$ be a system of subgroups such that $\Gamma = \Gamma_1$, $[\Gamma_i, \Gamma_j] \subset \Gamma_{i+j}$ and $\Gamma_i = \{1\}$ for almost all i. By an *f-generating set* of Γ (relative to the filtration (Γ_i)), we mean a subset E of Γ such that $E_i = E \cap \Gamma_i$ generates Γ_i for all i. Set $E'_i = E - E_{i+1}$. We define the *f-length* of a word in the elements of $E \cup E^{-1}$ as the increasing sequence $(n_1, n_2, \ldots)$, where n_i is the length (in the usual sense) of the contribution of $E'_i \cup E'^{-1}_i$ to the word. An element of Γ is said to be *of f-length* $\leqslant (r_1, r_2, \ldots)$, with $r_i \in \mathbf{R}_+$, if it can be expressed as a word of f-length $(n_1, n_2, \ldots)$, with $n_i \leqslant r_i$. Assuming E finite, we denote by ${}_fc(r_1, r_2, \ldots)$ the number of such elements. If ${}_fc'$ is the function defined in the same way as ${}_fc$ but starting from another finite f-generating set, there exist nonvanishing constants $a, b \in \mathbf{R}_+$ such that

$$ {}_fc(ar_1, ar_2, \ldots) \leqslant {}_fc'(r_1, r_2, \ldots) \leqslant {}_fc(br_1, br_2, \ldots) $$

(1) Numbers between brackets refer to the bibliography of [G].

for all sequences (r_i). This legitimates the following definition: we say that the group Γ has *polynomial f-growth of degree* $\leqslant d$ if there is a constant $c \in \mathbf{R}_+$ such that, for all $r \in \mathbf{R}_+$, one has

$$_{\mathrm{f}}c(r, r^2, r^3, \ldots) \leqslant cr^d + 1.$$

Proposition 1. — *If d_i denotes the rank of the abelian group Γ_i/Γ_{i+1}, then the group Γ has polynomial f-growth of degree $\leqslant id_i$.*

We proceed by double induction: a descending induction on $a = \sup\{i \mid \Gamma = \Gamma_i\}$ and (given a) an ordinary induction on the minimum m of the cardinalities of all generating sets of Γ_a/Γ_{a+1}. We choose the f-generating set E so that $\operatorname{Card} \mathrm{E}'_a = m$, that $[\mathrm{E} \cup \mathrm{E}^{-1}, \mathrm{E} \cup \mathrm{E}^{-1}] \subset \mathrm{E}$, and that, if an element of E'_a has a nontrivial power inside Γ_{a+1}, then its power with the smallest strictly positive exponent having that property also belongs to E. Let us choose $y \in \mathrm{E}'_a$ and denote by Γ' the subgroup of Γ generated by $\mathrm{E}'_a - \{y\}$. We first prove the following assertion, by induction on q:

(*) if w is a word of f-length $(n_1, n_2, \ldots)$ in the elements of $\mathrm{E} \cup \mathrm{E}^{-1}$ and if $(y_1, y_2, \ldots, y_p)$ is the contribution of $\{y, y^{-1}\}$ to w (thus, $y_i = y$ or y^{-1} for all i, and $p \leqslant n_a$), then, for $0 \leqslant q \leqslant p$, there is a word w_q representing the same element of Γ as w, with the same contribution of $\{y, y^{-1}\}$, starting by $y_1 y_2 \ldots y_q$, and whose f-length $(n'_1, n'_2, \ldots)$ satisfies the relations

$$n'_i \leqslant n_i + qn_{i-a} + \binom{q}{2} n_{i-2a} + \cdots$$

The assertion is obvious for $q = 0$. The induction hypothesis provides a word w_{q-1} starting with $y_1 \ldots y_{q-1}$ and of f-length $(n''_1, n''_2, \ldots)$ satisfying

$$n''_i \leqslant n_i + (q-1) n_{i-a} + \binom{q-1}{2} n_{i-2a} + \cdots$$

Now move y_q to the left by successive commutation. Its jumping over an element of E_i introduces a new element belonging to E_{i+a} (possibly the identity). Therefore, we eventually get a word w_q starting with $y_1 \ldots y_q$ and of f-length $(n'_1, n'_2, \ldots)$ such that $n'_i \leqslant n''_i + n''_{i+a}$, hence (*).

Assuming $n_i \leqslant r^i$ (for some $r \in \mathbf{R}_+$), making $q = p \leqslant r^a$, majorizing $\binom{q}{j}$ by q^j $(\leqslant r^{aj})$ and denoting by e the smallest value of i such that $\Gamma_i = \{1\}$, we deduce from (*) that

(**) every element $g \in \Gamma$ of f-length $\leqslant (r, r^2, \ldots)$ can be written as $g = y^s g'$, where $|s| \leqslant r^a$ and g' is an element of Γ' of f-length $\leqslant (er - |s|, er^2 - |s|, \ldots)$.

If Γ/Γ' is infinite, the induction hypothesis (on (a, m)) applied to Γ' implies the existence of a constant $c' \in \mathbf{R}_+$ such that the number of possible choices for g' is majorized by

$$c' r^{\Sigma(id_i) - a} + 1.$$

Since there are less than $2r^a+1$ admissible values for s, the proposition ensues. If Γ/Γ' has order $t<\infty$, we can rewrite g as $y^{s_1}g'_1$, where $0\leqslant s_1<t$, $g'_1\in\Gamma'$ and the f-length of g'_1 is $\leqslant(er, er^2, \ldots)$. This time, the induction hypothesis enables us to majorize the number of possible elements g' by $c'r^{\Sigma id_i}+1$ (for a suitable c') and the proposition again follows, since s_1 takes only finitely many values.

Lemma 1. — *Let* Γ *be a finitely generated nilpotent group of class* e. *Denote by* Z *the last nontrivial term of its descending central series, by* E *a finite generating subset of* Γ *and by* z *an element of* Z. *Then, there exists a constant* $c\in\mathrm{R}_+$ *such that, for all* $n\in\mathrm{N}$, z^n *can be expressed as a word of length* $c\sqrt[e]{n}$ *in the elements of* $\mathrm{E}\cup\mathrm{E}^{-1}$.

The proof will be by induction on e. It is clearly sufficient to consider the case where z is the commutator $[x,y]$ of an element x of E and an element y of the penultimate nontrivial term of the descending central series. Let $n\in\mathrm{N}$, let n_1 be the smallest integer which is strictly larger than $\sqrt[e]{n}$ and let a_1, a_2 be the integers defined by

$$n=a_1n_1^{e-1}+a_2,\qquad a_1<n_1,\qquad a_2\leqslant n_1^{e-1}.$$

The induction hypothesis applied to Γ/Z and $y \bmod \mathrm{Z}$ implies the existence of a constant $c'\in\mathbf{R}_+$ (independent of n) and of two elements y_1 and y_2 of length $\leqslant c'n_1$, respectively congruent to $y^{n_1^{e-1}}$ and $y^{a_2} \bmod \mathrm{Z}$. Now, the assertion follows from the fact that the length of

$$z^n=[x^{a_1},y^{n_1^{e-1}}]\,.\,[x,y^{a_2}]=[x^{a_1},y_1]\,.\,[x,y_2]$$

is $\leqslant 2a_1+4c'n_1+2$.

Proposition 2 (Bass-Wolf). — *Let* Γ *be a finitely generated nilpotent group and let* d_i *be the rank of the i-th quotient* Γ_i/Γ_{i+1} *of its descending central series* (Γ_i). *Set* $d=\Sigma id_i$. *Choose arbitrarily a finite generating set* E *of* Γ *and, for* $r\in\mathbf{R}_+$, *let* $c(r)$ *represent the number of elements of* Γ *which can be expressed as a word of length* $\leqslant r$ *in the elements of* $\mathrm{E}\cup\mathrm{E}^{-1}$. *Then, there exist constants* $c_1, c_2\in\mathbf{R}_+$ *such that*

$$c_1r^d\leqslant c(r)\leqslant c_2r^d+1$$

for all r; *in particular,* Γ *has polynomial growth of degree* d.

The existence of c_2 is an immediate consequence of proposition 1 applied to the descending central series. We prove the existence of c_1 by induction on the class e of Γ. Set $d'=d-ed_e$. The induction hypothesis implies the existence of a constant $c'_1\in\mathbf{R}_+$ such that, for $r\in\mathbf{R}_+$, there exists a subset of Γ of cardinality $\geqslant c'_1r^{d'}$ consisting of elements of length $\leqslant\frac{r}{2}$, pairwise non congruent mod Γ_e. On the other hand, by lemma 1, there exists $c''_1\in\mathbf{R}_+$ such that Γ_e has more than $c''_1r^{ed_e}$ distinct elements of length $\leqslant\frac{r}{2}$. Hence the claim.

(N.B. — The above proof of the existence of c_1, including lemma 1, essentially follows the line of the proof given by J. Wolf [10], who also obtains an upper bound for $c(r)$, sufficient to establish the polynomial growth but coarser than the upper bound of proposition 2, which is due to H. Bass [*loc. cit.*]. Our proof of the existence of c_2, obtained without prior knowledge of Bass' result, is different from his, at least formally: the introduction of the f-growth makes it less computational and, as it seems, somewhat shorter.)

A 2. A special case of the theorem of Milnor-Wolf

In Gromov's proof of his main theorem, the only reference to the theorem of Milnor-Wolf (stated in the introduction of [G]) is at the end of [G], § 4 (proof of the " algebraic lemma "), and the reader will easily convince himself that only the following very special case of the theorem is needed there (take for L the group Δ' of [G]):

Lemma 2. — *Let* L *be a group. Suppose that there exists a homomorphism* ξ *of* L *onto* $\mathbf{Z}$ *whose kernel is finitely generated and almost nilpotent. Then* L *itself is almost nilpotent or its growth is exponential.*

The following argument is extracted from [10] (where it is however somewhat hidden). We first prove:

Lemma 3. — *Let* Λ *be a free abelian group and let* $\alpha : \Lambda \to \Lambda$ *be an automorphism.*

(i) *If* α *(extended to* $\Lambda\otimes\mathbf{C}$*) is semi-simple and if all its eigenvalues have absolute value one, then* α *has finite order.*

(ii) *If* α *has an eigenvalue of absolute value* $\geqslant 2$, *there exists* $\lambda\in\Lambda$ *such that the elements*

$$\varepsilon_0\lambda + \varepsilon_1\alpha(\lambda) + \varepsilon_2\alpha^2(\lambda) + \dots \qquad (\varepsilon_i = 0 \text{ or } 1, \text{ and } = 0 \text{ for almost all } i)$$

are pairwise distinct.

(i) It suffices to observe that the orbits of $\{\alpha^z \mid z\in Z\}$ in $\Lambda\otimes\mathbf{C}$ have compact closures, from which follows that the orbits of that same group in Λ are finite.

(ii) If $\beta : \Lambda\to\mathbf{C}$ is a linear form such that $\beta\circ\alpha = \rho\beta$, with $|\rho|\geqslant 2$, then the assertion is true for every $\lambda\in\Lambda$ such that $\beta(\lambda)\neq 0$; indeed, one has

$$\beta\Big(\sum_{i=0}^{\infty}\varepsilon_i\alpha^i(\lambda)\Big) = \Big(\sum_{i=0}^{\infty}\varepsilon_i\rho^i\Big)\cdot\beta(\lambda)$$

and, in view of the inequality satisfied by ρ, the numbers $\sum_{i=0}^{\infty}\varepsilon_i\rho^i$ are pairwise distinct.

Proof of lemma 2. — Let z be an element of L such that $\xi(z) = 1$, and let L_1' be the greatest nilpotent normal subgroup of L_1; upon substituting L_1' for L_1 and the group generated by L_1' and z for L, we may—and shall—assume that L_1 is nilpotent. Let

$(L_i)_{1 \leqslant i \leqslant e}$ be a central series of L_1 normalized by z, and let α_i denote the automorphism of L_i/L_{i+1} induced by int z. We suppose the series (L_i) chosen in such a way that, whenever L_i/L_{i+1} is infinite, it is a free abelian group of which α_i is a semi-simple automorphism : any central series normalized by z clearly has a refinement satisfying that condition. If all α_i have finite order, there exists an integer $s \geqslant 1$ such that z^s centralizes each quotient L_i/L_{i+1}; then, the group generated by L_1 and z^s has finite index in L and is nilpotent, and the lemma is proved. Let us therefore assume that, for some j, α_j has infinite order. By lemma 3 (i), there exists an integer t such that α_i^t has an eigenvalue of absolute value $\geqslant 2$ and, by lemma 3 (ii), there exists $x \in L_j$ such that the elements

$$x^{\varepsilon_0}.(z^t x z^{-t})^{\varepsilon_1}.(z^{2t} x z^{-2t})^{\varepsilon_2}. \ldots \qquad (\varepsilon_i = 0 \text{ or } 1 \text{ and } = 0 \text{ for almost all } i)$$

are pairwise distinct. This implies that L has exponential growth, and the proof is complete.

(Note the similarity of the last argument with the proof of lemma (b) of [G], § 4.)

Collège de France,
11, place Marcelin-Berthelot,
75231 Paris Cedex 05.

Manuscrit reçu le 7 février 1981.