Abstract. This is the lecture note of 18.966: Geometry of Manifolds II at MIT spring 2019, taught by Professor Tristan Collins. We thank Tristan for the nice lectures. The problem sets of this course can be find at https://math.mit.edu/~tristanc/Math18.966.html.

Contents

0. Introduction 2
1. Review of Riemannian Geometry 2
   1.1. Jacobi Fields 4
   1.2. Geometry of Submanifolds 5
   1.3. Cut Locus 6
2. Comparison Theorems 9
   2.1. Generalized Distance Function 9
   2.2. Riccati Equation 10
   2.3. Riccati Comparison 13
   2.4. Toponogov’s Theorem 16
3. Ricci Spaces 19
   3.1. Bishop-Gromov Theorem 19
   3.2. Rigidity in Bishop-Gromov Theorem 21
   3.3. Rigidity of Myers Theorem 22
4. Gromov-Hausdorff Convergence 23
   4.1. Gromov-Hausdorff Distance 23
   4.2. Compactness of Gromov-Hausdorff Metric Space 24
   4.3. Tangent Cone 26
5. Elliptic PDE Theory 27
   5.1. Maximum Principle 29
   5.2. Elliptic Regularity 32
   5.3. Schauder Theory 32
6. Ricci Spaces II 33
   6.1. Cheeger-Gromoll Splitting Theorem 33
   6.2. Harmonic Coordinates 36
   6.3. $C^{k,\alpha}$ Topology 40

Date: May 14, 2019.
0. Introduction

This course will cover several aspects of comparison geometry and the convergence theory of manifolds as well as some applications. We will generally try to answer the following types of questions:

- How does curvature effect the global geometry of Riemannian manifold $(M, g)$?
- What can be said about the “class of manifolds” with certain geometric properties?

To get a sense of what kind of problems we will be studying, here are a few examples of such theme:

**Theorem 0.1.** (Cartan-Hadamard) If $(M, g)$ is a complete manifold with $K(g) \leq 0$, then for any point $p \in M$, the map $\exp_p : T_p M \rightarrow M$ is a covering map.

**Theorem 0.2.** (Myers) Let $(M, g)$ be a complete manifold with dimension $n$. If $\text{Ric}(g) \geq (n-1)k$ with $k > 0$, then $\text{diam}(M, g) \leq \frac{\pi}{\sqrt{k}} = \text{diam}(S^n, g_k)$

**Theorem 0.3.** (Cheng) In the setting of Myers theorem, if $\text{diam}(M, g) = \frac{\pi}{\sqrt{k}}$, then $(M, g) \cong (S^n, g_k)$.

**Theorem 0.4.** (Cheeger) Consider the set $(M)(d, V, k)$ of Riemannian manifolds $(M, g)$ s.t.

1. $\text{diam}(M, g) \leq d$.
2. $\text{Vol}(M) \leq V$.
3. $|K(g)| \leq k$.

Then $(M)(d, V, k)$ contains only finitely many diffeomorphism classes of manifolds.

All these theorems give some general properties of the manifolds when there are certain restrictions on the geometric quantity. The first two have bounds on sectional curvature and Ricci curvature respectively, while the last two are rigidity and compactness type of results.

1. Review of Riemannian Geometry

We review some basics of Riemannian geometry and fix certain notations throughout this course.

$(M, g)$ stands for a smooth manifold $M$ equipped with a Riemannian metric $g$. Unless stated otherwise, we will assume $(M, g)$ complete. $\nabla$ stands for the Levi-Civita connection on $(M, g)$. The curvature tensor $R$ can be written as:

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z$$
for any vector fields $X, Y, Z$ and $R(X, Y, Z, W) = \langle R(X, Y)Z, W \rangle$.

The Ricci tensor is

$$\text{Ric}(X, Y) = \sum_{i=1}^{n} R(e_i, X, Y, e_i)$$

where $\{e_i\}_{i=1}^{n}$ forms an orthonormal basis at a point.

Suppose for $p \in M$ we have a local coordinate system $(x_1, \ldots, x_n)$, we can express the metric and the curvatures in terms of this coordinates:

- $g_{ij} = \langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \rangle$.
- $R(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_l}) = R_{ijkl}$.
- $R(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}) = \frac{\partial}{\partial x_k} R_{ijkl}$.

Given $X, Y \in T_p M$, we have sectional curvature

$$K(X, Y) = \frac{R(X, Y, Y, X)}{|X|^2|Y|^2 - \langle X, Y \rangle^2}$$

If $X, Y$ are orthonormal, then $K(X, Y) = R(X, Y, Y, X)$. The sectional curvature determines the full curvature tensor $R$ (an exercise from Riemannian geometry class) and $\text{Ric}(X, X) = \sum_{i=1}^{n} K(e_i, X)$ where $e_i$ forms an orthonormal basis at a point.

Throughout the course, we will repetitively see curvature bounds. Here are what we mean by such bounds:

1. $K(g) \geq k$ means that $K(X, Y) \geq k$ for any $p \in M$ and $X, Y \in T_p M$.
2. $\text{Ric}(g) \geq (n - 1)k$ means that $\text{Ric}(X, X) \geq (n - 1)k|X|^2$, or equivalently, we can say $\text{Ric}(g) - (n - 1)K g \geq 0$. Note that $K(g) \geq k$ implies $\text{Ric}(g) \geq (n - 1)k$ since we can pick $\{e_i\}_{i=1}^{n}$ and $X/|X|$ as a set of orthonormal basis.

Recall that a geodesic $\gamma : [0, 1] \to M$ is a piecewise $C^1$ curve which is critical for the length functional, $l(\gamma) = \int_{0}^{1} |\gamma'(t)| \, dt$. For $p \in M$ and $V \in T_p M$, we can define the exponential map $\exp_p(Vt) = \gamma_V(t)$ where $\gamma_V$ is the geodesic with $\gamma_V(0) = p$ and $\gamma'_V(0) = V$. $(M, g)$ complete means that $\exp_p : T_p M \to M$ can be defined on the entire tangent space $T_p M$. More generally, $\exp_p$ is a diffeomorphism from $\{V \in T_p M : |V| < \delta \}$ to its image in $M$ for some $\delta > 0$.

Let $\{e_i\}_{i=1}^{n}$ be a set of orthonormal basis for $T_p M$ and identify $\mathbb{R}^n \cong T_p M$ by $(x_1, \ldots, x_n) \mapsto \sum_{i=1}^{n} x_i e_i$. If $r = \sqrt{\sum x_i^2} < \delta$, then $\exp_p : B_\delta(0) \subset \mathbb{R}^n \to M$ is a diffeomorphism. So $(x_1, \ldots, x_n)$ gives a local coordinates near $p$ and we call it normal coordinates. Let $\tilde{\omega} \in S^{n-1} \subset \mathbb{R}^n$, our next theorem follows immediately from the Gauss lemma.

**Theorem 1.1.** The pullback metric by the exponential map can be written as $\exp_p^* g = dr^2 + \tilde{g}(r, \tilde{\omega})$ where $\tilde{g}(r, \tilde{\omega})$ is a metric on $S^{n-1}$.
We define the geodesic ball \( B_R(p) = \exp_p(B_R(0) \subseteq T_pM) \) and the geodesic sphere \( S_R(p) = \partial B_R(p) = \exp_p(\{|x| = R\} \subseteq T_pM) \). Note that if \( q = \exp_p(y) \), then \( d(p,q) = |y| = r \) whenever \( \exp_p(ty) \) is length minimizing for \( t \in [0,1] \).

Let \((M,g)\) be a complete manifold. We denote the \textit{gradient} of a function \( f : M \to \mathbb{R} \) by \( \nabla f \) or \( \text{grad } f \), which is defined to be

\[
\langle \nabla f, X \rangle = df(X), \quad \forall X \in TM.
\]

In local coordinate, \( \nabla = g^{ij} \partial_i f \partial_j \).

We define the \textit{Hessian} of \( f \) by

\[
\text{Hess}_f = \nabla df \in (T^*M)^\otimes 2.
\]

In local coordinate, \( \text{Hess}_f = (\frac{\partial^2 f}{\partial x_i \partial x_j} - \Gamma^k_{ij} \frac{\partial f}{\partial x_k}) dx_i \otimes dx_j \). Equivalently, we can define Hessian which is the 2 tensor satisfying

\[
\text{Hess}_f(X,Y) = \langle \nabla_x \text{grad } f, Y \rangle, \quad \forall X,Y \in TM.
\]

It is easy to see that Hessian is symmetric.

We define the \textit{Laplacian} \( \Delta f \) to be the trace of the Hessian of \( f \), namely

\[
\Delta f = \sum_{i=1}^n \text{Hess}_f(e_i,e_i)
\]

where \( \{e_i\}_{i=1}^n \) is an orthonormal basis of \( T_pM \).

1.1. \textbf{Jacobi Fields.}

\textbf{Definition 1.2.} Suppose \( \gamma : [0,T] \to M \) is a geodesic. Then a vector field \( J(t) \) along \( \gamma \) is a \textit{Jacobi field} if

\[
\nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} J(t) - R(\dot{\gamma}, J(t))\dot{\gamma} = 0.
\]

The Jacobi field arise a the infinitesimal variations of geodesics. Let us sketch that how to see this. Given \( p \in M \) and \( V \in T_pM \), let \( \gamma(t) = \exp_p(Vt) \). Let \( W(s) : (-\epsilon,\epsilon) \to T_pM \) be a family of vectors, and assume \( W(0) = 0 \). For each \( s_0 \), define \( \gamma(s_0,t) = \exp_p(t(V + W(s_0))) \).

Now we define the variational vector field \( J(t) \) to be

\[
J(t) = \frac{d}{ds}|_{s=0}\gamma(s,t) = d(\exp_p)|_{t(V+W(0))}tW'(0).
\]

Let us check that \( J(t) \) is a Jacobi field. The first order derivative of \( J(t) \) satisfies

\[
\nabla_{\dot{\gamma}} J(t) - \nabla J(t) \dot{\gamma} = [\dot{\gamma}, J(t)]
\]

\[
= [d(\exp_p)|_{tV}V, d(\exp_p)|_{tV}tW'(0)]
\]

\[
= d(\exp_p)|_{tV}[V, tW'(0)] = 0.
\]

Here we use the assumption that \( W(0) = 0 \).
Then we have
\[ \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} J(t) = \nabla_{\dot{\gamma}} \nabla_{J(t)} \dot{\gamma} \]
\[ = R(\dot{\gamma}, J(t)) \dot{\gamma} + \nabla_{J(t)} \nabla_{\dot{\gamma}} \dot{\gamma} \]
\[ = R(\dot{\gamma}, J(t)) \dot{\gamma}. \]

So \( J(t) \) satisfies the Jacobi field equation.

Let us see some examples of Jacobi fields. Recall that we have three model spaces with constant sectional curvature \( K \), denoted by \((M_K, g_K)\):
- \( K < 0 \): the hyperbolic space \( \mathcal{H} = \{ (x_1, \ldots, x_n) \in \mathbb{R}^n : x_n > 0 \} \), with the metric \( g_k = (Kx_n^2)^{-1} \sum_{i=1}^n dx_i^2 \);
- \( K = 0 \): the standard Euclidean space \( \mathbb{R}^n \), with the Euclidean metric \( g_{\text{Euc}} = \sum_{i=1}^n dx_i^2 \);
- \( K > 0 \): the sphere \( S^n = \{ \sum_{i=1}^{n+1} x_i^2 = 1 \} \subset \mathbb{R}^{n+1} \), with the induced metric \( g = K^{-1} g_{\text{Euc}}|_{S^n} \).

The following Lemma (as an exercise) gives a description of the Jacobi fields in these constant curvature spaces.

**Lemma 1.3.** Let \( \gamma(t) \) be a geodesic in \((M_k, g_k)\), \( E_1(t), \cdots, E_{n-1}(t), E_n(t) = \dot{\gamma} \) be parallel orthonormal vector fields along \( \gamma(t) \). Then \( J(t) \) is a Jacobi fields along \( \gamma(t) \) with \( J(t) \perp \dot{\gamma}(t) \) if and only if:
- \( K < 0 \):
  \[ J(t) = \sum_{i=1}^{n-1} \left( a_i \frac{\sinh(\sqrt{-K}t)}{\sqrt{-K}} + b_i \cosh(\sqrt{-K}t) \right) E_i(t). \]
- \( K = 0 \):
  \[ J(t) = \sum_{i=1}^{n-1} (a_i t + b_i) E_i(t). \]
- \( K > 0 \):
  \[ J(t) = \sum_{i=1}^{n-1} \left( a_i \frac{\sin(\sqrt{K}t)}{\sqrt{K}} + b_i \cos(\sqrt{K}t) \right) E_i(t). \]

Note that \( J(0) = 0 \) is equivalent to \( b_i = 0 \) for all \( i \).

**Remark 1.4.** If \( K > 0 \), we can see that the conjugate points are those points at \( t = \pi l/\sqrt{K}, \ l \in \mathbb{Z} \). When \( K = 0 \) and \( K < 0 \), there are no conjugate points.

### 1.2. Geometry of Submanifolds.
Suppose \( N \subset (M, g) \) is a submanifold. Let \( g_N = g|_N \), and \( \nabla^N \) be the Levi-Civita connection of \( g_N \). Denote the \( g \)-orthogonal projection to the tangent bundle of \( N \) by \( \Pi_{TN} \).

**Exercise 1.5.** Check that \( \nabla^n \) is the projection of \( \nabla^M \). i.e. for any \( V, W \in T_pN \),
\[ \nabla^n V = \Pi_{TN}(\nabla^M V). \]

The (extrinsic) shape of the submanifold is determined by the second fundamental forms.
Definition 1.6. The second fundamental form of \( N \subset (M, g) \) is the map
\[
A : T_pN \times T_pN \longrightarrow (T_pN)^\perp \subset T_pM
\]
defined by 
\[
A(X,Y) = \langle \nabla^M_XY - \nabla^N_XY, F \rangle,
\]
where \( \nabla^M_XY = \nabla^M_YX + [X,Y] \) and \( X,Y \in T_pN \) implies that \( [X,Y] \in T_pN \), so 
\[
A(X,Y) = A(Y,X).
\]
i.e. \( A \) is symmetric.

Definition 1.7. The mean curvature \( H \) of \( N \) is the operator 
\[
H : N \longrightarrow (TN)^\perp
\]
defined by 
\[
H(p) = -\text{Tr}_g A.
\]
If \( \{E_i\}_{i=1}^k \) is a local orthonormal frame, then 
\[
H = -\sum_{i=1}^k A(E_i, E_i).
\]

Definition 1.8. Given a local section of \((TN)^\perp\), the shape operator \( S_F : T_pN \longrightarrow T_pN \) is defined by 
\[
\langle S_F(X), Y \rangle = \langle A(X, Y), F \rangle, \quad \forall X,Y \in T_pN.
\]
Equivalently, 
\[
S_F(X) = \Pi_{TN}(\nabla^M_X F).
\]

Example 1.9. If \( N \subset M \) is a codimension 1 hypersurface, then we can find \( \vec{v} \) the local unit normal vector field. Then 
\[
S_{\vec{v}}(X) = -\Pi_{TN}(\nabla^M_X \vec{v}).
\]
We also have 
\[
H = -\text{Tr}(S).
\]

Note that 
\[
\langle S_F(X), Y \rangle = \langle A(X, Y), F \rangle = \langle A(Y, X), F \rangle = \langle S_F(Y), X \rangle.
\]
Thus \( S_F \) is self-adjoint, hence \( S_F \) has eigenvalues \( \kappa_1, \ldots, \kappa_{n-1} \), with unit eigenvectors \( e_1, \ldots, e_{n-1} \). We can check that 
\[
\langle H, F \rangle = -\sum_{i=1}^{n-1} \kappa_i.
\]

1.3. Cut Locus. Let \((M, g)\) be a complete manifold. Given \( p \in M \), we define the cut locus of \( p \) to be 
\[
\text{Cut}(p) = \{ \gamma_V(l(V)) \} \subset M,
\]
where \( \gamma_V(t) \) is a geodesic such that \( \gamma_V(0) = p, \gamma'_V(0) = V \) and \( V \in T_pM \) such that \( |V| = 1; l(V) > 0 \) is the number such that \( \gamma_V \) is length minimizing (i.e. \( d(\gamma_V(t), p) = t \) on \([0, l(V)]\)), and it is not length minimizing on \([0, l(V) + \epsilon)\) for any \( \epsilon > 0 \).

Example 1.10. Let \( M = S^2 \) be the standard unit sphere. Let \( p \) be the north pole. Then for any unit vector \( V \in T_pS^2 \), 
\[
d(p, \gamma_V(t)) = t \text{ as long as } t \leq \pi.
\]
We can conclude that \( \text{Cut}(p) \) is just the south pole.
We also define the cut locus in the tangent space to be
\[ \text{Cut}(T_pM) = \{ l(V) : |V| = 1 \}. \]

Moreover, we define \( U_p \subset T_pM \) to be the set
\[ U_p = \{ sV : V \in T_pM, |V| = 1, s \leq l(V) \}. \]

Now we study some properties of the cut locus and the related geometric objects.

**Lemma 1.11.** For a complete manifold \( M \), \( M = \exp_p(U_p) \cup \partial \exp_p(U_p) \).

**Proof.** Let \( q \in M \) but not in \( \exp_p(U_p) \). \( (M, g) \) complete implies that there exists a length minimizing geodesic \( \exp_p(Vt) \), \( |V| = 1 \), such that \( \exp_p(V(d(p,q))) = q \). Thus for any \( t < d(p,q) \), \( tV \in U_p \). But \( q \) is not in \( \exp_p(U_p) \), which implies that \( d(p,q)V \notin U_p \). So \( q \in \partial \exp_p(U_p) \). \( \square \)

**Lemma 1.12.** If \( \tilde{V} \in \partial U_p = \text{Cut}(T_pM) \), then either:

1. there exists \( W \neq \tilde{V} \) such that \( \exp_p(W) = \exp_p(V) \),
or
2. \( D \exp_p \) is singular at \( \tilde{V} \). i.e. \( \exp_p(\tilde{V}) \) is conjugate to \( p \).

**Proof.** Let \( \gamma(t) = \exp_p(\tilde{V}t) \). For \( t > 1 \), \( \gamma(t) \) is not length minimizing. We can connect \( p \) to \( \exp_p(\tilde{V}t) \) by a length minimizing geodesic \( \sigma_t(s) \) for \( t > 1 \), such that \( \sigma_t(0) = p \), \( \sigma_t(1) = \gamma(t) \).

Since \( \gamma(t) \) is not minimizing for \( t > 1 \), we know \( \sigma_t'(0) \notin \{ \lambda \tilde{V} : \lambda > 0 \} \). Pick a subsequence \( t_n \downarrow 1 \) of \( t \) such that \( \sigma_{t_n}'(0) \) converges to some \( W \in T_pM \). Since the length of \( \sigma_{t_n} \) converges to length of \( \gamma|_{[0,1]} \), we have \( |W| = |\tilde{V}| \).

There are two cases. The first case is that \( W \neq \tilde{V} \). Then (1) holds.

The second case is that \( W = \tilde{V} \). We are going to show (2) holds. Let us argue by contradiction. If (2) does not hold, then \( D \exp_p \) is non-singular at \( \tilde{V} \), then by implicit function theorem \( \exp_p \) is a diffeomorphism in a neighbourhood of \( \tilde{V} \), say \( B_\epsilon(\tilde{V}) \). Since \( \sigma_{t_n}'(0) \) converges to \( \tilde{V} \), for \( t_n \) close to 1 enough, \( \sigma_{t_n}'(0) \) and \( t_n \tilde{V} \) both lie in \( B_\epsilon(\tilde{V}) \). Since under the exponential map the images of them are the same, we have \( \sigma_{t_n}'(0) = t_n \tilde{V} \). Thus \( \gamma(t) \) is length minimizing for \( t \in [0, t_n) \) where \( t_n > 1 \), which is a contradiction. \( \square \)

In other word, this Lemma tells us that \( U_p \subset T_pM \) is the largest set such that \( \exp_p|_{U_p} \) is a diffeomorphism.

Now let us study some specific geometric properties of the cut locus. Since it is a subset of \( M \), we can ask the general questions for a subset, such as whether it is closed, what is the size of it, etc. All the analysis are based on the following Lemma.

**Lemma 1.13.** The function \( V \in S^{n-1} \subset T_pM \rightarrow l(V) \in \mathbb{R}_{>0} \cup \{ \infty \} \) is continuous.

**Proof.** Let \( V_i \in S^{n-1} \) and suppose \( V_i \rightarrow V \). We are going to show \( l(V_i) \rightarrow l(V) \). We divide the proof into two parts.
(1) \(l(V) \geq \limsup_i l(V_i)\): Suppose \(t < l(V_i)\) for infinitely many \(i\), then \(d(p, \exp_p(tV_i)) = t\). Passing to limit we have \(d(p, \exp_p(tV)) = t\). Thus \(l(V) \geq t\), which leads to \(l(V) \geq \limsup_i l(V_i)\).

(2) \(l(V) \leq \liminf_i l(V_i)\): Suppose \(c < l(V)\), then \(\exp_p\) is a diffeomorphism in a neighbourhood of \(cV\). \(V_i \to V\) implies that for \(i\) sufficiently large \(\exp_p\) is also a diffeomorphism in a neighbourhood of \(cV_i\). We argue this for all \(\lambda < c\) and passing to subsequence (still denoted by \(V_i\)) to see that \(D \exp_p\) is non-singular at \(\lambda V_i\).

We are going to prove \(c \leq \liminf_i l(V_i)\), and we argue by contradiction. If \(c > \liminf_i l(V_i)\), there exists infinitely many \(i\) such that \(W_i \neq V_i\), \(|W_i| = 1\) and \(\exp_p(l(V_i)W_i) = \exp_p(l(V_i)V_i)\). Passing to a further subsequence we can find \(W_i \to W\). Note that \(W \neq V\), otherwise local diffeomorphism near \(\lambda V\) would implies \(W_i = V_i\) for \(i\) large enough, which is a contradiction. However, \(\exp_p(l(V_i)W_i) = \exp_p(l(V_i)V_i)\) implies that \(\exp_p(\liminf_i l(V_i)W) = \exp_p(\liminf_i l(V_i)V)\), which is a contradiction to \(c > \liminf_i l(V_i)\).

**Remark 1.14.** We can use the fact that \(\exp_p\) is a diffeomorphism not near a single \(\lambda V\), but in a neighbourhood of whole curve \(\exp_p(tV)\) for \(t < l(V)\).

A direct corollary is that \(U_p \subset T_p M\) is star shaped. Moreover we can obtain the following property of the cut locus:

**Corollary 1.15.** \(\text{Cut}(p) \subset M\) is closed and measure 0.

**Proof.** Let \(\Sigma\) be the set \(\{(l(V), V) \in \text{polar coordinate} : V \in S^{n-1}\} = \partial U_p\). Then by the continuity of \(l(V)\), it is clear that \(\Sigma\) is closed. Moreover, by Fubini theorem, \(\Sigma\) has measure 0.

Since exponential map is smooth, it is clear to see that \(\text{Cut}(p)\) is closed. Moreover, the exponential map is locally Lipschitz, thus \(\text{Cut}(p)\) has measure 0 once we have the following exercise.\(\square\)

**Exercise 1.16.** Suppose \(F : B_1(0) \to B_1(0)\) is Lipschitz, then \(F\) maps measure 0 set to a measure 0 set.

**Remark 1.17.** If \((M, g)\) is compact, then \(l(V) < \infty\) for all \(V \in S^{n-1}\). So \(U_p\) is homeomorphic to a ball. Thus \((M, g)\) is a compactification of \((U_p, dr^2 + g(r, \sigma))\).

Now let us discuss a main application of the study of the cut locus. Let us define the distance function \(d_p = d(p, x) : M \to \mathbb{R}_{\geq 0}\).

**Proposition 1.18.** There exists an open dense set \(U \subset M\) such that \(U^c\) has measure 0, \(d_p\) is a smooth function on \(U\) with \(|\nabla d_p| = 1\).

**Proof.** Just note that for \(p \in M\), \(r\) on \(U_p\) is a smooth function, and \(U_p\) gives the normal coordinate at \(p\). Then \(U = \exp_p(U_p) \setminus \{p\}\) is the desired domain, with \(d_p(x) = r(\exp_p^{-1}(x))\) is smooth (note that the distance function from \(p\) is not smooth at \(p\)). The gradient equality follows from the gradient of \(r\) has length 1 in \(U_p\). \(\square\).
Remark 1.19. The boundary of $U_p$ is continuous, but might not be smooth (actually it is just Lipschitz). However suppose $M$ is compact, then for any $\epsilon > 0$, we can always find $U_{p,\epsilon}$ such that $\partial U_{p,\epsilon}$ is a smooth submanifold, $\text{Vol}(U_p \setminus U_{p,\epsilon}) < \epsilon$, and $\frac{\partial}{\partial r}$ points into $U_p \setminus U_{p,\epsilon}$, i.e. $\langle \frac{\partial}{\partial r}, N_\epsilon \rangle > 0$. Here $N_\epsilon$ is the outward pointing normal vector to $\partial U_{p,\epsilon}$.

If $M$ is non-compact, then the above argument holds for $U_p \cap B_R$, $R > 0$.

The above fact is important in future. We will study the distance function $d_p$ in future, in particular we are interested in the Laplacian of $d_p$. However, $d_p$ is not smooth everywhere, which implies that we should study the “weak” version of the Laplacian, i.e. in the integral sense. Then integration by parts may involved, and the regularity of the boundary is important.

Definition 1.20. Given a point $p \in M$, the injective radius $\text{inj}$ at $p$ is defined to be

$$\text{inj}(p) := \inf \{ l(V) : V \in S^{n-1} \subset T_p M \}.$$ 

The injective radius of the whole manifold $M$ is defined to be

$$\text{inj}(M) := \inf_{p \in M} \text{inj}(p).$$

In other word, $B_{\text{inj}(p)}(0) \subset T_p M$ is the largest ball where the normal coordinate exists.

2. Comparison Theorems

2.1. Generalized Distance Function. Recall that the gradient of a function $f$ is defined to be the vector $\nabla f$ such that

$$\langle \nabla f, X \rangle = df(X)$$

hold for any vector field $X$. In the local coordinate $\nabla f = g^{ij} \partial_i f \partial_j$.

Definition 2.1. Let $U \subset M$ be an open set. A function $f : U \to \mathbb{R}$ is a (local) generalized distance function if $|\nabla f| = 1$.

One example of the generalized distance function is the classical distance function. Given $p \in M$, the distance function $d(x, p)$ on $U(p) \setminus \{p\} = M \setminus (\cup(p) \cup \{p\})$ is a generalized distance function.

The generalized distance function can be used to characterize the geodesics.

Lemma 2.2. If $\gamma$ is an integral curve of a generalized distance $f$, then $\gamma$ is a geodesic.

Proof. Let $V$ be a vector field on $U$, then

$$0 = V \langle \nabla f, \nabla f \rangle = 2 \langle \nabla f, \nabla_V \nabla f \rangle.$$ 

Now we claim that for any $X, Y \in T_p M$, $\langle X, \nabla_Y \nabla f \rangle = \langle Y, \nabla_X \nabla f \rangle$. Let us work in local normal coordinate at $p$, then

$$\langle X, \nabla_Y \nabla f \rangle = X^i g_{ik} (\nabla_y \nabla f)^k = X^i g_{ik} Y^l g^{kp} \partial_l \partial_p f = X^i Y^l \partial_l f.$$
Then by symmetry the claim is clear. Let $X = \nabla f$ and $Y = V$ we obtain
\[ 0 = 2\langle V, \nabla f \nabla f \rangle \]
for any $V$, thus $\nabla f \nabla f = 0$. So any curve $\gamma$ with $\gamma = \nabla f$ is a geodesic. \hfill \square

This lemma tells us that if $f$ is a generalized distance function, the level sets $\{ f^{-1}(c) \} \subset M$ are “parallel hypersurfaces” (at least locally).

Let us study the geometry of these level sets. Suppose $f : M \rightarrow \mathbb{R}$ is a generalized distance function, and let $N = \{ f = c \}$ is a codimension 1 submanifold. Then $\vec{v} = \nabla f$ is a unit normal vector to $N$.

If $X \in T_pN$, the shape operator is given by
\[ S(X) = -\Pi_{TN}(\nabla^M_X \nabla f). \]
Note $\langle \nabla^M_X \nabla f, \nabla f \rangle = 1/2 \nabla_X |\nabla f|^2$, hence $\nabla^M_X \nabla f$ is tangential. Thus
\[ S(X) = -\nabla^M_X \nabla f. \]

**Lemma 2.3.** If $f$ is a local generalized distance function, $N = f^{-1}(c)$. Let $p \in N$. Then the mean curvature of $N$ satisfies $H(p) = \Delta f(p)$.

**Proof.** Let $E_1, \ldots, E_{n-1}$ be an orthonormal basis of $T_pN$, then $E_1, \ldots, E_{n-1}, \nabla f$ is an orthonormal basis of $T_pM$. Recall that $\nabla_{\nabla f} \nabla f = 0$, so we have
\[ \Delta f = \sum_{i=1}^{n-1} \langle \nabla E_i \nabla f, E_i \rangle + \langle \nabla f \nabla f, \nabla f \rangle = -\text{Tr} S = H. \]
\hfill \square

**Example 2.4.** With this Lemma, it would be very easy to compute the mean curvature of the spheres in the Euclidean space. Let $r = \sqrt{\sum_{i=1}^{n} x_i^2}$, then $r$ is the distance function (from the origin). Then the sphere with radius $R$ is the level set $S^n_{n-1} = \{ r = R \}$.

Then the Lemma tells us that the mean curvature of the sphere is given by $\Delta r = (n - 1)/R$.

2.2. **Riccati Equation.** Now we do some computations which are related to the Jacobi Fields. Let $F : M \rightarrow \mathbb{R}$ be a local generalized distance function, then the time “$t$” flow along $\nabla f$ gives a diffeomorphism $E_t$. Note $E_t(\{ f = c \}) = \{ f = c + t \}$.

Last time we proved that the integral curves of $\nabla f$ are geodesics. Hence given a curve $p(s)$ in $\{ f = c \}$, $E_t(p(s))$ is a family of geodesics in $s$. So
\[ \left. \frac{d}{ds} \right|_{s=0} E_t(p(s)) = dE_t(p(0)) = J(t) \]
is a Jacobi field.

Let $\gamma(t) = E_t(p(0))$ is a geodesic. Let $N(p(s))$ be the unit normal to $\{ f = c \}$ at $p(s)$, i.e. $N = \nabla f$. 
Lemma 2.5. \( \nabla_{\dot{\gamma}} J(t) = -S_{\{f=c+t\}}(\dot{p}(0)). \)

Proof. Note that

\[
\nabla_{\dot{\gamma}} J(t) = \frac{d}{dt} |_{t=0} \frac{d}{ds} |_{s=0} \exp_{p(0)}(tN(p(s)))
\]

\[
= \frac{d}{ds} |_{s=0} \frac{d}{dt} |_{t=0} \exp_{p(0)}(tN(p(s)))
\]

\[
= \frac{d}{ds} |_{s=0} N(p(s)) = \nabla_{\dot{p}(0)} N
\]

\[
= -S(\dot{p}(0)).
\]

Now let us take the second order differential. Recall the Jacobi equation is

\[ \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} J(t) = R(\dot{\gamma}, J(t)) \dot{\gamma}. \]

The Lemma implies that \( \nabla_{\dot{\gamma}} J(t) = -S_t(J(t)) \), so we have

\[ \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} J(t) = -[(\nabla_{\dot{\gamma}} S)J(t) + S_t(\nabla_{\dot{\gamma}} J(t))], \]

by the Lemma again this equals

\[ S^2_t(J(t)) - (\nabla_{\dot{\gamma}} S)(J(t)). \]

So

\[ (\nabla_{\dot{\gamma}} S)(J(t)) = S^2_t(J(t)) - R(\dot{\gamma}, J(t)) \dot{\gamma} \]

as long as we stay away from the conjugate locus. Note that \( J(t) \) can be made arbitrary, so we conclude the following Lemma:

Lemma 2.6.

\[ \frac{D}{dt} S_t = S^2_t - R(\dot{\gamma}, \cdot) \dot{\gamma}. \]

Let us give another proof of the above computations.

Fix \( p \in N = N_c \), let \( \gamma(t) \) be the geodesic with \( \gamma(0) = p, \gamma'(0) = \text{grad } f \). Let \( V \in T_pN \), and let \( V(t) \) be the parallel transport of \( V \) along \( \gamma(t) \). Then \( V(t) \in T_{\gamma(t)} N_{c+t} \).

Let \( S_t \) be the shape operator of \( N_{c+t} \). Then we can compute

\[ \nabla_{\dot{\gamma}}(S_t(V(t))) = \nabla_{\dot{\gamma}}(-\nabla_{V(t)} \text{grad } f) = -\nabla_{\text{grad f}} \nabla_{V(t)} \text{grad } f. \]

Here we use that \( \dot{\gamma} = \text{grad } f \). Commute the derivatives we have

\[ \nabla_{\text{grad f}} \nabla_V = \nabla_V \nabla_{\text{grad f}} + \nabla_{[\text{grad f}, V]} + R(\text{grad f}, V). \]

Note \( \nabla_{\text{grad f}} \text{grad f} = \nabla_{\dot{\gamma}} \dot{\gamma} = 0 \) and

\[ [\text{grad f}, V] = \nabla_{\text{grad f}} V - \nabla_V \text{grad f} = -\nabla_V \text{grad f} = S_t(V), \]

so we get

\[ (\nabla_{\dot{\gamma}} S_t)(V(t)) = -R(\text{grad f}, V) \text{grad f} - \nabla_{S_t(V)} \text{grad f} = -R(\text{grad f}, V) \text{grad f} + S^2_t(V). \]
Remark 2.7. Since $S_t$ is a self adjoint operator, $S_t^2$ is positive in the sense that the eigenvalues are squared. Thus the equation

$$\frac{D}{dt} S_t = S_t^2 - R(N, \cdot)N$$

Can be understood as a Riccati equation. The general form of the Riccati equations is

$$f' = f^2 + k.$$  

The following corollary gives a precise way to see this.

Corollary 2.8. Suppose $(M, g)$ has sectional curvature $K(M, g) \geq k$ and $f$ is a local generalized distance function. Let $V(t)$ be a parallel vector fields along $\text{grad } f$, $V(t) \perp \text{grad } f$, $|V| = 1$. Then

$$\frac{d}{dt} \langle S_t(V(t)), V(t) \rangle \geq \langle S_t(V(t)), V(t) \rangle^2 + k.$$  

Proof.  

$$\frac{d}{dt} \langle S_t V(t), V(t) \rangle = \langle \frac{D}{dt} S_t V(t), V(t) \rangle$$

$$= \langle S_t^2 V, V \rangle - \langle R(\text{grad } f, V) \text{grad } f, V \rangle$$

$$= \langle S_t^2 V, V \rangle + K(\text{grad } f, V).$$

So we only need to prove $\langle S_t^2 V, V \rangle \geq \langle S_t V, V \rangle^2$ to conclude the inequality.

Recall that $S : T_pN \to T_pN$ is self adjoint. So we can find an orthonormal basis $e_1, \cdots, e_{n-1}$ in $T_pN$ such that $S e_i = \lambda_i e_i$. Suppose $V = \sum_{i=1}^{n-1} a_i e_i$. $|V| = 1$ implies that $\sum_{i=1}^{n-1} a_i^2 = 1$.

Then

$$\langle S V, V \rangle^2 = (\sum \lambda_i a_i e_i, e_i)^2$$

$$= (\sum \lambda_i a_i^2)^2$$

$$\leq (\sum \lambda_i^2 a_i^2)(\sum a_i^2) = \langle S^2 V, V \rangle.$$ 

Here the inequality comes from Cauchy-Schwartz. \hfill \Box

Now let us compute the evolution of another important quantity, the volume. Let $\{V_1(0), \cdots, V_{n-1}(0)\}$ be an orthonormal basis for $T_p\{f = c\}$, parallel translate them along $\gamma(t)$ ($\gamma'(t) = \text{grad } f$) to get an orthonormal basis $\{V_1(t), \cdots, V_{n-1}(t)\}$ for $T_{\gamma(t)}\{f = c + t\}$. Let $E(t)$ be the map sending $\gamma(0)$ to $\gamma(t)$ for every geodesics starting from $\{f = c\}$. Then we have a map

$$dE_t(p) : \mathbb{R}^{n-1} \to \mathbb{R}^{n-1},$$

where the first $\mathbb{R}^{n-1}$ is spanned by $\{V_1(0), \cdots, V_{n-1}(0)\}$, and the second $\mathbb{R}^{n-1}$ is spanned by $\{V_1(t), \cdots, V_{n-1}(t)\}$. 
Note that $E^*_t g_t = g_t(dE_t, dE_t)$, so the volume form $\sqrt{\det g_t}$ on $\{ f = c + t \}$ are given by $\det dE_t$ in this local frame. Recall that given a matrix $A(t) \in GL(n, \mathbb{R})$, we have

$$\frac{d}{dt} \det(A(t)) = \text{Tr}(A^{-1}(t) \frac{d}{dt} A(t)) \det(A(t)).$$

Recall in Lemma 2.5 we proved that

$$\frac{D}{dt} dE_t = -S_t,$$

So we have

(2.1) $$\frac{d}{dt} \sqrt{\det g_t} = H_t \sqrt{\det g_t}.$$ We can further differentiate the mean curvature to see that the mean curvature also satisfies a Ricatti equation.

$$\frac{d}{dt} H_t = \frac{d}{dt} (-\text{Tr}(S))$$

$$= - \text{Tr}(\frac{D}{dt} S)$$

$$= \text{Tr}(R(\text{grad } f, \cdot) \text{grad } f) - \text{Tr}(S^2)$$

$$= - \text{Ric}(\text{grad } f, \text{grad } f) - \text{Tr}(S^2).$$

Moreover, Cauchy-Schwartz inequality gives

$$(\text{Tr}(S))^2 \leq \text{Tr}(S^2)(n - 1).$$

Thus we conclude that

(2.2) $$\frac{d}{dt} H_t \leq - \text{Ric}(\text{grad } f, \text{grad } f) - \frac{1}{(n - 1)} H_t^2.$$ So the mean curvature also satisfies a Ricatti equation.

2.3. Riccati Comparison. Recall that a Ricatti equation is

$$g' = K + g^2.$$ It has a closed relation to the following second order equation

(2.3) $$f'' + kf = 0$$

Let $c_k$ be the solution to (2.3) with initial condition $c_k(0) = 1$, $c'_k(0) = 0$, and $s_k$ be the solution to (2.3) with initial condition $s_k(0) = 0$, $s'_k(0) = 1$. Then we have

<table>
<thead>
<tr>
<th>the sign of $k$</th>
<th>$c_k$</th>
<th>$s_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k &gt; 0$</td>
<td>$\cos(\sqrt{k}x)$</td>
<td>$\frac{\sin(\sqrt{k}x)}{\sqrt{k}}$</td>
</tr>
<tr>
<td>$k = 0$</td>
<td>1</td>
<td>$x$</td>
</tr>
<tr>
<td>$k &lt; 0$</td>
<td>$\cosh(\sqrt{-k}x)$</td>
<td>$\frac{\sinh(\sqrt{-k}x)}{\sqrt{-k}}$</td>
</tr>
</tbody>
</table>
One important fact is that if \( f \) solves (2.3), then
\[
\left( \frac{f'}{f} \right)' = \frac{f''}{f} - \left( \frac{f'}{f} \right)^2 = -k - \left( \frac{f'}{f} \right)^2.
\]

So \( g = -\frac{f'}{f} \) is a solution to a Ricatti equation.

Now we state the Ricatti comparison theorem.

**Theorem 2.9.** Suppose \( f, F \) are defined on some interval, satisfying
\[
f' \geq K + f^2,
F' \leq K + F^2.
\]
Then
\[
\left[ (f - F)e^{-f(f+F)} \right]' \geq 0
\]
The proof is straightforward so we omit it.

**Corollary 2.10.** We have the following estimates.

(a) If \( f(r_0) \geq F(r_0) \), then \( f(r) \geq F(r) \) for \( r \geq r_0 \);
(b) If \( f(r_0) \leq F(r_0) \), then \( f(r) \leq F(r) \) for \( r \leq r_0 \);
(c) if \( \lim_{r \to 0^+} f(r) = -\infty \), then \( f(r) \geq -\frac{s_k'(r)}{s_k(r)} \) as long as \( f(r) \neq -\infty \);
(d) if \( \lim_{r \to 0^+} F(r) = -\infty \), then \( F(r) \leq -\frac{s_k'(r)}{s_k(r)} \) as long as \( s_k'(r) \neq \infty \).

**Proof.** The proof of (a) and (b) are easy applications of the Ricatti comparison theorem, so we omit it here. (d) is the same as (c), so we only prove (c) here. The proof of (c) uses an idea of perturbation, which is very useful in future.

Let us argue by contradiction. Suppose there is \( r_0 > 0 \) such that
\[
f(r_0) < -\frac{s_k'(r_0)}{s_k(r_0)}.
\]
Then we can find \( \epsilon > 0 \) small enough such that
\[
f(r_0) \leq -\frac{s_k'(r_0 - \epsilon)}{s_k(r_0 - \epsilon)}.
\]
Then (b) implies that
\[
f(r) \leq -\frac{s_k'(r - \epsilon)}{s_k(r - \epsilon)}
\]
for all \( r \leq r_0 \). But let \( r \to \epsilon^+ \) we get a contradiction. \(\square\)

Now we are going to use Ricatti comparison theorem to give some applications to geometry. Fix \( p \in M \). Let us choose an open neighbourhood \( U \) of \( p \) which does not contain any conjugate points of \( p \). Then the distance function \( r = r|_U \) is a generalized distance function, and we can apply our comparison analysis to this local distance function.
Lemma 2.11. Suppose \((M, g)\) has sectional curvature \(K \geq \rho > 0\). Let \(\kappa_i(t)\) be the principal curvature of the level sets \(\{r = t\} \cap U\). Then

\[
\kappa_i(t) \geq -\frac{s'_\rho(t)}{s_\rho(t)}.
\]

Remark 2.12. Suppose \((M, g) = (M_\rho, g_\rho)\) is the space of constant curvature, then the principal curvatures of the distance spheres are exactly \(-\frac{s'_\rho(t)}{s_\rho(t)}\).

Remark 2.13. From this principal curvature estimate, we obtain a diameter estimate. When \(\rho > 0\), \(-\frac{s'_\rho(t)}{s_\rho(t)}\) blows up as \(t\) goes to \(\pi \sqrt{\rho}\). Thus the distance from any point in \(M\) to a conjugate point is at most \(\pi \sqrt{\rho}\). This can be viewed as one explanation of the Bonnet-Myers theorem.

Proof. Let \(\gamma(t)\) be a geodesic, \(U(t)\) be any vector field which is parallel along \(\gamma(t)\), \(|U| = 1\), \(U(t) \perp \dot{\gamma}(t)\). Then we previously showed

\[
\frac{d}{dt}\langle SU, U \rangle \geq -\langle R(\dot{\gamma}, U) \dot{\gamma}, U \rangle + (\langle SU, U \rangle)^2,
\]

so \(\langle SU, U \rangle\) satisfies the Ricatti inequality

\[
\frac{d}{dt}\langle SU, U \rangle \geq \rho + (\langle SU, U \rangle)^2.
\]

Moreover, \(\langle SU, U \rangle\) goes to \(-\infty\) as \(t\) goes to 0. This is because locally \(M\) is almost Euclidean, and the shape operator over the sphere with small radius is super negative. Thus by Corollary 2.10 (c) we conclude that

\[
\langle SU(t), U(t) \rangle \geq -\frac{s'_\rho(t)}{s_\rho(t)}.
\]

Since \(U\) is arbitrary, this implies that the principal curvature is always bounded from below by \(-\frac{s'_\rho(t)}{s_\rho(t)}\). \(\square\)

In the end of this section, we give a final application of the comparison theorem.

Theorem 2.14 (Generalized Rauch Theorem). Suppose \((M, g)\) has \(K(M, g) \geq \rho\). Suppose \(J(t)\) is a Jacobi field along a geodesic \(\gamma(t)\), \(|\gamma'| = 1\). Assume \(J(0) = 0\). Suppose \(\gamma(t)\) does not meet the conjugate locus of \(\gamma(0)\) for \(t < r_c\). Then

\[
\frac{|J(t)|}{s_\rho(t)}
\]

is non-increasing for \(0 < t < r_c\).

In particular we have Rauch theorem:

\[
|J(t)| \leq |J(0)| s_\rho(t).
\]
Proof. 

\[
\frac{d}{dt} \left| J(t) \right|_{s \rho(t)} = \frac{\langle J'(t), J(t) \rangle_{s \rho(t)}}{s \rho(t) |J(t)|} - \frac{s'_\rho(t) |J(t)|}{s^2_\rho(t)} = \frac{|J|}{s \rho(t)} \left( \frac{\langle J, J' \rangle}{|J|^2} - \frac{s'_\rho(t)}{s \rho(t)} \right).
\]

Recall \( J'(t) = S(J(t)) \) where \( S \) is the shape operator of \( \{r = t\} \). Recall that \( -S \) has eigenvalues \( \kappa_1, \ldots, \kappa_{n-1} \) and we proved in Lemma 2.11 that \( -\kappa_i \leq \frac{s'_\rho}{s \rho} \). Thus we have

\[
\frac{d}{dt} \left| J(t) \right|_{s \rho(t)} \leq 0
\]

and the monotonicity is clear. \( \square \)

2.4. Toponogov’s Theorem. We will end our discussion of comparison theorems with a remarkable result of Toponogov, which is also known as the triangle comparison theorem.

Let \( q \in (M, g) \) with \( K(M, g) \geq k \). The distance function \( r := d(q, \cdot) \) satisfies

\[
\text{Hess}_r(v, v)|_{r=r_0} \leq \frac{c_k(r_0)}{s_k(r_0)}
\]

where \( v \) is a unit vector perpendicular to \( \partial_r \). If written in matrix, one can view it as 0 on the first row and the first column and less than \( c_k s_k \cdot 1_{n-1} \) for the rest \((n-1) \times (n-1)\) block.

**Definition 2.15.** The modified distance function \( m_{dk}(r) \) is

\[
md_k(r) = \int_0^r s_k(t) \, dt = \begin{cases} \frac{1}{k} (1 - c_k(r)) & k \neq 0 \\ \frac{1}{2} r^2 & k = 0 \end{cases}.
\]

**Remark 2.16.** Note that

\[
\text{Hess}_{md_k(r)} = md'_k(r) \text{Hess}_r + md''_k(r) dr \otimes dr = s_k(r) \text{Hess}_r + c_k(r) dr \otimes dr.
\]

Together with the estimate for \( \text{Hess}_r \), this implies that \( \text{Hess}_{md_k(r)} \leq c_k(r) \cdot g \), which holds away from the \text{Cut}(q). The equality holds for model space \( M_k \).

**Definition 2.17.** A geodesic hinge \( c, c_0, \alpha \) in \( M \) consists of two non-constant geodesic segments, \( c, c_0 \) with some initial point making an angle \( \alpha \). A minimal connection \( c_1 \) between end points of \( c \) and \( c_0 \) is called a closing edge.

**Theorem 2.18.** (Toponogov’s triangle theorem) Let \( (M, g) \) with \( K(M, g) \geq k \).

A) Given \( p_0, p_1, q \in M \) with \( q = p_0 \) and \( q = p_1 \), a non-constant geodesic \( c \) from \( p_0 \) to \( p_1 \) and minimal geodesics \( c_i \) \((i = 0, 1)\) from \( q \) to \( p_i \). All curves parametrized by arc length. Assume \(|c| \leq |c_0| + |c_1| \) and \(|c| \leq \pi/\sqrt{k} \) if \( k > 0 \). Let \( \alpha_i \in [0, \pi] \) denote the angles at \( p_i \).
Then there exists a corresponding comparison triangle $\tilde{p}_0, \tilde{p}_1, \tilde{q}$ in the model space $M_k^2$ with geodesics $\tilde{c}_0, \tilde{c}_1$ and $\tilde{c}$ of lengths $|\tilde{c}| = |c|$ and $|\tilde{c}_i| = |c_i|$ $(i = 0, 1)$ such that

(a) the angles $\alpha_i \leq \alpha_i$ and $\tilde{\alpha} \leq \alpha$,
(b) $\text{dist}(q, \tilde{c}(t)) \leq \text{dist}(q, c(t))$ for any $t \in [0, |c|]$.

(B) Let $c, c_0, \alpha$ be a hinge in $M$ with $c_0$ minimal and assume $|c| \leq \pi/\sqrt{k}$ if $k > 0$. Let $c_1$ be the closing edge. Then, the closing edge $\tilde{c}_1$ of any hinge $\tilde{c}, \tilde{c}_0, \tilde{\alpha}$ in $M_k^2$ with $|\tilde{c}| = |c|$ and $|\tilde{c}_0| = |c_0|$ satisfies $|\tilde{c}_1| \geq |c_1|$.

We will prove the part (b) of theorem A. The rest are consequences of this result.

Proof. We first show the existence of comparison triangles. Assume, without the loss of generality, $|c| \leq \pi/\sqrt{k}$ and $|c_0| < \pi/\sqrt{k}$ if $k > 0$ (If equality holds, then by Cheng’s theorem $(M, g) \cong (S^n, g_k)$). Choose $\tilde{p}_0, \tilde{q} \in M_k^2$ with $d_k(\tilde{p}_0, \tilde{q}) = |c_0|$. Consider $B_{|c_1|}^M(\tilde{q})$ and $d_k(\tilde{p}_0, \tilde{q})|_{\partial B_{|c_1|}^M(\tilde{q})}$. Note that $\{c_0\}$. Note that $|c_0| \leq |c| + |c_1|$ since $c_0$ is minimal and $|c| \leq |c_0| + |c_1|$ ($\pi/\sqrt{k}$ if $k > 0$) by assumption. So there exists $\tilde{p}_1 \in \partial B_{|c_1|}^M(\tilde{q})$ such that $d(\tilde{p}_1, \tilde{p}_0) = |c|$. This shows that the comparison triangle exists.

We then prove part (b) assuming $c(t)$ avoids the cut locus $\text{Cut}(q)$ and assuming $|c| + |c_0| \leq |c_1| \leq \frac{2\pi}{\sqrt{k}}$ if $k > 0$. Let $r = d(q, \cdot)$ and $h = md_k(r)$ and $d_k, r_k, h_k$ be the corresponding function in the model space. Along $c(t)$, we have $h \leq c_k(r)$ and $h_k = c_k(r_k)$ along $\tilde{c}$. Note that they agree on end points. The estimate for the hessian of the modified distance function gives

$$\begin{cases} h'' + kh \leq 1 \\ h''_k + kh = 1 \end{cases}$$

Let $\lambda = h - h_k$. Then $\lambda'' \leq -h\lambda$ and $\lambda(0) = \lambda(|c|) = 0$. It then suffices to show that $\lambda$ is non-negative. This is simply various versions of maximum principle.

* $(k < 0)$ If $\lambda$ achieves a negative minimum $-\mu$ (with $\mu > 0$) at an interior point $t_0$. Then $0 \leq \lambda''(t_0) \leq k\mu < 0$ leads to a contradiction. Hence $\lambda \geq 0$.
* $(k = 0)$ Similarly, assume $\lambda$ achieves negative minimum $-2\mu$ at $t_0$. Then consider

$$\bar{\lambda}(t) = \lambda(t) + \frac{t(|c| - t)}{|c|^2} \mu.$$ 

This vanishes at endpoint and $\bar{\lambda}(t_0) \leq -\mu$. Hence $\bar{\lambda}'' \leq -\frac{2\mu}{|c|^2}$. The contradiction implies $\lambda \geq 0$.

* $(k > 0)$. By assumption $|c| + |c_0| + |c_1| < \frac{2\pi}{\sqrt{k}}$ and $|c| \leq |c_0| + |c_1|$, we can assume $|c| \leq \frac{\pi}{\sqrt{k}} - 2\varepsilon$. Set $\sigma_\varepsilon(t) = s_k(t + \varepsilon) - s_k(\varepsilon/2)$. Note that $\sigma_\varepsilon(t) > 0$ and $\sigma''_\varepsilon(t) = -k\sigma_\varepsilon$. Set $\tilde{\lambda}(t) = \frac{\lambda(t)}{\sigma_\varepsilon(t)}$. Then if $\lambda$ achieves a negative minimum at time $t_0$, then so is $\tilde{\lambda}$.

Apply the same argument as above, we can conclude that $\lambda \geq 0$.

This estimates only holds away from the cut locus. Next, we show how we extends our result without assuming that $c(t)$ avoids the cut locus. The idea is to replace the distance function with a new one that is smooth on cut locus.
Suppose in the above construction, \( \lambda \) (resp. \( \tilde{\lambda} \), \( \bar{\lambda} \)) achieve negative minimum at time \( t_0 \in (0, |c|) \) and \( c(t_0) \in \text{Cut}(q) \). Let \( \gamma \) be a unit speed minimal geodesic from \( q \) to \( c(t_0) \). For some small \( \eta > 0 \), define

\[
 r_\eta(x) = \text{dist}(x, \gamma(\eta)) + \eta.
\]

Clearly, we have \( r_\eta(c(t_0)) = r(c(t_0)) \) and \( r_\eta(x) \geq r(x) \) by triangle inequality. We claim that \( r_\eta \) is smooth in a neighbourhood of \( c(t_0) \).

Consider \( -\gamma \), the minimizing geodesic from \( c(t_0) \) to \( \gamma(\eta) \) is minimizing beyond \( \gamma(\eta) \). This implies that \( \gamma(\eta) \notin \text{Cut}(c(t_0)) \) and thus \( c(t_0) \notin \text{Cut}(\gamma(\eta)) \) which is equivalent to our claim.

Since this is a generalized distance function, we have

\[
 \text{Hess}_{r_\eta} \leq \frac{c_k}{s_k} \cdot g|_{T\{r_\eta=c\}}.
\]

Consider \( md_k(r_\eta) = h_\eta \). Then,

\[
 \text{Hess}_{md_k(r_\eta)} \leq c_k(r_\eta)g|_{T\{r_\eta=c\}} + s_k(r_\eta) \frac{c_k(r_\eta - \eta)}{s_k(r_\eta - \eta)} g|_{T\{r_\eta=c\}}.
\]

Note that \( s_k(\eta) = s_k(r - r_\eta)c_k(r_\eta) + c_k(\eta - r_\eta)s_k(r_\eta) \). This implies

\[
 \text{Hess}_{md_k(r_\eta)} \leq \left( c_k(r_\eta) + \frac{s_k(r_\eta)}{s_k(r_\eta - \eta)} \right) g
\]

Set \( \lambda_\eta = h_\eta - h \) and similarly \( \tilde{\lambda}_\eta \), \( \bar{\lambda}_\eta \) as before. Assume they achieve negative minimum \(-\mu\) with \( \mu > 0 \). One can check that, by taking \( \eta \) sufficiently small,

\[
 0 \leq \lambda_\eta''(t) \leq -2k\mu + \frac{s_k(\eta)}{s_k(r_\eta - \eta)} < 0
\]

\[
 0 \leq \tilde{\lambda}_\eta''(t) \leq -\frac{2\mu}{|c|^2} + \frac{s_k(\eta)}{s_k(r_\eta - \eta)} < 0
\]

\[
 0 \leq \bar{\lambda}_\eta''(t) \leq -\mu \frac{s_k(\varepsilon/2)}{\sigma_\varepsilon^2(t_0)} + \frac{s_k(\eta)}{s_k(r_\eta - \eta)} < 0
\]

Hence the assumption that \( c(t) \) avoids the cut locus of \( q \) can be relaxed. Next, we need to show the result for \( |c| + |c_0| + |c_1| \geq \frac{2\pi}{\sqrt{k}} \). If equality holds, then Cheng’s theorem implies that \((M, g) \cong (S^n, g_k)\) and the proof is trivial. So it remains to show that \( |c| + |c_0| + |c_1| > \frac{2\pi}{\sqrt{k}} \) is impossible. Take \( k > \delta > 0 \) such that \( |c| + |c_0| + |c_1| = \frac{2\pi}{\sqrt{k}} \). The comparison triangle on \( M_\delta^2 \) gives a a great circle since every curve has length less than \( \frac{\pi}{\sqrt{\delta}} \). Then the antipodal point \( \tilde{q} \) of \( q \) is a point on \( \tilde{c} \), say \( \tilde{c}(t_0) \). By the above result, we then have \( \text{dist}(q, c(t_0)) \geq \text{dist}(\tilde{q}, \tilde{c}(t_0)) = \frac{\pi}{\sqrt{\delta}} \).

This contradicts to the Myers theorem and hence completes the proof. \( \square \)
3. Ricci Spaces


**Theorem 3.1** (Bishop-Gromov Theorem). Assume $\text{Ric}(M, g) \geq (n-1)\rho$. Then for $r > 0$ the surface area ratios

$$\frac{\text{Vol}_{(n-1)}(\partial B^M_r(p))}{\text{Vol}_{(n-1)}(\partial B^M_\rho(\tilde{p}))}$$

and the ball volume ratios

$$\frac{\text{Vol}_{(n)}(B^M_r(p))}{\text{Vol}_{(n)}(B^M_\rho(\tilde{p}))}$$

are non-increasing. Both ratios go to 1 as $r$ goes to 0.

**Remark 3.2.** By the homogeneous property of $M_\rho$, $\tilde{p}$ can be an arbitrary point in the constant curvature space.

**Remark 3.3.** The Bishop-Gromov Theorem even holds for a conical region. We will prove for this general case.

**Proof.** The proof is based on the comparison theorem in the previous sections.

Let us first assume that $r$ is less than the distance from $p$ to the conjugate points. Let $\sigma(r, V)$ be the volume form on $M$ at $\exp_p(rV)$, and $\sigma_\rho(r, V)$ be the volume form on $M_\rho$ at $\exp_{\tilde{p}}(rV)$. Note that $\sigma_\rho(r, V) = \sigma_\rho(r)$ is independent of the vector $V$ by homogeneity of the model space.

Recall (2.1), (2.2):

(3.1)

$$\begin{cases}
\frac{d\sigma(r, V)}{dr} = H(r, V)\sigma(r, V) \\
\frac{dH(r, V)}{dr} \leq -\text{Ric}(\partial_r, \partial_r) - \frac{1}{(n-1)}H^2(r, V),
\end{cases}$$

$$\begin{cases}
\frac{d\sigma_\rho(r, V)}{dr} = H_\rho(r, V)\sigma_\rho(r, V) \\
\frac{dH_\rho(r, V)}{dr} = -(n-1)\rho - \frac{1}{(n-1)}H^2_\rho(r, V).
\end{cases}$$

Then we conclude that

$$\frac{d}{dr} \left( \frac{\sigma(r, V)}{\sigma_\rho(r, V)} \right) = (H(r, V) - H_\rho(r, V)) \left( \frac{\sigma(r, V)}{\sigma_\rho(r, V)} \right).$$

The Riccati equation for $H$ (i.e. (2.2)) implies that

$$\left( \frac{1}{(n-1)}H(r, V) \right)' \leq -\rho - \left( \frac{1}{(n-1)}H(r, V)^2 \right).$$
and $H(r) \sim (n-1)/r \rightarrow +\infty$ as $r \rightarrow 0^+$. Then the Ricatti comparison theorem shows

$$\frac{H(r, V)}{(n-1)} \leq \frac{s'(r)}{s(r)} = \frac{H_\rho(r, V)}{(n-1)} = \frac{H_\rho(r)}{(n-1)}.$$ 

So

$$\frac{d}{dr} \left( \frac{\sigma(r, V)}{\sigma_\rho(r, V)} \right) \leq 0.$$ 

For every $V \in S^{n-1} \subset T_p M$, we define the function

$$S(r, V) = \frac{\sigma(r, V)}{\sigma_\rho(r, V)}$$

is the ratio of the area elements evaluated at $\exp_\rho(rV)$. Then we have shown that $S(r, V)$ is decreasing as long as $rV \not\in \text{Cut}(p) \subset T_p M$ (i.e. $r < l(V)$). To avoid technical problems occuring at the cut locus, we use the cutoff function

$$\chi(r, V) = \begin{cases} 
1 & r < \ell(v) \\
0 & r \geq \ell(v) 
\end{cases}.$$ 

Then $\chi(r, V)S(r, V)$ is non-increasing in $r \in [0, +\infty)$. For convenience, we will also denote it by $S(r, V)$. By Myers theorem, $\ell(v) \leq \pi/\sqrt{\rho}$. We will take $r$ as the minimum of $r$ and $\pi/\sqrt{\rho}$ to ensure that we won’t “double count” the volume. Also note when $r \rightarrow 0^+$, $S(r, V) \rightarrow 1$. So $S(r, V) \leq 1$.

Let us define

$$\text{vol}(r, V) = \frac{\int_0^r \chi(s, V)\sigma(s, V)ds}{\int_0^r \sigma_\rho(s, V)ds} = \frac{\int_0^r S(s, V)\sigma_\rho(s, V)ds}{\int_0^r \sigma_\rho(s, V)ds},$$

which is the average of $S(r, V)$ with respect to the measure $\sigma_\rho(s)ds$. Hence $\text{vol}(r)$ is non-increasing.

Now let us consider $\Gamma \subset S^{n-1} \subset T_p M$ be a measurable set. Let us define

$$\text{Cone}(\Gamma) := \{\exp_\rho(rV) : V \in \Gamma, r \leq l(V)\},$$

and we define

$$\text{Vol}(\text{Cone}(\Gamma), R) = \text{Vol}(\text{Cone}(\Gamma) \cap B_R).$$

Then we have

$$\text{Vol}(\text{Cone}(\Gamma), R) = \int_{V \in \Gamma} \int_0^r \chi(s, V)\sigma(s, V)ds$$

$$= \int_{V \in \Gamma} \text{vol}(r, V) \int_0^r \sigma_\rho(s, V)ds.$$ 

Also note that

$$\text{Vol}_{\rho}(\text{Cone}(\Gamma), R) = |\Gamma| \int_0^r \sigma_\rho(s)ds.$$
Therefore
\[ \frac{\text{Vol}_{(n)}(\text{Cone}(\Gamma), r)}{\text{Vol}_{(n), M(\rho)}(\text{Cone}(\Gamma), r)} = \frac{1}{\Gamma} \int_{v \in \Gamma} \text{vol}(r, V), \]
which is non-increasing.

Define
\[ \text{Area}(\Gamma, r, M) \]
to be \( \text{Vol}_{(n-1)}(\text{Cone}(\Gamma) \cap \partial B_r(p)) \), then we have
\[ \frac{\text{Area}(\Gamma, r, M)}{\text{Area}(\Gamma, r, M(\rho))} = \frac{1}{|\Gamma|} \int_{V \in \Gamma} \sigma(r, V), \]
which is non-increasing. Then we proved Bishop-Gromov Theorem for cones. Taking \( \Gamma = S^{n-1} \) be the whole sphere we proved the Theorem. \( \square \)

**Remark 3.4.** In our proof, we implicitly used the fact that the cut locus has measure 0 (Corollary 1.15) and doesn’t count towards the volume. Bishop-Gromov Theorem holds even when the boundary \( \partial B_r(p) \) hit and pass the cut locus. In the proof, we extend \( S(r, V) \) the ratio to be 0 if \( r \geq l(V) \). This is reasonable because after we pass the cut locus, say \( r > l(V) \), then there must be a shorter path from \( p \) to \( \exp_p(rV) \), thus the area element at \( \exp_p(rV) \) is counted before.

**Remark 3.5.** When \( \rho > 0 \), note \( \sigma(\rho) \rightarrow 0 \) as \( r \rightarrow \frac{\pi}{\sqrt{\rho}} \). Since \( S(r, V) \) decreases, we know \( \sigma(r, V) \) also goes to 0 at or before \( r = \frac{\pi}{\sqrt{\rho}} \). This implies that \( l(V) \leq \frac{\pi}{\sqrt{\rho}} \). So we conclude the Myers Theorem:
\[ \text{diam}(M, g) \leq \frac{\pi}{\sqrt{\rho}}. \]

### 3.2. Rigidity in Bishop-Gromov Theorem
Let us consider what happens when the equality holds in Bishop-Gromov Theorem. I.e. If for any \( r \leq R \) we have
\[ \frac{\text{Vol}(B^M_r(p))}{\text{Vol}(B^{M(\rho)}_r)} = 1. \]
Then (3.1) implies \( H(r) = H_\rho(r) \) for all \( r \leq R \). In particular,
\[ \frac{dH(r)}{dr} = -(n-1)\rho - \frac{1}{(n-1)}(H(r))^2. \]
Also recall (2.2) and before we use the Cauchy-Schwartz inequality:
\[ \frac{dH(r)}{dr} = -\text{Ric}(\partial_r, \partial_r) - \text{Tr}(S_r)^2 \]
(Cauchy-Schwartz) \( \leq -\text{Ric}(\partial_r, \partial_r) - \frac{1}{(n-1)}(\text{Tr}(S_r))^2 \)
\[ = -\text{Ric}(\partial_r, \partial_r) - \frac{H^2}{(n-1)}. \]
Note $\text{Ric}(\partial_r, \partial_r) \geq (n-1)\rho$, we conclude that
$$\text{Ric}(\partial_r, \partial_r) = (n-1)\rho.$$ 

Also, we need to have equality at the step of Cauchy-Schwartz, so
$$S_r = \frac{-H(r)}{(n-1)} \mathbb{1},$$
which is a constant. Then we plug it into Lemma 2.6
$$\frac{d}{dt} S_r = S_r^2 - R(\dot{\gamma}, \cdot) \dot{\gamma}.$$ 

Then for any unit vector $w \perp \partial_r$, we have
$$\langle R(\partial_r, w), \partial_r \rangle = \rho.$$ 

Now, consider the Jacobi field along radial direction with $J(0) = 0$ and $J'(0) = w$. Solving the Jacobi equation (1.1), we have $J(t) = s_\rho(t)w$ where $s_\rho(t) = \sin(\sqrt{\rho}t)/\sqrt{\rho}$. Note that this is also the Jacobi field solution for the model space. By Gauss lemma,
$$\exp^* g = dr^2 + s_\rho(r)^2 g_{S^{n-1}}.$$ 

Hence, $B_r(p) \subset (M, g)$ is isometric to $B_r(0) \subset (M_\rho, g_\rho)$. Take $r = R$, then $B_R(p)$ is isometric to $B_R(0)$ in the model space.

Remark 3.6. If $R$, the maximum radius where the equality holds in the Bishop-Gromov theorem, equals the diameter of $M$, we can then conclude that $M$ is isometric to the model space by a concatenating curve argument.

3.3. Rigidity of Myers Theorem.

Theorem 3.7. (Cheng) If $(M^n, g)$ has ricci curvature bound $\text{Ric} \geq (n-1)\rho$ with $\rho > 0$ and $\text{diam}(M, g) = \pi/\sqrt{\rho}$, then $(M, g) \cong (S^n, g_\rho)$.

Proof. By scaling the metric, we can assume that $\rho = 1$. Take $p, q$ be two points in $M$ such that $d(p, q) = \pi$. Consider disjoint balls $B_r(p)$ and $B_{\pi-r}(q)$ in $M$ and denote $\tilde{B}_r$ be the ball of radius $r$ in the model space $M_\rho = (S^n, g_1)$. Define
$$\mu(r) = \frac{\text{vol}(\tilde{B}_r)}{\text{vol}(S^n)}.$$ 

Then, $\mu(r) + \mu(\pi - r) = 1$. Note that
$$\text{vol}(M) \geq \text{vol}(B_r(p)) + \text{vol}(B_{\pi-r}(q))$$
$$= \left( \frac{\text{vol}(B_r(p))}{\text{vol}(\tilde{B}_r)} \frac{\text{vol}(\tilde{B}_r)}{\text{vol}(S^n)} + \frac{\text{vol}(B_{\pi-r}(q))}{\text{vol}(\tilde{B}_{\pi-r})} \frac{\text{vol}(\tilde{B}_{\pi-r})}{\text{vol}(S^n)} \right) \text{vol}(s^n)$$

By Bishop-Gromov comparison theorem, we have
$$\frac{\text{vol}(B_r(p))}{\text{vol}(\tilde{B}_r)} \geq \frac{\text{vol}(M)}{\text{vol}(S^n)}$$

(3.2)
for any \( p \in M \) and \( r \leq \pi \). Plug this to the inequality above, we get

\[
\text{vol}(M) \geq \left( \mu(r) \frac{\text{vol}(M)}{\text{vol}(S^n)} + \mu(\pi - r) \frac{\text{vol}(M)}{\text{vol}(S^n)} \right) \text{vol}(S^n) = \text{vol}(M)
\]

In particular, we have the equality case for (3.2) for any \( p \) and any \( r \). By the rigidity case for Bishop-Gromov inequality, we know that \( B_\pi(p) \subset M \) is isometric \( B_\pi(N) \subset S^n \) where \( N \) is the north pole. It is then suffices to show that \( \partial B_\pi(p) = q \). Since \( M = B_r(p) + B_{\pi - r}(q) \), if \( x \in \partial B_r(p) \), then \( x \in \partial B_{\pi - r}(q) \). Hence \( x = \exp_p(rv) = \exp_q((\pi - r)w) \) for some \( v \in T_pM \) and \( w \in T_qM \). Then combining this curve gives a piecewise smooth minimizing curve between \( p, q \) and hence is a geodesic. It follows that all smooth geodesic from \( p \) reaches \( q \) in time \( \pi \). Therefore \( \partial B_\pi(p) = q \).

\[\Box\]

4. GROMOV-HAUSSDORFF CONVERGENCE

4.1. Gromov-Hausdorff Distance. For the next part of this course, we will discuss the convergence theory in Riemannian geometry. But first, our goal is to make sense of the convergence of metric spaces.

Remark 4.1. For two pointed metric spaces \((X, d_X, p)\) and \((Y, d_Y, q)\), we consider that they are the “same” if there is a map \( f : X \to Y \) such that

1. \( Y \subset f(X) \)
2. \( d_X(x_1, x_2) = d_Y(f(x_1), f(x_2)) \)
3. \( f(p) = q \)

Such a map is called an isometry. Then the convergence of metric spaces should capture the idea of “almost” isometry. Here’s how we will make this idea precise.

Definition 4.2. \( f : (X, d_X, p) \to (Y, d_Y, q) \) is an \( \varepsilon \)-isometry if

1. \( |d(x_1, x_2) - d_Y(f(x_1), f(x_2))| < \varepsilon \) for any \( x_1, x_2 \in B_{\varepsilon^{-1}}(p) \).
2. \( B_{\varepsilon^{-1}}(q) \subset B_{\varepsilon}(f(B_{\varepsilon^{-1}}(p))) \). (This is called \( \varepsilon \)-onto)
3. \( f(p) = q \).

Remark 4.3. The map \( f \) need not to be one-to-one and it is not necessarily continuous.

Remark 4.4. Usually, we will take \((X, d_X)\) to be a compact metric space and \((Y, d_Y)\) be a proper space (i.e. every closed ball is compact, which also implies that it is separable, complete and locally compact).

Definition 4.5. The Gromov-Hausdorff distance between two pointed metric spaces \((X, d_X, p)\) and \((Y, d_Y, q)\) is defined as

\[
d_{GH}(X, Y) = \inf\{ \varepsilon > 0 : \text{there exists } \varepsilon \text{-isometries } f, h \text{ with } f : X \to Y, h : Y \to X \}
\]

Remark 4.6. (exercise) If \( f : X \to Y \) is an \( \varepsilon \) isometry, there there exists a \( 3\varepsilon \)-isometry \( h : Y \to X \).
4.2. Compactness of Gromov-Hausdorff Metric Space.

**Theorem 4.7.**

(1) Let $\text{Met}$ be the space of compact metric spaces (modulo isometry). Then $(\text{Met}, d_{GH})$ is a compact metric space.

(2) Let $\text{Met}^p$ be the space of pointed proper metric spaces (modulo isometry). Then $(\text{Met}^p, d_{pGH})$ is complete.

**Proof.** (Sketch) The difficult part is to show that if $X$ and $Y$ are compact metric spaces and $d_{GH}(X,Y) = 0$, then $X$ and $Y$ are isometric. If this is true, then we can discretize the spaces and then pass to limit. We discuss more details later. □

**Remark 4.8.** There are many variations on the definitions of the Gromov-Hausdorff distance $d_{GH}$, especially in the pointed case. For example, we can define $(X_i, d_i, p_i) \xrightarrow{d_{GH}} (X_\infty, d_\infty, p_\infty)$ as for any $R > 0$, there exists $R_i \rightarrow R$ such that $(B_{R_i}(p_i), d_i, p_i) \xrightarrow{d_{GH}} (B_R(p_\infty), d_\infty, p_\infty)$.

However, all the definitions give the same topology. So when we discuss convergence, it is fine to use any one of the definitions.

The above compactness theorem for pointed metric spaces is not really a compactness theorem, because the space of all pointed metric spaces is still too large. We need to have some constraints on the pointed metric spaces to get a compactness theorem.

**Definition 4.9.** Let $(X,d)$ be a compact metric space. The **counting number** $N(X, \epsilon)$ is defined to be the smallest integer $N$ such that there exists $\{x_i\}_{i=1}^N$ with $X \subset \bigcup_{i=1}^N B_{\epsilon}(x_i)$.

**Example 4.10.** Suppose $(X, d)$ is $\mathbb{R}^n$ carrying the Euclidean metric. Then we have

$$N(B_r(0), \epsilon) \leq A_n \left(\frac{r}{\epsilon}\right)^n$$

for some universal constant $A_n$.

**Definition 4.11.** Let $C(r, \epsilon)$ be any positive function of $(r, \epsilon)$. Then we define the space $\mathcal{M}_C = \{(X,d,p) : N(B_r(p), \epsilon) \leq C(r, \epsilon)\}$.

**Theorem 4.12.** [Gromov] For any $C(r, \epsilon)$ the set $\mathcal{M}_C$ is pre-compact in Gromov-Hausdorff topology.

**Proof.** Let $(X_k, d_k, p_k) \subset M_C$ be a sequence of pointed metric spaces. Let $\{x_i^k\}_{i=1}^{\hat{C}(N,k)} \subset (X_k, d_k)$ be a sequence of points which are $1/N$ dense in $B_N(p_k)$. In other words,

$$B_N(p_k) \subset \bigcup_{i=1}^{\hat{C}(N,k)} B_{1/N}(x_i^k).$$

We can assume $\hat{C}(N,k) \leq C(N,1/N)$, so up to taking a subsequence, we may assume $\hat{C}(N,k) = \hat{C}(n)$. We can also pick $x_0^k = p_k$.

For any $i,j \leq \hat{C}(n)$, the distance $d_{ij}^k = d_k(x_i^k, x_j^k) \leq 2N$. So after taking a subsequence we can assume:
(1) $d_{ij}^k \to d_{ij}^\infty$,
(2) $|d_{ij}^l - d_{ij}^m| < 1/N$, for any $l,m$.

Now we define the discretized space $(\tilde{X}_k, d_k, p_k)$ to be

$$(\tilde{X}_k(N), d_k, p_k) = \{ x^k_i \}_{i=0}^C(n), d_k, p_k).$$

Then

$$d_{GH}[(\tilde{X}_k(N), d_k, p_k), (\bar{X}_l(N), d_l, p_l)] < \frac{1}{N}.$$  

We note that

$$d_{GH}[(X_k, d_k, p_k), (\tilde{X}_k(N), d_k, p_k)] < \frac{1}{N};$$

So we have

$$d_{GH}[(X_k, d_k, p_k), (\bar{X}_l(N), d_l, p_l)] < \frac{3}{N}.$$  

Repeat this argument for $N = N_1 < N_2 < \cdots$ and take a diagonal subsequence, we obtain a subsequence $(X_k, d_k, p_k)$ which is a Cauchy sequence. Then by completeness of the pointed metric space we can find a limit $(X_\infty, d_\infty, p_\infty)$ such that

$$(X_k, d_k, p_k) \xrightarrow{d_{GH}} (X_\infty, d_\infty, p_\infty).$$

So $\mathcal{M}_C$ is precompact.  

\begin{remark}
This discretization technique is very useful in the study of the convergence of metric spaces.
\end{remark}

With this compactness theorem, we can prove the compactness for a class the manifolds as metric spaces.

\begin{theorem}
For any $n \geq 1$, $k \in \mathbb{R}$, $D > 0$, the following classes of manifolds are precompact in Gromov-Hausdorff topology.

(1) Riemannian manifolds $(M, g)$ with Ricci lower bound $\text{Ric}(g) \geq (n-1)k$ and diameter bound $\text{diam} < D$.

(2) Pointed manifolds $(M, g, p)$ with Ricci lower bound $\text{Ric}(g) \geq (n-1)k$.

\end{theorem}

\begin{proof}
(1) is straightforward. We only prove (2). From the theorem above we only need to control the number of balls of radius $\epsilon$ needed to cover $\overline{B_r(p)}$.

Let us construct a cover in the following way:

\textbf{Step 1.} Take $x_0 = p$. Let $N_0 = B_r(x_0)$. If $\overline{B_r(p)} \subset N_0$, then we done.

\textbf{Step 2.} If $\overline{B_r(p)} \not\subset N_0$, then we take $x_1 \in \overline{B_r(p)} \cap \partial N_0$. Let $N_1 = N_0 \cup B_r(x_1)$. If $\overline{B_r(p)} \subset N_1$, then we done.

Then we repeat the above steps until we get $\{x_0, \cdots, x_L\}$ such that $\overline{B_r(p)} \subset \cup_{i=1}^L B_r(x_i)$. From the steps we know that $d(x_i, x_j) \geq \epsilon$ for any $i, j$. As a result, $B_{\epsilon/2}(x_i) \cap B_{\epsilon/2}(x_j) = \emptyset$.

\end{proof}
Now we show this process terminates after finite steps, and the number of steps \( L \) has uniform bound. If \( k \rho_\epsilon, B_{\epsilon/2}(x_i) \cap B_{\epsilon/2}(x_j) = \emptyset \) implies that

\[
\text{Vol}_M(B_{2\rho}(p)) \geq \sum_{i=1}^{L} \text{Vol}_M(B_{\epsilon/2}(x_i)) \geq (L + 1) \text{Vol}_M(B_{\epsilon/2}(\tilde{x})),
\]

where \( \tilde{x} \) is one of \( x_i \) such that

\[
\text{Vol}_M(B_{\epsilon/2}(\tilde{x})) = \min_{0 \leq i \leq L} \text{Vol}_M(B_{\epsilon/2}(x_i)).
\]

Note \( \text{Vol}_M(B_{2\rho}(p)) \leq \text{Vol}_M(B_{3\rho}(\tilde{x})) \). By Bishop-Gromov Theorem

\[
(L + 1) \leq \frac{\text{Vol}_M(B_{3\rho}(\tilde{x}))}{\text{Vol}_M(B_{\epsilon/2}(\tilde{x}))} \leq \frac{\text{Vol}_{M^k}(B_{3\rho})}{\text{Vol}_{M^k}(B_{\epsilon/2})} = C(r, \epsilon)
\]

\[\square\]

4.3. Tangent Cone. Consider a Riemannian manifold \((M, g)\). Let us consider pointed metric spaces \( \bar{M}_r = (M, g_r = r^{-2}g, p) \), where \( p \in M \) is a fixed point. Then

\[
\text{Ric}(g_r) = \text{Ric}(g).
\]

So by the compactness theorem in the previous section, a subsequence of \((M_r, g_r, p)\) will converge to a limit metric space \((\bar{M}_\infty, d, p)\) as \( r \to 0 \). In general, for a metric space \((X, d_X, p)\), if \((X, r^{-1}d_X, p) \xrightarrow{r \to 0} (Z, d_Z, 0)\), we say \((Z, d_Z, 0)\) is the tangent cone of \(X\) at \(p\).

**Example 4.15 (Cone Manifold).** Given a Riemannian manifold \((M, g_M)\), we can construct a manifold

\[
(C(M), \bar{g}) = (M \times \mathbb{R}_{\geq 0}, dr^2 + r^2g_M).
\]

We can check that the tangent cone at \(p\) is \(\mathbb{R}^n\), and the tangent cone at 0 is \((\overline{C(M)}, \bar{g})\).

Suppose \((M_i, g_i)\) is a sequence of Riemannian manifolds with \(\text{Ric}(g_i) \geq 0\), \(\text{diam}(M_i, g_i) \leq D < \infty\), \(\text{Vol}(M_i, g_i) \geq V > 0\). By Bishop-Gromov theorem, we have

\[
1 \geq \frac{\text{Vol}_{M_i}(B_r(p_i))}{\omega_n r^n} \geq \frac{V}{\omega_n D^n} = v.
\]

So \(\text{Vol}_{M_i}(B_r(p_i)) \geq v\omega_n r^n\). This tells us that \((M_i, g_i)\)’s are non-collapsed.

By the compactness theorem, \((M_i, g_i) \xrightarrow{d_{GH}} (M_\infty, g_\infty)\). Here \((M_\infty, g_\infty)\) is just a metric space. But right now let us assume it carries a volume structure, and we assume that the volume also converges

\[
\text{Vol}_{M_i}(B_r(p_i)) \to \text{Vol}_{M_\infty}(B_r(p_\infty)).
\]

Then if \(r < R\), Bishop-Gromov implies that

\[
\frac{\text{Vol}_{M_i}(B_r(p_i))}{\omega_n r^n} \geq \frac{\text{Vol}_{M_i}(B_R(p_i))}{\omega_n R^n} \geq v,
\]
and passing to limit gives
\[ \frac{\text{Vol}_{M_{\infty}}(B_r(p_{\infty}))}{\omega_n r^n} \geq \frac{\text{Vol}_{M_{\infty}}(B_R(p_{\infty}))}{\omega_n R^n} \geq v. \]

So, at least heuristically, \((M_{\infty}, g_{\infty})\) has some volume monotonicity and non-collapsing property.

Now let us consider the tangent cone of \((M_{\infty}, g_{\infty})\) at \(p_{\infty}\). For \((M_{\infty}, \lambda_i^{-2} g_{\infty}, p_{\infty})\), let \(\lambda_i \to \infty\) and we can check

\[ \text{Vol}_{(M_{\infty}, \lambda_i^{-2} g_{\infty})}(B_r(p_{\infty})) = \frac{1}{\lambda_i^n r^n} \text{Vol}_{(M_{\infty}, g_{\infty})}(B_{\lambda_i r}(p_{\infty})) r^n. \]

By monotonicity we may let

\[ A = \lim_{r \to 0} \frac{\text{Vol}_{M_{\infty}}(B_r(p_{\infty}))}{\omega_n r^n} \geq v > 0. \]

Then if \((Z, dz, 0)\) is the tangent cone, it satisfies

\[ \text{Vol}_{Z, dz}(B_r(0)) = Ar^n. \]

By the first pset problem, if \(Z\) is a smooth manifold, then it is a metric cone. So heuristically the tangent cones of a metric space which is the limit of smooth manifolds with non-collapsing property should be metric cones.

5. Elliptic PDE Theory

In this section, we will develop linear elliptic differential equation theory as a major tool for studying geometry. To begin with, here is a few examples of elliptic PDEs.

**Example 5.1.**

1. Let \(\Omega \subset \mathbb{R}^n\) be a domain and then

\[ \Delta u = \sum \frac{\partial^2 u}{\partial x_i^2} = f \]

is a linear elliptic PDE.

2. The Einstein’s equation \(\text{Ric}(g) = \lambda g\) in harmonic coordinates:

\[ g^{lm} \partial_l \partial_m g_{ij} = -2\lambda g_{ij} + 2(g, \partial g) \]

where the last term is a quadratic in \(g\) and \(\partial g\).

3. In hodge theory, we may encounter \(\Box \alpha\) where \(\alpha \in \Gamma^p T^*M\) and \(\Box\) is some certain elliptic operator.

**Remark 5.2.** Since we are only interested in local properties of elliptic PDE, we just need to consider equations on domains in \(\mathbb{R}^n\) for the most part.
Definition 5.3. Let $\Omega \subset \mathbb{R}^n$ be a domain and $D : \Omega \to \mathbb{R}^m$. A second order linear differential operator $L$ is

$$L\varphi = (L\varphi)^\alpha = A^{ij,\alpha}_{\beta} \partial_i \partial_j \varphi^\beta + b^i_{\beta,\alpha} \partial_i \varphi^\beta + c^\alpha_{\beta} \varphi^\beta$$

where $A^{ij,\alpha}_{\beta}$, $b^i_{\beta,\alpha}$ and $c^\alpha_{\beta}$ are functions for any $1 \leq i, j \leq n$ and $1 \leq \alpha, \beta \leq m$ (with summation convention applied). We say $L$ is uniformly elliptic if for any $\eta = \eta^\alpha_i \in \mathbb{R}^{mn}$,

$$\Lambda|\eta|^2 \geq A^{ij,\alpha}_{\beta} \eta^\alpha_i \eta^\beta_j \geq \lambda|\eta|^2.$$

The constants $0 < \lambda \leq \Lambda < +\infty$ are called ellipticity constants. Furthermore, we say $L$ is of divergence form if

$$(L\varphi)^\alpha = \partial_i (A^{ij,\alpha}_{\beta} \partial_j \varphi^\beta) + b^i_{\beta,\alpha} \partial_i \varphi^\beta + c^\alpha_{\beta} \varphi^\beta$$

Remark 5.4. If $A^{ij,\alpha}_{\beta}$ is $C^1$, then $L$ can always be written as divergence form after changing lower order terms. The divergence form will give us advantages when doing integration by part and it comes more naturally from geometry.

Let $\varphi^\alpha, f^\alpha : \Omega \to \mathbb{R}^m$ and consider elliptic equation $L\varphi = f$. We will try to answer the following questions regarding elliptic PDEs.

Question 5.5.

1. If $f \in W^{k,p}$, then is $\varphi \in W^{k,p}$? ($L^p$ theory)
2. If $f \in C^{k,\alpha}$, then is $\varphi \in C^{k,\alpha}$? (Schauder theory)

Question 5.6. Furthermore, if we add the boundary condition $u : \partial \Omega \to \mathbb{R}^m$, then can we find $\varphi$ such that

$$\begin{cases}
L\varphi = f \text{ in } \Omega \\
\varphi|_{\partial \Omega} = u
\end{cases}$$

? (Hodge theory)

We will first define the sobolev spaces and the maximum principle and then proceed to discuss the $L^p$ theory and Schauder theory.

Definition 5.7. Let $\Omega \subset \mathbb{R}^n$ be a domain. A function $f \in W^{k,p}(\Omega)$ if for any $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$ there is a function $V_\alpha \in L^p(\Omega)$ with the following property:

$$\int_\Omega V_\alpha \varphi = (-1)^{|\alpha|} \int_\Omega f \frac{\partial^{|\alpha|} \varphi}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} \quad \text{for } \forall \varphi \in C^\infty(\Omega) \text{ with supp}(\varphi) \subset \Omega.$$ 

We write

$$V_\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} f.$$

We also define

$$\|f\|_{W^{k,p}(\Omega)} = \sum_{m=0}^{k} \|D^m f\|_{L^p(\Omega)}.$$
Definition 5.8. Suppose $u : \Omega \to \mathbb{R}^m$ is $W^{1,p}$, $L$ is an elliptic operator of divergence form. We say $Lu = f$ weakly if for any $\varphi \in C^\infty_c(\Omega, \mathbb{R}^m)$, we have
\[
\int_\Omega -A_{ij}^{\alpha \beta} \partial_j u^\beta \partial_i \varphi^\alpha + b_i^{\alpha \beta} \partial_i u^\beta \varphi^\alpha + c_{ij}^{\alpha} u^\beta \varphi^\alpha = \int_\Omega f^\alpha \varphi^\alpha
\]
(summation convention is applied).

Exercise 5.9. If $u \in C^2(\Omega)$, then $Lu = f$ if and only if $Lu = f$ weakly.

Example 5.10. $\Omega \subset \mathbb{R}^n$ and $\Delta u = \text{div}(\nabla u)$. Then $\Delta u = f$ for $u = W^{1,p}$ if and only if for any $\varphi \in C^\infty_c(\Omega)$,
\[
-\int_\Omega \nabla u \cdot \nabla \varphi = \int_\Omega f \varphi.
\]

Example 5.11. $(M, g)$ Riemannian manifold.
\[
\Delta u := \frac{1}{\sqrt{\det g}} \partial_i (\sqrt{\det g} g^{ij} \partial_j u).
\]
Then $\Delta u = f$ if and only if for any $\varphi \in C^\infty_c(M)$,
\[
\int_M (\Delta u) \varphi \sqrt{\det g} = -\int_M (\nabla u, \nabla \varphi) g \sqrt{\det g} = \int_M f \varphi \sqrt{\det g}.
\]

5.1. Maximum Principle. Let $\Omega$ be a open bounded domain in $\mathbb{R}^n$ and we consider the second order differential operator:
\[
L \varphi = a^{ij} \partial_i \partial_j \varphi + b^i \partial_i \varphi + c \varphi
\]
where $L$ is uniformly elliptic with elliptic constants $0 < \lambda \leq \Lambda < \infty$ and $b$ and $c$ are uniformly bounded in $\Omega$.

Lemma 5.12. (Weak maximum principle) Suppose $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$, $Lu > 0$ and $c \equiv 0$ (resp. $c \leq 0$). Then if $u$ has a maximum (resp. non-negative max) in $\Omega$, then this maximum cannot be achieved in $\Omega$.

Proof. For a contradiction, suppose $x_0 \in \Omega$ and $u(x_0) = \max_{\Omega} u$. This implies that $\nabla u = 0$ and $D^2 u(x_0) \leq 0$. Then, if $c \equiv 0$, we have
\[
Lu(x_0) = a^{ij} u_{ij}(x) + b^i u_i(x_0) \leq 0
\]
which contradicts to $Lu > 0$. Similarly if $u$ reaches a non-negative maximum at $x_0$ and $u \leq 0$. Then $cu(x_0) \leq 0$ and so is $Lu$, which leads to contradiction. \qed

Lemma 5.13. (Slightly stronger maximum principle) Suppose $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$, $Lu \geq 0$ and $c \equiv 0$ (resp. $c \leq 0$). Then if $u$ attains its maximum (resp. non-negative maximum) on $\Omega$, it must achieve it’s max (resp. non-negative max) on $\partial \Omega$, that is $u(x) \leq \max_{\partial \Omega} u$.

Remark 5.14. This is different from the previous lemma as we relaxed the condition to $Lu \geq 0$. The result is therefore weaker than the previous one since $u$ can achieve maximum in the interior but it can’t exceeds the maximum on the boundary.
Proof. We use an auxiliary function, \( W_\varepsilon = \varepsilon e^{\alpha x_1} + u \) where \( \varepsilon > 0 \) and \( \alpha \) some large number. By assumption, we have \( a_{11} \geq \lambda \) and \( b \) and \( c \) are uniformly bounded. Hence we can choose \( \alpha \) such that \( a_{11} \alpha^2 + b \alpha + c > 0 \). By the last result, we have

\[
W_\varepsilon \leq \max_{\partial \Omega} W_\varepsilon.
\]

Sending \( \varepsilon \) to 0, we then obtain \( u(x) \leq \max_{\partial \Omega} u \). \( \square \)

Remark 5.15. This lemma is often refered as comparison principle in the following sense. Suppose \( Lu \geq 0 \) in \( \Omega \) and \( c \leq 0 \). If \( u|_{\partial \Omega} \leq 0 \), then \( u \leq 0 \) in \( \Omega \).

Here is one corollary of this comparison principle.

Corollary 5.16. Suppose \( c \leq 0 \), \( Lu = Lw \) and \( u|_{\partial \Omega} = w|_{\partial \Omega} \), then \( u = w \).

Proof. Apply comparison principle to \( u - v \) to get \( u \leq v \) and then to \( v - u \) for the other direction. \( \square \)

Lemma 5.17. (Hopf lemma) Let \( B_r \) be a ball of radius \( r \). Assume \( u \in C^2(B_r) \cap C^0(B_r \cup \{x_0\}) \) with \( Lu \geq 0 \) in \( B_r \) and \( x_0 \in \partial B_r \). Further assume \( c \equiv 0 \) (resp. \( c \leq 0 \)) and \( u(x) < u(x_0) \) for any \( x \in B_r \) (resp. assume further that \( u(x_0) \geq 0 \)). Then for every outward pointing normal vector \( \nu \) to \( \partial B_r \) at \( x_0 \), we have

\[
\liminf_{t \to 0^+} \frac{u(x_0) - u(x_0 - t\nu)}{t} > 0
\]

Proof. We will use the comparison principle and a barrier function to bound the limit.

First, we can assume \( u(x_0) = 0 \), otherwise take \( \tilde{u} = u - u(x_0) \). Then \( \tilde{L} \tilde{u} = Lu - cu(x_0) \geq Lu \geq 0 \).

Consider \( h(x) = e^{-\alpha \rho^2} - e^{-\alpha r^2} \). Then \( h|_{\partial B_r} = 0 \) and

\[
Lh = (-2a^{ij}x_i x_j + 4\alpha^2 a^{ij}x_i x_j - 2b^i x_i) e^{-\alpha \rho^2} + c h.
\]

If \( |x| > \rho \), then \( 4\alpha^2 a^{ij}x_i x_j \geq 4\lambda |x|^2 \alpha^2 \geq 4\lambda \rho^2 \alpha^2 \). So we can choose \( \alpha \) large enough (depending on \( \rho, \lambda, r, |b|, |c| \)) such that \( Lh > 0 \) in the annulus \( \{ \rho < |x| < r \} \).

Take \( \rho = r/2 \) and consider the annulus \( \Sigma = \{ \rho < |x| < r \} \). We claim that there exists \( \varepsilon > 0 \) such that \( -\varepsilon h \geq u \) in \( \Sigma \). Indeed, for \( |x| = \rho \), we have, by assumption, \( u|_{\partial B_\rho} < u(x_0) - \delta = -\delta \) for some \( \delta > 0 \). So we can choose \( \varepsilon \) so that \( -\varepsilon h > -\delta \). For \( |x| = r \), \( u(x) \leq u(x_0) = 0 \) and \( -\varepsilon h|_{\partial B_r} = 0 \). Hence, we can find \( \varepsilon > 0 \) such that \( -\varepsilon h > u \) on \( \partial \Sigma \). Use comparison principle to conclude that \( -\varepsilon h > u \) in \( \Sigma \) and \( u(x_0) = -\varepsilon h(x_0) = 0 \). Therefore,

\[
\liminf_{t \to 0^+} \frac{u(x_0) - u(x_0 - t\nu)}{t} > \liminf_{t \to 0^+} \frac{-\varepsilon h(x_0) + \varepsilon h(x_0 - t\nu)}{t} > 0.
\]

\( \square \)

Remark 5.18. In a more general domain \( \Omega \), we may have to require \( u \in C^1(\Omega) \). For reference, see Evans’ book on PDE. This version of Hopf lemma suffices to prove the strong maximum principle.
Theorem 5.19. (Strong maximum principle) Suppose \( u \in C^2(\Omega) \cap C^0(\overline{\Omega}) \), \( Lu \geq 0 \) and \( c \equiv 0 \) (resp. \( c \leq 0 \)). Then the maximum of \( u \) (resp. non-negative maximum) can only be achieved on \( \partial \Omega \) unless \( u \) is constant, that is, \( u \leq \max_{\partial \Omega} \) where equality holds if and only if \( u \) is constant.

**Proof.** Let \( M = \sup_{\Omega} u \) and consider \( \Sigma = \{ x \in \Omega : u(x) = M \} \). Clearly, this set is closed. Suppose \( \Sigma \neq \emptyset, \Omega \). Then \( \Omega \setminus \Sigma \) is open. We can choose \( B_r \subset \Omega \setminus \Sigma \) such that \( \partial B_r \cap \partial \Sigma = \{ x_0 \} \in \Omega \). Since \( u \) achieves its maximum at \( x_0 \), then \( Du(x_0) = 0 \). But by Hopf lemma, \( \frac{\partial u}{\partial \nu}(x_0) > 0 \). Contradiction implies that \( \Sigma = \emptyset, \Omega \). \( \square \)

Corollary 5.20. (Comparison theorem) Suppose \( Lu \geq 0 \) and \( c \leq 0 \). If \( u|_{\partial \Omega} \leq 0 \), then \( u < 0 \) or \( u \equiv 0 \) in \( \Omega \).

Here is an application of the strong maximum principle.

Corollary 5.21. Let \((M,g)\) be a compact Riemannian manifold. Then \( \Delta_g u \equiv 0 \) if and only if \( u \) is constant.

**Proof.** Since \( M \) compact, there exists \( p \in M \) such that \( u(p) = \sup_M p = L \). Take local coordinate of a ball centered at \( p \). We have,

\[
\begin{cases}
\Delta_g u = 0 \quad \text{in } B \\
|_{\partial B} \leq L \\
u(p) = L
\end{cases}
\]

Then the strong maximum principle implies that \( u \equiv L \) in \( B \). Do it on all charts of \( M \), we then conclude that \( u \equiv L \). The reverse direction is trivial. \( \square \)

Some other notions of weak solutions to the elliptic equation also give similar maximum principle.

**Definition 5.22.** We say a \( C^2 \) function \( \varphi \) is an upper (resp. lower) barrier for \( u \) at \( p \) if there exists open set \( p \in V \subset \Omega \) such that \( \varphi \geq u \) on \( V \) (resp. \( \varphi \leq u \)) and \( \varphi(p) = u(p) \).

We say \( Lu \leq c \) (resp. \( Lu \geq c \)) in a barrier sense if for any \( \varepsilon > 0 \), there exists an upper (resp. lower) barrier \( \varphi_\varepsilon \) such that \( L\varphi(p) \leq c + \varepsilon \) (resp. \( L\varphi(p) \geq c - \varepsilon \)).

**Theorem 5.23.** Suppose \( L \) uniformly elliptic, \( c \equiv 0 \) and \( Lu \geq 0 \) in the barrier sense. If \( u \) has an interior maximum, then \( u \) is constant.

**Definition 5.24.** We say a \( C^2 \) function \( \varphi \) is an upper (resp. lower) test function for \( u \) at \( p \) if there exists open neighbourhood \( V \subset \Omega \) of \( p \) such that \( \varphi \geq u \) on \( V \) (resp. \( \varphi \leq u \)) and \( \varphi(p) = u(p) \).

We say \( Lu(p) \leq c \) (resp. \( Lu(p) \geq c \)) in the viscosity sense if for every lower (resp. upper) test function \( \varphi \) for \( u \) at \( p \), we have \( L\varphi(p) \leq c \) (resp. \( L\varphi(p) \geq c \)).

Similarly as before, the maximum principle holds if \( Lu \geq 0 \) in the viscosity sense.
5.2. Elliptic Regularity.

**Theorem 5.25.** Suppose $\Omega \subset \mathbb{R}^n$ is a domain. $L = \partial_i (a_{ij}^{ij} \partial_j) + b_i^{\alpha} \partial_i = c^0_\alpha$ is elliptic with ellipticity constants $0 < \lambda \leq \Lambda < \infty$. Suppose $u \in W^{1,2}(\Omega)$ solves $Lu = f$ in $\Omega$ with $f \in L^2(\Omega)$.

Then, for any compact $V \subset \Omega$, there exists a constant $C$ depending on $d(V, \partial \Omega)$, $\lambda$, $\Lambda$, $\|a_{ij}^{ij}\|_{C^1(\Omega)}$, $\|b_i^{\alpha}\|_{L^\infty(\Omega)}$ and $\|c\|_{L^\infty(\Omega)}$ such that

1. $u \in W^{2,2}(\Omega)$.
2. $\|u\|_{W^{2,2}(V)} \leq C(\|u\|_{W^{1,2}(\Omega)} + \|f\|_{L^2(\Omega)})$.

**Remark 5.26.**

1. We can replace the $W^{1,2}$ of $u$ with $L^2$ in part (2) of the theorem.
2. If $b = c = 0$, this works even if $u \in L^2(\Omega)$ and $Lu = f$ weakly. The proof uses Fourier analysis.

**Theorem 5.27.** (Higher regularity) Suppose in the above theorem $a_{ij}, b_i, c \in C^m$, and $f \in W^{m,2}$. If $u \in W^{1,2}$ solve $Lu = f$ weakly, then $u \in W^{m+2,2}$ and

$$\|u\|_{W^{m+2,2}} \leq C(\|f\|_{L^2} + \|u\|_{W^{1,2}}).$$

For the proofs, we refer to [Eva10].

5.3. Schauder Theory. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded domain. For a function $u : \Omega \to \mathbb{R}$, and any $\alpha \in (0, 1]$, we define the pseudonorm to be

$$[u]_{C^\alpha(\Omega)} = \sup_{x \neq y, x, y \in \Omega} \frac{|u(x) - u(y)|}{|x - y|^\alpha}.$$

**Definition 5.28.** The $C^{k,\alpha}$ norm is defined as

$$\|u\|_{C^{k,\alpha}(\Omega)} = \sum_{|\ell| \leq k} \sup_{x \in \Omega} |\partial^\ell u| + \sum_{|\ell| = k} [\partial^\ell u]_{C^\alpha(\Omega)}.$$

The Hölder space is $C^{k,\alpha}(\Omega) = \{ u : \Omega \to \mathbb{R} : \|u\|_{C^{k,\alpha}(\Omega)} < +\infty \}$.

Note that $C^{k,\alpha}(\Omega)$ is a Banach space. If $\{u_m\} \subseteq C^{k,\alpha}$ is a uniform bounded sequence, there is a convergent subsequence in $C^{k,\beta}$ for $\beta < \alpha$. This is essentially Arzelà–Ascoli theorem.

Now back to our discussion of elliptic equation. Assume $Lu = a^{ij}u_{ij} + b^iu_i + cu$ with $a^{ij}, b^i, c \in C^{k,\alpha}(\Omega)$ and $L$ is uniformly elliptic with constants $\lambda, \Lambda$.

**Theorem 5.29.** (Interior Schauder estimate) If $Lu = f$ in $\Omega$ and $f \in C^{k,\alpha}$ and $\Omega' \subset \subset \Omega$.

Then there exists $C$ depending on $\Omega, \Omega', \alpha, k$ and $C^{k,\alpha}$ norms of $a$, $b$ and $c$, such that

$$\|u\|_{C^{k+2,\alpha}(\Omega')} \leq C \left( \|f\|_{C^{k,\alpha}(\Omega)} + \|u\|_{C^{k}(\Omega)} \right)$$

**Remark 5.30.** We will first give the a priori estimate, i.e. assuming that $u$ is already in $C^{k,\alpha}$. For this we will use a standard argument by contradiction with the following "standard" steps:
(1) Assume no constant suffices for the estimate.
(2) "Blow up" to get a non-constant harmonic function on $\mathbb{R}^n$ with sublinear growth using rescaling and Arzela-Ascoli.
(3) Prove that such function doesn’t exist.

When later we are dealing with problems in geometry such argument can be employed with Gromov-Hausdorff compactness in place of Arzela-Ascoli.

For simplicity, we will assume $b = c = 0$ and it suffices to show for $k = 0$. Note that the proof works for systems as well. For general results, one can consult Gilbarg and Trudinger’s book.

6. Ricci Spaces II

6.1. Cheeger-Gromoll Splitting Theorem. Recall that if $(M, g)$ is complete, $\text{Ric}(g) \geq 0$, then for any $p \in M$ the distance function $r(x) = d(p, x)$ is smooth away from $\text{Cut}(p) \cup \{p\}$. At any $x \not\in \text{Cut}(p) \cup \{p\}$, it also satisfies the laplacian comparison:

$$\Delta_g r \leq \frac{n - 1}{r}$$

Similar result holds for ricci spaces $\text{Ric}(g) \geq (n - 1)k$.

**Theorem 6.1.** $(M, g)$ complete with $\text{Ric}(g) \geq 0$, then for any $p \in M$, the distance function $r(x) = d(x, p)$ satisfies $\Delta_g r \leq \frac{n - 1}{r}$

(1) in the weak sense.
(2) in the barrier sense.

**Proof.** (1) In the weak sense, observe that $\nabla r$ is defined on $M \setminus (\text{Cut}(p) \cup \{p\})$ with $|\nabla r| = 1$. Since $\text{Cut}(p)$ has measure 0, $r \in W^{1,p}_\text{loc}(M)$.

Since $\text{Cut}(p)$ may not be smooth, we approximate by its $\varepsilon$-neighborhood $N_\varepsilon$. More precisely, let $l : S^{n-1} \subset T_p M \to \mathbb{R}_{>0} \cup \{+\infty\}$ be the distance from $p$ to $\text{Cut}(p)$ along vector $v$. Note that $l$ is continuous and we hence can choose $C^\infty$ functions $l_\varepsilon(v) < r < l(v)$ for any $\varepsilon > 0$. Set $N_\varepsilon = \{vr : l_\varepsilon(v) < r < l(v)\}$. Then $\partial_\varepsilon$ points into this set.

Given $\varphi \in C_0^\infty(M, \mathbb{R})$, $\varphi > 0$, we evaluate

$$\int_M r\Delta_g \varphi = -\int_M \langle \nabla r, \nabla \varphi \rangle = -\lim_{\varepsilon \to 0} \int_{M \setminus N_\varepsilon} \langle \nabla r, \nabla \varphi \rangle$$

On $M \setminus N_\varepsilon$, $r$ is smooth. Thus,

$$\int_{M \setminus N_\varepsilon} \langle \nabla r, \nabla \varphi \rangle = \int_{\partial N_\varepsilon} \varphi \langle \nabla r, -\nu_\varepsilon \rangle - \int_{M \setminus N_\varepsilon} (\Delta_g r) \varphi$$

where $\nu_\varepsilon$ is outward pointing normal of $N_\varepsilon$. So $\langle \nabla r, -\nu_\varepsilon \rangle > 0$.

Take this back to first equation and drop the negative term, we have

$$\int_M r\Delta_g \varphi \leq \lim_{\varepsilon \to 0} \int_{M \setminus N_\varepsilon} (\Delta_g r) \varphi \leq \int_M \frac{n - 1}{r} \varphi$$
(2) In the barrier sense, recall that if \( x \in \text{Cut}(p) \) and \( \gamma \) is a minimizing geodesic from \( p \) to \( x \), then for any \( \eta > 0 \), we can define
\[
\eta \eta = \eta + \text{dist}(\cdot \cdot, \gamma(\eta))
\]
Note that \( \eta \eta \) is smooth near \( x \), \( \eta \eta(x) = x \) and \( \eta \eta(y) \geq r(y) \).

Then \( \eta \eta \) is an upper barrier for \( r \) at \( x \). As \( \eta \to 0 \)
\[
\Delta g \eta \eta(x) \leq \frac{n - 1}{\text{dist}(x, \eta \eta))} \to \frac{n - 1}{r}
\]

Now we will use laplacian comparison to prove the Cheeger-Gromoll splitting theorem

**Theorem 6.2.** (Cheeger-Gromoll) \( (M^n, g) \) complete with \( \text{Ric}(g) \geq 0 \). Suppose \( M \) contains a line, i.e. a geodesics \( \gamma : (-\infty, +\infty) \to M \) unit speed and such that \( d(\gamma(s) - \gamma(t)) = |t - s| \) for any \( s, t \in \mathbb{R} \). Then,
\[
(M^n, g) \cong (N^{n-1} \times \mathbb{R}, g_N + dt^2)
\]

We will show first that to prove this theorem, it is sufficient to find a global function with unit gradient and vanishing Hessian. More precisely,

**Proposition 6.3.** Assuming \( (m, g) \) complete with non-negative Ricci curvature and admits a line, then there’s a function \( h : M \to \mathbb{R} \) such that \( |\nabla h| = 1 \) and \( \text{Hess} h = 0 \).

**Proof of Theorem 6.2.** Assuming Proposition 6.3, and let \( N = h^{-1}(0), \varphi_t \) be the time \( t \) flow along \( \nabla h \).

Consider \( \Phi : N \times \mathbb{R} \to M \), with \( (n, t) \to \varphi_t(n) \). Then this gives a diffeomorphism between \( N \times R \) and \( M \).

Furthermore, note that if \( V \in T_p N \), then \( \langle \nabla h, V \rangle = V h = 0 \). So, \( g = g_N + dt^2 \) along \( N \). But the Lie derivative of \( g \), \( \mathcal{L}_{\nabla h} g = 2 \text{Hess} h = 0 \). Hence \( \nabla h \) generates isometries, \( (M, g) \cong (N \times \mathbb{R}, g_N + dt^2) \).

Now we need to construct such a function in Proposition 6.3. Note that if \( \text{Ric}(g) \geq 0, h \) is harmonic and \( |\nabla h| = 1 \), then by Bochner formula, we can conclude that \( \text{Hess} h = 0 \) and \( \text{Ric}(\nabla h, \nabla h) = 0 \). Hence, we are actually looking for a harmonic function with \( |\nabla h| = 1 \).

To this end, we introduce the Busemann function. Let \( \gamma \) be a line in \( M \) and fix a point \( p = \gamma(0) \). Then we define
\[
b^t_p(x) = d(x, \gamma(t)) - d(p, \gamma(t)) = d(x, \gamma(t)) - t
\]

We make a few observations of this function. First, \( |\nabla b^t_p| = 1 \) whenever it is well defined (away from \( \text{Cut}(\gamma(t)) \)). By triangle inequality, we know that \( |b^t_p(x) - b^t_p(z)| \leq d(x, z) \) and thus \( b^t_p \) is bounded on compact sets. Moreover, if \( 0 < s < t \) or \( t < s < 0 \), then \( b^s_p(x) \geq b^t_p(x) \). This follows from the triangle inequality and the fact that \( \gamma \) is a line.
For any \( t \in \mathbb{R} \) and any \( \eta > 0 \), there exists \( y^t_\eta \) such that \( d(x, y^t_\eta) = \eta \) and
\[
\frac{|b^t_p(x) - b^t_p(y^t_\eta)|}{d(x, y^t_\eta)} = 1.
\]
Indeed, if \( \sigma^t(s) \) be the unit speed minimizing geodesic, then take \( y^t_\eta = \sigma^t(\eta) \).

Finally, by global Laplacian comparison, we have
\[
\Delta b^t_p \leq \frac{n - 1}{d(x, \gamma(t))}
\]
in the barrier/weak sense.

Define
\[
b^\pm(x) := \lim_{t \to \pm \infty} b^t_p(x)
\]
Note that \( b^t_p \) is monotone for \( t > 0 \) or \( t < 0 \), is bounded on compact sets and equicontinuous. So \( b^t_p \to b^\pm \) uniformly on compact sets by Arzela-Ascoli.

We make two claims about the functions \( b^\pm \).

**Claim 6.4.** For any \( x \in M \), any \( \eta \in (0, 1) \), there exists \( y^\pm_\eta \) such that \( d(x, y^\pm_\eta) = \eta \) and
\[
\frac{|b^\pm(x) - b^\pm(y^\pm_\eta)|}{d(x, y^\pm_\eta)} = 1.
\]
Naturally, take \( y^\pm_\eta = \lim_{t \to \pm \infty} y^t_\eta \). This converges as they are in a ball at \( x \) of radius \( \eta \).

**Claim 6.5.** \( \Delta b^\pm \leq 0 \) in the barrier sense and in the weak sense.

**Proof.** In the distribution (weak) sense, we have that, for any \( t \) and any \( \varphi \in C^\infty_0(M) \),
\[
\int_M b^t_p(x) \Delta \varphi \leq \int_M \frac{n - 1}{d(x, \gamma(t))} \varphi.
\]
Note \( b^t_p(x) \to b^\pm \) uniformly on compact sets and \( \frac{n - 1}{d(x, \gamma(t))} \to 0 \). Taking the limit, then we have \( \int_M b^\pm(x) \Delta \varphi \leq 0 \).

In the barrier sense, we show that upper barrier exists for \( b^+ \) (Similarly for \( b^- \)).

Let \( t_i \to \infty \) be a sequence of time and \( \sigma_i(s) \) be the unit speed geodesic flow from \( x \) to \( \gamma(t_i) \). Let \( v_i = \sigma_i'(0) \) with \( |v_i| = 1 \). Then \( v_i \to v \in T_xM \) with \( |v| = 1 \). Set \( \sigma \) be the geodesic from \( x \) with initial vector \( v \) and
\[
f_{\sigma,s}(y) := d(y, \sigma(s)) - d(x, \sigma(s))
\]
We claim that \( \bar{f}_{\sigma,s}(y) = f_{\sigma,s}(y) + b^+(x) \) is an upper barrier for \( b^+ \) at \( x \).

First, we note that
\begin{enumerate}
  \item \( \bar{f}_{\sigma,s} \) is \( C^\infty \) near \( x \).
  \item \( \bar{f}_{\sigma,s}(x) = b^+(x) \).
\end{enumerate}
(3) $\Delta \tilde{f}_{\sigma,s}(y) \leq \frac{n-1}{d(y,\sigma(s))} \to 0$ as $s \to \infty$. This follows from the triangle inequality, $d(y,\sigma(s)) \geq d(x,\sigma(s)) - d(x,y) - s + d(x,y)$. So it is then sufficient to show $\tilde{f}_{\sigma,s} \geq b^+(x)$ in a neighborhood of $x$. For each $i$, let

$$\tilde{f}_{\sigma,s}(y) = f_{\sigma,s}(y) + b^i(y) = d(y,\sigma_i(s)) - d(x,\sigma_i(s)) + d(x,\gamma(t_i)) - d(p,\gamma(t_i))$$

By triangle inequality, for any $0 \leq s \leq d(x,\gamma(t_i))$, we have

$$d(y,\gamma(t_i)) \leq d(y,\sigma_i(t)) + d(x,\gamma(t_i)) - d(x,\sigma_i(t)).$$

Subtract both sides by $d(p,\gamma(t_i))$, we obtain that $b^i(y) \leq \tilde{f}_{\sigma,s}(y)$ for $s \in [0,d(x,\gamma(t_i))]$. Take a limit as $i$ tends to infinity. We then have $\tilde{f}_{\sigma,s}(y) \geq b^+(y)$ for any $s \in [0,\infty)$.

Thus, $\Delta b^+ \leq 0$ in the barrier sense. \qed

Now we are in the position to prove Proposition 6.3.

Proof of Proposition 6.3. Consider $b^+ + b^-$. Clearly $b^+_p + b^-_p \geq 0$ by triangle inequality and the fact $\gamma$ is a line. Take the limit, we conclude that $b^+ + b^- \geq 0$. Furthermore, $b^+ + b^-|_\gamma = 0$.

We have the setting:

$$\begin{cases} 
\Delta(b^+ + b^-) \leq 0 \text{ in the weak sense} \\
b^+ + b^- \geq 0 \\
b^+ + b^-|_\gamma = 0
\end{cases}$$

By strong maximum principle, $b^+ + b^- = 0$. Since $0 \geq \Delta b^+ \geq -\Delta b^- \geq 0$, we conclude that $\Delta b^+ = 0$. Use elliptic regularity and Sobolev embedding we have $b^+ \in C^\infty$. Note that $|\nabla b^+| = 0$, hence it is the function we are looking for. \qed

6.2. Harmonic Coordinates.

Definition 6.6. Let $U \subset (M,g)$ be an open subset, $x = (x_1,...,x_n) : U \to B_r(0)$ is a coordinate system on $U$. We call $x$ a $C^m$ harmonic coordinate chart with $\|x\|_r \leq 1$ if

1. $\Delta x_i = 0$.
2. $g_{ij} = \langle \nabla x_i, \nabla x_j \rangle$ on $B_r(0) \subset \mathbb{R}^n$.

and satisfies

$$|g_{ij} - \delta_{ij}| + \sum_{k=1}^{m} r^k |\partial^\alpha g_{ij}| < 10^{-6}$$

Remark 6.7.

1. Most of the time, we only care about $x \in C^1$;
2. $r^k$ appears in the above expression so that norm is scale invariant.

Remark 6.8. Why do we care about the harmonic coordinates? Because we want to study the convergence of metrics, and we want to study the regularity of the metric.

Let us see an heuristic example. Consider the standard Euclidean metric in the usual coordinate $g = \delta_{ij} = dx^i \otimes dx^j$. This is a very nice coordinate, for instance it is $C^\infty$, bounded,
etc. However, if we write it in spherical coordinate $dr^2 + r^2g_{S^{n-1}}$, then the coordinate is not that nice, for instance it is degenerate at the origin.

This example implies that the metric behaves very differently in different coordinate system. So even under certain coordinate the metric is very bad, it does not mean that the metric itself is bad. Thus we hope to find some nice coordinate system to write down the metric, which is the harmonic coordinate system.

**Remark 6.9.** Now we briefly show why harmonic coordinate is a “nice” coordinate. Suppose $(M_i, g_i)$ is a sequence of Ricci flat manifolds, each one is covered by $N$ harmonic coordinate charts with $\|x\|_r \leq 1$.

In each harmonic coordinate chart $B_r$, Ric$(g_i) = 0$ implies that

$$\Delta g_{ij} = Q(g, \partial g),$$

where $Q$ is a quadratic function, and we treat $g_{ij}$ as functions on $B_r$. Since we are working in a harmonic coordinate chart, $|g|$ and $|\partial g|$ are bounded, then the elliptic regularity tells us that $g_{ij} \in C^{1,\alpha}$ for some $\alpha$. Then $g, \partial g$ lie in $C^{0,\alpha}$, then Schauder theory implies that $g \in C^{2,\alpha}$. Then a bootstrap argument gives that $g_{ij}$ are actually smooth functions with uniformly $C^{k,\alpha}$ for any $k$. Moreover, we can take limit $g_i \to g_\infty$ as matrix valued functions on $B_r$.

As a result, we can passing to subsequence such that $M_i \to M_\infty$ as smooth Ricci flat manifolds. This is some nice convergence we hope to see for manifolds.

The scale that the harmonic coordinate exists is called the harmonic radius.

**Definition 6.10.** Let $(M, g)$ be a Riemannian manifold and fix $p \in M$. Given $Q > 1, k \in \mathbb{N}$, $\alpha \in (0, 1)$, define

$$r_H(Q, k, \alpha)(p) := \sup\{r : \text{there exists } Q \text{ bounded } C^{k,\alpha} \text{ harmonic coordinate chart on } B_r(p)\}.$$

i.e. $r_H(Q, k, \alpha)(p)$ is the smallest $r$ such that there exists coordinate $\{x_1, \cdots, x_n\}$ on $B_r(p)$ such that

1. $\Delta g_{ij} = 0$;
2. $Q^{-1} g_{ij} \leq g_{ij} = \langle \nabla x_i, \nabla x_j \rangle \leq Q \delta_{ij}$;
3. $$\sum_{1 \leq |\beta| \leq k} r^{|\beta|}|\partial^\beta g_{ij}| + \sum_{|\beta| = k} \frac{r^{k+\alpha}|\partial^\beta g_{ij}(x) - \partial^\beta g_{ij}(y)|}{(d_g(x, y))^\alpha} < Q - 1, \quad \forall x, y \in B_r(p).$$

**Definition 6.11.** We define the harmonic radius $r_H(Q, k, \alpha)$ to be

$$r_H(Q, k, \alpha) := \inf_{p \in M} r_H(Q, k, \alpha)(p).$$

Let us study how the harmonic radius scale. Given $(M, g)$, if we replace $g$ by $g_\lambda = \lambda^2 g$, then we have $d_{g_\lambda}(p, q) = \lambda d_g(p, q)$. As a result, $B^p_\lambda(p) = B^{d_g}_\lambda(p)$ as a set.
Let \((x_1, \ldots, x_n)\) be \((Q, k, \alpha)\) bounded harmonic coordinates on \(B^2_\alpha(p)\). Define \(\tilde{x}_i = \lambda x_i\), then \(\tilde{x}_i\) are harmonic coordinates on \(B^2_\lambda(p)\). Moreover,

\[
(g_\lambda)_{ij}(\tilde{x}) = \lambda^2 \left( \frac{1}{\lambda} \frac{\partial}{\partial x_i}, \frac{1}{\lambda} \frac{\partial}{\partial x_j} \right)_g = g_{ij}(\tilde{x}).
\]

So we have

\[
Q^{-1} \delta_{ij} \leq (g_\lambda)_{ij} \leq Q \delta_{ij}
\]

in \(B^2_\lambda(p)\). Similar computation shows that \(\tilde{x}_i\) also satisfies the \((Q, k, \alpha)\) bound. So we conclude that

\[
r^g_H(Q, k, \alpha) = \lambda r^g_H(Q, k, \alpha).
\]

In the rest of this section, we are going to prove that at any point \(p \in M\), there always exists a harmonic coordinate chart around \(p\).

**Theorem 6.12.** Given \((M, g)\) a Riemannian manifold and \(p \in M\), there exists \(r > 0\) such that \(B_r(p)\) has a harmonic coordinate system.

**Proof.** We will use the notation \(A \lesssim B\) to indicate that there is a constant \(C\) only depending on \((M, g)\) such that \(|A| \leq C|B|\).

Let \(Y = (y_1, \ldots, y_n)\) be normal coordinates at \(p\) defined on \(B_{R_0}(p)\). Since \(Y\) is normal coordinates,

\[
\Delta_g y_k = \frac{1}{\sqrt{\det g}} \partial_i (\sqrt{\det g} g^{ij} \partial_j y_k) = \frac{1}{\sqrt{\det g}} \partial_i (\sqrt{\det gg^{ik}}),
\]

so \(\Delta_g y_k(0) = 0\), and \(|\Delta_g y_k(y)| \sim \rho\) for \(|y| \leq \rho\).

Given \(\rho \in (0, R_0)\), let \(x_k = x^\rho_k\) be the solutions to the Dirichlet problem

\[
\begin{align*}
\Delta_g x_k &= 0, \quad \text{in } B_\rho(p), \\
x_k &= y_k, \quad \text{on } \partial B_\rho(p).
\end{align*}
\]

Then \(x_k\)’s are all harmonic functions.

It suffices to show that \((x_1, \ldots, x_k)\) form a coordinate system for \(\rho\) small. Consider \(x_k = x_k(y_1, \ldots, y_n)\), we want to show when \(\rho\) small enough \(\left(\frac{\partial x_k}{\partial y}\right)\) is invertible.

Rescale \(y \to \rho z\), i.e. \(z_i = \rho^{-1} y_i\), and \((\tilde{g})_{ij}(z)dz_i \otimes dz_j = (g_z)_{ij}(z)dz_i \otimes dz_j = \rho^{-2} \Delta g_{ij}(\rho y) dy_i \otimes dy_j\). Then \(B_\rho(p) = B^\rho_1(p)\), and

\[
\Delta_z x_k(z) = \frac{1}{\sqrt{\det \tilde{g}}} \partial_i ^\rho (\sqrt{\det \tilde{g}} \partial_j ^\rho x_k)(z) = \rho^2 \frac{1}{\sqrt{\det g}} \partial_i (\sqrt{\det g} \rho^2 g^{ij} \partial_j x_k)(y) = \rho^2 \Delta x_k(y) = 0,
\]

where \(\tilde{g}\) is the rescaled metric.
\[ \Delta_{z}y_k(z) = \frac{1}{\sqrt{\det g}} \partial_{i}^{z} (\sqrt{\det g g^{ij} \partial_{j}^{z} y_k})(z) \]
\[ = \rho^{2} \frac{1}{\sqrt{\det g}} \partial_{l}^{z} (\sqrt{\det g \rho^{2} g^{ij} \partial_{j}^{z} y_k})(\rho^{-1} y) \]
\[ = \rho^{2} \Delta y_k(y) \leq \rho^{3}. \]

As a result, \( \Delta_{z}(x_k - y_k) \lesssim \rho^{3} \) for \( z \in B_{\rho}^{z}(p) \).

Now we apply the Schauder estimate Theorem 5.29:
\[ \|x_k - y_k\|_{C^{2,\alpha}(B_{1/2}(p))} \lesssim \|\Delta_{z}(x_k - y_k)\|_{C^{\alpha}(B_{1})} + \|x_k - y_k\|_{C^{0}(B_{1})} \]
\[ \lesssim \|\Delta_{z}y_k\|_{C^{\alpha}(B_{1})} + \|x_k - y_k\|_{C^{0}(B_{1})}. \]

We estimate each terms on the right hand side.
\[ \|\Delta_{z}y_k\|_{C^{\alpha}(B_{1}^{z})} = \sup_{z_1, z_2 \in B_{1}^{z}} \frac{|\Delta_{z}y_k(z_1) - \Delta_{z}y_k(z_2)|}{|z_1 - z_2|^{\alpha}} \]
\[ = \sup_{z_1, z_2 \in B_{1}^{z}} \rho^{2+\alpha} \frac{|\Delta y_k(\rho z_1) - \Delta y_k(\rho z_2)|}{|\rho z_1 - \rho z_2|^{\alpha}} \]
\[ = \rho^{2+\alpha} \|\Delta y_k\|_{C^{\alpha}(B_{\rho})}. \]

We claim \( \|x_k - y_k\|_{C^{0}(B_{1}^{z})} \lesssim \rho^{3} \).

**Proof.** Let us consider a function \( F(z) = x_k - y_k - \tilde{C} \tilde{d}_{g}(z, p)^{2} \rho^{3} \), where \( \tilde{C} \) is to be determined. Assume that \( |\Delta_{z}(x_k - y_k)| \leq C \rho^{3} \). Then if \( \tilde{C} \) is sufficiently large, we have
\[ \Delta_{z}F(z) \leq C \rho^{3} - 2 \tilde{C} \text{Tr} (\tilde{g}) \rho^{3} < 0. \]

Then by maximum principle,
\[ x_k - y_k - \tilde{C} \tilde{d}_{g}(z, p)^{2} \rho^{3} \geq \inf_{\partial B_{1}} x_k - y_k - \tilde{C} \tilde{d}_{g}(z, p)^{2} \rho^{3} = \tilde{C} \rho^{3}. \]

So \( x_k - y_k \geq C' \rho^{3} \) for some constant \( C' \). Similarly we have \( y_k - x_k \geq C' \rho^{3} \). As a result, \( |x_k - y_k| \lesssim \rho^{3} \).

So now we know that \( \|x_k - y_k\|_{C^{2,\alpha}(B_{1/2}^{z})} \lesssim \rho^{2+\alpha} \). In particular,
\[ \left| \frac{\partial x_k}{\partial z_{l}} - \frac{\partial y_k}{\partial z_{l}} \right| \leq C \rho^{2+\alpha}, \text{ in } B_{1/2}^{z}. \]

Change of variables we have
\[ \left| \frac{\partial x_k}{\partial y_{l}} - \delta_{kl} \right| \leq C \rho^{1+\alpha}, \text{ in } B_{\rho/2}. \]

So if \( \rho \) is sufficiently small, \( (\frac{\partial y_k}{\partial y_{l}}) \) is invertible. Hence \( \{x_1, \ldots, x_n\} \) are coordinates. \( \square \)
6.3. $C^{k,\alpha}$ Topology. In this section, we discuss a notion of convergence which is stronger than the notion of Gromov-Hausdorff convergence.

**Definition 6.13.** A sequence of (pointed) complete Riemannian manifolds $(M_i, g_i, p_i)$ is said to converge in **pointed $C^{k,\alpha}$ topology**, written as

$$(M_i, g_i, p_i) \xrightarrow{C^{k,\alpha}} (M, g, p),$$

if for any $R > 0$, there exists a domain $\Omega \supset B(p, R)$ in $M$ and embeddings $F_i : \Omega \to M_i$ for all $i$ large enough such that

1. $F_i(p) = p_i$;
2. $F_i(\Omega) \supset B_{M_i}(p_i, R)$;
3. $F_i^*g_i \to g$ on $\Omega$ in $C^{k,\alpha}$ topology.

**Remark 6.14.** If $(M_i, g_i)$ have diameter uniformly bounded, then $F_i$’s are diffeomorphisms. So $M_i \cong M \cong M_j$ for sufficiently large $i, j$.

**Example 6.15.** A sequence of surfaces $M_i$ where $M_i$ has $i$ genus can not converge in $C^{k,\alpha}$ topology.

**Remark 6.16.** Even if $M_i = M$, convergence in $C^{k,\alpha}$ topology is NOT the same as the $C^{k,\alpha}$ convergence of the metrics $g_i \to g$. It is up to a self-diffeomorphism.

**Example 6.17.** Let $S^2 \subset \mathbb{R}^3$ be the unit sphere. Let $x, y, z$ be the coordinates of $\mathbb{R}^3$, and let $V = \pi_{TS^2}(-\partial_z)$. Let $\varphi_t$ be the flow of $V$. $\varphi_t$ can be viewed as the height descent flow from the north pole to the south pole. Let $g_t = \varphi_t^*g_{S^2}$. Then

$$g_t \to \begin{cases} g_{S^2}, & \text{at the north pole and the south pole;} \\ 0, & \text{elsewhere.} \end{cases}$$

So the metric does not converge in $C^{k,\alpha}$ topology, but the manifolds converge in $C^{k,\alpha}$ topology.

Now we study an Arezela-Ascoli type convergence theorem in $C^{k,\alpha}$ topology for manifolds.

**Theorem 6.18** (Fundamental Convergence Theorem, i.e. Arezela-Ascoli Theorem for manifolds). Given $Q > 1$, $n \geq 2$, $k \geq 0$, $\alpha \in (0, 1]$ and $r > 0$. Let

$$M^{k,\alpha}(Q, r, n) := \{(M^n, g) : r_H(Q, k, \alpha) > r \}.$$

Then $M^{k,\alpha}(Q, r, n)$ is compact in a pointed $C^{m,\beta}$ topology for $m + \beta < k + \alpha$.

**Proof.** The proof will be divided into several steps.

**Step 1.** We first show a sequence of $(M_i, g_i, p_i) \subset M^{k,\alpha}(Q, r, n)$ has a subsequence (still denoted by $(M_i, g_i, p_i)$) converging to a metric space $(X, d_X, p)$ in Gromov-Hausdorff topology.
Proof. Recall Theorem 4.12, it suffices to show that for any $R > 0$ and $\epsilon > 0$, there exists $C(R, \epsilon)$ such that $B^M_i(p_i, R)$ can be covered by $C(R, \epsilon)$ number of balls of radius $\epsilon$.

Let us first consider the case $R < r, \epsilon < r$. Then we are working in a harmonic coordinate chart, say $\{x_1, \ldots, x_n\}$. So we can treat $B_R(p_i)$ as a subset of the Euclidean ball $B_{R}^{\mathbb{R}^n}(0, QR)$, with metric $g_i$ not the Euclidean metric. Similarly for any $x \in B_R(p_i)$, we can treat $B(x, \epsilon)$ as a subset of $B_{QR}^{\mathbb{R}^n}(x, Q\epsilon)$. Note $B(x, \epsilon)$ also contains $B_{QR}^{\mathbb{R}^n}(0, Q-1\epsilon)$ as a subset. So we have

$$C(R, \epsilon) \leq \tilde{C}(n) \left(\frac{QR}{Q-1}\right) = \tilde{C}(n) \frac{Q^{2n} R^n}{\epsilon^n}.$$}

Now we use an induction argument to consider the case that $R$ may be larger than $r$. We only need to consider the case that $\epsilon < r$. Let us first consider $B(p, \frac{lr}{2})$ for $l \in \mathbb{N}$. Suppose $B(p, \frac{lr}{2})$ is covered by $N$ balls of radius $r/4$ centered at $x_1, \ldots, x_N$. Then by triangle inequality we have

$$B(p, \frac{(l+1)r}{2}) \subset \bigcup_{i=1}^{N} B(x_i, r).$$

But we have already shown that $B(x_i, r)$ can be covered by $P$ balls with radius $r/4$, here $P = C(n)Q^{2n}4^n$. Then we conclude that $B(p, \frac{(l+1)r}{2})$ can be covered by $NP$ balls with radius $r/4$. Then we can run induction to conclude that $B(p, lr/2)$ can be covered by $P^l$ balls of radius $r/4$.

Given $R$, let us choose $l$ such that $lr/2 < R \leq (l+1)r/2$. Then let us cover $B(p, R)$ by $M$ balls of radius $2\epsilon$, say $B(x_i, 2\epsilon)$ for $i = 1, \ldots, M$, such that $B(x_i, \epsilon)$ are disjoint. Then

$$\text{Vol}(B(p, R)) \leq N^{l+1} \max_{x \in B(p, R)} \text{Vol}(B(x, r/4)) \leq N^{l+1} Q^n r^n,$$

and when $\epsilon < r$, we have

$$\text{Vol}(B(p, R)) \geq M \min_{1 \leq i \leq M} \text{Vol}(B(x_i, \epsilon)) \geq C(n) MQ^{-n} \epsilon^n.$$

So

$$M \leq N^{l+1} Q^{2n} \left(\frac{r}{\epsilon}\right)^n.$$

Since $l \sim R/r$, we get

$$C(R, \epsilon) \sim N^{R/r} Q^{2n} \left(\frac{r}{\epsilon}\right)^n.$$

□

Step 2. Next we construct a $C^{m, \beta}$ atlas for $(X, d_X, p)$.

Proof. Given $q \in X$, let $q_i \in M_i$ such that $q_i \to q$ (See Remark 6.21). Suppose $(M_i, g_i) \in M^{k, \alpha}$, then there exists a homeomorphism covered by harmonic coordinates $\varphi_i : B(0, r) \to U_i \subset M_i$, where 0 is sent to $q_i$. Moreover, the pull back metric satifies the bound

$$Q^{-1}\delta_{jk} < (\varphi_i^* g_i)_{ij} < Q\delta_{jk}.$$
By the bound of the pull back metric we have
\[ d_{M_i}(\varphi_i(x_1), \varphi_i(x_2)) \leq Q|x_1 - x_2|, \]
so \( \varphi_i \) are uniformly Lipschitz. Therefore, there is a subsequence (still denoted by \( \varphi_i \)) converging to \( \varphi_\infty : B(0, r) \to (X, d_X) \), which send 0 to \( q \) (See Remark 6.22).

Moreover,
\[ d_{M_i}(\varphi_i(x_1), \varphi_i(x_2)) \geq Q^{-1} \min\{|x_1 - x_2|, 2r - |x_1| - |x_2|\}. \]
Here we note that on \( M_i \), the shortest distance between \( \varphi_i(x_1) \) and \( \varphi_i(x_2) \) can be realized by the length of a curve either inside \( \varphi_i(B(0, r)) \), or first goes outside of \( \varphi_i(B(0, r)) \) then goes back. Then we conclude that
\[ Q^{-1} \min\{|x_1 - x_2|, 2r - |x_1| - |x_2|\} \leq d_X(\varphi_\infty(x_1), \varphi_\infty(x_2)) \leq Q|x_1 - x_2|. \]
So \( \varphi_\infty \) is a homeomorphism onto its image.

Next we define a metric on \( \varphi_\infty(B(0, r)) \). Since \( \varphi_i^* g_i \) are \( C^{k,\alpha} \) Riemannian metrics on \( B(0, r) \), by Arezela-Ascoli Theorem, a subsequence of \( \varphi_i^* g_i \) (still denoted by \( \varphi_i^* g_i \)) converge to a limit \( g_\infty \) in \( C^m,\beta \) topology, on \( B(0, r) \). Moreover, we still have
\[ Q^{-1}\delta_{jk} \leq (g_\infty)_{jk} \leq Q\delta_{jk}, \]
and \( g_\infty \) also has the \( C^{m,\beta} \) bound with the constant \( Q \). Then we can use \( g_\infty \) as a Riemannian metric on \( \varphi_\infty(B(0, r)) \), and still denote it by \( g_\infty \). By taking limit we can see that \( g_\infty \) defines a local distance which is the same as \( d_X \).

\[ \square \]

**Step 3.** Next we show \( X \) has a differentiable structure by showing the transition functions are \( C^1 \).

**Proof.** If \( \varphi, \psi \) are two coordinate patches constructed as in step 2, then \( \varphi^{-1} \circ \psi \) is locally Lipschitz with respect to the Euclidean metric. Also
\[ \varphi^{-1} \circ \psi : (B(0, r), (g_\psi^\infty)_{ij}) \to (B(0, r), (g_\psi^\infty)_{ij}) \]
is an isometry. Here \( g_\psi^\infty, g_\psi^\infty \) are at least \( C^{0,\beta} \).

Then the following Theorem improves the regularity:

**Theorem 6.19** (Calabi-Hartman, [CH70] Theorem (ii)). If \( F : (B(0, r), g_1) \to (B(0, r), g_2) \) is a distance preserving map between \( C^\beta \) Riemannian metrics, then \( F \) is \( C^{1,\beta/2} \).

Although this is not exactly the same as the statement of Theorem (ii) in [CH70], but one can check the proof of Theorem (ii) in [CH70] and see that the statement here is true.

Thus \( X \) is actually a differentiable manifold with \( C^{1,\beta/2} \) atlas. Then the following theorem of Whitney shows that \( X \) is actually a smooth manifold.

**Theorem 6.20** (Whitney, [Whi36] Page 654, Section 8, Theorem 1). A manifold with a \( C^1 \) atlas has a compatible \( C^\infty \) structure.

So \( X \) is a \( C^\infty \) manifold with \( C^m,\beta \) metric \( g_\infty \).

\[ \square \]
Step 4. It remains to show that \((M_i, g_i, p_i)\) converge to \((X, g_\infty, p)\) in \(C^{m, \beta}\) topology.

Proof. Suppose \(\varphi_s : B(0, r) \rightarrow U_s \subset X\) is a countable atlas such that

\[
\varphi_s = \lim_{i \rightarrow \infty} \varphi_{is},
\]

where \(\varphi_{is} : B(0, r) \rightarrow U_{is} \subset M_i\), and

\[
\varphi^{*}_{is}(g_i) \xrightarrow{C^{m, \beta}} \varphi^{*}_{s} g_{\infty}.
\]

Define

\[
f_{is} = \varphi_{is} \circ \varphi^{-1}_{is} : U_s \rightarrow U_{is}.
\]

We need to glue \(f_{is}\) to construct a global mapping.

Let us do the gluing for two charts and the general case is similar. Suppose \(\mu_1, \mu_2\) is a partition of unity of \(U_1, U_2\), define

\[
F_i = \begin{cases} 
\varphi_{i2} \circ (\mu_1(x)\varphi^{-1}_{i1} \circ f_{i1} + \mu_2(x)\varphi^{-1}_{i2} \circ f_{i2}), & \text{on } U_2; \\
 f_{i1}, & \text{on } U_1.
\end{cases}
\]

The \(F_i\) is the desired glued map. \(\square\)

Finally, the Schauder theory can be used to show that \((X, g_\infty)\) \(\in C^{k, \beta}(Q, r, n)\). So we conclude the whole proof. \(\square\)

Remark 6.21. Although the convergence in Gromov-Hausdorff topology does not imply the
convergence of points in the metric space, we can still pick a sequence of points converging
to a point in the limit, in the following sense: Suppose \((X_i, d_{X_i}, p_i)\) converge to \((X, d_X, p)\) in
Gromov-Hausdorff topology, then there exist metric spaces \((Z_i, d_{Z_i})\) and isometric embedding
\(F_i : (X_i, d_{X_i}, p_i) \rightarrow (Z_i, d_{Z_i})\) and \(G_i : (X, d_X, p) \rightarrow (Z_i, d_{Z_i})\), with \(F_i(p_i) = G_i(p)\), and the
Hausdorff distance of \(F_i(X_i)\) and \(G_i(X)\) in \(Z\) is less than \(\epsilon_i\), and \(\epsilon_i \rightarrow 0\) as \(i \rightarrow \infty\). Therefore,
given \(q \in X\), we may pick \(q_i \in X_i\) such that \(d_{Z_i}(F_i(q_i), G_i(q)) \leq \epsilon_i\). Then we say the sequence
\(q_i \in X_i\) converging to \(q \in X\).

Remark 6.22. Let us clarify the (subsequence) convergence of \(\varphi_i : B(0, r) \rightarrow M_i\) if they
are uniformly Lipschitz, and \(\varphi_i(0) = q_i \rightarrow q\) as \(i \rightarrow \infty\).

Recall how we prove the classical Arezela-Ascoli Theorem: we first pick a countable dense
subset and prove that the functions pointwise convergence in this dense subset, then using
a denseness argument to show the convergence on the whole space. Uniformly Lipschitz is
used to get a uniform bound to prove pointwise convergence in the dense subset, and the
denseness argument.

Here we use the similar idea. Let \(S = \{x_j\}_{j=1}^{\infty}\) be a countable dense subset of \(B(0, r)\), for
instance we may pick all the rational points. Then \(d_{M_i}(x_j)(\varphi_i(x_j), \varphi_i(0)) \leq Q|x_j|\) implies
that \(d_{M_i}(x_j)(\varphi_i(x_j), q_i)\) are uniformly bounded. Again \((M_i, g_i, p_i)\) converge to \((X, d_X, p)\) in
Gromov-Hausdorff topology implies that there exist metric spaces \((Z_i, d_{Z_i})\) and isometric embedding
\(F_i : (M_i, d_{g_i}, p_i) \rightarrow (Z_i, d_{Z_i})\) and \(G_i : (X, d_X, p) \rightarrow (Z_i, d_{Z_i})\), with \(F_i(p_i) = G_i(p)\),
and the Hausdorff distance of $F_i(M_i)$ and $G_i(X)$ in $Z_i$ is less than $\epsilon_i$, and $\epsilon_i \to 0$ as $i \to \infty$. So we may pick a point $y_i^j \in M_i$ such that $F_i(\varphi_i(x_j))$ and $G_i(y_i^j)$ has distance $d_{Z_i}$ less than $\epsilon_i$. Then
\[
    d_X(q, y_i^j) = d_{Z_i}(G_i(q), G_i(y_i^j)) \\
    \leq d_{Z_i}(G_i(q), F_i(q_i)) + d_Z(F_i(q_i), F_i(\varphi_i(x_j))) + d_Z(F_i(\varphi_i(x_j)), G_i(y_i^j)) \\
    \leq Q|x_j| + \epsilon_i.
\]
So $d_X(q, y_i^j)$ is uniformly bounded by $Q|x_j| + \epsilon_0$. Moreover, $B^X(q, Q|x_j| + \epsilon_0)$ is compact because it is the Gromov-Hausdorff limit of $B^M(q, Q|x_j| + \epsilon_0)$ (see Theorem 4.7 and Remark 4.8). So a subsequence of $y_i^j$ converges to a limit $y_j$.

A standard diagonal trick allows us to pick a subsequence of $\varphi_i$ (still denoted by $\varphi_i$) converging to a limit $\varphi_\infty$ on each $x_j$, in the above sense. Now we claim that for $x_j, x_k \in S$,
\[
    d_X(\varphi_\infty(x_j), \varphi_\infty(x_k)) = \lim_{i \to \infty} d_M(\varphi_i(x_j), \varphi_i(x_k)).
\]
This is an easy consequence of the isometric embedding argument we used before:
\[
d_X(\varphi_\infty(x_j), \varphi_\infty(x_k)) = d_X(y_j, y_k) \leq d_X(y_j, y_i^j) + d_X(y_i^j, y_k) + d_X(y_i^j, y_k) \\
= d_X(y_j, y_i^j) + d_{Z_i}(G_i(y_i^j), G_i(y_k)) + d_X(y_k, y_k) \\
\leq d_X(y_j, y_i^j) + d_Z(F_i(\varphi_i(x_j)), F_i(\varphi_i(x_k))) + 2\epsilon_i + d_X(y_k, y_k) \\
= d_X(y_j, y_i^j) + d_M(\varphi_i(x_j), \varphi_i(x_k)) + 2\epsilon_i + d_X(y_k, y_k),
\]
and $d_X(y_j, y_i^j) \to 0$, $d_X(y_k, y_k) \to 0$ and $\epsilon_i \to 0$.

Now for any point $x \in B(0, r)$, we can define
\[
    \varphi_\infty(x) = \lim_{x_j \to x, x_j \in S} \varphi_\infty(x_j).
\]
Note that $d_X(\varphi_\infty(x_j), \varphi_\infty(x_k)) = \lim_{i \to \infty} d_M(\varphi_i(x_j), \varphi_i(x_k))$ implies that $d_X(\varphi_\infty(x_j), \varphi_\infty(x_k)) \leq Q|x_j - x_k|$, so $\varphi_\infty(x)$ is uniquely well-defined. Moreover, we have
\[
    d_X(\varphi_\infty(\tilde{x}_1), \varphi_\infty(\tilde{x}_2)) = \lim_{i \to \infty} d_M(\varphi_i(\tilde{x}_1), \varphi_i(\tilde{x}_2))
\]
by passing to limit.

Then we get the desired construction of the Arezela-Ascoli limit for these class of uniformly Lipschitz maps.

6.4. Lower Bound of Harmonic Radius. There are several applications of the Fundamental Convergence Theorem we developed in the previous section. We first discuss a theorem by Anderson on the lower bound of the harmonic radius of manifolds.

**Theorem 6.23** (Anderson). Given $n \geq 2$, $\alpha \in (0, 1)$, $\Lambda > 0$, $R > 0$, and fixed $Q > 1$. Then there exists $r = r(n, \alpha, \Lambda, R)$ such that if $(M^n, g)$ is compact, $|\text{Ric}(g)| \leq \Lambda$, $\text{inj}(g) \geq R$, then $r_H(Q, 1, \alpha) \geq r$. 
Proof. We argue by contradiction. Suppose not, then there exists a sequence of manifolds \((M_i, g_i)\) and \(p_i \in M_i\) such that \(r_H^{(M_i, g_i)}(Q, 1, \alpha)(p_i) = r_H^{(M_i, g_i)}(Q, 1, \alpha) \in (0, 1/i)\). i.e. there exists \((Q, 1, \alpha)\) bounded harmonic coordinates on \(B(p_i, r_i) \subset (M_i, g_i)\) where \(r_i \in (0, 1/i)\), but there is no such bounded harmonic coordinate on a larger ball.

Let us define \(g_i = r_i^{-2}g_i\). Then \(M_i, g_i\) satisfies
\[
\begin{aligned}
\{ |\text{Ric}(g_i)| &\leq r_i^2 \Lambda; \\
\text{inj}(g_i) &\geq r_i^{-1} R,
\end{aligned}
\]
and \(r_H^{(M_i, g_i)}(Q, 1, \alpha) = 1\).

Claim 1. Given \(\gamma \in (\alpha, 1)\), there exists \(C\) depending on \(\gamma, \Lambda, n, Q\) such that \(r_H^{(M_i, g_i)}(CQ, 1, \gamma) > 1/2\).

Proof. By assumption, for any \(q \in (M, \bar{g})\), there exists harmonic coordinates on \(B(q, 1)\) such that
\[
\begin{aligned}
&\text{(1)} \quad Q^{-1}\delta_{ij} < g_{ij} < Q\delta_{ij}, \\
&\text{(2)} \quad \sup_{B_1} |\partial \bar{g}| + |\partial \bar{g}|_C^\alpha \leq Q - 1, \\
&\text{(3)} \quad -r_i^2\Lambda g_{jk} \leq \text{Ric}(\bar{g}) \leq r_i^2\Lambda g_{jk}.
\end{aligned}
\]
Recall that we have the following equation
\[
\Delta g_{jk} + P(\bar{g}, \partial g) = -\text{Ric}(\bar{g})_{jk},
\]
here \(P\) is a quadratic function. By (1) we have \(\text{Ric}(\bar{g})\) is bounded, by (2) we have \(P(\bar{g}, \partial g)\) is bounded; also (1) implies that \(\Delta g = (\bar{g})^{jk}\partial_j \partial_k\) is uniformly elliptic with eigenvalues bounded between \(Q^{-1}\) and \(Q\), and (1) also implies that \((\bar{g})^{jk}\) are \(C^{1,\alpha}\).

Then by the elliptic regularity, we have
\[
\|g_{jk}\|_{W^{2,p}(B_{1/2})} \leq C(P, Q, \Lambda)(Q + 1).
\]
Take \(p\) large such that \(\gamma = 1 - n/p\), the Sobolev embedding shows that \(g\) is controlled in \(C^{1,\gamma}(B_{1/2})\). \(\square\)

From Claim 1 and Fundamental Convergence Theorem, we have a subsequence of \((M_i, \bar{g}_i)\) converging to \((M_\infty, g_\infty)\) in \(C^{1,\alpha}\) topology, where \(g_\infty \in C^{1,\alpha}\) satisfies the elliptic equation
\[
\Delta g_\infty(g_\infty)_{ij} + P(g_\infty, \partial g_\infty) = 0
\]
weakly in \(W^{1,2}\). By elliptic regularity, \(g_\infty\) is smooth and \(\text{Ric}(g_\infty) = 0\).

Claim 2. \((M_\infty, g_\infty) \cong (\mathbb{R}^n, g_{\text{Euc}})\).
Proof. Pick $p_i \in M_i$ and $V_i \in T_{p_i}M_i$, $|V_i| = 1$. Since $(M_i, g_i)$ converge to $(M, g)$ in $C^{1,\alpha}$ topology, we can assume $f_i : U \to M_i$ is an embedding from an open subset $U$ of $M_i$ to $M_i$, and $f_i^*(g_i)$ converge to $g$ in $C^{1,\alpha}$ topology. Then let $q_i \in M_i$ such that $p_i = f_i^{-1}(q_i)$ converge to $p$, and $V_i \in T_{q_i}M_i$ such that $|V_i| = 1$ and $f_i^*(V_i)$ converge to $V$.

Let $\gamma_i$ be the geodesic in $(M_i, g_i)$, through $p_i$, with $\gamma'_i(0) = V_i$. Then the injective radius bound implies that $\gamma_i$ is length minimizing on $(-r_i^{-1}R, r_i^{-1}R)$. Then the $C^{1,\alpha}$ theory of ODE implies that $\gamma_i \to \gamma$, where $\gamma$ is a geodesic in $M$. Moreover, $\gamma(0) = p$ and $\gamma'(0) = V$, and $\gamma$ is length minimizing on $(-\infty, \infty)$.

By Cheeger-Gromoll splitting theorem, we have $(M, g) \cong (N \times \mathbb{R}, dt^2 + \langle g_N \rangle)$. Repeat this argument $(n - 1)$ times we conclude that $(M, g) \cong (\mathbb{R}^n, g_{\text{Euc}})$.

From Claim 2 we know that the limit manifold $(M, g)$ is isometric to $(\mathbb{R}^n, g_{\text{Euc}})$, and in particular it has global harmonic coordinates. Now we argue as the same way as the proof of Theorem 6.12, i.e. the existence of harmonic coordinates to show that for $i$ large, there exists $(Q, 1, \alpha)$ bounded harmonic coordinates on $B(p_i, 2) \subset (M, g_i)$.

Let $\Omega$ be a domain in $\mathbb{R}^n$ such that $F_i(\Omega) \supset B(M, g_i)(p_i, 3)$, $F_i$ is an embedding. Suppose $x$ is the harmonic coordinate on $\Omega$, then $y_i = x \circ F_i^{-1}$ serves as an “almost harmonic” coordinate, i.e. $\Delta g_i y_i \sim 0$. Solve the Dirichlet problem

$$\begin{cases}
\Delta g_i z_i = 0, & \text{in } B(p_i, 2), \\
z_i = y_i, & \text{on } \partial B(p_i, 2).
\end{cases}$$

Similar to the argument in the existence of harmonic coordinate, by Schauder theory when $i$ large, $z_i$ are harmonic coordinates, with small $C^{1,\alpha}$ bound. However, this is a contradiction to $r_H^{(M, g_i)}(Q, 1, \alpha) = 1$. □

**Corollary 6.24.** Given $n \geq 2$, $\Lambda > 0$, $D > 0$, $R > 0$. Then the set of $(M^n, g)$ satisfying $|\text{Ric}| \leq \Lambda$, $\text{diam}(M) \leq D$, $\text{inj}(g) \geq R$ contains only finitely many diffeomorphism types.

**Remark 6.25.** The proof of this theorem actually illustrate a general principle. Roughly speaking we have “rigidity results give rise to ‘corresponding’ compactness results”. For example, Cheeger-Gromoll Theorem implies that some limit had to be $\mathbb{R}^n$, which gives us the lower bounds for $r_H$.

We will see some further application of this principle.
6.5. Gap Theorem of Euclidean Space. We have another rigidity result from Bishop-Gromov Theorem. Suppose \((M^n, g)\) has \(\text{Ric}(g) \geq 0\). Suppose that
\[
\frac{\text{Vol}(B(p, r))}{\omega_n r^n} = \nu
\]
is true for any \(r > 0\), then we have
- \(\nu = 1\) implies that \((M^n, g) \cong (\mathbb{R}^n, g_{\text{Euc}})\);
- \(\nu < 1\) implies that \((M^n, g) \cong \text{a cone over } p\).

We can use this rigidity result to prove the following Theorem.

**Theorem 6.26** (Anderson, Gap Theorem of Euclidean Space). Let \((M^n, g)\) be a complete Riemannian manifold with \(\text{Ric}(g) = 0\). Then there exists \(\epsilon = \epsilon(n) > 0\) such that if
\[
\frac{\text{Vol}(B(p, r))}{\omega_n r^n} \geq 1 - \epsilon,
\]
then \((M^n, g) \cong (\mathbb{R}^n, g_{\text{Euc}})\).

**Proof.** Clearly we have \(\text{diam}(M, g) = \infty\). Let us argue by contradiction. If the theorem fails, then there exists a sequence of manifolds \((M_i, g_i)\) such that \(\text{Ric}(g_i) = 0\), and \(p_i \in M_i\) such that
\[
\frac{\text{Vol}(B^{(M_i, g_i)}(p_i, r))}{\omega_n r^n} \geq 1 - \epsilon_i,
\]
where \(\epsilon_i \to 0\), but \((M_i, g_i) \neq (\mathbb{R}^n, g_{\text{Euc}})\).

By Bishop-Gromov, (*) is equivalent to
\[
limit_{r \to \infty} \frac{\text{Vol}(B^{(M_i, g_i)}(p_i, r))}{\omega_n r^n} \geq 1 - \epsilon_i.
\]

Therefore, we conclude that if (*) holds for \(p_i\), it must holds for any \(q \in M\), i.e.
\[
\frac{\text{Vol}(B^{(M_i, g_i)}(q, r))}{\omega_n r^n} \geq 1 - \epsilon_i.
\]

Next, let us fix \(p_i\) in \(M_i\). We claim \(\text{inj}_M(p_i) < \infty\). In fact, if \(\text{inj}_M(p_i) = \infty\), then Cheeger-Gromoll splitting Theorem implies that \((M_i, g_i) \cong (\mathbb{R}^n, g_{\text{Euc}})\), which is a contradiction. Let \(R_i = \text{inj}_{M_i}(p_i) < \infty\).

Take \(x_i\) in \(B(p_i, iR_i)\) be the point minimizing the function
\[
C_i(x) = \frac{\text{inj}_{M_i}(x)}{\text{dist}_{M_i}(x, \partial B(p_i, iR_i))}.
\]

Then we have
\[
C_i(x_i) \leq c_i(p_i) = \frac{R_i}{iR_i} = 1 \to 0, \quad \text{as } i \to \infty.
\]

For the simplicity of the notation, we use \(A_i\) to denote \(\partial B(p_i, iR_i)\). Let \(\overline{g}_i = (\text{inj}_{g_i}(x_i))g_i\). Then \(\overline{g}_i\) has the following properties:
(1) $\text{inj}_{\overline{g}_i}(x_i) = 1$;

(2) \[
\frac{\text{Vol}(M_i, g_i)(B(r))}{\omega_n r^n} \geq (1 - \epsilon_i);
\]

(3) \[
d_{\overline{g}_i}(x_i, A_i) = (\text{inj}_{\overline{g}_i}(x_i))^{-1}d_{\overline{g}_i}(x_i, A_i) = \frac{1}{C_i(x_i)} \to \infty;
\]

(4) More generally, for fixed $\overline{R} > 0$ and $\overline{x}$ such that $d_{\overline{g}_i}(x_i, \overline{x}) < \overline{R}$, we have

\[
\text{inj}_{\overline{g}_i}(\overline{x}) = C_i(\overline{x})d_{\overline{g}_i}(\overline{x}, A_i)
\]

by scaling invariance of $C_i$. After rescaling we have

\[
d_{\overline{g}_i}(\overline{x}, A_i) \geq d_{\overline{g}_i}(x_i, A_i) - d_{\overline{g}_i}(\overline{x}, x_i)
\]

\[
\geq d_{\overline{g}_i}(x_i, A_i) - \overline{R}.
\]

Since $d_{\overline{g}_i}(x_i, A_i) \to \infty$ by (3), so for $i$ sufficiently large (depending on $\overline{R}$) we have

\[
d_{\overline{g}_i}(\overline{x}, A_i) \geq \frac{1}{2}d_{\overline{g}_i}(x_i, A_i).
\]

Therefore

\[
\text{inj}_{\overline{g}_i}(\overline{x}) \geq C_i(\overline{x})d_{\overline{g}_i}(\overline{x}, A_i) \geq \frac{1}{2} \text{inj}_{\overline{g}_i}(x_i) = \frac{1}{2}.
\]

(5) $\text{Ric}(\overline{g}_i) = 0$.

Thus, for any $\overline{R} > 0$, for $i$ sufficiently large, $(B_{\overline{g}_i}(x_i, \overline{R}), \overline{g}_i)$ satisfies

\[
\text{inj}_{\overline{g}_i}(x_i) = 1, \quad \text{inj}(\overline{g}_i) \geq \frac{1}{2}, \quad \text{Ric}(\overline{g}_i) = 0.
\]

Now we can apply a slightly variant version of Theorem 6.23 to this sequence, i.e. the interior lower bound of harmonic radius for manifolds with boundary. The proof is similar to the compact case.

What we can conclude is that around every points on $B_{\overline{g}_i}(x_i, \overline{R}/2)$, there exists $C^{1,\alpha}$ bounded harmonic coordinates on a ball of fixed size in $B_{\overline{g}_i}(x_i, \overline{R})$. Since $\text{Ric}(\overline{g}_i) = 0$, in harmonic coordinate we have

\[
\Delta_{\overline{g}_i} \overline{g}_i + Q(\overline{g}_i, \overline{\nabla} \overline{g}_i) = 0.
\]

So elliptic regularity theory and Schauder theory implies that $\overline{g}_i$ are $C^{k,\alpha}$ bounded for any $k$. Thus by taking a sequence with $\overline{R} \to \infty$, $i \to \infty$, we get $C^{k,\alpha}$ convergence subsequence

\[
(M_i, \overline{g}_i, x_i) \to (M_\infty, g_\infty, x_\infty),
\]

where $(M_\infty, g_\infty, x_\infty)$ is a complete Ricci flat manifold.

From $C^{1,\alpha}$ convergence (which implies the convergence of the length of the geodesics), we have

\[
\frac{\text{Vol}(B_\infty(r))}{\omega_n r^n} = 1.
\]
So by the rigidity part of Bishop-Gromov, we conclude that \((M_\infty, g_\infty) \cong (\mathbb{R}^n, g_{\text{Euc}})\). However, 
\(g_i \to g_\infty\) in \(C^2\) implies that the injective radius also convergence, i.e.
\[
\text{inj}_{g_i}(X_i) \to \text{inj}_{g_\infty}(x_\infty).
\]
So \(\text{inj}_{g_\infty}(x_\infty) = 1\), which contradicts to \(g_\infty\) being the Euclidean metric on \(\mathbb{R}^n\).

\[\square\]

**Remark 6.27.** In the proof we use the fact that \((\mathbb{R}^n, g_{\text{Euc}})\) has infinite injective radius. This fact is always used to generate a contradiction in the study of convergence and rigidity of manifolds.

**Remark 6.28.** \(\text{Ric}(g) = 0\) is important, as opposed to \(\text{Ric}(g) \geq 0\). There are many examples of \(\text{Ric}(g) \geq 0\) and volume growth arbitrarily close to 1 which are not the standard Euclidean spaces.

As an application of Anderson’s Gap theorem, we prove a rigidity theorem of sphere by Anderson.

**Theorem 6.29** (Anderson, Rigidity Theorem of Sphere). Suppose \((M^n, g)\) is a compact manifold with \((n - 1) \leq \text{Ric}(g) \leq C\). Then there exists \(\epsilon = \epsilon(n, C)\) such that if 
\[
\text{Vol}(M^n, g) \geq (1 - \epsilon) \text{Vol} S^n(1),
\]
then \(M^n\) is diffeomorphic to sphere, and \(g\) is \(C^{1,\alpha}\) close to the round metric with sectional curvature equals to 1.

**Proof.** Let us first prove the following claim: there exists \(\epsilon = \epsilon(n, C)\) such that if \((n - 1) \leq \text{Ric}(g) \leq C\) and \(\text{Vol}(M^n, g) \geq (1 - \epsilon) \text{Vol} S^n(1)\), then for any \(Q > 1\), \(r_H(Q, 1, \alpha) \geq r_0 > 0\).

Before we prove this claim, let us make the following observation. By Myer’s Theorem 0.2, the diameter of \((M^n, g)\) is no more than \(\pi = \text{diam}(S^n(1))\). So by Bishop-Gromov Theorem, if we have \(\text{Vol}(M^n, g) \geq (1 - \epsilon) \text{Vol} S^n(1)\), then
\[
\frac{\text{Vol}_g(B(\pi))}{\text{Vol}_{S^n}(B(\pi))} = \frac{\text{Vol}(M^n, g)}{\text{Vol}(S^n(1))} \geq (1 - \epsilon),
\]
so for any \(0 < r < \pi\) we have
\[
(\star\star) \quad \frac{\text{Vol}_g(B(r))}{\text{Vol}_{S^n}(B(r))} \in [1 - \epsilon, 1].
\]

**Proof of Claim.** Suppose not. Then we can find a sequence of manifolds \((M_i, g_i)\) which satisfy \((n - 1) \leq \text{Ric}(g_i) \leq C\), \(\text{Vol}(M_i, g_i) \geq (1 - \epsilon_i) \text{Vol} S^n(1)\), and \(r_H^i = r_H^{(M_i, g_i)}(Q, 1, \alpha) \leq 1/i\), where \(\epsilon_i \to 0\).

Let us rescale the metric \(\overline{g}_i = (r_H^i)^{-2} g_i\), then the new metrics satisfy \((n - 1)(r_H^i)^2 \leq \text{Ric}(\overline{g}_i) \leq C(r_H^i)^2\), and \(r_H^i(\overline{g}_i) = 1\). So we have \(C^{1,\alpha}\) convergence 
\[
(M_i, \overline{g}_i, p_i) \to (M_\infty, g_\infty)
\]
to a limit manifold. Moreover, note the volume of a ball under the rescaling satisfies
\[
B_{\overline{g}_i}((r_H^i)^{-1} R) = (r_H^i)^{-n} \text{Vol} B_{g_i}(R),
\]
in the limit we have
\[
\frac{\Vol_{g_\infty}(B(r))}{\omega_n r^n} \in [1 - \epsilon, 1].
\]

Then by the Gap Theorem 6.26, we conclude that \((M_\infty, g_\infty)\) is isomorphic to \((\mathbb{R}^n, g_{\text{Euc}})\). However, passing to limit we obtain \(r_H(g_\infty) = 1\), which is a contradiction. \(\square\)

Now we prove the original theorem by contradiction. Suppose not. Then by the claim, we can find a sequence of manifolds \((M_i, g_i)\) such that \(r_H(g_i) \geq r_0\) ans \((**\) holds. Again, use Arezela-Ascoli Theorem of \(C^{1,\alpha}\) topology can find a \(C^{1,\alpha}\) limit \((M_\infty, g_\infty)\), and it satisfies
\[
\frac{\Vol_{g_\infty}(B(r))}{\Vol_{\mathbb{S}^n}(B(r))} = 1.
\]

If we know \(\text{Ric}(g_\infty) \geq (n-1)\), then \((M_\infty, g_\infty)\) is isomorphic to the standard sphere by the rigidity of Bihop-Gromov Theorem. However, \(g_i \to g_\infty\) only in \(C^{1,\alpha}\), so \(g_\infty\) may only be \(C^{1,\alpha}\), then we can not even define \(\text{Ric}(g_\infty)\)!

So we are going to prove \(g_\infty\) is actually smooth. Since \(g_i\) converge to \(g_\infty\) in \(C^{1,\alpha}\), the volume also converges. Thus we have
\[
\int_{M_i} |\text{Ric}(g_i) - (n-1)g_i|^p d\Vol_{g_i} \to 0.
\]

So \(g_\infty\) solves \(\text{Ric}(g_\infty) = (n-1)g_\infty\) weakly. Then by \(C^{1,\alpha}\) bounds and elliptic regularity we have \(g_\infty\) is actually smooth, and solves \(\text{Ric}(g_\infty) = (n-1)g_\infty\) classically.

Therefore we have proved that \((M_\infty, g_\infty)\) is the standard sphere. Thus \((M_i, g_i)\) are diffeomorphic to a standard sphere with metric \(g_i\) is \(C^{1,\alpha}\) closed to the round metric, which is a contradiction. \(\square\)

7. Volume Stability

In this section, we will prove a volume stability result due to Colding. It states that if a manifold with non-negative Ricci curvature is Gromov-Hausdorff close to Euclidean space, then the volume is close to Euclidean volume. More precisely, we will prove the following theorem.

**Theorem 7.1.** For any \(\epsilon > 0\), there exists \(\delta = \delta(\epsilon)\) such that the following always holds: if \((M, g)\) is a complete manifold with \(\text{Ric}(g) \geq 0\) and \(x \in M\), \(r > 0\) and assume
\[
d_{GH} \left((B^M(x, r), x), (B^{\mathbb{R}^n}(0, r), 0)\right) < \delta r
\]
then
\[
\text{Vol}(B^M(x, \delta r)) > (\omega_n - \epsilon)(\delta r)^n.
\]

We will prove the theorem under a slightly general assumption: instead of assuming non-negative Ricci curvature, we will assume
\[
\text{Ric}(g) \geq -(n-1) \left(\frac{\delta}{r}\right)^2 g.
\]
Remark 7.2. The techniques in the proof can also prove the Cheeger-Gromoll splitting theorem for limit spaces. More precisely, suppose
\[(M_i, g_i, p_i) \xrightarrow{GH} (X, d_X, p)\]
and \(\text{Ric}(g_i) \geq \epsilon_i \to 0\), then we say \((X, d_X)\) is a length space. Moreover, if it contains a line, then \((X, d_X) \cong (Z \times \mathbb{R}, d_Z + d_\mathbb{R})\) for some metric space \(Z\).

The rough idea of the proof goes as follows: by rescaling, we can assume \(r = 1\). Take standard orthonormal basis \(e_j\) in \(\mathbb{R}^n\). By Gromov-Hausdorff closeness assumption, we will find corresponding points in \(\tilde{e}_j \in B^M(x, 1)\) such that \(d_{GH}(\tilde{e}_j, e_j) < \delta\). Let \(\tilde{b}_j(p) = d(\tilde{e}_j, p) - d(\tilde{e}_j, x)\) be the analog of Busemann function in the proof of Cheeger-Gromoll splitting. Solve Dirichlet problem:
\[
\begin{aligned}
\Delta b_j &= 0 \quad \text{on } B^M(x, 1) \\
b_j &= \tilde{b}_j \quad \text{on } \partial B^M(x, 1)
\end{aligned}
\]

We will show that this gives an “almost isometry” \((b_1, ..., b_n) : B^M(x, 1) \to \mathbb{R}^n\) in the sense that \(\langle \nabla b_j, \nabla b_k \rangle \approx \delta_{jk}\).

Before the proof, we first need some integral estimates.

Definition 7.3. Let \((M, g)\) be a complete Riemannian manifold. Define \(SM = \{v \in TM : |v|_g = 1\}\) be the sphere bundle equipped with the standard measure that gives unit sphere of volume 1.

Given \(v \in TM\), with a slight abuse of notation, denote \(\gamma_v(t)\) the geodesic with \(\gamma_v'(0) = 0\).

Lemma 7.4. For any \(t \in \mathbb{R}\), the map \(SM \to SM\) by \(v \mapsto \gamma'_v(t)\) is measure preserving.

Lemma 7.5. Let \((M, g)\) be a complete manifold, \(p \in M\) and \(l, r > 0\). Then for any \(C^2\) function \(f : B(p, r + l) \to \mathbb{R}\) and any \(t \in [0, l]\), we have
\[
(7.1) \quad \int_{SB(p, r)} \left| (f \circ \gamma_v)'(t) - \frac{f(\gamma_v(l)) - f(\gamma_v(0))}{l} \right| \leq l \int_{B(p, r+l)} |\nabla^2 f|\]
\[
= \int_{SB(p, r)} \left| \langle \nabla f, v \rangle - \frac{f(\gamma_v(l)) - f(\gamma_v(0))}{l} \right| \leq l \int_{B(p, r+l)} |\nabla^2 f|
\]

Proof. It is sufficient to prove the first estimate. To get the second inequality, simply apply \(t = 0\) in inequality (7.1).

By mean value theorem, there exists \(s \in [0, l]\) such that
\[
(f \circ \gamma_v)'(s) = \frac{f \circ \gamma_v(l) - f \circ \gamma_v(0)}{l}
\]

Therefore for any \(t \in [0, l]\), we have
\[
\left| (f \circ \gamma_v)'(t) - \frac{f(\gamma_v(l)) - f(\gamma_v(0))}{l} \right| \leq \int_0^l (f \circ \gamma_v)'(s) \, ds \leq \int_0^l |\nabla^2 f|(\gamma_v(s)) \, ds
\]
Integrate both side on $SB(p, r)$ and Lemma 7.4 implies

$$
\text{LHS of (7.1)} \leq \int_0^t \int_{SB(p, r)} |\nabla^2 f|\left(\gamma_v(s)\right) \, ds \\
\leq \int_0^t \int_{SB(p, r)} |\nabla^2 f|\left(\gamma_v(0)\right) \, ds \\
\leq l \int_{SB(p, r+l)} |\nabla^2 f| \\
\Box
$$

Next, we want integral estimates of the laplacian of “Busemann” functions. Here is an important principle to keep in mind when proving some integral estimates like

$$
\int_{B(x, R)} \Delta f \leq C
$$
given $\Delta f \geq 0$. It is suffices to understand

1. $\|f\|_{L^\infty(B(x,2R))}$
2. The existence of “good” cutoff functions.

For example, if there exists $\varphi : M \to [0, 1]$ such that equals 1 in $B(x, R)$ and 0 outside $B(x, 2R)$. Then,

$$
\int_{B(x, R)} \Delta f \leq \int_{B(x, 2R)} \varphi \Delta f = \int_{B(x, 2R)} f \Delta \varphi
$$

If we can control $L^\infty$ norm of $f$ and the Hessian of the cutoff function, we then easily get an integral estimate for $f$.

Here’s an example of this idea. Suppose $\text{Ric}(g) \geq 0$ and $\Delta u = 0$. Then by Bochner’s formula, $\Delta \frac{1}{2} |\nabla u|^2 = \text{Ric}(\nabla u, \nabla u) + |\text{Hess} u|^2$. If $|\nabla u|^2 - 1$ is small in integral sense or $L^\infty$ sense and good cutoff function exists, then by previous observation, integrating Bochner formula gives

$$
\int |\text{Hess} u|^2 \leq \int \frac{1}{2} (|\nabla u|^2 - 1) \Delta \varphi.
$$

Combining this with Lemma 7.5, we conclude that $u$ is almost linear “on average”.

In fairly general settings, such function indeed exists.

**Proposition 7.6.** For any $\rho \in (0, 1)$, there exists $c = c(\rho) < +\infty$ such that if $(M, g)$ has $\text{Ric}(g) \geq -(n-1)$ and $p \in M$, $0 < r \leq 1$, then there exists $\varphi \in C^\infty(M, \mathbb{R})$ with the following properties:

1. $\text{supp} \varphi \subset B(p, r)$.
2. $\varphi \equiv 1$ on $B(p, \rho r)$.
3. $|\nabla \varphi| \leq c/r$.
4. $|\nabla^2 \varphi| \leq c/r^2$. 
Assuming this, we can prove the following Lemma.

**Lemma 7.7.** For any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon)$ such that the following holds. If $(M, g)$ is a complete Riemannian manifold with $\text{Ric}(g) \geq -(n-1)\frac{\delta^2}{r} g$, $r > 0$, $p \in M$ and there exists $q_{\pm} \in M$ with $d(p, q_{\pm}) < \delta r$ and further assume that for any $x \in B(p, 2r)$

$$|b_+(x) + b_-(x)| < \delta r$$

where $b_{\pm}(x) = d(x, q_{\pm}) - d(p, q_{\pm})$. Then,

$$r \int_{B(p, r)} |\Delta b_{\pm}| < \varepsilon$$

**Proof.** By rescaling, $\tilde{g} = r^{-2} g$, we have $\text{Ric}(\tilde{g}) = \text{Ric}(g) \geq -(n-1)(\delta/r)^2 g = -(n-1)\delta^2 \tilde{g}$. Hence we can assume, without the loss of generality, that $r = 1$.

By assumption, $|b_+ + b_-| < \delta$ on $B(p, 1)$. Since $d(x, q_{\pm}) > \frac{1}{10} \delta^{-1}$ on $B(p, 2)$, we have $\Delta b_{\pm} \leq 100(n-1)\delta$ on $B(p, 2)$ by Laplacian comparison. Let $\varphi$ be the cutoff function from Proposition 7.6 with $r = 2$ and $\rho = 1/2$. Then,

$$\left| \int_{B(p, 2)} \varphi \Delta (b_+ + b_-) \right| = \left| \int_{B(p, 2)} (b_+ + b_-) \Delta \varphi \right| \leq C\delta$$

But we also have bound $\Delta b_{\pm} \leq 100(n-1)\delta$. Therefore the non-negative part of $\Delta (b_+ + b_-)$ also has estimate

$$\int_{B(p, 2)} \chi_{\Delta (b_+ + b_-) \geq 0} \cdot \varphi \Delta (b_+ + b_-) \leq C\delta.$$

Combining this with the previous estimate, we then obtain

$$\int_{B(p, 1)} |\Delta (b_+ + b_-)| \leq \int_{B(p, 2)} \varphi |\Delta (b_+ + b_-)| \leq C\delta$$

Use $\Delta b_{\pm} \leq 100(n-1)\delta$ again to get integral estimate for $\Delta b_{\pm}$. Then we can choose $\delta$ small to make the right hand side less than $\varepsilon$. \(\square\)

Next, we make precise the notion of “almost” splitting as an approximate analog of Cheeger-Gromoll splitting.

**Definition 7.8.** ($\varepsilon$-splitting) Let $\varepsilon > 0$, $(M, g)$ be a Riemannian manifold with $p \in M$ and a scale $r > 0$. A map $u = (u^1, \ldots, u^n) : B(p, r) \to \mathbb{R}^n$ is an $\varepsilon$-splitting if it has the following properties.

1. $u^l$ is harmonic for $l = 1, \ldots, n$.
2. For any $x \in B(p, r)$, we have

$$|\nabla u^l(x)| := \sup_{v \in T_x M, |v| = 1} |\nabla_v u^l| < 1 + \varepsilon$$
(3) For all \( l_1, l_2 \in \{1, ..., n\} \), we have
\[
\left| \langle \nabla u^{l_1}, \nabla u^{l_2} \rangle - \delta_{l_1 l_2} \right| < \varepsilon^2
\]

(4) For all \( l \in \{1, ..., n\} \), we have
\[
r^2 \int_{B(p,r)} \left| \nabla^2 u^l \right|^2 < \varepsilon
\]

We also say \( u \) is a weak \( \varepsilon \)-splitting if it satisfies (1), (3) and (4).

For manifolds with non-negative Ricci curvature and is Gromov-Hausdorff closed to Euclidean space, there exists an \( \varepsilon \)-splitting that simultaneously realizes Gromov-Hausdorff Hausdorff approximation. The following proposition is a partial version of this statement.

**Proposition 7.9.** For any \( \varepsilon > 0 \), there exists \( \delta = \delta(\varepsilon) > 0 \) such that the following holds.

If \((M, g)\) is a complete Riemannian manifold with \( \text{Ric}(g) \geq -(n-1)(\delta/r)^2 g \), \( p \in M \), \( r > 0 \) and assume Gromov-Hausdorff closeness

\[
d_{GH}(\left[ B^M(p, \delta^{-1} r), p \right], [B^{\mathbb{R}^n}(0, \delta^{-1} r), 0]) < \delta r
\]

Then there is a \( C^\infty \) mapping \((u) = (u^1, ..., u^n) : B^M(p, r) \to \mathbb{R}^n\) such that

1. \( u \) is an \( \varepsilon r \)-Gromov-Hausdorff approximation onto \( B^{\mathbb{R}^n}(0, r) \),
2. \( u^l \) is harmonic for all \( l \in \{1, ..., n\} \), and
3. for each \( l \), we have
\[
\int_{B(p,r)} \left| \nabla u^l \right|^2 - 1 < \varepsilon^2
\]

The proof of this proposition uses Poincaré inequality and Cheng-Yau gradient estimate, which we will state here and omit the proof.

**Theorem 7.10.** (Poincaré inequality) If \((M, g)\) has \( \text{Ric}(g) \geq -1 \), \( p \in M \) and \( 0 < r < A \) a radius such that \( B(p, 2r) \neq M \). Then for any \( q \geq 1 \), there exists constant \( C = C(A, q) \) such that for any \( f \in C^1_c(B(p, r)) \), we have
\[
\int_{B(p,r)} |f|^q \, dg \leq Cr^q \int_{B(p,r)} |\nabla f|^q \, dg
\]

**Theorem 7.11.** (Cheng-Yau gradient estimate) Suppose \((M, g)\) complete and \( \text{Ric}(g) \geq -(n-1)K \) with \( K \geq 0 \). Suppose \( u \) solves \( \Delta u = 0 \) on \( B(p, 2R) \) and \( u \geq 0 \). Then there exists a universal constant \( C = C_n \) depending only on the dimension \( n \) such that
\[
\frac{|\nabla u|}{u} \leq C_n \frac{1 + R\sqrt{K}}{R} \quad \text{on} \quad B(p, R)
\]

Assuming these two estimates, we now prove the previous proposition.
Proof of Proposition 7.9. By rescaling argument as before, we can fix a scale \( r = 1 \). Throughout the proof, we will use the notation \( \Phi(\delta) \) to denote some positive continuous function depending on the universal data and \( \Phi(\delta) \to 0 \) when \( \delta \to 0 \). It may vary from equations to equations.

Let \( e^j = (0, \ldots, 0, 1, 0, \ldots, 0) \) be the standard basis in \( \mathbb{R}^n \). Since \( B^M(p, \delta^{-1}) \) is \( \delta \) Gromov-Hausdorff close to \( B^{\mathbb{R}^n}(0, \delta^{-1}) \), we can find points \( q^j_\pm \in M \) corresponding to \( \pm \frac{1}{2} \delta^{-1} e^j \) under Gromov-Hausdorff approximation. Picking \( \delta \) small, we can have \( q^j_\pm \in B^M(p, \frac{1}{2} \delta^{-1} + \delta) \subset B^M(p, \delta^{-1}) \).

Define the Busemann functions, \( b^j_\pm(x) = d(x, q^j_\pm) - d(p, q^j_\pm) \). Then for any \( x \in B(p, 10) \) we have

\[
|b^j_+(x) + b^j_-(x)| < 10 \delta
\]

This is not immediate in \( M \) but it is trivial in \( \mathbb{R}^n \). But using Gromov-Hausdorff closeness we can add error term \( \delta \) for each distance function appeared in the estimate.

Then by Lemma 7.7, we have estimate

\[
\int_{B(p, 10)} |\Delta b^j_\pm| < \Phi(\delta)
\]

Let \( u^j \) solves

\[
\begin{cases}
\Delta u^j = 0 & \text{on } B(p, 10) \\
u^j = b^j_+ & \text{on } \partial B(p, 10)
\end{cases}
\]

Note that, if \( \partial B(p, 10) \) is not smooth, we will use a slightly larger smoothly bounded domain. For simplicity, we will assume it has smooth boundary.

We can view \( u^j \) as harmonic approximation of \( b^j_+ \). In particular, we have

\[
\int_{B(p, 10)} |\nabla (u^j - b^j_+)|^2 = -\int_{B(p, 10)} (u^j - b^j_+)(\Delta (u^j - b^j_+)) = \int_{B(p, 10)} (u^j - b^j_+)(\Delta b^j_+).
\]

By maximum principle, \( \sup_{B(p, 10)} |u^j| = \sup_{\partial B(p, 10)} |u^j| = \sup_{\partial B(p, 10)} |b^j_+| < 10 \) while the last step uses the triangle inequality. By estimate (7.2), we have

\[
\int_{B(p, 10)} |\nabla (u^j - b^j_+)|^2 < \Phi(\delta)
\]
Since $|\nabla b^j_+| = 1$ almost everywhere, we can write in the integral sense,
\[
\int_{B(p,10)} |\nabla u^l|^2 - 1 = \int_{B(p,10)} (u^l + b^j_+) \cdot \nabla (u^l - b^j_+)
\]
\[=
\int_{B(p,10)} |\nabla (u^l - b^j_+)|^2 + 2\nabla b^l_+ \cdot \nabla (u^l - b^j_+)
\]
\[< \Phi(\delta) + 2 \left( \int_{B(p,10)} |\nabla (u^l - b^j_+)|^2 \right)^{\frac{1}{2}} < \Phi(\delta)
\]
This concludes part (3) of Proposition 7.9. It remains to show that \(u = (u^1, ..., u^n)\) is a Gromov-Hausdorff approximation of \(B^M(p, 1)\) to \(B^{R^n}(0, 1)\).

We first show that \(b = (b^1_+, ..., b^n_+)\) gives a \(\Phi(\delta)\) Gromov-Hausdorff approximation from \(B^M(p, 10)\) to \(B^{R^n}(0, 10)\).

Clearly, we have \(\Phi(\delta)\) on \(B(p, 1)\). Fix a \(\delta\) Gromov-Hausdorff approximation \(\varphi : B^M(p, \delta^{-1}) \to B^{R^n}(0, \delta^{-1})\) which maps \(p\) to 0. Then
\[
b^j_+ (x) = d^M(x, q^j_+) - d^M(p, q^j_+)
\]
\[\approx d^{R^n}(\varphi(x), \frac{1}{2}\delta^{-1}e^j) - d^{R^n}(0, \frac{1}{2}\delta^{-1}e^j) \pm 4\delta
\]
\[\approx [\varphi(x)]_j \pm \Phi(\delta)
\]
That is \(|b^j_+ (x) - [\varphi(x)]_j| < \Phi(\delta)|j\) where \([\varphi(x)]_j\) is the \(j\)-th coordinate of \(\varphi(x)\). Since \(\varphi\) gives a \(\delta\) Gromov-Hausdorff approximation on \(B^M(p, \delta^{-1})\), by choosing \(\delta\) small, \(b\) gives a \(\Phi(\delta)\) approximation from \(B^M(p, 10)\) to \(B^{R^n}(0, 10)\).

It is then sufficient to show that \(|u - b| < \Phi(\delta)| on \(B(p, 1)\). This goes by contradiction.

Since \(u^l (x) \geq -10\) on \(B(p, 10)\) by using maximum principle as before, Cheng-Yau estimate (Theorem 7.11) implies that
\[
\frac{|\nabla u^l|}{u^l + 10} \leq \frac{C_n(1 + 5\sqrt{\delta})}{5} = C_n
\]
Thus \(|\nabla u^l| \leq C_n because \(|u^l| \leq 10|.

Now assume for a contradiction that \(|u^l (x) - b^j_+ (x)| > a\) at some \(x \in B(p, 4)\), then \(|u^l - b^j_+| > a/2\) on \(B(x, \frac{a}{2C_n}) \cap B(p, 5)\) by the gradient estimate. Poincare inequality and previous integral estimate on \(|\nabla (u^l - b^j_+)|^2| implies that
\[
\int_{B(p,10)} |u^l - b^j_+|^2 \leq C \int_{B(p,10)} |\nabla (u^l - b^j_+)|^2 < \Phi(\delta)
\]
But this contradicts to the volume growth on \(M\). More specifically, we have
\[
\int_{B_{p,5}} |u^l - b^j_+|^2 \geq \frac{a^2}{4} \text{Vol} \left(B \left(x, \frac{a}{2C_n}\right)\right)
\]
Then,
\[
\frac{a^2 \text{Vol} \left( B \left( x, \frac{a}{2C_n} \right) \right)}{4 \frac{\text{Vol}(B(p,10))}{\text{Vol}(B(p,10))}} < \Phi(\delta)
\]

But by Bishop-Gromov,
\[
\frac{\text{Vol} \left( B \left( x, \frac{a}{2C_n} \right) \right)}{\text{Vol}(B(p,10))} \geq \frac{\text{Vol} \left( B \left( x, \frac{a}{2C_n} \right) \right)}{\text{Vol}(B(x,14))} \geq \frac{1}{14^n} \left( \frac{a}{2C_n} \right)^n
\]

Picking \( \delta \) small and this yields contradiction. Hence \( |u^i - b^i_+| < \Phi(\delta) \) on \( B(p,4) \) and this concludes part (1) in the proposition. \( \square \)

Theorem 7.9 assures the bound for the average of the first order derivatives of the splitting map \( u \). Next Lemma provides a bound for the average of the second order derivatives.

**Lemma 7.12.** There exists \( C > 0 \) such that the following holds: Suppose \( (M, g) \) is complete, \( p \in M \), \( r > 0 \) with \( \text{Ric}(g) \geq -(n-1)(\delta/r)^2 g \), and \( u \) is a harmonic function on \( B(p,2r) \). Then
\[
2 \int_{B(p,2)} |\nabla^2 u|^2 \varphi \leq \left( \int_{B(p,2)} \left( |\nabla u|^2 - a \right) \delta^2 (n-1) \varphi \right) + 2a\delta^2(n-1) + \left( \int_{B(p,2)} (|\nabla u|^2 - a) \Delta \varphi \right)
\]
\[
\leq \left( C \int_{B(p,2)} |\nabla u|^2 - a \right) + 2a\delta^2(n-1).
\]

Proof. Note the statement of the lemma implies that the inequality holds for any \( a \in \mathbb{R} \). Without loss of generality assume \( r = 1 \). Recall the Bochner formula
\[
\Delta(|\nabla u|^2 - a) \geq -2\delta^2(n-1)|\nabla u|^2 + 2|\nabla^2 u|^2.
\]
Let \( \varphi \) be the cut-off function from Proposition 7.6, i.e. \( \varphi \) satisfies
\[
\begin{cases}
\varphi = 1, \text{ on } B(p,1) \\
\varphi = 0, \text{ outside } B(p,2) \\
|\nabla \varphi| + |\Delta \varphi| \leq C.
\end{cases}
\]

Then we have
\[
2 \int_{B(p,2)} |\nabla^2 u|^2 \varphi \leq \left( \int_{B(p,2)} (|\nabla u|^2 - a) \delta^2 (n-1) \varphi \right) + 2a\delta^2(n-1) + \left( \int_{B(p,2)} (|\nabla u|^2 - a) \Delta \varphi \right)
\]
\[
\leq \left( C \int_{B(p,2)} |\nabla u|^2 - a \right) + 2a\delta^2(n-1).
\]

On the other hand, we have
\[
\int_{B(p,2)} |\nabla^2 u|^2 \varphi \geq \frac{1}{\text{Vol}(B(p,2))} \int_{B(p,1)} |\nabla^2 u|^2 \geq C \int_{B(p,1)} |\nabla^2 u|^2
\]

By Bishop-Gromov Theorem. Then we conclude the Lemma. \( \square \)

Now we start proving the volume stability Theorem. We first prove that there exists a good \( \varepsilon \)-splitting if \( M \) is Gromov-Hausdorff closed to \( \mathbb{R}^n \).
Theorem 7.13 (Colding). For any $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that the following holds.

Suppose $(M, g)$ is a complete Riemannian manifold, $p \in M$, $\text{Ric}(g) \geq -(n-1)(\delta/r)^2 g$, $r > 0$. Then if

$$d_{GH}([B^M(p, \delta^{-1}r), p], [B^{\mathbb{R}^n}(0, \delta^{-1}r), 0]) < \delta r,$$

there exists an $\epsilon$-splitting $u : B^M(p, r) \to \mathbb{R}^n$, with

$$||\nabla u(x) - d(p, x)|| < \epsilon$$

for any $x \in B^M(p, r)$.

**Sketch of Proof.** Without loss of generality, let us assume $r = 1$. Proposition 7.9 shows that on $B(p, 16)$, there exists $u : B^M(p, 16) \to \mathbb{R}^n$ such that

1. $u$ is an $\Psi(\delta)$-Gromov-Hausdorff approximation, and $u(p) = 0$;
2. $u$ is harmonic;
3. for any $l$,

$$\int_{B^M(p, 16)} |\nabla u|^2 - 1 \leq \Psi(\delta).$$

Moreover, Lemma ?? implies that

$$(\star) \quad \int_{B^M(p, 8)} |\nabla^2 u|^2 \leq \Psi(\delta).$$

Here $\Psi(\delta)$ is a function depending on $\delta$ such that $\Psi(\delta) \to 0$ as $\delta \to 0$.

The idea is to show that $u = (u_1, \cdots, u_l)$ is a coordinate which is almost Euclidean coordinate. The main issue is to show that

$$\int_{B^M(p, 4)} \langle \nabla u_{l_1}, \nabla u_{l_2} \rangle^2 \leq \Psi(\delta), \quad l_1 \neq l_2.$$

By Lemma 7.5, we have

$$\int_{SB(p, 4)} |\langle \nabla u_{l_1}, V \rangle - (u_{l_1}(\gamma_V(1)) - u_{l_1}(\gamma_V(0)))| \leq \int_{B(p, 5)} |\nabla^2 u_{l_1}|^2.$$

$(\star)$ implies that the right hand side of the above formula is bounded by

$$(\star\star) \quad \Psi(\delta) \text{Vol}(B(p, 5)) \leq \tilde{\Psi}(\delta) \text{Vol}(B(p, 4)),$$

By Bishop-Gromov.

Now we sketch the rest part of the proof. It is not hard to make it rigorous. By Proposition 7.9, $|\nabla u_{l_1}|^2 \approx 1$ on average. So if we pick $V = \nabla u_{l_1}$, the above estimate implies that

$$|u_{l_1}(\gamma_V(1)) - u_{l_1}(\gamma_V(0))| \sim 1.$$
But $d(\gamma_V(1), \gamma_V(0)) \leq 1$, and

$$d(\gamma_V(1), \gamma_V(0)) \sim \sqrt{\sum_{i=1}^{n} (u_i(\gamma_V(1)) - u_i(\gamma_V(0)))^2}.$$ 

Then we conclude that

$$|u_{i_2}(\gamma_V(1)) - u_{i_2}(\gamma_V(0))| \sim 0.$$ 

Then the (**) bound implies that

$$|\langle \nabla u_{i_2}, V \rangle| \sim 0.$$ 

Note $V \sim \nabla u_1$, so

$$|\langle \nabla u_{i_2}, \nabla u_1 \rangle| \sim 0.$$ 

In conclusion, we have proved the existence of weak $\varepsilon$-splitting. To show it is a $\varepsilon$-splitting, we only need to show the gradient estimate of $u$. This follows from the gradient estimate of Cheng-Yau. We omit it here, but provide the reference [Yau75].

Finally, the closeness of $u(x)$ and $d(p, x)$ comes from the closeness of $u$ and the coordinates. □

Now we are going to prove the volume stability Theorem 7.1. Let us briefly discuss a scenario which we want to rule out, which is called $\delta \mathbb{Z}$-web.

**Example 7.14 ($\delta \mathbb{Z}$-web).** Assume $r = 1$. $\delta \mathbb{Z}$-web is a manifold $M^n$ be isometric embedded into $\mathbb{R}^n$ with the following image: the image consists of many solid cylinders, forming a web, where the radius of the cylinders are roughly $\delta^2$, while the “holes” of the web has side length $\delta$. See Figure 1. The red part is a $\delta \mathbb{Z}$-web.

![Figure 1. A $\delta \mathbb{Z}$-web.](image)
\( \delta \mathbb{Z}\)-web is \( \delta \)-dense in \( B^{\mathbb{R}^n}(0,1) \), so the Gromov-Hausdorff distance from \( M \) to \( B^{\mathbb{R}^n}(0,1) \) is less than \( \delta \). One can compute the volume of \( M \) is \( \sim \delta \times \delta^{2(n-1)} \times \delta^{-n} \sim \delta^{n-1} \). Then we can see that the volume of \( M \) does not satisfy the desired lower bound of volume stability.

From the stability theorem, we actually know that \( \delta \mathbb{Z}\)-web can not have the desired Ricci curvature lower bound. Moreover, from the proof we will see that, \( \delta \mathbb{Z}\)-web can not be an \( \epsilon \)-splitting.

So if we want to prove the volume stability theorem, we need to rule out this bad scenario. Of course, this scenario is not the only bad scenario, but it indicates an important feature that a bad scenario has. Roughly speaking, the \( \delta \mathbb{Z}\)-web consists of too many “big holes”, which make the volume too small. Hence, we need to show that not too many big hole can appear. This is a result of the following Lemma.

**Lemma 7.15.** There exists \( \epsilon_0 > 0 \) such that the following holds. Suppose \((M, g)\) is complete, \( p \in M, r > 0, \text{Ric}(g) \geq -(n-1)(\delta/r)^2 g \). Then if \( u : B^M(p, r) \to \mathbb{R}^n \) is an \( \epsilon \)-splitting for \( \epsilon \leq \epsilon_0 \), \( u(p) = 0 \), and \( y \in \mathbb{R}^n \) with \( |y| \geq r \). Then \( u(B^M(p, r)) \cap B^{\mathbb{R}^n}(y, |y|) \neq \emptyset \).

**Sketch of Proof.** Without loss of generality, assume \( r = 1 \). Since \( u \) is an \( \epsilon \)-splitting, we have \( \langle \nabla u, \nabla u \rangle \approx \delta u x \). Therefore, we can find a point \( x \in B(p, 1/10) \) and \( V \in T_xM, |V| = 1 \), such that \( d\mu(V) \) is almost parallel to \( y \). In fact, Lemma 7.5 implies that

\[
|\langle \nabla u, V \rangle - 2(u(\gamma_V(1/2)) - u(\gamma_V(0)))| \leq \Phi(\epsilon) \text{ in average.}
\]

So we can find \( V \in T_xM \) such that

\[
\sum_{l=1}^n \langle \nabla u, V \rangle y^l > (1 - \Phi(\epsilon))|y|.
\]

Then we can show that \( u(\gamma_V(1/2)) \in B(y, |y|) \).

Finally, we prove the volume stability Theorem by proving the following Lemma.

**Lemma 7.16.** For every \( \epsilon > 0 \), there exists \( \delta = \delta(\epsilon) > 0 \) such that the following holds. Suppose \((M, g)\) is complete, \( p \in M, r > 0, \text{Ric}(g) \geq -(n-1)(\delta/r)^2 g \), and suppose \( u : B^M(p, r) \to \mathbb{R}^n, u(p) = 0 \) is an \( \delta \)-splitting with

\[
||u(x)| - d(p, x)| < \delta r.
\]

Then

\[
\text{Vol}(B^M(p, r)) > (\omega_n - \epsilon) r^n.
\]

**Sketch of Proof.** Without loss of generality, assume \( r = 1 \). As we see in the \( \delta \mathbb{Z}\)-web example, we need to rule out too many big holes. We will apply Lemma 7.15 to a ball centered at \( x_i \), where \( x_i \) lies on the edge of a big hole. If the hole has size \( r_i \), then Lemma 7.15 implies that \( u|_{B^M(x_i, r_i)} \to \mathbb{R}^n \) can not be an \( \epsilon_0 \)-splitting.

Thus,

\[
\int_{B(x_i, r_i)} |\langle \nabla u_1, \nabla u_2 \rangle - \delta x_1 x_2|^2 + r_i^2 \sum_{l=1}^n \int_{B(x_i, r_i)} |\nabla^2 u_l|^2 > \epsilon_0^2 \text{Vol}(B(x_i, r_i)).
\]
If we have many such disjoint balls as $\delta \mathbb{Z}$ (we can pick these balls by Vitali covering technique), by summation we have

$$
(7.3) \quad \varepsilon_0^2 \sum_i \text{Vol}(B^M(x_i, r_i)) \leq \varepsilon_0^2 \sum_i \text{Vol}(B^M(p, 1)) \left[ \int_{B^M(p, 1)} |\nabla u_{i_1}, \nabla u_{i_2} - \delta_{i_1, i_2}| + \sum_{i=1}^n \int_{B^M(p, 1)} |\nabla^2 u_i|^2 \right] < \delta^2 \text{Vol}(B^M(p, 1)).
$$

Thus the total volume contained in these balls is about $\sim \delta^2/\varepsilon_0^2 \text{Vol}(B(p, 1))$. By Bishop-Gromov, the volume of the ball

$$
\text{Vol}(B(x_i, r_i)) \geq C \text{Vol}(B(p, 1)) r_i^2.
$$

So we conclude that

$$
\sum_i r_i^n \leq C \frac{\delta^2}{\varepsilon_0^2}.
$$

Note the left hand side is in the same scale as the the volume of the “holes”. Therefore, we conclude that if $\delta$ is sufficiently small we have

$$
|\mathbf{u}(B(p, 1))| \geq \omega_n - \varepsilon.
$$

Since $\mathbf{u}$ is an $\delta$-splitting, $|\det Du - 1| \approx 0$. Thus we have

$$
\text{Vol}(B(p, 1)) \approx |\mathbf{u}(B(p, 1))| \geq \omega_n - \varepsilon.
$$

$\square$

**References**


