LIPSCHITZ CONVERGENCE OF RIEMANNIAN MANIFOLDS

R. E. Greene and H. Wu

Let $\mathscr{C} \equiv \mathscr{C}(n, \Lambda, \delta_0, V_0)$ be the set of all connected compact C^∞ n-dimensional Riemannian manifolds with |sectional curvature| $< \Lambda^2$, diameter $< \delta_0$, and volume $> V_0$. The main result of this paper is that this class $\mathscr{C}$ has certain compactness, or more precisely, precompactness properties. The class $\mathscr{C}$ consists of only finitely many diffeomorphism classes so the precompactness properties can be thought of as dealing with the set of metrics satisfying the class $\mathscr{C}$ requirements on a fixed differentiable manifold. The main theorem of this paper is then that a sequence of such metrics always has a subsequence which, after application of suitable diffeomorphisms of M, converges to a limit metric. The regularity of the limit can be taken to be $C^{1,\alpha}$, for all α with $0 < \alpha < 1$ and the convergence to be in the $C^{1,\alpha}$ norm.

In [12], M. Gromov stated, in terminology that will be explained in a moment, a striking convergence theorem concerning the class $\mathscr{C}$ defined above (Theorems 8.25 and 8.28 of [12]):

(∗)   Given a sequence $\{ M^l\colon l = 1, 2, 3, \ldots \}$ in $\mathscr{C}$, there exists a subsequence $\{ M^k \}$ and a $D^{1,1}$-Riemannian manifold M such that $\{ M^k \}$ converges to M in the Lipschitz distance.

Here a $C^{1,1}$-manifold is by definition a C^1-manifold with coordinate transition functions having Lipschitz continuous first derivatives (note that the notion of Lipschitz continuity used here is that of Euclidean space to Euclidean space, so no choice of metric is involved). A $D^{1,1}$-Riemannian manifold M is a $C^{1,1}$-manifold with a continuous Riemannian metric having the additional property that, for each point x in M, the distance function ρ_x determined by the metric is C^1 (within the cut locus) and that the derivatives of ρ_x are Lipschitz continuous with a Lipschitz constant that is independent of the point x. The Lipschitz distance $d_L(M_1, M_2)$ between two homeomorphic compact metric spaces M_1, M_2 is defined to be the infimum of

$$\left| \log \operatorname{dil}(f) \right| + \left| \log \operatorname{dil}(f^{-1}) \right|$$

where f ranges over the homeomorphisms of M_1 to M_2 and by definition,
$$\operatorname{dil} f = \sup d(f(x_1), f(x_2))/d(x_1, x_2)$$
where the supremum is taken over all $x_1, x_2 \in M$, $x_1 \neq x_2$, and $\operatorname{dil} f^{-1}$ is defined in the obviously analogous way. Two non-homeomorphic compact metric spaces will be considered to have Lipschitz distance $+\infty$ from each other.

Note that what we called $D^{1,1}$-Riemannian manifold in the above is our ad hoc terminology for what Gromov calls a "$C^{1,1}$-Riemannian manifold." The latter terminology is unfortunate since such a Riemannian manifold need not, by definition, possess a $C^{1,1}$-Riemannian metric tensor. Note that in this notion of Lipschitz convergence, the underlying C^∞ structure of the manifolds plays no role whatever; only the underlying metric space structure enters into the discussion.

The result (∗) implies the finiteness theorem of Cheeger [5] and Peters [17] via the result of Shikata [18]: with convergence as in (∗), [18] implies that all but a finite number of the M^k are diffeomorphic to M so that the M^l could not have been mutually nondiffeomorphic. The most remarkable part of (∗) is the assertion of convergence: By Shikata [18], the convergence of $\{M^k\}$ to M with respect to d_L implies that, for all large enough k, M is (C^1) diffeomorphic to M^k; it would then be natural to treat metric convergence by looking for C^1-diffeomorphisms $f_k: M \to M^k$ such that, if g and g_k are respectively the Riemannian metrics of M and M^k, then $\{f_k^* g_k\}$ converges uniformly to g on M. Now, by Ascoli's Theorem, the convergence of $\{f_k^* g_k\}$ to g would necessarily take place (for some subsequence) if the coefficients of the $f_k^* g_k$'s were uniformly bounded and equicontinuous. Since |sectional curvature| $< \Lambda^2$, Λ independent of k, the uniform boundedness of the coefficients in exponential normal coordinates follows from the Rauch Comparison Theorem. However, to achieve equicontinuity in these same coordinates, one needs something like a uniform bound on the first derivatives of $f_k^* g_k$; such a bound would seem to require a uniform bound on the *covariant derivative* of the curvature tensors of the (M_k, g_k); cf. Kaul [16] and the example in [14]. Such a covariant derivative bound is an assumption which has been studiously avoided in the discussion thus far!

In [12], Gromov gave an outline of an argument to justify (∗). Since then, Katsuda [15] has supplied details to all the parts of Gromov's outline except one: it remains unclear that the Hausdorff limit of a sequence in $\mathscr{C}$ must possess a $D^{1,1}$-Riemannian structure. (We are indebted to Professor T. Sakai for pointing this out to us.) More precisely, the proof of 8.25 in [12] has been shown by Katsuda to be entirely correct,

but the proof of 8.27 in [12] remains unclear. In the meantime, Berger [1], Katsuda [15] and Brittain [4] have made use of (sometimes an even stronger version of) the statement (∗) to prove generalized pinching theorems. In particular, the arguments of Berger [1] would seem to require that the Riemannian metric of M in (∗) is at last C^1, whereas a priori the metric of M is only continuous.

The main purpose of this paper is to clarify this situation by proving in detail a stronger version of (∗), by argument along lines completely different from Gromov's original scheme. The precise theorem we shall prove is as follows:

THEOREM. *Given a sequence* $\{ M^l\colon l = 1, 2, 3, \ldots \}$ *in* $\mathscr{C}$ *and an* $\alpha \in (0, 1)$, *there exists a subsequence* $\{ M^k \}$ *together with a* C^∞-*manifold* M *on which is defined a* $C^{1,\alpha}$-*Riemannian metric such that* $\{ M^k \}$ *converges to* M *with respect to the Lipschitz distance.*

The result of Shikata [18] again implies that all but a finite number of the M^k are diffeomorphic to M. The diffeomorphisms are in fact constructed explicitly in our proof of the Theorem without appeal to [18]. By choosing successive α converging to 1 and repeatedly passing to subsequences, and applying the standard diagonal argument, one can deduce, from the Theorem as stated, the slightly stronger corollary:

COROLLARY. *Given a sequence* $\{ M^l\colon l = 1, 2, 3, \ldots \}$ *in* $\mathscr{C}$, *there exists a subsequence* $\{ M^k \}$ *together with a* C^∞-*manifold* M *on which is defined a metric that is* $C^{1,\alpha}$ *for every* $\alpha \in (0, 1)$ *such that* $\{ M^k \}$ *converges to* M *with respect to the Lipschitz distance.*

Both in this Corollary and in the Theorem, a $C^{1,\alpha}$-metric is by definition a metric with g_{ij} coefficients of that regularity; the metric regularity is thus not defined in terms of the distance function directly as in Gromov's definition of $D^{1,1}$-Riemannian manifold. S. Peters has shown (preprint) that $C^{1,\alpha}$ is optimal here. He has also shown (independently of the present authors) the validity of the Theorem above (preprint, to appear *Compositio Math.*). For the chronology of these events, see the note at the end of this section.

The Theorem and a fortiori the Corollary suffice for the needs of Berger [1].

We shall employ the basic theorems on harmonic coordinates due to Jost-Karcher [14]. Basically, our proof consists of reworking Peter's proof of the finiteness theorem [17] using "harmonic balls" in place of geodesic

balls. In the process, we give yet another proof of the finiteness theorem, although we should point out explicitly that for the finiteness theorem itself, our proof does not yield more information than [17]. Throughout the proof of the theorem given below, we make implicit use of the following result of Cheeger [5]: there is a positive constant $i_0 = i_0(n, \Lambda, \delta_0, V_0)$ such that each $M \in \mathscr{C}$ possesses an injectivity radius $\geq i_0$.

It should be mentioned that the original $D^{1,1}$ assertion of Gromov in (∗) follows immediately from our work. Indeed, the proof of the Theorem (see the end of §2 below) shows the limit metric g on M to be the limit in the $C^{1,\alpha}$-norm of a sequence of C^∞-Riemannian metrics g_k on M such that (M, g_k) is isometric to M^k for each k. By the Hessian comparison theorem, the Hessian $D^2\rho_k$ of the distance function ρ_k of g_k is uniformly bounded (in a small neighborhood of each point) by a constant depending only on Λ. By Ascoli's theorem, a subsequence $\{\rho_{k_i}\}$ must converge uniformly to a $C^{1,1}$-function ρ'. However, since $g_k \to g$ in the $C^{1,\alpha}$-topology and g is $C^{1,\alpha}$, the limit of the distance functions of g_k equals the distance function of g. Thus the distance function of g must be ρ', which is $C^{1,1}$. (The reader should take note that we are here essentially repeating part of the proof of 8.27 in [12].)

After the completion of the research of this paper in May of 1985, there appeared the research announcement of O. Durumeric ([8]) which seems to prove, *assuming the validity of* (∗), that the metric of M in (∗) is in fact C^1 (rather than just C^0). This would also, suffice to justify the arguments in Berger [1]. It must be emphasized that, precisely because of the difficulty concerning the proofs previously suggested for (∗), our proof of the Theorem stated starts from scratch and does not rely on any of the arguments of Gromov [12] or Katsuda [15].

In the Differential Geometry Workshop held in May of 1985 at the Mathematical Sciences Research Institute of Berkeley, we circulated an earlier version of this paper (which differed from this version only in minor details of exposition). On that version, Jeff Cheeger made the comment that the proof of Theorem 2.5 in the joint paper [6] by himself and Gromov would also suffice to prove 8.25 of [12] and, in particular, also the finiteness theorem in the form proved by Peters [17]; this then overlaps the paper [15] of Katsuda. (We wish to remark that if the relevance of [6] to the convergence theorem has escaped the attention of most people, it is because Theorem 2.5 of [6] is a technical result imbedded in a long paper which does not mention in either its title or its introduction anything about the finiteness or convergence theorem.) Cheeger also pointed out that Theorem 2.5 of [6] should also suffice to

prove a stronger version of (∗) with $C^{0,\alpha}$ ($\alpha < 1$) regularity of the limit metric. As pointed out above, $C^{0,\alpha}$ ($\alpha < 1$) regularity is not adequate for applications. Furthermore, the only way known to us at the time of the Workshop to complete the $C^{0,\alpha}$ argument was to repeat the argument of this paper, with harmonic coordinates replaced by another kind of coordinates. Also during that Workshop, C. Pugh outlined an elementary method to get at the $D^{1,1}$ statement of (∗). Subsequently, he carried the method to completion and, in the process, showed a second way to prove the above-mentioned $C^{0,\alpha}$ ($\alpha < 1$) regularity of the limit metric by using Theorem 2.5 of [6]. See his forthcoming paper entitled: "The $C^{1,1}$ conclusion in Gromov's theory". As already noted, this conclusion follows immediately from our work.

The idea of using harmonic coordinates, especially the work of Jost and Karcher [14], for a possible proof of (∗) has apparently occurred to many people. In particular, at the July 1984 geometry meeting at Luminy, France, organized by M. Berger, approaching the convergence theorem through harmonic coordinates was the subject of considerable group discussion. The first-named author, who participated in these discussions, expresses his gratitude to the others who did so also and in particular to H. Karcher for his explanations of the harmonic coordinates construction in [14]. Both authors, however, feel that the Luminy participants should in no way be held responsible for any errors that may occur in the proof given below because the evolution of ideas has taken us a considerable distance from the approaches contemplated in that meeting.

Note As mentioned above, the present manuscript differs only slightly from the version distributed in May, 1985 in the Differential Geometry Workshop of MSRI (Berkeley); it was put in final form in July, 1985 and mailed to other workers in this field. We received the preprint of S. Peters on the convergence of Riemannian manifolds in May, 1986.

1. Harmonic coordinates. A coordinate system $\{(h^1, \ldots, h^n)\}$ defined on an open subset of a Riemannian manifold M is called a *harmonic coordinate system* if and only if $\Delta h^i \equiv 0$, $i = 1, 2, \ldots, n$. For each point x in M, there is always a harmonic coordinate system defined on some neighborhood of x; according to Bers, John, and Schechter [2, p. 228], one can choose harmonic functions defined locally near x with $h(x)$ and $dh(x)$ both prescribed (cf. Greene-Wu [9, 10]). We shall need, however, to have much more detailed information than simply this existence statement. In particular, we shall need among other things control over the size of the coordinate neighborhood. The following specific result will be

needed for our purposes:

(∗∗)   There exist positive numbers $R_0 = R_0(n, \Lambda, \delta_0, V_0)$ and $C = C(n, \Lambda, \delta_0, V_0, \alpha)$ such that, for each $M \in \mathscr{C}$ and $x \in M$, there is a harmonic coordinate system $\{h^i\colon\ i = 1, 2, \ldots, n\}$ on $B(x, R_0)$ satisfying

$$\|g_{ij}\|_{C^{1,\alpha}} < C$$

where

$$g_{ij} = g\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right)$$

are the metric coefficients relative to the h^i's and the $(1, \alpha)$ norm is taken in the h^i coordinates. (Hereafter, α is always assumed to satisfy $0 < \alpha < 1$.)

Note that in geodesic coordinates, the Rauch Comparison Theorem gives a bound

$$\|g_{ij}\| < C(n, \Lambda, \delta_0, V_0),$$

but there is no bound available on the derivatives of the g_{ij}'s so (∗∗) is a gain of $(1, \alpha)$ derivatives. The fact that harmonic coordinates are best for estimates of this type was originally pointed out by de Turck and Kazdan [7]. The important point that a precise $(1, \alpha)$ estimate exists on balls of fixed radius is due to Jost and Karcher [14]. We briefly recall some facts from [13] and [14].

Given $m \in M$ and $u \in M_m$ with $|u| = 1$, we first define the *almost linear function* $l_u(x) \equiv l(x)$ associated with u as follows: Let

$$r(x) = d(x, m),$$
$$p(x) = \exp_m r(x) u,$$
$$q(x) = \exp_m\{-r(x)u\},$$

then set

$$l_u(x) = \frac{1}{4r(x)}\{d^2(x, q(x)) - d^2(x, p(x))\}.$$

In $\mathbf{R}^n$, l_u would be the linear functional determined by inner product with u. If $\{u_1, \ldots, u_n\}$ is an orthonormal basis of M_m and $l^1, \ldots, l^n$ are the associated almost linear functions, then $\{l^1, \ldots, l^n\}$ gives local coordi-

nates on a ball of fixed radius which depends only on the four standard parameters n, Λ, δ_0, V_0. Call this radius ρ; then on the geodesic ball $B(m, \rho)$ we solve the following Dirichlet problem for each $i = 1, \ldots, n$:

$$\Delta h^i = 0 \text{ in } B(m, \rho), \qquad h^i|_{\partial B(m, \rho)} = l^i|_{\partial B(m, \rho)}.$$

These n harmonic functions $\{h^1, \ldots, h^n\}$ are a coordinate system on $B(m, \rho)$ when $\rho(n, \Lambda, \delta_0, V_0)$ is taken to be sufficiently small. By letting $R_0 = \frac{1}{2}\rho$ (for example), the estimate in $(**)$ above would be valid.

Another property of these harmonic coordinates $\{h^1, \ldots, h^n\}$ crucial for our purpose is the following: if f is a harmonic function defined on $B(m, \rho)$, then in $B(m, \rho/2)$,

$$(\natural) \qquad \left\langle \left\| \frac{\partial f}{\partial h^i} \right\|_{C^\alpha}, \left\| \frac{\partial^2 f}{\partial h^i \partial h^j} \right\|_{C^\alpha} \right\rangle \le C^* \| f \|_{C^0},$$

where C^* is a constant depending on ρ, n, Λ, δ_0, V_0 and α, and as in $(**)$, the C^α-norm is taken in the $\{h^i\}$ coordinates.

For later reference, we now state and prove the following elementary fact.

LEMMA. *Suppose* $f: \mathbf{R}^n \to \mathbf{R}^n$ *is* C^∞, $f(0) = 0$, $f = (f_1, \ldots, f_n)$ *are such that for all* $x \in B(0, \rho)$ *the* L^2-*norm of* df *satisfies* $\| df_x - I \| < 1/\sqrt{n}$; *then* f *is a diffeomorphism on* $B(0, \rho)$.

Proof of the Lemma. The fact that $\| df_x - I \| < 1/\sqrt{n}$ for all $x \in B(0, \rho)$ implies that $|df_i(x)| < 1/\sqrt{n}$ for all i and $x \in B(0, \rho)$. Now we set

$$A(p^1, \ldots, p^n) = \begin{bmatrix} df_1(p^1) \\ \vdots \\ df_n(p^n) \end{bmatrix}$$

for $p^1, \ldots, p^n \in B(0, \rho)$; here $df_1(p^1)$ means df_1 evaluated at p^1, not applied to p^1. Then $\| A(p^1, \ldots, p^n) - I \| < 1$. Thus $A(p^1, \ldots, p^n)$ is invertible. But if x, $x^1 \in B(0, \rho)$, then there exist $p^1, \ldots, p^n \in B(0, \rho)$ such that

$$f(x) - f(x^1) = A(p^1, \ldots, p^n) \cdot (x - x^1)$$

by the Mean Value Theorem. The matrix A being invertible and $x - x^1 \ne 0$ then imply that $f(x) - f(x^1) \ne 0$. $\qquad \square$

2. The proof of the convergence theorem. To prove the convergence theorem (the Theorem as stated in the introduction), we shall use special coverings of the manifolds by open sets that are balls in certain harmonic coordinate systems. The first order of business is to establish the definition and basic properties of these balls and coverings by them. The proof proper of the Theorem begins later in the section.

Given $M \in \mathscr{C} = \mathscr{C}(n, \Lambda, \delta_0, V_0)$ and $m \in M$, let

$$R_1 = R_1(n, \Lambda, \delta_0, V_0)$$

be a positive number such that on $B(m, R_1)$ there is a coordinate system of almost linear functions $\{l^1, \ldots, l^n\}$ corresponding to an orthonormal basis $\{v_1, \ldots, v_n\}$ of M_n. We shall put successive restriction on R_1 in the following discussion.

Recall from [13] and [14] that if $v_t(x)$ denotes the parallel translation of v_1 to x via the radial geodesic joining m to x, then, for all i and all x in $B(m, R_1)$,

$$(1) \qquad \left| \operatorname{grad} l^i(x) - v_t(x) \right| \leq \left(2 \frac{\sinh 2\Lambda R_1}{\sin 2\Lambda R_1} \right) \Lambda^2 r^2(x)$$

where $r(x) \equiv d(m, x)$. In particular,

$$\left| \operatorname{grad} l^i(x) - v_i(x) \right| \leq C_1(\Lambda R_1) \Lambda^2 R_1^2$$

where $C_1(\Lambda R_1)$ denotes a constant depending only on ΛR, which remains bounded as $R_1 \to 0^+$. Our next requirement on R_1 is that for all i, and all $x \in B(m, R_1)$,

$$(2) \qquad \left| \operatorname{grad} l^i(x) - v_i(x) \right| \leq 1/n.$$

Yet another requirement of R_1 is that there exists on $B(m, R_1)$ a harmonic coordinate system $\{h^1, \ldots, h^n\}$ obtained by solving the Dirichlet problem

$$\Delta h^i \equiv 0, \quad h^i = l^i \quad \text{on } \partial B(m, R_1).$$

Specifically we have estimates

$$(3) \qquad \left| h^i - l^i \right| < C_2(n, \Lambda R_1) \Lambda^2 R_1^3,$$

$$(4) \qquad \left| \operatorname{grad} h^i(x) - v_i(x) \right| < C_3(n, \Lambda R_1) \Lambda^2 R_1^2.$$

Also, by the Lemma of §1, we see that $\{h^1, \ldots, h^n\}$ would be a coordinate system in $B(m, R_1)$ if we had $|\operatorname{grad}(h^i - l^i)| < 1/n$. So we require

$$(5) \qquad C_3 \Lambda^2 R_1^2 < 1/n.$$

It is a straightforward continuity argument to see that there is a (unique) point $m' \in B(m, R_1)$ with $h^i(m') = 0$ for all $i = 1, 2, \ldots, m$. Define for each $\rho < 0$ the sets in $B(m, R_1)$

$$H(m', \rho) = \left\{ x \in B(m, R_1): \sum_i \left[h^i(x) \right]^2 \le \rho^2 \right\}.$$

LEMMA 1. *Let* $\lambda_1 \equiv (\sqrt{n} - 1)/n \in (0, 1)$. *Then for each* R_1 *satisfying* (2) *and* (5), *the set* $H(m', \rho)$ *is disjoint from* $\partial B(m, R_1)$ *if* $\rho \le \lambda_1 R_1$.

Proof. Let $f \equiv \sqrt{\sum_i (h^i)^2}$. We shall show that $f > \lambda_1 R_1$ on $\partial B(m, R_1)$, which will then prove the Lemma. Let $f_1 \equiv \sqrt{\sum_i (l^i)^2}$. It suffices to show that $f_1 > \lambda_1 R_1$ on $\partial B(m, R_1)$ since $h^i = l^i$ on $\partial B(m, R^1)$. Let $\gamma: [0, R_1] \to B(m, R_1)$ be a normal geodesic such that $\gamma(0) = m$ and $\gamma(R_1) = x \in \partial B(m, R_1)$. Since $\sum_i \langle \dot{\gamma}(0), v_i \rangle^2 = 1$, there is some j with $\langle \dot{\gamma}(0), v_j \rangle^2 \ge 1/n$. For this j, since $l^j(m) = 0$,

$$l^j(x) = \int_0^{R_1} \frac{d}{dt} l^j(\gamma(t)) \, dt = \int_0^{R_1} dl^j(\dot{\gamma}(t)) \, dt$$

$$= \int_0^{R_1} \langle \operatorname{grad} l^j, \dot{\gamma}(t) \rangle \, dt = \langle \operatorname{grad} l^j(\gamma(t_0)), \dot{\gamma}(t_0) \rangle R_1$$

for some $t_0 \in [0, R_1]$. Also, for all t

$$\left| \langle \operatorname{grad} l^j, \dot{\gamma}(t) \rangle \right| = \left| \langle v_j(\gamma(t)), \dot{\gamma}(t) \rangle + \langle \operatorname{grad} l^j - v_j(\gamma(t)), \dot{\gamma}(t) \rangle \right|$$

$$> \frac{1}{\sqrt{n}} - \frac{1}{n}$$

by (2). (We have strict inequality because either $\langle \gamma(0), v_i \rangle^2 = 1/n$ for all i or $\langle \gamma(0), v_j \rangle^2 > 1/n$ for some j. In the latter case, we are done. In the former case, we get strict inequality from the Schwarz inequality). Hence

$$f_1(x) \ge \left| l^j(x) \right| > \frac{\sqrt{n} - 1}{n} R_1. \qquad \square$$

When $H(m', \rho)$ is disjoint from $\partial B(m, R_1)$, we call $H(m', \rho)$ a *harmonic ball* associated with $B(m, R_1)$. In this case, $B(m, R_1)$ is considered to come equipped with harmonic coordinates constructed in the specific way described.

We shall need to shift from metric balls to harmonic balls and vice-versa. Lemma 1 shows one direction to be possible. The following lemma gives the other direction.

LEMMA 2. *Let* $\lambda_2 \equiv 1 - 1/n \in (0,1)$. *Then for all positive* $\rho < \lambda_1 R_1$, *the ball* $B(m', \lambda_2\rho)$ *is disjoint from* $\partial H(m', \rho)$.

Proof. Recall that on $B(m, R_1)$ if $g = \sum_{i,j} g_{ij} dh^i dh^j$ then

$$\left| g_{ij} - \delta_{ij} \right| \le nC_4\Lambda^2 R_1^2 / \left(1 - nC_4\Lambda^2 R_1^2 \right)$$

for some constant $C_4(n, \Lambda R_1)$, which remains bounded as $R_1 \to 0$. We now require in addition to previous requirements that R_1 be sufficiently small so that on $B(m, R_1)$:

$$(6) \qquad\qquad \left| g_{ij} - \delta_{ij} \right| < \frac{1}{n^2}$$

for all i, j. The Schwarz Inequality shows:

$$\left| \sum_{i,j} (g_{ij} - \delta_{ij}) x^i x^j \right| \le \left| \sum_j \left(\sum_i (g_{ij} - \delta_{ij})^2 x_i^2 \right) \right|^{1/2} \cdot \left| \sum_j (x^j)^2 \right|^{1/2}$$

$$< \frac{1}{n} |x|^2$$

so that

$$\sum g_{ij} x^i x^j > \left(1 - \frac{1}{n} \right) |x|^2.$$

Hence any curve from m' to $\partial H(m', \rho)$ has length $> (1 - 1/n)\rho = \lambda_2\rho$. □

LEMMA 3. $d(m, m') < (1 - \lambda_1\lambda_2) R_1$.

Proof. By Lemma 2, $B(m', \lambda_1\lambda_2 R_1)$ is disjoint from $\partial H(m', \lambda_1 R_1)$ and by Lemma 1 $H(m', \lambda_1 R_1)$ is itself disjoint from $\partial B(m, R_1)$. □

From now on, a *harmonic ball* $H(m', \rho)$ will always be a harmonic ball associated with some $B(m, R_1)$. We now impose one more requirement on R_1: From (3), we have that on $B(m, R_1)$

$$\left| h^i - l^i \right| < \left(C_2\Lambda^2 R_1^2 \right) R_1.$$

The new requirement is that

$$(7) \qquad\qquad C_2\Lambda^2 R_1^2 < \frac{1}{2n}\lambda_1.$$

Then

$$\left| h^i(m) \right| = \left| h^i(m) - l^i(m) \right| < \frac{1}{2n}\lambda_1 R_1$$

so that

$$\sqrt{\sum_l \left(h^l(m) \right)^2} < \frac{1}{2}\lambda_1 R_1$$

and hence

$$m \in H(m', \Lambda_1 R_1).$$

Since $H(m', \lambda_1 R_1)$ is a harmonic ball, Lemma 1 implies the following:

LEMMA 4. *Every $x \in M$ is in a harmonic ball $H(m', \frac{1}{2}\lambda_1 R_1)$ such that $H(m', \lambda_1 R_1)$ is itself a harmonic ball.*

LEMMA 5. *Let $\lambda_3 = n/(n + 1) \in (0, 1)$. If $H(m', \rho)$ is a harmonic ball, then $H(m', \rho) \subseteq B(m', \rho/\lambda_3)$.*

Proof. As in the proof of Lemma 2, the inequality (6) leads to

$$\sum_{i,j} g_{ij} x^i x^j < \left(1 + \frac{1}{n} \right) |x|^2$$

so that any radial curve (in the harmonic coordinate system) from m' to $\partial H(m', \rho)$ has length $< (1 + 1/n)\rho = \rho/\lambda_3$.   □

LEMMA 6. *Let $\lambda_4 \equiv \frac{1}{3}\lambda_1\lambda_2\lambda_3 \in (0, 1)$. If $H_1(p_1, \lambda_4 R_1)$ and $H_2(p_2, \lambda_4 R_1)$ are two harmonic balls associated with $B(m_1, R_1)$ and $B(m_2, R_1)$ such that $H_1 \cap H_2 \neq \varnothing$, then*

$$H_1(p_1, \lambda_4 R_1) \cup H_2(p_2, \lambda_4 R_1) \subseteq B(m, R_1) \cap B(m_2, R_1).$$

Proof. By Lemma 5, it suffices to show that

$$B\left(p_1, \frac{\lambda_4}{\lambda_3} R_1 \right) \cup B\left(p_2, \frac{\lambda_4}{\lambda_3} R_1 \right) \subseteq B(m_1, R_1) \cap B(m_2, R_1).$$

Let

$$x \in B\left(p_2, \frac{\lambda_4}{\lambda_3} R_1 \right), \qquad y \in B\left(p_1, \frac{\lambda_4}{\lambda_3} R_1 \right) \cap B\left(p_2, \frac{\lambda_4}{\lambda_3} R_1 \right).$$

Then

$$d(x, m_1) < d(x, p_2) + d(p_2, y) + d(y, p_1) + d(p_1, m_1)$$

$$< \frac{3\lambda_4}{\lambda_3} R_1 + (1 - \lambda_1\lambda_2) R_1 = R_1.$$ □

We can summarize the previous discussion as follows: Let $\lambda = \min(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, so $0 < \lambda < 1$. Let R_1 be sufficiently small in the sense of the previous requirements (2), (5), (6), and (7). Then

(i) Each harmonic ball $H(m', \lambda R_1)$ is diffeomorphic with $B(0, \lambda R_1) \subset R^n$.

(ii) For each $x \in M$, there is a harmonic ball $H(m', \lambda R_1)$ with $x \in H(m', \frac{1}{2}\lambda R_1)$.

(iii) If $H(m', \rho)$ is a harmonic ball, then $B(m', \lambda\rho) \subseteq H(m', \rho) \subseteq B(m', \rho/\lambda)$.

(iv) If $H(m_1', \rho)$ is a harmonic ball in $B(m_1, R_1)$, $H(m_2', \rho)$ a harmonic ball in $B(m_2, R_1)$ and if $H(m_1', \rho) \cap H(m_2', \rho) \neq \varnothing$, and $\rho \leq \lambda R_1$, then

$$H(m_1', \rho) \cup H(m_2', \rho) \subseteq B(m_1, R_1) \cap B(m_2, R_1).$$

We now begin the proof proper of the Theorem. Let $R = \lambda^4 R_1/100$. Given a sequence $\{M^k\}$ in $\mathscr{C}$, let $Q(k)$ be the maximum number of disjoint harmonic balls of radius $R/4$ in M^k. By (iii), each such harmonic ball contains a geodesic ball of radius $\lambda R/4$ so there are at least $Q(k)$ disjoint geodesic balls of this radius in M^k. By the standard Bishop-Gromov packing argument (cf. [12], pp. 65–66), $Q(k)$ is thus bounded by a constant independent of k. By passing to a subsequence if necessary, we assume $Q(k) \equiv Q \in Z^+$ for all k. Note that Q has an upper bound depending only on n, Λ, δ_0 and R (see again [12], pp. 65–66).

Now fix k, and let $\{H(z_i, R/4)\}$, $i = 1, \ldots, Q$ be a maximal family of disjoint harmonic balls of radius $R/4$.

Claim. $\{H(z_i, R/2\lambda^2)\}$, $i = 1, \ldots, Q$ is a covering of M^k.

Proof of Claim. Suppose not. By (iii), $B(z_i, R/2\lambda) \subseteq H(z_i, R/2\lambda^2)$ so that $\{B(z_i, R/2\lambda)\}$ is then a disjoint family of geodesic balls in M^k which do not together cover M^k. Let $x \in M$, $x \notin$ union of these geodesic balls. Then $d(x, z_i) \geq R/2\lambda$ for all $i = 1, \ldots, Q$. By (iii) again, $H(z_i, R/4) \subseteq B(z_i, R/4\lambda)$ for each i, so that

$$B(x, R/4\lambda) \cap H(z_i, R/4) = \varnothing.$$

Since $H(x, R/4) \subseteq B(x, R/4\lambda)$ by (iii) again, it follows that $H(x, R/4) \cap H(z_i, R/4) = \varnothing$ for all i, contradicting the maximality of the family $\{H(z_i, R/4)\}$. $\square$

Now let $R_0 \equiv R/\lambda^2$. Then $\{H(z_i, R_0)\}$, $1 \leq i \leq Q$, is a covering of M^k for each k; in fact, $\{H(z_i, R_0/2)\}$ is already a covering; and, moreover, $\{H(z_i, \lambda^2 R_0/4)\}$, $1 \leq i \leq Q$, is a disjoint family. From now on,

we denote $H(z_i, R_0) \subseteq M^k$ by H_i^k, since dependence on k needs to be made explicit.

Since the number of H_i^k, k fixed, is independent of k, we can, by passing to a subsequence if necessary, assume that, for fixed i and j either $H_i^k \cap H_j^k = \varnothing$ for all k *or* $H_i^k \cap H_j^k \neq \varnothing$ for all k. We now identify each $H(z_i, 3R_0/\lambda)$ with $B(0, 3R_0/\lambda)$ in $\mathbf{R}^n$ via the given harmonic coordinates in $H(z_i, 3R_0/\lambda)$. Call this identification map

$$\eta_i \colon B(0, 3R_0/\lambda) \to H(z_i, 3R_0/\lambda)$$

or, if reference to M^k is needed, η_i^k.

Now suppose $H_i^k \cap H_j^k \neq \varnothing$; then

(v) $H_i^k \cup H_j^k \subseteq H(z_i, 3R_0/\lambda) \cap H(z_j, 3R_0\lambda)$.

The proof of (v) is the same as (iv), using (iii). Thus we have C^∞-embeddings

(vi) $\eta_j^{-1} \circ \eta_i \colon B(0, R_0) \to B(0, 3R_0/\lambda)$.

Note that $3R_0/\lambda < \lambda R_1$ so that $H(z^i, 3R_0/\lambda)$ is a harmonic ball, for every i. We claim further that: There is a positive constant $C(n, \Lambda, \delta_0, V_0)$ such that

(vii) $\|\eta_j^{-1} \circ \eta_i\|_{C^{2,\alpha}} < C(n, \Lambda, \delta_0, V_0)$.

(Here $\alpha \in (0, 1)$ is fixed.) Note that the C^0-norm is bounded universally because $\eta_j^{-1} \circ \eta_i$ maps into $B(0, 3R_0/\lambda)$.

The reason that its first derivatives are uniformly bounded is more subtle: Let the harmonic coordinates of H_i and H_j be, respectively,

$$\{h_i^\mu\}, 1 \le \mu \le n \quad \text{and} \quad \{h_j^\nu\}, \quad 1 \le \nu \le n.$$

The question is whether the partial derivatives $\partial h_j^\nu / \partial h_i^\mu$, all ν, μ, are uniformly bounded by some $C(n, \Lambda, \delta_0, V_0)$. By (v), the h_j are functions defined on $H(z_i, 6R_0/\lambda)$ and on $H(z_i, 6R_0/\lambda)$ the harmonic coordinates h_i^μ are also valid. Thus the usual Schauder-estimates argument (see ($\natural$) at the end of §1) gives the desired bound on partial derivatives on $H(z_k, 3R_0/\lambda)$ and hence on $H(z_i, R_0) \equiv H_i$. Similarly we get a universal bound on the $C^{1,\alpha}$-norm

$$\left\| \frac{\partial h_j^\nu}{\partial h_i^\mu} \right\|_{C^{1,\alpha}}$$

of partials, proving (vii).

By (vii) and the Ascoli Theorem (and by passing to a subsequence if necessary), we can arrange that, for each fixed pair of indices i, j, the functions $(\eta_j^k)^{-1} \circ \eta_i^k$ converge:

(viii) $(\eta_j^k)^{-1} \circ \eta_i^k \overset{k}{\to} \eta_{ji} \colon B(0, R_0) \to B(0, 3R_0/\lambda)$.

By an elementary argument,

$$\|\eta_{ji}\|_{C^{2,\alpha}} \le C(n, \lambda, \delta_0, V_0).$$

(It is very important here that the estimate is independent of i, j, and k.) In particular, we have for each fixed i, j and any preassigned $\varepsilon > 0$:

(ix) $\|(\eta_j^k)^{-1} \circ \eta_i^k - (\eta_j^l)^{-1} \circ \eta_i^l\|_{C^{2,\alpha}} < \varepsilon$

for all k, l sufficiently large, say $k, l > N_\varepsilon \in Z^+$. We shall choose ε later on.

Now let $M = M_{N_\varepsilon}$ and $\tilde{M} = $ any one of the M^k's, $k > N_\varepsilon$. Let the harmonic balls on M and $\tilde{M}$ be $\{H_i\}$ and $\{\tilde{H}_i\}$, respectively. Recall that

$$B(0, R_0/\lambda) \overset{\eta_i}{\to} H(z_i, 3R_0/\lambda),$$

$$B(0, R_0/\lambda) \overset{\tilde{\eta}_i}{\to} \tilde{H}(\tilde{z}_i, 3R_0/\lambda).$$

Define, for each i, a diffeomorphism $f_i \colon \tilde{H}_i \to H_i$ by $f_i \equiv \eta_i \circ \tilde{\eta}_i^{-1}$; then f_i is exactly the identification of the harmonic balls H_i and $\tilde{H}_i$ via the given harmonic coordinates

$$\{h_i^1, \ldots, h_i^n\} \quad \text{and} \quad \{\tilde{h}_i^1, \ldots, \tilde{h}_i^n\}.$$

Since $H_i \cup H_j \subseteq H(z_j, 3R_0/\lambda)$, we may compare the two maps

$$f_i, f_j \colon \tilde{H}_i \cap \tilde{H}_j \to H(z_j, 3R_0/\lambda).$$

We claim that relative to the given harmonic coordinates on $\{\tilde{H}_i\}_{1 \le i \le Q}$ and on $\{H(z_j, 3R_0/\lambda)\}_{1 \le j \le Q}$, there is a $C_1 = C_1(\Lambda, n, \delta_0, V_0, R_0, \alpha)$ such that

(x) $\|f_i - f_j\|_{C^{2,\alpha}} \le C_1\varepsilon$

where ε is as in (ix). To verify this claim, note first that both f_i and f_j can be extended to be defined on $\tilde{H}_i \cup \tilde{H}_j$ and consider

$$f_i, f_j \colon \tilde{H}_i \to H(z_j, 3R_0/\lambda).$$

By definition, we have to show that the Euclidean $C^{2,\alpha}$-norm satisfies

$$\left\|\eta_j^{-1} \circ f_i \circ \tilde{\eta}_i - \eta_j^{-1} \circ f_j \circ \tilde{\eta}_i\right\|_{C^{2,\alpha}} \le \varepsilon$$

(here C_1 is not needed). But again by definition this equals

$$\left\|\eta_j^{-1} \circ \eta_i - \tilde{\eta}_j^{-1} \circ \tilde{\eta}_i\right\|_{C^{2,\alpha}}$$

so the desired inequality follows from (ix).

We must now show that (x) holds if $H(z_j, 3R_0/\lambda)$ is replaced by $H(z_i, 3R_0/\lambda)$, i.e.,

$$f_i, f_j \colon \tilde{H}_i \to H(z_i, 3R_0/\lambda).$$

In this case,

$$\eta_i^{-1} \circ f_i \circ \tilde{\eta}_i - \eta_i^{-1} \circ f_j \circ \tilde{\eta}_i = I - \left(\eta_i^{-1} \circ \eta_j\right) \circ \left(\tilde{\eta}_j^{-1} \circ \tilde{\eta}_i\right)$$

$$= \left(\tilde{\eta}_i^{-1} \circ \tilde{\eta}_j - \eta_i^{-1} \circ \eta_j\right) \circ \left(\tilde{\eta}_j^{-1} \circ \tilde{\eta}_i\right).$$

In view of (vii) and (ix), the inequality

$$\left\| \eta_i^{-1} \circ f_i \circ \tilde{\eta}_i - \eta_i^{-1} \circ f_j \circ \tilde{\eta}_i \right\|_{C^{2,\alpha}} \le C_1 \varepsilon$$

follows from the following elementary observation: If $R^m \overset{F}{\to} R^n \overset{G}{\to} R^k$ are maps with

$$\|F\|_{C^{2,\alpha}} \le A \quad \text{and} \quad \|G\|_{C^{2,\alpha}} \le B$$

then (assuming $A, B < 1$)

$$\|G \circ F\|_{C^{2,\alpha}} \le B\left(1 + 3A + 2A^2 + A^{1+\alpha} + A^{2+\alpha}\right).$$

(*Proof.*

$$\|G \circ F\|_{C^0} \le B, \qquad \|D(G \circ F)\|_{C^0} \le AB,$$

$$\|D^2(G \circ F)\|_{C^0} \le A^2 B + AB,$$

$$\|D^2(G \circ F)\|_{C^\alpha} \le A^{2+\alpha} B + A^2 B + A^{1+\alpha} B + AB.)$$

With these preliminaries disposed of, we are now going to construct a diffeomorphism $f \colon \tilde{M} \to M$ from the local mappings f_i by using the standard center of mass technique ([11]; [3], [17]). The possibility of doing so will depend on the ε, and hence N_ε, of earlier in this section being chosen sufficiently small (and large, respectively). To avoid circular reasoning, we shall of course choose ε depending only on n, Λ, δ_0, V_0, R_1, R_0 and the various estimation constants already determined. In particular, ε will not depend on N_ε nor $\tilde{M}$ nor M, except inasmuch as these latter two must belong to $\mathscr{C}$.

To construct the diffeomorphism from $\tilde{M}$ to M, we begin by choosing a partition of unity on $\tilde{M}$ as follows: Take C_0^∞ functions ψ_i, $i = 1, 2, \ldots, Q$ such that

$$0 \le \psi_i \le 1,$$

$$\operatorname{supp} \psi_i \subseteq \tilde{H}(\tilde{z}_i, 3R_0/4),$$

$$\psi_i | \tilde{H}(\tilde{z}_i, R_0/2) \equiv 1,$$

$$\|d\psi_i\|_{C^{2,\alpha}} \le C(n, \Lambda, \delta_0, V_0, R, R_0).$$

These conditions can be realized by taking each ψ_i to be a fixed (independent of i) function of $\sum_j (\tilde{h}^j)^2$ where the $\tilde{h}$'s are the harmonic coordinates associated to the ith harmonic ball. Now set $\phi_i \equiv \psi_i/(\sum_j \psi_j)$. The $\{\phi_i\}$

form a well-defined partition-of-unity, since $\Sigma_j \psi_j \geq 1$ everywhere (be-cause $H(z_i, R_0/2)$ is a covering). Also because $\Sigma_j \psi_j \geq 1$ everywhere, it follows that $\|\phi_i\|_{C^{2,\alpha}}$ are also bounded by a bound depending only on n, Λ, δ_0, V_0, R_1 and R_0.

Now we add one more requirement on R: that $B(x, 10R_0/\lambda^2)$ is convex for all $x \in M$, $M \in \mathscr{C}$.

We are now prepared to define the diffeomorphism $f: \tilde{M} \to M$ out of the $\{f_i\}$. Without loss of generality (by renumbering if necessary) we define $f(x)$ for $x \in \tilde{H}_1$ with $x \in \tilde{H}_1 \cap \cdots \cap \tilde{H}_l$ and x in no other $\tilde{H}$'s. Then

$$\{f_1(x), \ldots, f_l(x)\} \subseteq H(z_1, 3R_0/\lambda).$$

We use the fact that the maps

$$f_1, \ldots, f_l: \tilde{H}_1 \to H(z_1, 3R_0/\lambda)$$

are C^1-close (because ε is small; how small is needed will be discussed in a moment). We set $f(x) =$ the center of mass in M of $f_1(x), \ldots, f_l(x)$ with weights respectively $\phi_1(x), \ldots, \phi_l(x)$, i.e., $f(x)$ is the unique point in $B(z_1, 3R_0/\lambda^2)$, which contains $H(z_1, 3R_0/\lambda)$, such that

$$(8) \qquad \sum_{j=1}^{l} \phi_j(x)\exp_{f(x)}^{-1}f_j(x) = 0.$$

(With R small, depending only on n, Λ, δ_0, V_0, this is well-defined and depends C^∞ on x, by standard results [11].) We now assert:

(xi) There exists a positive constant $C_2 = C_2(n, \Lambda, \delta_0, V_0, R_0, \alpha)$ such that for all $\tilde{x} \in \tilde{M}$ and for all unit $\tilde{v} \in \tilde{M}_{\tilde{x}}$, $|df(\tilde{v})| \geq C_2$.

Proof of (xi). We begin with the elementary observation that if we have C^∞-mappings $M \xrightarrow{F} X \times Y \xrightarrow{G} N$ between manifolds such that $F(m) = (x, y)$ and $G(x, y) = n$ and such that for $v \in M_m$, $dF(v) = (v_1, v_2) \in (X \times Y)_{(x,y)}$ and $G_1: X \to N$, $G_2: Y \to N$ are C^∞-mappings defined by $G_1(x') = G(x', y)$, $G_2(y') = G(x, y')$ for all $x' \in X$, $y' \in Y$, then

$$(9) \qquad d(G \circ F)(v) = dG_1(v_1) + dG_2(v_2).$$

This is just the chain rule.

Notation as above, fix $\tilde{x} \in \tilde{M}$ and define maps

$$\tilde{M} \xrightarrow{\xi_i} M \times M \xrightarrow{\Xi} TM$$

(TM denotes the tangent bundle of M) by: $\xi_i(x) = (f(x), f_i(x))$ and $\Xi(x_1, x_2) = \exp_{x_1}^{-1} x_2$. Furthermore, let $\Xi_i^1: M \to TM$ and $\Xi^2: M \to TM$ be defined by $\Xi_i^1(x) = \exp_x^{-1} f_i(\tilde{x})$ and $\Xi^2(x) = \exp_{f(\tilde{x})}^{-1} x$ respectively.

With $\tilde{v} \in \tilde{M}_{\tilde{x}}$ and $|\tilde{v}| = 1$ as given, we have by (9) that

(10) $\qquad d(\Xi \circ \xi_i)(\tilde{v}) = d\Xi_i^1(df(\tilde{v})) + d\Xi^2(df_i(\tilde{v})).$

Differentiating (8) and using (10), we obtain

(11) $\left| \sum_j \phi_j(\tilde{x}) \cdot d\Xi_j^1(df(\tilde{v})) \right|$

$$= \left| \sum_j \phi_j(\tilde{x}) \cdot d\Xi^2(df_j(\tilde{v})) + \sum_j d\phi_j(\tilde{v}) \cdot (\Xi \circ \xi_j)(\tilde{x}) \right|$$

$$\geq \left| \sum_j \phi_j(\tilde{x}) \cdot d\Xi^2(df_j(\tilde{v})) \right| - \left| \sum_j d\phi_j(\tilde{v}) \cdot (\Xi \circ \xi_j)(\tilde{x}) \right|.$$

We shall estimate the three sums separately.

We claim that there exists a positive constant $C_3 = C_3(n, \Lambda, \delta_0, V_0)$ such that if $x \in M$ and $v \in M_x$, then:

(12) $\qquad\qquad\qquad\qquad \left| d\Xi_i^1(v) \right| \leq C_3|v|.$

To see this, let $\bar{x} = f_i(\tilde{x})$ and let $\gamma: [0, b] \to M$ be a normal geodesic (i.e., $|\dot{\gamma}| \equiv 1$) such that $\dot{\gamma}(0) = v$. Furthermore, let $U(t)$ be the vector field along γ such that $U(t) = \exp_{\gamma(t)}^{-1} \bar{x}$. Then $\Xi_i^1(\gamma(t)) = U(t)$ and $d\Xi_i^1(v) = \dot{U}(0)$, the latter denoting the initial tangent vector of the curve $t \mapsto U(t)$ in TM. Recall that the Riemannian connection on TM induces an orthogonal decomposition of each tangent space of TM into a vertical component V along the fibres and a horizontal component H transversal to the fibres. Thus relative to this metric on TM:

(13) $\qquad \left| \dot{U}(0) \right|^2 = \left| \dot{U}(0)^H \right|^2 + \left| \dot{U}(0)^V \right|^2 = \left| \dot{\gamma}(0) \right|^2 + \left| D_{\dot{\gamma}(0)} U \right|^2,$

the last equality being a well-known interpretation of the connection D. Now consider the 1-parameter family of geodesics $\zeta_t: [0, 1] \to M$ such that $\zeta_t(s) = \exp_{\gamma(t)}(1 - s)U(t)$, and let T be the transversal vector field of this family along the base curve ζ_0 (which joins $\bar{x}$ to x). Note that $T(1) = \dot{\gamma}(0)$, so that

(14) $\qquad D_{\dot{\gamma}(0)} U = D_T U \big|_{s=1, t=0} = D_U T \big|_{s=1, t=0} = \dot{T}(1),$

where $\dot{T}(s) = D_{\zeta_0(s)} T$. We now proceed to estimate $|\dot{T}(1)|$ by the method of §6 in Buser-Karcher [3]. By construction, we have

(15) $\qquad\qquad\qquad T(0) = 0, \qquad \dot{T}(0) = d\exp_{\bar{x}}^{-1} v.$

Let $E(s)$ be any unit parallel vector field along ζ_0. Then

$$\langle s\dot{T}(s) - T(s), E(s) \rangle' = s\langle \ddot{T}(s), E(s) \rangle = -s\langle R_{\zeta_0 T}\dot{\zeta}_0, E(s) \rangle,$$

because $T(s)$ is a Jacobi field along ζ_0. Hence

$$\langle \dot{T}(1) - T(1), E(1) \rangle = \int_0^1 -s \langle R_{\zeta_0 T} \dot{\zeta}_0, E \rangle,$$

$$|\langle \dot{T}(1) - T(1), E(1) \rangle| \le \Lambda^2 \|T\| |\dot{\zeta}_0|^2 = \Lambda^2 \|T\| d^2(x, \bar{x}),$$

where $\|T\| \equiv \max_{0 \le s \le 1} |T(s)|$. By Rauch's comparison theorem and (15), we have

$$\|T\| \le \frac{1}{\Lambda d(x, \bar{x})} \cdot |\exp_{\bar{x}}^{-1} v| \cdot \sinh(\Lambda d(x, \bar{x}))$$

$$\le \frac{\sinh \Lambda d(x, \bar{x})}{\sin \Lambda d(x, \bar{x})} \cdot |v|.$$

This together with the fact that $E(1)$ is arbitrary show that

$$|\dot{T}(1) - T(1)| \le \Lambda^2 d(x, \bar{x}) \cdot \frac{\sinh \Lambda d(x, \bar{x})}{\sin \Lambda d(x, \bar{x})} \cdot |v|$$

$$\le \Lambda^2 \delta_0 \frac{\sinh \Lambda \delta_0}{\sin \Lambda \delta_0} \cdot |v|,$$

$$|\dot{T}(1)| \le \Lambda^2 \delta_0 \frac{\sinh \Lambda \delta_0}{\sin \Lambda \delta_0} \cdot |v| + \frac{\sinh \Lambda \delta_0}{\sin \Lambda \delta_0} \cdot |v|$$

$$= \left[(1 + \Lambda^2 \delta_0) \frac{\sinh \Lambda \delta_0}{\sin \Lambda \delta_0} \right] \cdot |v| \equiv c' |v|.$$

By (13) and (14), we obtain finally

$$|d\Xi_i^1(v)|^2 = |\dot{U}(0)|^2 \le (1 + (c')^2) \cdot |v|^2,$$

which proves (12).

We next analyze $d\Xi^2(df_i(\tilde{v}))$. By (x), we see that given ε_1, we may assume that for all unit tangent vectors $\tilde{v}$ in $\tilde{M}$,

(10) $$|df_i(\tilde{v}) - df_j(\tilde{v})| < \varepsilon_1$$

for all i and j, where the inequality is understood in the sense that if $\{h^1, \ldots, h^n\}$ are the harmonic coordinates in $B(z_1, 3R_0/\lambda^2)$, and if

$$df_i(\tilde{v}) = \sum_k A_i^k \frac{\partial}{\partial h^k}(f_i(x)),$$

$$df_j(\tilde{v}) = \sum_k B_j^k \frac{\partial}{\partial h^k}(f_j(x)),$$

then $\sum_k (A_i^k - B_j^k)^2 < \varepsilon_1^2$. This can be achieved by taking ε sufficiently small in (ix). We now impose an additional condition on the size of R_1, namely, it should be so small that for all $p, q \in B(z_1, 3R_0/\lambda^2)$, for all $z_1 \in M$, and for all $X \in M_p$, $Y \in M_q$ such that $|X - Y| < \varepsilon_1$ in the sense of (16), the inequality

$$(17) \qquad\qquad |\mathbf{P}X - \mathbf{P}Y| < \varepsilon_2$$

should hold for a preassigned ε_2, where $\mathbf{P}$ denotes parallel translation along the shortest geodesic from p or q to any point $x' \in B(z_1, 3R_0/\lambda^2)$. This can be accomplished if R_0 is sufficiently small (cf. [3], §6.2). The precise requirement on the size of ε_2 will be made presently. In any case, we have that with $x' = f(\tilde{x})$,

$$\left| \mathbf{P}(df_i(\tilde{v})) - \mathbf{P}(df_j(\tilde{v})) \right| < \varepsilon_2$$

holds for all unit vectors $\tilde{v}$ and for all i, j. In view of the known fact that inside $B(x', R)$,

$$(18) \qquad\qquad \left| d\exp_{x'}^{-1} X - \mathbf{P}X \right| \le \left(\frac{\Lambda R}{\sin \Lambda R} - 1 \right) \cdot |X|$$

(cf. [3], 6.4.2), where $\mathbf{P}$ denotes parallel translation back to x' along the shortest geodesic, we see that

$$(19) \qquad \left| d\exp_{x'}^{-1}(df_i(\tilde{v})) - d\exp_{x'}^{-1}(df_j(\tilde{v})) \right|$$

$$\le \left(|df_i(\tilde{v})| + |df_j(\tilde{v})| \right) \left(\frac{\Lambda R}{\sin \Lambda R} - 1 \right) + 2\varepsilon_2.$$

Since $\Xi^2(x) = \exp_{f(\tilde{x})}^{-1} x$, the left side of (19) is just $|d\Xi^2(df_i(\tilde{v})) - d\Xi^2(df_j(\tilde{v}))|$. Moreover, we know by (vii) that for some positive constant $C_4 = C_4(\Lambda, n, \delta_0, V_0)$,

$$|df_i(\tilde{v})| \le C_4 |\tilde{v}| = C_4$$

for all i. Thus given any $\varepsilon_3 > 0$, we may choose ε_2 and R to be so small (but still depending only on $\Lambda, n, \delta_0, V_0$) that, from (19),

$$(20) \qquad\qquad \left| d\Xi^2(df_i(\tilde{v})) - d\Xi^2(df_j(\tilde{v})) \right| \le \varepsilon_3$$

for all $i, j = 1, \dots, l$. Now we claim that there exists a positive integer Φ, Φ depending only on n, Λ and δ_0, such that for each $\tilde{x} \in \tilde{M}$, there are at most Φ of the indices j with the property that $\phi_j(\tilde{x}) \ne 0$. Indeed, given any $r \le \delta_0$, the usual Bishop-Gromov packing argument ([12], pp. 65–66) shows that there is such a constant Φ with the property that any geodesic ball $B(\tilde{x}, r)$ of radius r in $\tilde{M}$ contains at most Φ disjoint geodesic balls of radius $\lambda^4 r/8$ (λ being as in (i)–(iv)). Now suppose for some j, $\phi_j(\tilde{x}) \ne 0$,

then $\tilde{x} \in \tilde{H}(\tilde{z}_j, R_0) \subset B(\tilde{z}_j, R_0/\lambda)$ (by (iii)). This implies $\tilde{H}(\tilde{z}_j, R_0) \subset B(\tilde{x}, 2R_0/\lambda)$. But by the choice of R_0, the $\{\tilde{H}(\tilde{z}_k, \lambda^2 R_0/4)\}$ are disjoint and each $\tilde{H}(\tilde{z}_k, \lambda^2 R_0/4)$ contains $B(\tilde{z}_k, \lambda^3 R_0/4)$ by (iii) again). Thus for all these j's with $\phi_j(\tilde{x}) \neq 0$, the geodesic balls $\{B(\tilde{z}_j, \lambda^3 R_0/4)\}$ form a disjoint family in $B(\tilde{x}, 2R_0/\lambda)$. By the choice of Φ, there can be at most Φ of the former, thereby proving our claim.

Now for a fixed i,

$$\sum_j \phi_j(\tilde{x}) \cdot d\Xi^2\big(df_j(\tilde{v})\big)$$

$$= \bigg(\sum_j \phi_j(\tilde{x})\bigg) \cdot d\Xi^2\big(df_i(\tilde{v})\big)$$

$$+ \sum_{j \neq i} \phi_j(\tilde{x})\big\{ d\Xi^2\big(df_j(\tilde{v})\big) - d\Xi^2\big(df_i(\tilde{v})\big)\big\}$$

so that by (20) and the fact that $\sum_j \phi_j = 1$,

$$(21) \qquad \bigg|\sum_j \phi_j(\tilde{x}) \cdot d\Xi^2\big(df_j(\tilde{v})\big)\bigg| \geq \big|d\Xi^2\big(df_i(\tilde{v})\big)\big| - \Phi\varepsilon_3.$$

Now because of (6), $|\tilde{v}| = 1$ implies that $|df_i(\tilde{v})| \geq C_5$ for a positive constant C_5 depending only on n, Λ, δ_0, V_0, and for all i. Moreover, it follows from (18) that if R is sufficiently small (the smallness of R being to insure the smallness of the right-hand side of (18), and hence still dependent only on n, Λ, δ_0, V_0), $|d\Xi^2(v)| \geq \frac{1}{2}|v|$ for all v. Hence

$$d\Xi^2\big(df_i(\tilde{v})\big) \geq \frac{1}{2}C_5,$$

so that if ε_3 is chosen to be smaller than $\frac{1}{4}(C_5/\Phi)$, we would have from (21) that

$$(22) \qquad \bigg|\sum_j \phi_j(\tilde{x}) \cdot d\Xi^2\big(df_j(\tilde{v})\big)\bigg| \geq \tfrac{1}{4}C_5(n, \Lambda, \delta_0, V_0),$$

where we have written $C_5(n, \Lambda, \delta_0, V_0)$ in place of C_5 for emphasis.

We finally deal with the last sum $|\sum_j d\phi_j(\tilde{v}) \cdot (\Xi \circ \xi_j)(\tilde{x})|$ of (11). The following argument is standard (see [17], for example). By the definition of $\{\phi_j\}$, there is a constant C_6 depending only on n, Λ, δ_0, V_0 such that $|d\phi_j| \leq C_6/R_0$ for all j. Furthermore, by (x), $\|f_i(\tilde{x}) - f_j(\tilde{x})\|_{C^0} \leq C_1\varepsilon$, so that by (6), $d(f_i(\tilde{x}), f_j(\tilde{x})) \leq C_7(n, \Lambda, \delta_0, V_0)\varepsilon$ for some constant C_7 depending only on the parameters indicated. The center of mass $f(\tilde{x})$ is therefore at most of distance $C_7\varepsilon$ from any of the $f_1(\tilde{x}), \ldots, f_l(\tilde{x})$. Hence

$|(\Xi \circ \xi_j)(\tilde{x})| = |\exp^{-1}_{f(\tilde{x})} f_j(\tilde{x})| \le C_7 \varepsilon$. It follows that if we impose one last requirement on ε, namely,

$$\varepsilon < \frac{C_5 R_0}{\Phi C_6 C_7},$$

then we can insure that

(23)
$$\left| \sum_j d\phi_j(\tilde{v}) \cdot (\Xi \circ \xi_j)(\tilde{x}) \right| < \frac{1}{8} C_5(n, \Lambda, \delta_0, V_0).$$

To conclude the proof of (xi), we combine (11), (12), (22) and (23) to obtain:

$$C_3 \Phi |df(\tilde{v})| \ge \left| \sum_j \phi_j(\tilde{x}) \cdot d\Xi^1_j(df(\tilde{v})) \right| \ge \tfrac{1}{4} C_5 - \tfrac{1}{8} C_5 = \tfrac{1}{8} C_5.$$

Hence (xi) holds with $C_2 \equiv C_5(8 C_3 \Phi)^{-1}$. $\square$

In the course of proving (xi), we have specified the size requirement for $\varepsilon = \varepsilon(n, \Lambda, \delta_0, V_0)$. Now fix ε, and fix once and for all $M = M^{N_\varepsilon}$, but consider $\tilde{M} = M^k (k > N_\varepsilon)$ as variable, and when k is varying, write $f = f^k$.

Each $f = f^k \colon M^k \to M$ is an immersion, as noted. In fact, more is true: each f is a diffeomorphism. To see this, note that, by interchanging the roles of M and $\tilde{M}$ in the previous discussion we obtain an immersion $\tilde{f}^k \colon M \to \tilde{M} = M^k$. The composition $f \circ \tilde{f} \colon M \to M$ is of course also an immersion and hence a covering map. By construction, this covering is easily seen to be a homotopy equivalence and hence a diffeomorphism so that f is also necessarily a diffeomorphism.

The diffeomorphisms $\{f^k\}$ are $C^{2,\alpha}$-uniformly bounded in harmonic coordinates. This follows from the facts that the $f_1, \ldots, f_Q$ and the $\phi_1, \ldots, \phi_Q$ are $C^{2,\alpha}$-bounded and (and this is important) that the center of mass construction is being carried out on the *fixed* C^∞ manifold M with fixed C^∞ metric.

We want to analyze the metrics $(f^{-1})^* g_{\tilde{M}}$ on M. For this purpose, we need to estimate the derivative boundedness properties of f^{-1}. First, note that (as usual, in harmonic coordinates):

$$\frac{\partial (f^{-1})^j}{\partial h^i} = \left(\text{polynomial in } \left\{ \frac{\partial f^l}{\partial \tilde{h}^s} \right\} \right) \Big/ \det(df)$$

where $\{h^i\}$ are harmonic coordinates on M, $\{\tilde{h}^s\}$ on M, $\det(df) =$ the Jacobian of f relative to h and $\tilde{h}$ coordinates, and the formula itself just

comes from matrix inversion. Hence

$$\left| \frac{\partial (f^{-1})^j}{\partial h^l} \right| \le C(n, \Lambda, \delta_0, V_0),$$

by the C^1-boundedness of f and the lower bound (xi) on $\|df\|$.

By differentiating the expression for

$$\frac{\partial (f^{-1})^j}{\partial h^i}$$

and applying (xi) and the $C^{2,\alpha}$-estimates of f, we obtain a $C^{2,\alpha}$-bound (in harmonic coordinates) on f^{-1}, this bound depending only on n, Λ, δ_0, V_0 and α.

Altogether, we have proved that $(f^k)^{-1}\colon M \to M^k$ is $C^{2,\alpha}$-bounded relative to harmonic coordinates with a bound independent of $k > N_\varepsilon$. Let g^k be the metric of M^k and for convenience write $F \equiv (f^k)^{-1}$. Then for each $H(z_l, R) \subseteq M$, $1 \le l \le Q$, with harmonic coordinates $\{h^l\}$:

$$F^* g^k = F^* \left(\sum_{i,j} g^h_{ij} d\tilde{h}^i d\tilde{h}^j \right) = \sum_{p,q} \left(\sum_{i,j} g^k_{ij} \frac{\partial F^i}{\partial h^p} \frac{\partial F^j}{\partial h^q} \, dh^p \, dh^q \right)$$

where $F^i = \tilde{h}^i(F)$, $\{\tilde{h}^i\}$ being harmonic coordinates in $M = M^k$. Each of g^k_{ij}, $\partial F^i / \partial h^p$, $\partial F^j / \partial h^q$ is bounded in $C^{1,\alpha}$ norm by a constant depending only on n, Λ, δ_0, V_0, and α. Hence by passing to a subsequence Q times, $F^* g^k$ converges to a $C^{1,\alpha}$ metric. $\qquad\square$

REFERENCES

[1]   M. Berger, *Sur les variétés riemanniennes princées juste au-dessous de 1/4*, Ann. Inst. Fourier (Grenoble), **33** (1983), 135–150.

[2]   L. Bers, F. John and M. Schechter, *Partial Differential Equations*, Wiley-Interscience, New York, 1964.

[3]   P. Buser and H. Karcher, *Gromov's almost flat manifolds*, Astérisque, **81** (1981).

[4]   D. Brittain, *Lecture at Luminy*, France, July 1984.

[5]   J. Cheeger, *Finiteness theorems for Riemannian manifolds*, Amer. J. Math., **92** (1970), 61–74.

[6]   J. Cheeger and M. Gromov, *On the Characteristic Numbers of Complete Manifolds of Bounded Curvature and Finite Volume*, Differential Geometry and Complex Analysis: H. E. Rauch Memorial Volume, I. Chavel and H. M. Farkas, editors, Springer-Verlag, New York-Berlin-Heidelberg, 1985. 115–154.

[7]   D. De Turck and J. Kazdan, *Some regularity theorems in Riemannian geometry*, Ann. Sci. École Norm. Sup. Paris 4e, serie **14** (1981), 249–260.

[8]   O. Durumeric, *A generalization of Berger's almost 1/4-pinched theorem*, I, Bull. Amer. Math. Soc., **12** (1985), 260–264.

[9]   R. Greene and H. Wu, *Embedding of open Riemannian manifolds by harmonic functions*, Ann. Inst. Fourier (Grenoble), **25** (1975), 215–235.

[10] _____, *Whitney's imbedding theorem by solutions of elliptic equations and geometric consequences*, Proc. Sym. Pure Math. Amer. Math. Soc., **27** (1975), 287–295.

[11] K. Grove and H. Karcher, *How to conjugate C^1 close group actions*, Math. Z., **132** (1973), 11–20.

[12] M. Gromov, *Structures métriques pour les variétés riemanniennes*, redige par J. Lafontaine et P. Pansu, Textes math. n^0 1 Cedic-Nathan, Paris, 1981.

[13] J. Jost, *Harmonic mappings between Riemannian manifolds*, Proc. Centre for Math. Analysis, Australian Nat. Univ., **4** (1983).

[14] J. Jost and H. Karcher, *Geometrische methoden zur Gewinnung von a-priori-Schranker für harmonische Abbildungen*, Manuscripta Math., **40** (1982), 27–77.

[15] A. Katsuda, *Gromov's convergence theorem and its application*, Nagoya Math. J., **100** (1985), 11–48.

[16] H. Kaul, *Schranken für die Christoffelsymbole*, Manuscripta Math., **19** (1976), 261–273.

[17] S. Peters, *Cheeger's finiteness theorem for diffeomorphism classes of Riemannian manifolds*, J. Reine Angew. Math., **349** (1984), 77–82.

[18] Y. Shikata, *On a distance function on the set of differentiable structures*, Osaka J. Math., **3** (1966), 65–79.

UNIVERSITY OF CALIFORNIA
LOS ANGELES, CA 90024

AND

UNIVERSITY OF CALIFORNIA
BERKELEY, CA 94720