THE BIEBERBACH CASE IN GROMOV'S ALMOST FLAT MANIFOLD THEOREM

Peter Buser* and Hermann Karcher**

1. Introduction (and abstract)

In 1976 M. Gromov has shown that every compact Riemannian manifold with normalized diameter whose sectional curvature is sufficiently close to zero is covered by a compact nilmanifold (= quotient of a nilpotent Lie group). [3] . This theorem, known as the almost flat manifold theorem has soon become famous not only because of its content but also because of the many unconventional methods Gromov has introduced to Riemannian geometry to get the proof.

The aim of the present notes is to explain how the ideas from Gromov's proof of the almost flat manifold theorem can be specialized to give a proof of the Bieberbach theorem. Since this specialization is much more accessible than the almost flat manifold theorem, one can very nicely explain some of Gromov's ideas in this context. It is also interesting to compare this new proof with older proofs of Bieberbach's theorem.

2. The Bieberbach theorem

We fix some notation. A euclidean motion $\alpha: \mathbb{R}^n \to \mathbb{R}^n$ is given by $\alpha x = Ax + a$, $A \in O(n)$, $a \in \mathbb{R}^n$. We call $A = r(\alpha)$ the rotational part and $a = t(\alpha)$ the translational part of the motion. To each rotation A corresponds an orthogonal decomposition

$$\mathbb{R}^n = E_0 \oplus E_1 \oplus \ldots \oplus E_k$$

such that A restricted to E_i is a rotation through the angle θ_i ; in the orientation reversing case E_k is eigenspace of A for the eigenvalue -1, we include this in the case $\theta_k = \pi$. Then

$$\measuredangle\ (x, Ax) = \theta_i \quad \text{for all} \quad x \in E_i \ .$$

These so called main rotational angles are arranged in increasing order:

$$0 = \theta_0 < \theta_1 < \ldots < \theta_k \ .$$

---

* Supported by the Swiss National Science Foundation

** written under the programm "Sonderforschungsbereich Theoretische Mathematik", Bonn University.

The dimension of E_0 may be zero. The main rotational angles give rise to the following biinvariant distance function (Finsler metric) in the orthogonal group:

$$\|A\| := \theta_k = \max_{\|x\|=1} \measuredangle (x,Ax) \ , \quad d(A,B) = \|AB^{-1}\| \ .$$

From this metric we derive a distance function for the entire group of motions by

2.1 $\|\alpha\| := \max\{\|r(\alpha)\|, \text{const}\cdot|t(\alpha)|\} \ , \quad d(\alpha,\beta) = \|\alpha\beta^{-1}\|$.

There is a degree of freedom in the choice of the constant. It will be fixed later according to the momentary needs.

A crystallographic group is a discrete group of euclidean motions with compact fundamental domain.

2.2 Theorem (Bieberbach) [1]. Let G be a crystallographic group in $\mathbb{R}^n$.

(i) Each $\alpha \in G$ has either $A = \mathrm{id}$ or $d(A,\mathrm{id}) \geq \frac{1}{2}$.

(ii) The group Γ of pure translations in G is a normal subgroup of finite index. G/Γ has order $\leq 2\cdot(4\pi)^{\dim SO(n)}$.

(iii) In addition to (i): If $\alpha \in G$, $r(\alpha) \in SO(n)$ and $0<\theta_1<\ldots<\theta_k$

are the main rotational angles of $A = r(\alpha)$ then

$$\theta_\varkappa \geq \tfrac{1}{2}(4\pi)^{\varkappa-k} \ , \quad \varkappa = 1,\ldots,k \ .$$

The original version of Bieberbach's theorem consists only of the statement that G/Γ has finite index. It was used by Bieberbach to solve the 18^{th} Hilbert problem:

2.3 Corollary (Bieberbach) [1]. For each n there exist only finitely many isomorphism classes of crystallographic groups in $\mathbb{R}^n$.

In the formulation 2.2 of the Bieberbach theorem the most important part is 2.2 (i): The translations in G are those motions which have a rotational part smaller than $\frac{1}{2}$. This characterization is Gromov's discovery; the proof depends as all other proofs of the Bieberbach theorem on commutator estimates, but Gromov combines these with the a priori bound 2.5 on the length of nontrivial commutators. The further statements 2.2 (ii) and 2.2 (iii) follow rather easily in 2.9 and 2.10. In particular the bound 2.2 (ii) on the order of G/Γ implies that there are only finitely many possibilities for the group of rotational parts; this is the main part of the finiteness theorem 2.3. The remaining part is a group cohomology argument dealing with nonisomorphic extensions of $\mathbb{Z}^n$ by finite groups.

Proof of the Bieberbach theorem (Following Gromov).
We introduce the finite subset

$$G_\rho^\varepsilon = \{\alpha \in G \,|\, \|r(\alpha)\| < \varepsilon \ , \ |t(\alpha)| < \rho\}$$

of G, where $0 < \varepsilon \leq \frac{1}{2}$ and $\rho > 0$ (large), and denote by $\langle G_\rho^\varepsilon \rangle$ the smallest subgroup of G which contains G_ρ^ε. The working tools will be lemmas 3,4,5 in section 1
The proof is divided into two parts:
2.4 For any $R > 0$ we can find some $\rho \geq R$ such that for all $x \in \mathbb{R}^n$ with $|x| \leq \frac{3}{4}\rho$ there is $\alpha \in G_\rho^\varepsilon$ with $|t(\alpha) - x| \leq \rho/4$.

2.5 $\langle G_\rho^\varepsilon \rangle$ is d-nilpotent with $d \leq 3^{n^2}$.

By d-nilpotent we mean that all d-fold commutators $[\ldots[\beta_1,\beta_2],\ldots,\beta_d]$ are trivial ($[\alpha,\beta] = \alpha\beta\alpha^{-1}\beta^{-1}$).

Hence instead of showing that the pure translations provide a vector space basis of $\mathbb{R}^n$, it is first shown that (the translational parts of) the almost translations ($=G_\rho^\varepsilon$) do, and instead of showing commutativity one starts with nilpotency. The reason why this procedure carries over to more general situations is that both, 2.4 and 2.5 are proved by means of estimates rather than by equations.

2.4 and 2.5 together suffice to show 2.2 (i) and in particular that G_ρ^ε is in fact a set of pure translations.

Assume there is $\gamma \in G$ with $r(\gamma) = C$, $t(\gamma) = c$ such that $\|C\| = \theta \in (0,\frac{1}{2})$. Then decompose $\mathbb{R}^n$ into $E \oplus E^\perp$ where E is an invariant plane of maximal rotation and let $x = x^E + x^\perp$ be the corresponding decomposition of vectors in $\mathbb{R}^n$. Put $\varepsilon = \frac{1}{10}(\sin\frac{\theta}{2})^d$ and choose $\rho \geq 2|c|$ in 2.4 so that one can find $\alpha \in G$ with $\|A\| \leq \varepsilon$ and $|a - x| \leq \frac{\rho}{4}$ for $x \in E$, $|x| = \frac{3}{4}\rho$; consequently $|c| \leq |a| \leq 2|a^E|$. Consider the iterated commutators

$$\alpha_k = [\ldots[\alpha,\gamma],\ldots,\gamma] \quad \text{(k-fold)} \ , \quad k = 1,\ldots,d \ .$$

From 4.3 we have the estimate

$$\|A_{k+1}\| = \|[A_k,C]\| \leq 2\|A_k\|\cdot\|C\| \leq \|A_k\| \leq \cdots < \|A\| \leq \varepsilon$$

which we use in the decomposition

$$a_{k+1} = (\mathrm{id} - C)a_k^E + (\mathrm{id} - C)a_k^{\perp} + (\mathrm{id} - [A_k,C])Ca_k + A_kC(\mathrm{id} - \bar{A}_k^{1})\bar{C}^{1}c$$

to obtain first inductively

$$|a_{k+1}| \leq (\|C\| + \|A_{k+1}\|)\cdot|a_k| + \|A_k\|\cdot|c| \leq |a| \quad .$$

Then, since $E,E^{\perp}$ are invariant under C we can use the last two estimates to obtain

$$\begin{aligned} |a_d^E| &\geq |(\mathrm{id} - C)a_{d-1}^E| - \|[A_{d-1},C]\|\cdot 1a_{d-1}\| - \|A_{d-1}\|\cdot|c| \\ &\geq 2|a_{d-1}^E|\sin\frac{\theta}{2} - 2\epsilon|a| \\ &\geq (2\sin\frac{\theta}{2})^d\cdot|a^E| - 2\epsilon|a|\cdot\sum_{k=0}^{d-1}(2\sin\frac{\theta}{2})^k \\ &\geq |a^E|(\sin\frac{\theta}{2})^d \quad . \end{aligned}$$

Now $|a_d| > 0$, which contradicts 2.5 and proves 2.2 (i) .

2.6 A pigeon hole argument (Proof of 2.4).

Put $\rho_i = (R+r)\cdot 10^{i+1}$, $i = 0,\dots,2\cdot\mathrm{int}(2\pi/\epsilon)^{\dim SO(n)} = N(\epsilon)$, where r is the diameter of the fundamental domain of G .(This is the only point in the proof where compactness of $\mathbb{R}^n/G$ is used). Define $\mathcal{U}_i = \{\alpha\in G|\ |t(\alpha)| < \rho_i\}$. For each $x\in\mathbb{R}^n$, $|x| \leq \frac{3}{4}\rho_i$ choose $\alpha_i\in G$ with $a_i = t(\alpha_i)$ next to x ; then $|a_i - x| \leq r$ and $r+\rho_{i-1} < \frac{1}{4}\rho_i$ imply for all $\beta\in\mathcal{U}_{i-1}$ that $|t(\alpha_i\circ\bar{\beta}^1) - x| < \frac{1}{4}\rho_i$ and $|t(\alpha_i\circ\bar{\beta}^1)| < \rho_i$. Therefore if 2.4 were false for all the ρ_i we would have for each i some $\alpha_i\in G$ with $\|r(\alpha_i\circ\bar{\beta}^1)\| > \epsilon$ for all $\beta\in\mathcal{U}_{i-1}$. In particular we get $N(\epsilon)+1$ elements $r(\alpha_i)\in O(n)$ with pairwise distance $>\epsilon$, contradicting lemma 4.4.

2.7 The short basis (Proof of 2.5)

We fix the constant in 2.1 to be ϵ/ρ . Then $\|\alpha\| < \epsilon$ for all $\alpha\in G_\rho^\epsilon$, and 4.3 implies

$$\|[\alpha,\beta]\| < \min\{\|\alpha\|,\|\beta\|\} \qquad (\alpha,\beta\in G_\rho^\epsilon) \quad .$$

A <u>short basis</u> $\{\alpha_1,\dots,\alpha_d\}$ is defined inductively by choosing a nontrivial

$\alpha_1 \in G_\rho^\epsilon$ with minimal $\|\alpha_1\|$,

$\alpha_{i+1} \in G_\rho^\epsilon - \langle\{\alpha_1,\dots,\alpha_i\}\rangle$ with minimal $\|\alpha_{i+1}\|$.

($\langle\{\alpha_1,\dots,\alpha_i\}\rangle$ is the smallest subgroup of G containing $\{\alpha_1,\dots,\alpha_i\}$.)

The basis is finite since G_ρ^ϵ is finite. The important point is that d has an upper bound which is independent of ρ and ϵ: If we could find $i<j\leq d$ such that $d(\alpha_i,\alpha_j)<\|\alpha_j\|$, then $\|\alpha_j\bar{\alpha}_i^{-1}\|<\|\alpha_j\|<\epsilon$ hence $\alpha_j\bar{\alpha}_i^{-1}\in G_\rho^\epsilon$ and also $\alpha_j\bar{\alpha}_i^{-1}\in\langle\{\alpha_1,\dots,\alpha_{j-1}\}\rangle$ since $\|\alpha_j\|$ is minimal in the complement. Now $\alpha_j=(\alpha_j\bar{\alpha}_i^{-1})\circ\alpha_i\in\langle\{\alpha_1,\dots,\alpha_{j-1}\}\rangle$ is impossible. Hence the elements of a short basis satisfy the pairwise distance condition of 4.5 so that $d\leq 3^{n+\dim SO(n)}$.

This d is also a bound on the length of nonvanishing commutators since $\|[\alpha_i,\alpha_j]\|<\min(\|\alpha_i\|,\|\alpha_j\|)$ implies first

$$(2.8)\qquad [\alpha_i,\alpha_j]\in\langle\{\alpha_1,\dots,\alpha_{i-1}\}\rangle\qquad (i<j)\ .$$

Then use induction on the wordlength based on the formulas $[\alpha\,\beta,\gamma]=[\beta,\gamma]\cdot[[\gamma,\beta],\alpha]\cdot[\alpha,\gamma]$ and $[\bar{\alpha}^{-1},\gamma]=[\bar{\alpha}^{-1},[\gamma,\alpha]]\cdot[\gamma,\alpha]$ and an induction on i to show that $\langle\{\alpha_1,\dots,\alpha_i\}\rangle$ is i-nilpotent.

2.9 Proof of 2.2 (ii)

The translations in G - clearly a normal subgroup, have been described as the set of all α with $\|A\|<\frac{1}{2}$. From 2.4 $\mathbb{R}^n/\Gamma$ is compact hence G/Γ is finite. The homomorphism $r\colon G\to O(n)$ induces an isomorphism between G/Γ and a discrete subgroup of $O(n)$ whose elements satisfy the pairwise distance condition of 4.4 with $\epsilon=\frac{1}{2}$. Therefore G/Γ has order $\leq N(\frac{1}{2})$ (2.6.)

2.10 Proof of 2.2 (iii)

Consider pairwise orthogonal 2-planes $R_1\subseteq E_1,\dots,R_k\subseteq E_k$ through the origin such that A as restricted to R_i is a rotation by θ_i. Let S_i^1 be the unit circle in R_i. Fix $\varkappa\leq k-1$. Then A acts isometrically on the flat torus $T_\varkappa=S^1_{\varkappa+1}\times\dots\times S^1_k$ not only with respect to the Riemannian but also with respect to the Finsler distance $d(x,y)=\max\{\measuredangle(x_i,y_i)\mid i=\varkappa+1,\dots,k\}$, ($x_i$ = orthogonal projection of x to R_i). The function $d(x,Ax)$ is constant on $T_\varkappa$. Since each torus has the same volume as the Finsler ball of radius π in its tangent space and since points of pairwise distance $\frac{1}{2}$ give disjoint balls of radius $\frac{1}{4}$ we have the volume ratio $m_\varkappa=\operatorname{int}(4\pi)^{k-\varkappa}$ as a bound on the number of such points. It follows that for some power A^m, $0<m\leq m_\varkappa$, we have $d(x,A^mx)<\frac{1}{2}$ for all $x\in T_\varkappa$, which implies $|\measuredangle(x,A^mx)|<\frac{1}{2}$ for all $x\in E_{\varkappa+1}\oplus\dots\oplus E_k$. Therefore, if we had

$0<\theta_\varkappa<\frac{1}{2}(4\pi)^{\varkappa-k}$ it would follow from $0<\theta_1<\ldots<\theta_\varkappa<\frac{1}{2}m_\varkappa$ that $\sphericalangle(x,A^m x)<\frac{1}{2}$ for all $x\in R^n$, i.e. $\|A^m\|<\frac{1}{2}$ but $A^m \neq \mathrm{id}$, a contradiction to 2.2 (i).

3. Earlier proofs

In this section we sketch Bieberbach's original proof [1] and the one given in Wolf's book [4]. Both use commutator estimates though in different form. To simplify the description we use 4.3.

3.1 The structure of Bieberbach's approach consists of the following observations (X p. 317 and XII p. 328, Math. Ann. 70 (1911)).

(i) All main rotation angles occurring in G are rational $(\in \pi\mathbb{Q})$.

(ii) An infinite discrete group of motions has elements without fixed points.

The two propositions are proved independently. From (i) it follows that each infinite subgroup of G contains translations, and by a not too complicated induction argument Bieberbach then concludes:

(iii) If all translations of G are contained in a subspace E of R^n, then also all translational parts are contained in E.

At this point the proof is complete: Since G has compact fundamental domain E must be R^n. While the proof of (ii), based on the commutator estimate (Hilfssatz on p. 328) makes no trouble we like to comment on (i), which is the heart of Bieberbach's arguments. The way of proving (i) is by showing that irrational angles would imply the existence of <u>infinitesimal sequences</u>, i.e. sequences in G which do not contain the identity but which converge to it. First $\alpha\in G$ is chosen with the maximal possible number of irrational angles $\theta_1,\ldots,\theta_\lambda$. By taking powers it is achieved that all other angles are zero. By a change of origin there is a 2λ-dimensional invariant subspace $E\subseteq R^n$ such that $t(\alpha)\in E^\perp$ and $r(\alpha)|E^\perp = \mathrm{id}$. Since G has compact fundamental domain there is $\gamma\in G$ with $t(\gamma)\notin E^\perp$. This γ does not commute with any power $\alpha^m (m\neq 0)$. Certainly one can construct a set of powers of α such that the rotational parts form an infinitesimal sequence. The problem is to have the translational parts converge also. This is achieved together with γ in the following way.

By Minkowski's theorem on simultaneous rational approximation there exist for all $j=1,2,\ldots$ integers $x_1(j),\ldots,x_\lambda(j)$ and $n(j)$ such that simultaneously

$$\left|2\pi\frac{x_\iota(j)}{n(j)}-\theta_\iota\right|\le\frac{1}{j\cdot n(j)}\qquad \iota=1,\ldots,\lambda\ .$$

Now for each fixed m (which serves as parameter) Bieberbach considers the sequence of m-fold commutators

$$\gamma_m^{(i)} = [\ldots[\gamma,\alpha^{n(j)}],\ldots,\alpha^{n(j)}] \ , \qquad j = 1,2,\ldots$$

Due to Minkowski's inequality the powers $\alpha^{n(j)}$ have small rotational angles, and from this by an involved calculation the following orders of magnitude are shown

$$\|r(\gamma_m^{(j)})\| = O(j^{1-m}) \ , \quad |t(\gamma_m^{(j)})| = O(j^{\lambda+2-m}), \quad m \geq 2 \ .$$

Therefore the proof of (i) is complete if for $m = \lambda + 3$ the sequence $\{\gamma_m^{(j)}\}_{j=1}^{\infty}$ does not contain the identity. Now by the particular choice of α and γ one finds these sequences free from the identity for $m = 2$ and 3. Yet there may be a minimal $m \geq 4$ such that $\{\gamma_m^{(j)}\}_{j=1}^{\infty}$ is not infinitesimal. If this happens, various cases must be considered. If $m = 4$ and

$$[\gamma_2^{(j)},\gamma_3^{(j)}] \neq 1 \ (j \geq j_0) \ , \text{ then } \{[\gamma_3^{(j)},\gamma_1^{(j)}\gamma_2^{(j)}]\}_{j=1}^{\infty}$$

is infinitesimal instead. If $m = 4$ and $[\gamma_2^{(j)},\gamma_3^{(j)}] = 1$ then $\{\gamma_3^{(j)}\}$ is infinitesimal. For $m = 5$ one can take $\{[\gamma_3^{(j)},\gamma_4^{(j)}\gamma_1^{(j)}]\}$. And finally for $m \geq 6$ the sequence looked for is $\{\gamma_{m-1}^{(j)} \cdot (\gamma_{m-2}^{(j)})^{-1}\}$.

It is interesting, how a little more information about γ simplifies the proof of (i). From the pigeon hole argument 2.6 one can choose γ such that in addition $\|r(\gamma)\| < \frac{1}{2}$. Then $\gamma_m^{(j)} \neq 1$ for all m, $(j \geq 3)$; for otherwise by the lemma below, $\gamma_2^{(j)}$ and a fortiori each further $\gamma_m^{(j)}$ is a translation which due to the choice of α and γ has always a nonzero component in E, a contradiction. Hence $\{\gamma_m^{(j)}\}_{j=1}^{\infty}$ is always infinitesimal, in particular for $m = \lambda + 3$. However there is a still simpler argument: Look at the series $\{\gamma_m^{(j)}\}_{m=1}^{\infty}$ instead of $\{\gamma_m^{(j)}\}_{j=1}^{\infty}$. As mentioned, it does not contain the identity. By the commutator estimate (c.f. 2.7) it converges to the identity. Thus it is infinitesimal.

3.2 Bieberbach's proof succeeded by extracting translations from G by means of powers (based on the non-existence of irrational angles). The logical structure od Gromov's proof is different. He first <u>defines</u> a subgroup $(\langle G_\rho^{\frac{1}{2}}\rangle)$ of finite index (by the pigeon hole argument) in G and then proves that the subgroup is already a group of translations (by the short basis trick). Wolf's proof also starts by defining a suitable normal subgroup $G^* = r^{-1}(T) \subseteq G$ where $T \subseteq SO(n)$ is the identity component of the closure of $r(G)$. G/G^* is almost immediately finite: Since T is closed and $SO(n)$ is compact, only finitely many different sets $r(\gamma)\cdot T$ occur as γ runs through G . The task is again to show that G^* is purely translational: First one observes ([4] p. 100)

<u>Lemma</u> If $A,B \in SO(n)$, $\|A\|$, $\|B\| < \pi/2$ then

$$[[A,B],B] = 1 \text{ implies } [A,B] = 1 .$$

(This lemma is not used by Gromov since due to occurring homotopy errors there is no analogue for non flat situations). Together with the commutator estimate (c.f. 2.7) one finds T toral in $SO(n)$. Hence the subspace $W = \{x \in \mathbb{R}^n | T(x) = x\}$ is characterized as the fixed point set of a single rotation $r(\gamma_0)$, $\gamma_0 \in G^*$, and by a change of origin one may assume $t(\gamma_0) \in W$. Since T is abelian, one checks that $t(\gamma) \in W$ for all further $\gamma \in G^*$ also. Since $\mathbb{R}^n/G^*$ is compact this is possible only if $W = \mathbb{R}^n$. Hence $T = r(G^*) = \{id\}$, i.e. G^* is a set of translations.

4. The group of motions.

The lemmas of this section will be proved with differential geometric techniques. We recall the following facts:

4.1 The orthogonal group $O(n)$ is a Lie group with identity component $SO(n)$. Its Lie algebra $so(n)$ is the space of skewsymmetric matrices $X,Y,\ldots$ and is canonically identified with the space of left invariant vector fields, using that the brackets of left invariant vector fields are left invariant.

(1) $$\operatorname{ad}X\,(Y) := [X,Y] = XY - YX .$$

The exponential map $\exp: so(n) \to SO(n)$, $\exp X = id + \sum_{k=1}^{\infty} \frac{X^k}{k!}$ relates to ad and conjungation K_A: $B \to ABA^{-1}$ as follows:

(2) $$\exp Y \cdot \exp X \cdot \exp(-Y) = \exp(dK_{\exp Y} X)$$

(3) $$\operatorname{Exp}(t\operatorname{ad}Y) := id + \sum_{k-1}^{\infty} \frac{1}{k!} (t\operatorname{ad}Y)^k = (d\,K_{\exp tY})_{id} .$$

Denote by D^L the left invariant connection for which left invariant vector fields are parallel, then

$$D_X Y := D^L_X Y + \frac{1}{2}\ [X,Y]$$

defines a torsion free biinvariant connection with parallel curvature tensorfield R:

$$R(X,Y)Z = \frac{1}{4}[Z,[X,Y]] \quad .$$

Obviously

$$R(J): = D^L_{\dot{c}}\, D^L_{\dot{c}}\, J + D^L_{\dot{c}}[\dot{c},J] = D_{\dot{c}}\, D_{\dot{c}}\, J + R(J,\dot{c})\dot{c}$$

for vector fields J along geodesics $t \to c(t) = \exp tX$. The solutions of $R(J) = 0$ are the Jacobifields and are either obtained as

$$(4) \qquad J(t) = dL_{c(t)} \cdot k^L(t) \ , \quad k^L\colon \mathbb{R} \to so(n) \ , \quad \ddot{k}^L + [X,\dot{k}^L] = 0 \ ,$$

$(L_A(B): = A \cdot B)$ where $dL_{c(t)}$ is parallel translation along c with respect to the connection D^L, or as

$$(5) \qquad J(t) = P_t \cdot k(t),\ k\colon\ \mathbb{R} \to so(n),\quad \ddot{k} - \frac{1}{4}\,(\mathrm{ad}X)^2 k = 0$$

where

$$(6) \qquad P_t: = dL_{c(t)} \circ \operatorname{Exp}(-\frac{t}{2}\mathrm{ad}X)$$

is parallel translation along c with respect to the connection D. The differential $d\exp$ can be described with Jacobi fields as follows

$$(7) \qquad (d\exp)_{tX} Y = \frac{1}{t}\, J(t) \quad \text{if} \quad J(0) = 0,\ \frac{D}{dt}\, J(0) \quad (= \frac{D^L}{dt}\, J(0)) = Y \quad .$$

4.2 If for $S \in so(n)$ we put

$$\|S\| = \max\{|Sv|;\ v \in \mathbb{R}^n\ ,\ |v| = 1\} \quad ,$$

then from (1)

$$(8) \qquad \|[S,T]\| \leq 2\|S\| \cdot \|T\| \ .$$

By left translating this norm to all other tangent spaces we obtain a Finsler metric for $O(n)$ whose distance function

$$d(A,B) = \max\{\ \sphericalangle\,(v,Av) \mid \ v \in \mathbb{R}^n\ ,\ |v| = 1\}\ .$$

has already been introduced in section 2. The diameter of $SO(n)$ and the injectivity radius of exp with respect to the Finsler metric are π. Since the distance function is biinvariant, $(dK_A)_{id}\colon so(n) \to so(n)$ is a norm isometry and it follows from (3)

$$(9) \qquad \|\mathrm{Exp}\,(\mathrm{ad}Y)\cdot X\| = \|X\| \;, \quad X,Y \in so(n) \quad .$$

Hence both parallel translations $dL_{c(t)}$ (by definition) and P_t (by (6) and (9)) are norm preserving.

If $J(t) = dL_{c(t)}\, k^L(t)$ is a Jacobifield (4) , then k^L satisfies $\dot{k}^L(t) = \mathrm{Exp}\,(t\mathrm{ad}X)\cdot\dot{k}(0)$, $\|\dot{k}^L(t)\| = \|\dot{k}(0)\|$ (9) and therefore

$$\|J(t)\| = \|k^L(t)\| \le t\| \frac{D^L}{dt} J(0)\| = t\|Y\| \quad .$$

$$(10) \quad \|(d\exp)_{tX} Y\| \le \|Y\| \quad ,$$

i.e. exp does not increase lengths in the Finsler metric.

4.3 Lemma

$$d([A,B],id) \le 2d(A,id)\cdot d(B,id)\,, \quad A,B \in SO(n) \quad .$$

Proof Let $A = \exp X$, $B = \exp Y$ and connect A with BAB^{-1} by the curve (c.f. (2) and (3))

$$t \to \gamma\,(t) = \exp\,(\mathrm{Exp}\,(t\,\mathrm{ad}\,Y)\cdot X)\,, \qquad t \in [0,1] \quad .$$

From the biinvariance of the Finsler metric

$$d([A,B],id) = d(A,BAB^{-1}) \le \int_0^1 \|\dot\gamma(t)\|dt \quad .$$

Since exp does not increase lengths (10)

$$\|\dot\gamma(t)\| \le \| \frac{d}{dt}\,(\mathrm{Exp}(t\,\mathrm{ad}\,Y)\cdot X\,\| \overset{(3)}{=} \|\mathrm{Exp}(t\,\mathrm{ad}\,Y)\cdot[Y,X]\|$$

$$\overset{(9)}{=} \|[X,Y]\| \overset{(8)}{\le} 2\|X\|\cdot\|Y\| = 2\,d(A,id)\cdot d(B,id) \quad .$$

4.4 Lemma

For $\varepsilon > 0$ there exist at most $N(\varepsilon) = 2\,\mathrm{int}\,(2\pi/\varepsilon)^{\dim SO(n)}$ rotations in $O(n)$ with pairwise distances $\ge \varepsilon$.

Proof. It suffices to prove $N(\varepsilon)/2$ as upper bound on $SO(n)$. Since metric balls $B_{\varepsilon/2}$ of radius $\varepsilon/2$ around the considered elements have pairwise disjoint interior and equal volumes, it follows from $B_\pi = SO(n)$ that

$$\operatorname{vol} B_\pi / \operatorname{vol} B_{\varepsilon/2}$$

is an upper bound. To get it explicitly we estimate $\det(d\exp)_{tX}$ in the standard Riemannian metric (which provides the volume function on $SO(n)$); the Levi-Civita connection is D. Norms with respect to the Riemannian metric are denoted by $|\cdot|$. We use an orthonormal Basis $\{Y_1,\dots,Y_m\} \subseteq so(n)$ of eigenvectors with eigenvalues $\lambda_1^2,\dots,\lambda_m^2$ of the nonnegative symmetric operator $-(\operatorname{ad}X)^2$, $m = \dim so(n)$. If J is the Jacobifield (7) for $Y = Y_i$, then in (5) obviously $k(t) = \frac{2}{\lambda_i} \sin \frac{t}{2} \lambda_i$ is a solution. Therefore since the Levi-Civita parallel translation P_t perserves $|\cdot|$ we conclude from (6) and (7) that the Jacobifields corresponding to $Y_1,\dots,Y_m$ are pairwise orthogonal along $c(t) = \exp tX$ and satisfy

$$|(d\exp)_{tX} \cdot Y| = t^{-1}|J(t)| = t^{-1}|k(t)| = \left(\frac{t\lambda_i}{2}\right)^{-1} \sin t\lambda_i/2 , \quad t \leq \pi .$$

Since $\|\operatorname{ad} X\| \leq 2\,\|X\| \leq 2\pi$ (1), the eigenvalues of $-(\operatorname{ad} X)^2$ are $\leq 4\pi^2$. Hence

$$\det(d\exp)_{tX} = \prod_{i=1}^{m} \frac{\sin}{\mathrm{id}}\left(\frac{t}{2}\lambda_i\right)$$

is not increasing and $\operatorname{vol} B\pi/\operatorname{vol} B_t \leq (\pi/t)^m$, q.e.d.

4.5 Lemma

There exist at most $3^{n+\dim SO(n)}$ euclidean motions $\alpha, \beta, \dots$ which pairwise satisfy the condition (c.f.2.1)

$$d(\alpha,\beta) \geq \max\{\|\alpha\|, \|\beta\|\} .$$

Proof Consider m such motions α_i and corresponding pairs $w_i = (S_i, a_i) \in so(n) \times \mathbb{R}^n$ where $\exp S_i = A_i = r(\alpha_i)$, $a_i = t(\alpha_i)$. Introducing the norm $\|(S,a)\| = \max\{\|S\|, c\cdot|\;|\}$ (the constant is irrelevant) in the vectorspace $so(n) \times \mathbb{R}^n$ we find m points $\tilde{w}_i = \|w_i\|^{-1} w_i$ on the unit sphere satisfying

$$\|\tilde{w}_i - w_j\| \geq \|w_j\|^{-1}\|w_i - w_j\| - \|\,\|w_j\|^{-1} w_i - \tilde{w}_i\| \geq 1$$

(if w.l.o.g. $\|w_j\| \leq \|w_i\|$), since by (10)

$$\|w_i - w_j\| \geq d(\alpha_i,\alpha_j) \geq \max\{\|\alpha_i\|, \|\alpha_j\|\} = \max\{\|w_i\|,\|w_j\|\} \quad .$$

It follows that the open balls of radius $1/2$ around the $\tilde{w}_i$ are pairwise disjoint and contained in a ball of radius $3/2$. Now m cannot exceed the volume ratio $3^{\dim(so(n) \times R^n)}$. q.e.d.

Remark. There is no finite bound if the condition is replaced by $d(\alpha,\beta) \geq \epsilon \max\{\|\alpha\|,\|\beta\|\}$ $\epsilon < 1$. In many cases as e.g. in the proof of Gromov's theorem, it is desirable to have an open condition. One such condition is

$$d(\alpha,\beta) \geq \max\{\|\alpha\| - \epsilon\|\beta\|, \|\beta\| - \epsilon\|\alpha\|\} \quad .$$

The number of motions is then bounded above by $(\frac{3-\epsilon}{1-\epsilon})^{n+\dim SO(n)}$, the proof is the same.

References

1. Bieberbach, L., Ueber die Bewegungsgruppen der Euklidischen Räume, I: Math. Ann. 70 (1911) 297-336; II: Math. Ann. 72 (1912) 400-412.

2. Buser, P., Karcher, H., Gromov's almost flat manifold theorem (to appear).

3. Gromov, M., Almost flat manifolds, J. Differential Geometry 13 (1978) 231-241.

4. Wolf, J., Spaces of constant curvature. McGraw Hill, New York (etc.) 1967.

Other proofs of the Bieberbach theorem

5. Auslander, L., Bieberbach's theorems on space groups and discrete uniform subgroup of Lie groups I: Ann. of Math. 71 (1960) 579-590; Amer. J. Math. 83 (1961), 276-280.

6. Frobenius, G. Über die unzerlegbaren diskreten Bewegungsgruppen, Sitz. ber. der Preuss. Akad. Wissen., Berlin, (1911), 654-665.

7. Zassenhaus, H., Beweis eines Satzes über diskrete Gruppen, Abh. Math. Sem. Univ. Hamburg, 12 (1938), 289-312.