A DIRECT PROOF OF ONE GROMOV'S THEOREM

YU. D. BURAGO, S. G. MALEV, D. I. NOVIKOV[‡]

Abstract. We give a new proof of the Gromov theorem: For any $C > 0$ and integer $n > 1$ there exists a function $\Delta_{C,n}$ such that if the Gromov–Hausdorff distance between complete Riemannian n-manifolds V and W is not greater than δ, absolute values of their sectional curvatures $|K_\sigma| \le C$, and their injectivity radii $\ge 1/C$, then the Lipschitz distance between V and W is less than $\Delta_{C,n}(\delta)$ and $\Delta_{C,n} \to 0$ as $\delta \to 0$.

1. Introduction

Denote by $\mathfrak{M}(\rho, C, n)$ the class of complete n-dimensional Riemannian manifolds V with section curvatures $|K_\sigma| \le C < \infty$ and the injective radii $r_{in}(V) \ge \rho$, where C, ρ are some positive constants. There are two well-known metrics on this class: the Lipschitz metric and the Gromov–Hausdorff metric.

Recall that the Lipschitz distance $\mathrm{d}_{Lip}(X, Y)$ between metric spaces X, Y is defined as

$$\mathrm{d}_{Lip}(V, W) = \ln \inf\{k : \mathrm{Bilip}_k(V, W) \neq \emptyset\},$$

where $\mathrm{Bilip}_k(V, W)$ denotes the class of all bi-Lipschitz homeomorphisms between V and W with bi-Lipschitz constant $k \ge 1$. By bi-Lipschitz constant of a homeomorphism ζ we mean maximum of Lipschitz constants for ζ and ζ^{-1}.

Instead of the Gromov–Hausdorff metric, we use a metric, equivalent to it (see, for instance, [1]). We preserve notation d_{GH} for this metric. By definition, the distance $\mathrm{d}_{GH}(V, W)$ is the infimum of all $\delta > 0$ with the property that there exists a mapping $\chi : V \to W$ such that $\chi(V)$ is a δ-net in W and χ changes distances by at most on δ:

$$|\mathrm{d}_W(\chi(x), \chi(y)) - \mathrm{d}_V(x, y)| < \delta$$

for any points $x, y \in V$. Note that χ is not supposed to be continuous.

The purpose of this paper is to give a direct proof for the following Gromov theorem.

Theorem 1 (Gromov [4], page 379)**.** *For given $\rho > 0$, $C > 0$ and an integer $n > 1$, there exists a positive function $\Delta = \Delta_{(C,n,\rho)}$ such that $\Delta(\delta) \to 0$ as $\delta \to +0$ and if $V, W \in \mathfrak{M}(\rho, C, n)$ satisfy the condition $\mathrm{d}_{GH}(V, W) < \delta$ then*

$$\mathrm{d}_{Lip}(V, W) < \Delta(\delta).$$

In contrast to Gromov's proof using axillary embeddings of the manifolds V, W into an Euclidean space of a large dimension, we directly construct a bi-Lipschitz diffeomorphism $h(x)$ between V and W with a required bi-Lipschitz constant $\Delta(\delta)$. The map $h(x)$ is obtained by gluing together "local maps" φ_i with help of partition of unity. The maps φ_i are defined on some balls $B_{2\varepsilon}(v_i) \subset V$ which form a locally

The first author is partially supported by RFBR grant 05-01-00939.
The third author is supported by Samuel M. Soref & Helene K. Soref Foundation.

finite covering of V. This gluing is based on Karcher's center of mass technique [6]. The resulting map turns out to be bi-Lipschitz with the required constant since the mappings φ_i are C^1-close one to another on the intersection of their domains.

To justify publishing our proof, note that though ideas of Gromov's proof explained very clear in his book, some details are omitted in his exposition.

Later on C denotes different constants depending on $n = \dim V = \dim W$ only. We always assume δ to be sufficiently small, $\delta < \delta_0$, where δ_0 depends on n only. All these constants can be computed explicitly, if such a need arises.

2. Preparations

By a suitable rescaling of V, W, one can get rid of one parameter and assume that the absolute values of section curvatures smaller than δ and the injectivity radii are bigger than δ^{-1}. Also we can assume that and $d_{GH} < \delta$.

We always suppose that $0 < \delta \ll 1$ and denote

$$\varepsilon^2 = \delta \ll 1. \tag{1}$$

2.1. **ε-Orthonormal base.** We say that a basis $\{e_1, \dots, e_n\} \subset \mathbb{R}^n$ is ε-orthonormal if $|(e_i, e_j) - \delta_{ij}| < \varepsilon$ for all $i, j = 1, \dots, n$, where δ_{ij} is the Kronecker symbol. A linear map $L\colon M \to N$ of two Euclidean spaces is ε-close to isometry if $\|L - Q\| < \varepsilon$ for some isometry $Q\colon M \to N$.

Lemma 1. *Let $\{\xi_i\}$ be an ε-orthonormal base of $\mathbb{R}^n$, $\varepsilon < \frac{1}{2n}$. Let $L : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator.*

If $\|L(\xi_i)\| < \delta$, $i = 1, \dots, n$ then $\|L\| < 2\sqrt{n}\delta$.

Assume that $8n\sqrt{n}\delta < 1$ and

$$|\langle L\xi_i, L\xi_j\rangle - \langle \xi_i, \xi_j\rangle| < \delta, \quad i, j = 1, .., n.$$

Then L is $8n\sqrt{n}\delta$-close to an isometry.

Proof. Lemma could be easily proved by straightforward calculation. We give a proof to the first part, the second part can be obtained the same way.

Let $x = \sum x_i \xi_i$ be a unit vector. Then

$$1 = \langle x, x\rangle = \sum_{i,j} x_i x_j \langle \xi_i, \xi_j\rangle \geq \sum_{i=1}^n x_i^2 - \varepsilon \left(\sum_{i=1}^n |x_i|\right)^2 \geq (1 - n\varepsilon) \sum_{i=1}^n x_i^2.$$

Therefore $\sum x_i^2 \leq (1 - \varepsilon n)^{-1} \leq 2$. Thus

$$\|Lx\| = \|\sum_{i=1}^n x_i L(\xi_i)\| \leq \delta \sum_{i=1}^n |x_i| \leq \delta \sqrt{n \sum x_i^2} \leq 2\delta\sqrt{n}.$$

□

2.2. **On comparison theorems and exponential mappings.** Denote by J a Jacobi vector field along a geodesic $\gamma\colon [0, 2] \to V$. As usually, $\dot{J}(t) = \frac{D}{dt} J$ means the covariant derivative of J along γ. We need a known comparison theorem of Rauch-style, see [3], 7.4, or the original Karcher's paper [6]. We formulate the theorem in the form adopted to our case.

Theorem 2. *Suppose that all section curvatures $|K| \le \delta$ and $0 < \delta \ll 1$ (it is enough to have $\delta < 10^{-2}$).*

Then

$$|J(0)|\cos(\sqrt{\delta}t) + |\dot{J}(0)|\delta^{-1/2}\sin(\sqrt{\delta}t) \le |J(t)| \le |J(0)|\cosh(\sqrt{\delta}t) + |\dot{J}(0)|\delta^{-1/2}\sinh(\sqrt{\delta}t). \quad (2)$$

As usually, we apply this estimate to the exponential and logarithmic mappings. The latter, $\log_v$, is the inverse of the exponential mapping $\exp_v : T_vV \to V$. By assumptions, it is well defined on balls of the radius $\delta^{-1} \gg 1$.

Denote by $\tau_{v_1,v_2} \colon T_{v_1}V \to T_{v_2}V$ the parallel translation along the minimal geodesic (which is always unique in our conditions) joining points $v_1, v_2 \in V$. Let $a, \xi, \eta \in T_vV$, with $\|a\| = r < 2$, and denote $v' = \exp_v a$. Let γ be the geodesic connecting v with v', $\gamma(t) = \exp_v(ta/r)$. Obviously, $d_a \exp_v(\xi) = J(r)$, where J is the Jacobi field along γ with the initial values $J(0) = 0, \dot{J}(0) = \xi/r$.

Denote by $\eta(t)$ the parallel translation of η to $\gamma(t)$ along γ, so $\dot{\eta} = 0$, and denote $f(t) = \langle \eta(t), J(t)\rangle$. Evidently, $f(0) = 0$, $\dot{f}(0) = \langle \xi/r, \eta\rangle$, and

$$|\ddot{f}(t)| = \langle \eta(t), \ddot{J}(t)\rangle = \langle \eta(t), R(J,\dot{\gamma})\dot{\gamma}\rangle \le \delta\ \|\eta\|\ \|\xi\| \left(1 + \frac{\delta}{2}t^2\right) \le 2\delta\ \|\eta\|\ \|\xi\|$$

by Theorem 2. This implies that

$$|f(r) - \langle \xi, \eta\rangle| \le \delta r^2\ \|\eta\|\ \|\xi\|.$$

Therefore, $\|d_a \exp_v(\xi) - \tau_{v,v'}\xi\| \le \delta r^2 \|\xi\|$, which proves the following Lemma.

Lemma 2. *Let $a \in T_vV$, $v' = \exp_v a$. If $r = \mathrm{dist}(v,v') < 2$, then the parallel translation $\tau_{v,v'} \colon T_vV \to T_{v'}V$ and the differential $(d\exp_v)_a \colon T_vV \to T_{v'}V$ are $r^2\delta$-close one to the other. In particular, $\exp_v$ is $(1+\delta r^2)$-bi-Lipschitz on balls $B_r(v) \in V$ of radius $r < 2$.*

Similar statements hold for $d\log_w$ and $\tau_{x,w}$.

Lemma 3. *Let $x, y, z \in B_2(v)$. Then*

$$\|\log_z y - \log_z x - \tau_{x,z}\log_x y\| < C\,\delta \quad (3)$$

Proof. This lemma is a simple consequence of one Karcher's estimate ([6], inequality (C2.3) on the page 540).

Indeed, substituting $a = \log_z x$, $v = \log_z y - \log_z x$, $p = z$ to formula (C2.3)[6], and taking into account that section curvatures $K_\sigma < \delta$, one gets

$$d(y, \exp_x(\log_z y - \log_z x) \le C\delta.$$

Since logarithmic mapping has very small distortion (Lemma 2), we obtain desired inequality by replacing points to their $\log_x$-images in the last inequality .

□

3. Local maps $\varphi_i(x)$

3.1. **Construction of φ_i.** Let $\{v_i\}$ be a ε-separated ε-net in V. By definition of the Gromov–Hausdorff metric, there exists a mapping $\chi : \{v_i\} \to W$ such that $U_\delta(Im\ \chi) = W$ and

$$|\mathrm{dist}_V(v_i, v_j) - \mathrm{dist}_W(\chi(v_i), \chi(v_j))| < \delta \quad (4)$$

for all i, j. We will call such mappings by ε-approximations.

The construction of mappings φ_i is the same for all i, so we choose some $v = v_i$ and drop the index i till the end of the subsection.

(1) Choose $\mathfrak{B} = \{b_1, b_2, \dots, b_n\} \subset T_vV$ be some orthonormal basis of the Euclidean space T_vV.

(2) Pick $e_k \in \{v_j\}$ such that $\operatorname{dist}(e_k, \exp_v b_k) < \varepsilon$ for $k = 1, \dots, n$. Denote by $\mathfrak{E}$ the basis $\{\log_v e_k\}$ of T_vV, and let $\mathfrak{F}$ be the basis $\{\log_w f_k\}$ of T_wW, where $w = \chi(v)$, $f_k = \chi(e_k) \in W$

(3) Denote by L the linear mapping $T_vV \to T_wW$ such that
$$L(\log_v(e_k)) = \log_w(f_k), \quad k = 1, \dots, n. \tag{5}$$
In the other words, L is the linear extension of the restriction to the basis $\mathfrak{E}$ of the mapping $\log_w \circ \chi \circ \exp_v : T_vV \to T_wW$ (the latter is defined on a discreet set of points only).

(4) We define mapping φ as:
$$\varphi = (\exp_w \circ L \circ \log_v)|_{B_{4\varepsilon}(v)}, \tag{6}$$
where $B_{4\varepsilon}(v) \subset V$ is a ball of radius 4ε with center in v.

Evidently, $L = d_v\varphi$. Also, $\varphi(e_k) = f_k$ for $k = 1, \dots, n$.

Lemma 4. *The base $\mathfrak{E}$ and $\mathfrak{F}$ are $C\varepsilon$-orthonormal. The mapping $L : T_vV \to T_wW$ is $C\delta$-close to an isometry.*

Let $e' \in \{v_k\} \subset V$ be a point in the net $\{v_i\}$, $\operatorname{dist}(e', v) < 2$, and let $f' = \chi(e')$. Then $\|L(\log_v e') - \log_w f'\| \le C\delta$.

Proof. The norms of the vectors $\log_v e_k - b_k$ are less than 2ε by the choice of e_k and Lemma 2. Therefore $\mathfrak{E}$ is $C\varepsilon$-orthonormal basis.

The pairwise distances between v, e_k, e' change by at most $C\delta$ by $\log_w \circ \chi$, due to Lemma 2 and the main property (4) of χ. This implies that the scalar products $\langle \log_v e_k, \log_v e_l\rangle$ differ from the scalar products $\langle \log_w f_k, \log_w f_l\rangle$ by at most $C\delta$,
$$|\langle \log_v e_k, \log_v e_l\rangle - \langle \log_w f_k, \log_w f_l\rangle| < C\delta, \tag{7}$$
due to the cosine theorem
$$2\langle a, b\rangle = \|a\|^2 + \|b\|^2 - \|a - b\|^2. \tag{8}$$
By the same reasons,
$$|\langle \log_v e', \log_v e_l\rangle - \langle \log_w f', \log_w f_l\rangle| < C\delta. \tag{9}$$

The (7) means that the difference of the Gram matrices $G_{\mathfrak{E}}, G_{\mathfrak{F}}$ of the base $\mathfrak{E}$ and $\mathfrak{F}$ correspondingly, is $C\delta$-small,
$$\|G_{\mathfrak{E}} - G_{\mathfrak{E}}\| < C\delta,$$
and, by $C\varepsilon$-orthogonality of $\mathfrak{E}$, this implies that both matrices are $C\varepsilon$-close to the identity matrix. This implies that the basis $\mathfrak{F}$ is a $C\varepsilon$-orthonormal basis.

Also, it implies that their inverses $G_{\mathfrak{E}}^{-1}, G_{\mathfrak{F}}^{-1}$ are $C\delta$-close as well, so L is $C\delta$-close to an isometry.

Together with (9), this implies that the coefficients for decomposition of vectors $\log_v e', \log_w f'$ in bases $\mathfrak{E}$ and $\mathfrak{F}$ correspondingly are $C\delta$-close, since one can restore these coefficients from the tuples $\{(\log_v e', \log_v e_l)\}$ and $\{(\log_w f', \log_w f_l)\}$ using $G_{\mathfrak{E}}^{-1}, G_{\mathfrak{F}}^{-1}$ correspondingly. Since $L\mathfrak{E} = \mathfrak{F}$ by definition, the coefficients of decomposition of $L(\log_v e')$ and $\log_w f'$ in the basis $\mathfrak{F}$ are $C\delta$-close, which proves the Lemma. $\square$

Corollary 1. *There is some constant $C > 0$ which depends on n only, such that for each i, φ_i is a $(1+C\delta)$-bi-Lipschitz map and its differential is $C\delta$-close to isometry.*

This follows immediately from the previous Lemma and Lemma 2.

3.1.1. *Maps $\{\varphi_i\}$ are C^1-close one to another.* Let $v_1, v_2 \in \{v_i\}$ be two points of the ε-net on V, and assume that $\operatorname{dist}(v_1, v_2) < 4\varepsilon$. Here we prove that the mappings φ_1 and φ_2 are $C\delta$-close in C^1-sense in $B_{2\varepsilon}v_1 \cap B_{2\varepsilon}v_2$.

Lemma 5. *For every $x \in B_{4\varepsilon}(v_1) \bigcap B_{4\varepsilon}(v_2)$*

$$\operatorname{dist}(\varphi_1(x), \varphi_2(x)) < C\delta. \tag{10}$$

Moreover, the parallel translation $\tau_{\varphi_1(x),\varphi_2(x)} \circ d_x\varphi_1$ of $d_x\varphi_1$ is $C\delta$-close to $d_x\varphi_2$.

Proof. By Lemma 2:

$$\|\tau_{w_1,x} \circ L_1 \circ \tau_{x,v_1} - d\varphi(x)\| < C\delta \tag{11}$$

The vector $L_1(\tau_{x,v_1} \log_x e_i)$ is $C\delta$-close to

$$L_1(\log_{v_1} x - \log_{v_1} e_i) = \log_{w_1} \varphi_1(x) - \log_{w_1} f_i,$$

by Lemma 3. The parallel translation of the right part of the last equation to the point $\varphi_1(x)$ by $\tau_{w_1,\varphi_1(x)}$ is $C\delta$-close to $\log_{\varphi_1(x)} f_i$, again by the same Lemma 3. Summing up, we get

$$\|d_x\varphi_1(\log_x e_i) - \log_{\varphi_1(x)} f_i\| < C\delta,$$

and similarly for $d_x\varphi_2$.

Therefore by Lemma 3 the vectors

$$\tau_{\varphi_1(x),\varphi_2(x)} \circ d_x\varphi_1(\log_x e_i) - d_x\varphi_2(\log_x e_i), \quad i = 1, \dots, n,$$

are $C\delta$-close to $\log_{\varphi_1(x)} \varphi_2(x)$; i.e., to zero, if the $\operatorname{dist}(\varphi_1(x), \varphi_2(x)) < C\delta$. Since $\{\log_x e_i\}$ is a $C\varepsilon$-orthonormal basis of T_xV, this means that the first claim of the Lemma implies the second.

Now, by Lemma 4,

$$\operatorname{dist}(\varphi_1(v_2), \varphi_2(v_2)) < C\delta. \tag{12}$$

Therefore, $\|\tau_{\varphi_1(v_2),w_2} \circ d_{v_2}\varphi_1 - L_2\| < C\delta$.

The parallel translation along a geodesic triangle $\triangle w_1 w_2 \varphi_1(v_2)$ is $C\delta$-close to identity, so, using (11), we conclude that

$$\|\tau_{w_1,w_2} \circ L_1 \circ \tau_{v_2,v_1} - L_2\| < C\delta.$$

Applying this to the vector $\log_{v_2} x \in T_{v_2}V$ and using Lemma 3 and (12) we get

$$\| \log_{w_2} \varphi_2(x) - \log_{w_2} \varphi_1(x)\| < C\delta,$$

which, by Lemma 2, proves the first statement of the Lemma as well. $\square$

4. Gluing together local mappings

Here we glue the mappings φ_i into one mapping $h : V \to W$ using the center of mass construction of Karcher.

Suppose that a Riemannian manifold V is covered by balls B_i of radius 2ε, and that for every i there exists a mapping $\varphi_i : B_i \to W$ which is δ-close to isometry. Assume that the images of φ_i cover W and are δ-close in C^1 sense on intersections of their domains. We prove that there exists a bi-Lipschitz mapping between V and W which is C^1 close to φ_i on B_i.

4.0.2. *Partition of unity: construction and estimates.* The partition of unity $\{\psi_i(x)\}$ is constructed in a standard way. We need some estimates on the norms of their differentials, so we repeat this classical construction here.

Let $\Psi(r)$ be any non-negative monotonic and $C^\infty-$smooth function such that it is equal to 1 for $r \leq 1$ and 0 for $r \geq 2$. Consider functions $\tilde{\psi}_i(x) = \Psi(\varepsilon^{-1}|xv_i|)$. We define

$$\psi_i = \frac{\tilde{\psi}_i}{\sum\limits_k \tilde{\psi}_k}$$

It is easy to estimate the differential of ψ_i:

Lemma 6. *For every $x \in V$*

1. *at least one of the numbers $\tilde{\psi}_i(x)$ is equal to 1,*
2. *at most 7^n of $\psi_i(x)$ are different from zero assuming that δ is sufficiently small,*
3. $\| d_x\psi_i \| \leq C\varepsilon^{-1}$ *with some absolute constant C.*

Proof. The first statement follows because $\{v_i\}$ is an ε-net.

The Bishop-Gromov upper estimate for the number of non-vanishing $\psi_i(x)$ is $\mathrm{Vol}_{-\delta} B_{5\varepsilon/2}/\mathrm{Vol}_\delta B_{\varepsilon/2}$, where $\mathrm{Vol}_c B_r$ denotes the volume of a ball of radius r in the space form of curvature c. For $\delta \ll 1$ this ratio is smaller than 7^n, and the second claim follows.

Let us estimate the differentials:

$$\|d\psi_i\| \leq \left\| \frac{d\tilde{\psi}_i}{\sum \tilde{\psi}_k} \right\| + \left\| \frac{\tilde{\psi}_i \sum d\tilde{\psi}_i}{(\sum \tilde{\psi}_k)^2} \right\| \leq \left\| d\tilde{\psi}_i \right\| + \left\| \tilde{\psi}_i \cdot \sum d\tilde{\psi}_i \right\|, \tag{13}$$

because $\sum\limits_k \tilde{\psi}_k \geq 1$.

But $\| \partial\tilde{\psi}_k \| \leq C\varepsilon^{-1}$, where C is some absolute constant. So the estimate for $\| d\psi_i \|$ follows from the second claim of the Lemma. □

4.1. **Definition of Φ.** Define the function $\Phi : V \times W \to \mathbb{R}$ as

$$\Phi(x, y) = \frac{1}{2} \sum_i \psi_i(x) \mathrm{dist}(\varphi_i(x), y)^2,$$

This is a smooth function because if $\mathrm{dist}(\varphi_i(x), y)^2$ is big, than the corresponding coefficient is zero. We define the mapping $h(x)$ by the condition

$$\Phi(x, h(x)) = \min_{y \in W} \Phi(x, y).$$

In the other words, $h(x)$ is the center of mass for the points $\varphi_i(x)$ with weights $\psi_i(x)$.

Lemma 7 (Karcher [6])**.** *The function $h(x)$ is well-defined.*

The reason is that for a fixed $x \in V$, the points $\varphi_i(x)$ corresponding to non-zero $\psi_i(x)$ are $C\delta$-close one to another by Lemma 5. Therefore $\Phi(x, y) \geq 4C^2\delta^2$ for y outside a ball $B_{3C\delta}\varphi_i(x) \subset W$ of radius $3C\delta$ centered at one of $\varphi_i(x)$, and is less that $C^2\delta^2$ for this $\varphi_i(x)$. Therefore the minimum lies in $B_{3C\delta}\varphi_i(x)$. On the other hand, one can show that $\Phi(x, \exp(\cdot))$ is a convex function in $B_{4\delta}(0) \subseteq T_{\varphi_i(x)}W$, similar to Corollary 2 below, so the minimum is unique. In particular, we see that

$$\mathrm{dist}(h(x), \varphi_i(x)) < 3C\delta. \tag{14}$$

Since for each $x \in V$, the function Φ reaches minimum at $y = h(x)$, we have

$$d_*\Phi = 0, \tag{15}$$

at all points $(x, h(x))$, where d_* means the restriction of $d\Phi$ to the subspace $\{0\} \times TW \subset T(V \times W)$. Below we will prove that the restriction $\mathrm{Hess}_*\Phi$ of the hessian of Φ to $TW \times TW$ is not degenerate at points satisfying (15). Then the function h is smooth by the implicit function theorem. Substituting $y = h(x)$ in (15) and differentiating, we obtain

$$d^2\Phi(a, b) + d^2\Phi(dh(a), b) = 0.$$

As a result

$$dh = d_*^2\Phi^{-1} \circ d_{**}^2\Phi, \tag{16}$$

where $d_*^2\Phi$ and $d_{**}^2\Phi$ mean the restrictions of $d^2\Phi$ to $TW \times TW$ and $TV \times TW$, accordingly, understood as mappings $d_*^2\Phi : TW \to T^*W$ and $d_{**}^2\Phi : TV \to T^*W$.

Lemma 8. *The mapping $h : V \to W$ is smooth.*

Moreover, its differential is non-degenerate for each $x \in V$ and has a bi-Lipschitz constant $\Delta(\delta) = O(\varepsilon)$, where $\varepsilon^2 = \delta$.

Theorem 1 easily follows from Lemma 8, see Section 5 below.

The first assertion of the lemma is already proved up to non-degeneracy of $\mathrm{Hess}_*\Phi$.

So to prove the lemma, it is enough to check that at points $(x, h(x))$, both $d_{**}^2\Phi$ and $d_*^2\Phi$ are $C\varepsilon$-close to isometries.

Denote $\mathrm{dist}(x, y)$ by $|x, y|$. We claim that it is enough to prove that, first, the similarly defined $d_{**}^2|\varphi_i(x), y| : TV \to T^*W$ are all $C\varepsilon$-close and also $C\varepsilon$-close to an isometry, and, second, that the same statements hold for $d_*^2|\varphi_i(x), y| : TW \to T^*W$.

Indeed, for $a \in T_xV$ we have $d_{**}^2\Phi(a) \in T^*W$, and

$$d_{**}^2\Phi(a) = \sum_{i=1}^{N} d_x\psi_i(a) \cdot d_y|\varphi_i(x), y| + \sum_{i=1}^{N} \psi_i(x) d_{**}^2|\varphi_i(x), y|(a), \tag{17}$$

where $y = h(x)$. If all $d_{**}^2|\varphi_i(x), y|$ are all $C\varepsilon$-close to some isometry then the second sum, being their convex combination, is also $C\varepsilon$-close to the same isometry. But the first sum in (17) is $C\varepsilon$-small: by Lemma 6 there are no more than 7^n non-zero terms in the first sum, and $|d_x\psi_i(a)| \le C\delta^{-1/2}$, $|\varphi_i(x), h(x)| \le C\delta$ and $\|d_y|\varphi_i(x), h(x)|\| \le 1$ in each of them.

For $d_*^2\Phi$ the proof is similar.

So Lemma 8 follows from the next lemma, and its Corollary.

Lemma 9. *Let $|x, y|$ be the distance function of a Riemannian manifold M. Suppose, that points x, y are joined by a unique geodesic $\gamma\colon [0, 1] \to M$ and no one point of γ is conjugate with x, y. Let $a \in T_yM$, $b \in T_yM$, $\tilde{b}$ is obtained by parallel translation of b along γ to x. Then at the point $(x, y) \in M \times M$ we have:*

(i) $d_*^2|x, y|(a)(a) = \langle J(1), J'(1)\rangle$, where J is the Jacobi field along γ with initial data $J(0) = 0$, $J(1) = a$.

(ii) $d_{**}^2|x, y|(a)(\tilde{b}) = \langle J(t), J'(t)\rangle|_0^1$, where the Jacobi field J satisfies the conditions $J(0) = \tilde{b}$, $J(1) = a$.

In the lemma, notations d_*^2, d_{**}^2 have the same sense as in formula (16).

Proof. In the case (i) $d^2_*|x,y|(a)(a)$ is equal to the second variation of energy for geodesic variation σ of γ such that $\sigma(0,\tau)\equiv\gamma(0)$ and $\frac{D\sigma}{\partial\tau}=a$, and $\frac{D^2\sigma}{\partial\tau^2}=0$ at $\gamma(1)$. It is well known that this second variation is equal $\langle J(1),J'(1)\rangle$.

The case (ii) is proved similarly. □

Corollary 2. *$d^2_*|x,y|(a)(a)$ and $d^2_{**}|x,y|(a)(\tilde b)$ are $C\delta$-close to $\|a\|^2$ and $\langle a,b\rangle$, correspondingly.*

Indeed, in the case of the Euclidean metric $\langle J(1),J'(1)\rangle$ is equal a^2 in the first case. The Rauch theorem shows that$\langle J(1),J'(1)\rangle$ is $C\delta$-close to a^2 in our assumptions.

The second case is similar.

Corollary 3. *The mappings $d^2_*|\varphi_i(x),y|:TW\to T^*W$ are all $C\varepsilon$-close one to another and $C\varepsilon$-close to some isometry. The same holds for $d^2_{**}|\varphi_i(x),y|:TV\to T^*W$.*

Indeed, $d^2_*|\varphi_i(x),y|$ are $C\delta$-close to $b\to\langle\cdot,b\rangle$ by previous corollary, so the first claim is trivial.

Similarly, $d^2_{**}|\varphi_i(x),y|$ are $C\delta$-close to $b\to\langle\cdot,d_x\varphi_i(b)\rangle$, which are $C\delta$-close by Lemma 5 and $C\delta$ close to isometry by Corollary 1.

5. Proof of Theorem 1

The differential of the mapping $h:V\to W$ defined by Lemma 7 is $C\varepsilon$-close to isometry, in particular non-degenerate by Lemma 8. This means that $h(V)$ is open, and, as V is compact, also closed. Therefore $h(V)=W$, and h is a covering.

Now, if $y=h(x_1)=h(x_2)$, then choosing v_i such that $\operatorname{dist}(v_i,x_i)<\varepsilon$, we get $\operatorname{dist}(h(v_i),y)<3\varepsilon/2$, which means that $\operatorname{dist}(v_1,v_2)<2\varepsilon$, so $\operatorname{dist}(v_1,x_2)<3\varepsilon$.

The mapping h is smoothly homotopic to φ_1 on the ball $B_{7\varepsilon/2}(v_1)$: one can smoothly deform the partition of unity in such a way that ψ_1 becomes identical 1 on $B_{7\varepsilon/2}(v_1)$. During this homotopy the mapping h remains smooth, and also $h(\partial B_{7\varepsilon/2}(v_1))$ remains $C\delta$-close to $\varphi_1(\partial B_{7\varepsilon/2}(v_1))$, by Corollary 2 and (14) hold for any partition of unity. This means that the number of preimages of $y\in B_{2\varepsilon}(\chi(v_1))$ in $B_{7\varepsilon/2}(v_1)$ remains the same during the homotopy. For φ_1 this number is equal to 1, as φ_1 is a diffeomorphism on $B_{4\varepsilon}(v_1))$, so $x_1=x_2$.

Therefore h is a diffeomorphism, which, together with Lemma 8, proves Theorem 1.

References

[1] Burago, D. Yu.; Burago, Yu. D.; Ivanov, S. *A course in metric geometry. Graduate Studies in Mathematics*, 33. American Mathematical Society, Providence, RI, 2001. xiv+415 pp. ISBN: 0-8218-2129-6

[2] Burago, Yu. D.; Zalgaller, V. A. *An introduction to Riemannian geometry* (Russian). "Nauka", Moscow, 1994, 319 pp. ISBN: 5-02-024606-9

[3] Chavel, I. *Riemannian geometry — a modern introduction*. Cambridge Tracts in Mathematics, 108. Cambridge University Press, Cambridge, 1993. xii+386 pp. ISBN: 0-521-43201-4;

[4] Gromov, M.*Metric Structures for Riemannian and Non-riemannian Spaces*, 1997, 379-383.

[5] Jost, Ju. *Riemannian geometry and geometric analysis.* Third edition. Universitext. Springer-Verlag, Berlin, 2002. xiv+532 pp. ISBN: 3-540-42627-2 53-02 (58-02)

[6] Karcher, H. *Riemannian center of mass and mollifier smoothing*. Comm. Pure Appl. Math. **30** (1977), no. 5, 509-541.

Steklov Mathematical Institute at St.Petersburg, Russia
E-mail address: `yuburago@pdmi.ras.ru`

Department of Mathematics, Weizmann Institute of Science, Rehovot, Israel
E-mail address: `sergey.malev@weizmann.ac.il`

Department of Mathematics, Weizmann Institute of Science, Rehovot, Israel
E-mail address: `dmitry.novikov@weizmann.ac.il`