CHAPTER 9

Ricci Curvature Comparison

In this chapter we shall prove some of the fundamental results for manifolds with lower Ricci curvature bounds. Two important techniques will be developed: Relative volume comparison and weak upper bounds for the Laplacian of distance functions. With these techniques we shall show numerous results on restrictions of fundamental groups of such spaces and also present a different proof of the estimate for the first Betti number by Bochner.

We have already seen how variational calculus can be used to obtain Myers' diameter bounds and also how the Bochner technique can be used. In the 50s Calabi discovered that one has weak upper bounds for the Laplacian of distance function given lower Ricci curvature bounds even at points where this function isn't smooth. However, it wasn't until around 1970, when Cheeger and Gromoll proved their splitting theorem, that this was fully appreciated. Around 1980, Gromov exposed the world to his view of how volume comparison can be used. The relative volume comparison theorem was actually first proved by Bishop in [**13**]. At the time, however, one only considered balls of radius less than the injectivity radius. Later, Gromov observed that the result holds for all balls and immediately put it to use in many situations. In particular, he showed how one could generalize the Betti number estimate from Bochner's theorem using only topological methods and volume comparison. Anderson then refined this to get information about fundamental groups. One's intuition about Ricci curvature has generally been borrowed from experience with sectional curvature. This has led to many naive conjectures that haven proven to be false through the construction of several interesting examples of manifolds with nonnegative Ricci curvature. On the other hand, much good work has also come out of this, as we shall see. The reason for treating Ricci curvature before the more advanced results on sectional curvature is that we want to break the link between the two. The techniques for dealing with these two subjects, while similar, are not the same.

1. Volume Comparison

1.1. The Fundamental Equations. Throughout this section, assume that we have a complete Riemannian manifold (M, g) of dimension n. Furthermore, we are given a point $p \in M$ and with that the distance function $r(x) = d(x,p)$. We know that this distance function is smooth on the image of the interior of the segment domain. In analogy with the fundamental equations for the metric:

$$\begin{aligned}
&(1) \quad L_{\partial_r} g = 2\mathrm{Hess} r,\\
&(2) \quad (\nabla_{\partial_r}\mathrm{Hess} r)(X,Y) + \mathrm{Hess}^2 r(X,Y) = -R(X,\partial_r,\partial_r,Y),
\end{aligned}$$

we also have a similar set of equations for the volume form.

PROPOSITION 39. *The volume form* dvol *and Laplacian* Δr *of* r *are related by:*

(tr1) $L_{\partial_r} d\mathrm{vol} = \Delta r d\mathrm{vol}$,

(tr2) $\partial_r \Delta r + \frac{(\Delta r)^2}{n-1} \leq \partial_r \Delta r + |\mathrm{Hess} r|^2 = -\mathrm{Ric}(\partial_r, \partial_r)$.

PROOF. The way to establish the first equation is by first selecting orthonormal 1-forms θ^i. The volume form is then given by

$$d\mathrm{vol} = \theta^1 \wedge \cdots \wedge \theta^n.$$

As with the metric g, we also have that $d$vol is parallel. Next observe that

$$\begin{aligned}
\left(L_{\partial_r} \theta^i\right)(X) &= \partial_r \left(\theta^i (X)\right) - \theta^i (L_{\partial_r} X) \\
&= \partial_r \left(\theta^i (X)\right) - \theta^i (\nabla_{\partial_r} X) + \theta^i (\nabla_X \partial_r) \\
&= \left(\nabla_{\partial_r} \theta^i\right)(X) + \theta^i (\nabla_X \partial_r).
\end{aligned}$$

This shows that

$$\begin{aligned}
L_{\partial_r} d\mathrm{vol} &= L_{\partial_r} \left(\theta^1 \wedge \cdots \wedge \theta^n\right) \\
&= \sum \theta^1 \wedge \cdots \wedge L_{\partial_r} \theta^i \wedge \cdots \wedge \theta^n \\
&= \sum \theta^1 \wedge \cdots \wedge \nabla_{\partial_r} \theta^i \wedge \cdots \wedge \theta^n \\
&\quad + \sum \theta^1 \wedge \cdots \wedge \theta^i \circ \nabla . \partial_r \wedge \cdots \wedge \theta^n \\
&= \nabla_{\partial_r} \left(\theta^1 \wedge \cdots \wedge \theta^n\right) + \mathrm{tr}(\nabla . \partial_r)\, \theta^1 \wedge \cdots \wedge \theta^n \\
&= \nabla_{\partial_r} d\mathrm{vol} + \mathrm{tr}(\nabla . \partial_r)\, d\mathrm{vol} \\
&= \Delta r d\mathrm{vol}.
\end{aligned}$$

To establish the second equation we take traces in (2). Thus we select an orthonormal frame E_i, set $X = Y = E_i$ and sum over i. We can in addition assume that $\nabla_{\partial_r} E_i = 0$. We already know that

$$\sum_{i=1}^{n} R(E_i, \partial_r, \partial_r, E_i) = \mathrm{Ric}(\partial_r, \partial_r).$$

On the left hand side we get

$$\begin{aligned}
\sum_{i=1}^{n} (\nabla_{\partial_r} \mathrm{Hess} r)(E_i, E_i) &= \sum_{i=1}^{n} \partial_r \mathrm{Hess} r (E_i, E_i) \\
&= \partial_r \Delta r
\end{aligned}$$

and

$$\begin{aligned}
\sum_{i=1}^{n} \mathrm{Hess}^2 r (E_i, E_i) &= \sum_{i=1}^{n} g(\nabla_{E_i} \partial_r, \nabla_{E_i} \partial_r) \\
&= \sum_{i,j=1}^{n} g(\nabla_{E_i} \partial_r, g(\nabla_{E_i} \partial_r, E_j) E_j) \\
&= \sum_{i,j=1}^{n} g(\nabla_{E_i} \partial_r, E_j)\, g(\nabla_{E_i} \partial_r, E_j) \\
&= |\mathrm{Hess} r|^2
\end{aligned}$$

Finally we need to show that

$$\frac{(\Delta r)^2}{n-1} \leq |\mathrm{Hess}r|^2 .$$

To this end we also assume that $E_1 = \partial_r$. Then

$$\begin{aligned} |\mathrm{Hess}r|^2 &= \sum_{i,j=1}^{n} \left(g\left(\nabla_{E_i}\partial_r, E_j\right)\right)^2 \\ &= \sum_{i,j=2}^{n} \left(g\left(\nabla_{E_i}\partial_r, E_j\right)\right)^2 \\ &\leq \frac{1}{n-1}\left(\sum_{i=2}^{n} g\left(\nabla_{E_i}\partial_r, E_i\right)\right)^2 \\ &= \frac{1}{n-1}(\Delta r)^2 . \end{aligned}$$

The inequality

$$|A|^2 \leq \frac{1}{k}|\mathrm{tr}\,(A)|^2$$

for a $k \times k$ matrix A is a direct consequence of the Cauchy-Schwarz inequality

$$\begin{aligned} |(A, I_k)|^2 &\leq |A|^2 |I_k|^2 \\ &= |A|^2 k, \end{aligned}$$

where I_k is the identity $k \times k$ matrix. □

If we use the polar coordinate decomposition $g = dr^2 + g_r$ and let $d\mathrm{vol}_{n-1}$ be the standard volume form on $S^{n-1}(1)$, then we have that

$$d\mathrm{vol} = \lambda(r,\theta)\, dr \wedge d\mathrm{vol}_{n-1},$$

where θ indicates a coordinate on S^{n-1}. If we apply (tr1) to this version of the volume form we get

$$\begin{aligned} L_{\partial_r} d\mathrm{vol} &= L_{\partial_r}\left(\lambda(r,\theta)\, dr \wedge d\mathrm{vol}_{n-1}\right) \\ &= \partial_r(\lambda)\, dr \wedge d\mathrm{vol}_{n-1} \end{aligned}$$

as both $L_{\partial_r} dr = 0$ and $L_{\partial_r} d\mathrm{vol}_{n-1} = 0$. We can therefore simplify (tr1) to

$$\partial_r \lambda = \lambda \Delta r.$$

In constant curvature k we know that

$$g_k = dr^2 + \mathrm{sn}_k^2(r)\, ds_{n-1}^2,$$

thus the volume form is

$$\begin{aligned} d\mathrm{vol}_k &= \lambda_k(r)\, dr \wedge d\mathrm{vol}_{n-1} \\ &= \mathrm{sn}_k^{n-1}(r)\, dr \wedge d\mathrm{vol}_{n-1}, \end{aligned}$$

this conforms with the fact that

$$\begin{aligned} \Delta r &= (n-1)\frac{\mathrm{sn}_k'(r)}{\mathrm{sn}_k(r)}, \\ \partial_r\left(\mathrm{sn}_k^{n-1}(r)\right) &= (n-1)\frac{\mathrm{sn}_k'(r)}{\mathrm{sn}_k(r)}\mathrm{sn}_k^{n-1}(r) . \end{aligned}$$

1.2. Volume Estimation. With the above information we can prove the estimates that are analogous to our basic comparison estimates for the metric and Hessian of r assuming lower sectional curvature bounds (see chapter 6).

LEMMA 34. (Ricci Comparison Result) *Suppose that* (M, g) *has* $\mathrm{Ric} \geq (n-1)\cdot k$ *for some* $k \in \mathbb{R}$. *Then*

$$\begin{aligned} \Delta r &\leq (n-1)\frac{\mathrm{sn}_k'(r)}{\mathrm{sn}_k(r)}, \\ d\mathrm{vol} &\leq d\mathrm{vol}_k, \end{aligned}$$

where dvol_k *is the volume form in constant sectional curvature* k.

PROOF. Notice that the right-hand sides of the inequalities correspond exactly to what one would get in constant curvature k.

For the first inequality, we use that

$$\partial_r \Delta r + \frac{(\Delta r)^2}{n-1} \leq -(n-1)\cdot k$$

dividing by $n-1$ and using λ_k this gives

$$\partial_r\left(\frac{\Delta r}{n-1}\right) + \left(\frac{\Delta r}{n-1}\right)^2 \leq -k = \partial_r(\lambda_k) + (\lambda_k)^2$$

Separation of variables then yields:

$$\frac{\partial_r \frac{\Delta r}{n-1}}{k + \left(\frac{\Delta r}{n-1}\right)^2} \leq \frac{\partial_r \lambda_k}{k + (\lambda_k)^2}.$$

Thus

$$F(\lambda(r)) \leq F(\lambda_k(r)),$$

where F is the antiderivative of $\frac{1}{\lambda^2+k}$ satisfying $\lim_{\lambda\to\infty} F(\lambda) = 0$. Since F has positive derivative we can conclude that $\lambda(r) \leq \lambda_k(r)$.

For the second inequality we now know that

$$\partial_r \lambda \leq (n-1)\frac{\mathrm{sn}_k'(r)}{\mathrm{sn}_k(r)}\lambda$$

while

$$\partial_r \lambda_k = (n-1)\frac{\mathrm{sn}_k'(r)}{\mathrm{sn}_k(r)}\lambda_k.$$

In addition the metrics g and g_k agree at p. Thus also the volume forms agree at p. This means that

$$\begin{aligned} \lim_{r\to 0}(\lambda - \lambda_k) &= 0, \\ \partial_r(\lambda - \lambda_k) &\leq (n-1)\frac{\mathrm{sn}_k'(r)}{\mathrm{sn}_k(r)}(\lambda - \lambda_k). \end{aligned}$$

Whence the volume form inequality follows. □

Our first volume comparison gives the obvious upper volume bound coming from our upper bound on the volume density.

LEMMA 35. *If* (M, g) *has* $\mathrm{Ric} \geq (n-1) \cdot k$, *then*

$$\mathrm{vol} B(p, r) \leq v(n, k, r),$$

where $v(n, k, r)$ *denotes the volume of a ball of radius* r *in the constant-curvature space form* S_k^n.

PROOF. Above, we showed that in polar coordinates around p we have

$$d\mathrm{vol} \leq d\mathrm{vol}_k.$$

Thus

$$\begin{aligned} \mathrm{vol} B(p, r) &= \int_{\mathrm{seg}_p \cap B(0,r)} d\mathrm{vol} \\ &\leq \int_{\mathrm{seg}_p \cap B(0,r)} d\mathrm{vol}_k \\ &\leq \int_{B(0,r)} d\mathrm{vol}_k \\ &= v(n, k, r). \end{aligned}$$

□

With a little more technical work, the above absolute volume comparison result can be improved in a rather interesting direction. The result one obtains is referred to as the relative volume comparison estimate. It will prove invaluable in many situations throughout the rest of the text.

LEMMA 36. (Relative Volume Comparison, Bishop-Cheeger-Gromov, 1964-1980) *Suppose* (M, g) *is a complete Riemannian manifold with* $\mathrm{Ric} \geq (n-1) \cdot k$. *Then*

$$r \to \frac{\mathrm{vol} B(p, r)}{v(n, k, r)}$$

is a nonincreasing function whose limit is 1 as $r \to 0$.

PROOF. We will use exponential polar coordinates. The volume form $\lambda(r, \theta) dr \wedge d\theta$ for (M, g) is initially defined only on some star-shaped subset of

$$T_p M = \mathbb{R}^n = (0, \infty) \times S^{n-1},$$

but we can just set $\lambda = 0$ outside this set. The comparison density λ_k is defined on all of $\mathbb{R}^n$ for $k \leq 0$ and on $B\left(0, \pi/\sqrt{k}\right)$ for $k > 0$. We can likewise extend $\lambda_k = 0$ outside $B\left(0, \pi/\sqrt{k}\right)$. Myers' theorem says that $\lambda = 0$ on $\mathbb{R}^n - B\left(0, \pi/\sqrt{k}\right)$ in this case. So we might as well just consider $r < \pi/\sqrt{k}$ when $k > 0$.

The ratio of the volumes is

$$\frac{\mathrm{vol} B(p, R)}{v(n, k, R)} = \frac{\int_0^R \int_{S^{n-1}} \lambda dr \wedge d\theta}{\int_0^R \int_{S^{n-1}} \lambda_k dr \wedge d\theta},$$

and we know that

$$0 \leq \lambda(r, \theta) \leq \lambda_k(r, \theta) = \mathrm{sn}_k^{n-1}(r)$$

everywhere.

Differentiation of this quotient with respect to R yields

$$\begin{aligned}
& \frac{d}{dR}\left(\frac{\mathrm{vol}B(p,R)}{v(n,k,R)}\right) \\
= & \frac{\left(\int_{S^{n-1}}\lambda\left(R,\theta\right)d\theta\right)\left(\int_0^R\int_{S^{n-1}}\lambda_k\left(r,\theta\right)dr\wedge d\theta\right)}{\left(v(n,k,R)\right)^2} \\
& -\frac{\left(\int_{S^{n-1}}\lambda_k\left(R,\theta\right)d\theta\right)\left(\int_0^R\int_{S^{n-1}}\lambda\left(r,\theta\right)dr\wedge d\theta\right)}{\left(v(n,k,R)\right)^2} \\
= & \left(v(n,k,R)\right)^{-2}\cdot\int_0^R\left[\left(\int_{S^{n-1}}\lambda\left(R,\theta\right)d\theta\right)\cdot\left(\int_{S^{n-1}}\lambda_k\left(r,\theta_{n-1}\right)d\theta\right)\right. \\
& \left.-\left(\int_{S^{n-1}}\lambda_k\left(R,\theta\right)d\theta\right)\left(\int_{S^{n-1}}\lambda\left(r,\theta\right)d\theta\right)\right]dr.
\end{aligned}$$

So to see that

$$R\rightarrow\frac{\mathrm{vol}B(p,R)}{v(n,k,R)}$$

is nonincreasing, it suffices to check that

$$\frac{\int_{S^{n-1}}\lambda\left(r,\theta\right)d\theta}{\int_{S^{n-1}}\lambda_k\left(r,\theta\right)d\theta}=\frac{1}{\omega_{n-1}}\int_{S^{n-1}}\frac{\lambda\left(r,\theta\right)}{\lambda_k\left(r,\theta\right)}d\theta$$

is nonincreasing. This follows from

$$\begin{aligned}
\partial_r\left(\frac{\lambda\left(r,\theta\right)}{\lambda_k\left(r,\theta\right)}\right) &= \frac{\lambda_k\partial_r\lambda-\lambda\partial_r\lambda_k}{\lambda_k^2} \\
&\leq \frac{\lambda_k\left(n-1\right)\frac{\mathrm{sn}_k'(r)}{\mathrm{sn}_k(r)}\lambda-\lambda\left(n-1\right)\frac{\mathrm{sn}_k'(r)}{\mathrm{sn}_k(r)}\lambda_k}{\lambda_k^2} \\
&= 0.
\end{aligned}$$

□

1.3. Maximal Diameter Rigidity. Given Myers' diameter estimate, it is natural to ask what happens if the diameter attains it maximal value. The next result shows that only the sphere has this property.

THEOREM 62. (S. Y. Cheng, 1975) *If* (M,g) *is a complete Riemannian manifold with* $\mathrm{Ric}\geq(n-1)k>0$ *and* $\mathrm{diam}=\pi/\sqrt{k}$, *then* (M,g) *is isometric to* S_k^n.

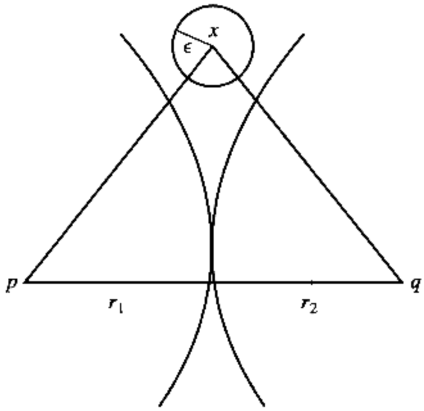

Figure 9.1

PROOF. Fix $p, q \in M$ such that $d(p,q) = \pi/\sqrt{k}$. Define $r(x) = d(x,p)$, $\tilde{r}(x) = d(x,q)$. We will show that

(1) $r + \tilde{r} = d(p,x) + d(x,q) = d(p,q) = \pi/\sqrt{k},\ x \in M$.
(2) $r, \tilde{r}$ are smooth on $M - \{p,q\}$.
(3) $\mathrm{Hess}r = (\mathrm{sn}_k'/\mathrm{sn}_k)\, ds_{n-1}^2$ on $M - \{p,q\}$.
(4) $g = dr^2 + \mathrm{sn}_k^2 ds_{n-1}^2$.

We know that (3) implies (4) and that (4) implies M must be S_k^n.

Proof of (1): The triangle inequality shows that

$$d(p,x) + d(x,q) \geq \pi/\sqrt{k},$$

so if (1) does not hold, we can find $\varepsilon > 0$ such that (see Figure 9.1)

$$d(p,x) + d(x,q) = 2 \cdot \varepsilon + \frac{\pi}{\sqrt{k}} = 2 \cdot \varepsilon + d(p,q).$$

Then the metric balls $B(p, r_1),\ B(q, r_2)$, and $B(x, \varepsilon)$ are pairwise disjoint, when $r_1 \leq d(p,x),\ r_2 \leq d(q,x)$ and $r_1 + r_2 = \pi/\sqrt{k}$. Thus,

$$\begin{aligned} 1 &= \frac{\mathrm{vol}M}{\mathrm{vol}M} \geq \frac{\mathrm{vol}B(x,\varepsilon) + \mathrm{vol}B(p,r_1) + \mathrm{vol}B(q,r_2)}{\mathrm{vol}M} \\ &\geq \frac{v(n,k,\varepsilon)}{v\left(n,k,\frac{\pi}{\sqrt{k}}\right)} + \frac{v(n,k,r_1)}{v\left(n,k,\frac{\pi}{\sqrt{k}}\right)} + \frac{v(n,k,r_2)}{v\left(n,k,\frac{\pi}{\sqrt{k}}\right)} \\ &= \frac{v(n,k,\varepsilon)}{v\left(n,k,\frac{\pi}{\sqrt{k}}\right)} + 1, \end{aligned}$$

which is a contradiction.

Proof of (2): If $x \in M - \{q,p\}$, then x can be joined to both p and q by segments σ_1, σ_2. The previous statement says that if we put these two segments together, then we get a segment from p to q through x. Such a segment must be smooth, and thus σ_1 and σ_2 are both subsegments of a larger segment. This implies from our characterization of when distance functions are smooth that both r and $\tilde{r}$ are smooth at $x \in M - \{p,q\}$.

Proof of (3): We have $r(x)+\tilde{r}(x)=\pi/\sqrt{k}$, thus $\Delta r=-\Delta\tilde{r}$. On the other hand,

$$\begin{aligned}
(n-1)\frac{\mathrm{sn}_k'(r(x))}{\mathrm{sn}_k(r(x))} &\geq \Delta r(x)\\
&= -\Delta\tilde{r}(x)\\
&\geq -(n-1)\frac{\mathrm{sn}_k'(\tilde{r}(x))}{\mathrm{sn}_k(\tilde{r}(x))}\\
&= -(n-1)\frac{\mathrm{sn}_k'\left(\frac{\pi}{\sqrt{k}}-r(x)\right)}{\mathrm{sn}_k\left(\frac{\pi}{\sqrt{k}}-r(x)\right)}\\
&= (n-1)\frac{\mathrm{sn}_k'(r(x))}{\mathrm{sn}_k(r(x))}.
\end{aligned}$$

This implies,

$$\Delta r=(n-1)\frac{\mathrm{sn}_k'}{\mathrm{sn}_k}$$

and

$$\begin{aligned}
-(n-1)k &= \partial_r(\Delta r)+\frac{(\Delta r)^2}{n-1}\\
&\leq \partial_r(\Delta r)+|\mathrm{Hess}r|^2\\
&\leq -\mathrm{Ric}(\partial_r,\partial_r)\\
&\leq -(n-1)k.
\end{aligned}$$

Hence, all inequalities are equalities, and in particular

$$(\Delta r)^2=(n-1)|\mathrm{Hess}r|^2.$$

Recall that this gives us equality in the Cauchy-Schwarz inequality $|A|^2\leq k\left(\mathrm{tr}A\right)^2$. Thus $A=\frac{\mathrm{tr}A}{k}I_k$. In our case we have restricted $\mathrm{Hess}r$ to the $(n-1)$ dimensional space orthogonal to ∂_r so on this space we obtain:

$$\begin{aligned}
\mathrm{Hess}r &= \frac{\Delta r}{n-1}g_r\\
&= \frac{\mathrm{sn}_k'}{\mathrm{sn}_k}g_r.
\end{aligned}$$

□

We have now proved that any complete manifold with $\mathrm{Ric}\geq(n-1)\cdot k>0$ has diameter $\leq\pi/\sqrt{k}$, where equality holds only when the space is S_k^n. A natural perturbation question is therefore: Do manifolds with $\mathrm{Ric}\geq(n-1)\cdot k>0$ and $\mathrm{diam}\approx\pi/\sqrt{k}$, have to be homeomorphic or diffeomorphic to a sphere?

For $n=2,3$ this is true, when $n\geq4$, however, there are counterexamples. The case $n=2$ will be settled later, while $n=3$ goes beyond the scope of this book (see [**85**]). The examples for $n\geq4$ are divided into two cases: $n=4$ and $n\geq5$.

Example 46. *(Anderson, 1990) For $n=4$ consider metrics on $I\times S^3$ of the form*

$$dr^2+\varphi^2\sigma_1^2+\psi^2(\sigma_2^2+\sigma_3^2).$$

If we define

$$\varphi(r)=\begin{cases}\frac{\sin(ar)}{a} & r\le r_0,\\ c_1\sin(r+\delta) & r\ge r_0,\end{cases}$$

$$\psi(r)=\begin{cases}br^2+c & r\le r_0,\\ c_2\sin(r+\delta) & r\ge r_0,\end{cases}$$

and then reflect these function in $r=\pi/2-\delta$, we get a metric on $\mathbb{C}P^2\sharp\bar{\mathbb{C}}P^2$. For any small $r_0>0$ we can now adjust the parameters so that φ and ψ become C^1 and generate a metric with $\mathrm{Ric}\ge(n-1)$. *For smaller and smaller choices of r_0 we see that $\delta\to 0$, so the interval $I\to[0,\pi]$ as $r_0\to 0$. This means that the diameters converge to π.*

EXAMPLE 47. *(Otsu, 1991) For $n\ge 5$ we only need to consider standard doubly warped products:*

$$dr^2+\varphi^2\cdot ds_2^2+\psi^2 ds_{n-3}^2$$

on $I\times S^2\times S^{n-3}$. Similar choices for φ and ψ will yield metrics on $S^2\times S^{n-2}$ with $\mathrm{Ric}\ge n-1$ *and diameter $\to\pi$.*

In both of the above examples we actually only constructed C^1 functions φ,ψ and therefore only C^1 metrics. The functions are, however, concave and can easily be smoothed near the break points so as to stay concave. This will not change the values or first derivatives much and only increase the second derivative in absolute value. Thus the lower curvature bound still holds.

2. Fundamental Groups and Ricci Curvature

We shall now attempt to generalize the estimate on the first Betti number we obtained using the Bochner technique to the situation where one has more general Ricci curvature bounds. This requires some knowledge about how fundamental groups are tied in with the geometry.

2.1. The First Betti Number. Suppose M is a compact Riemannian manifold of dimension n and $\tilde{M}$ its universal covering space. The fundamental group $\pi_1(M)$ acts by isometries on $\tilde{M}$. Recall from algebraic topology that

$$H_1(M,\mathbb{Z})=\pi_1(M)/[\pi_1(M),\pi_1(M)],$$

where $[\pi_1(M),\pi_1(M)]$ is the commutator subgroup. Thus, $H_1(M,\mathbb{Z})$ acts by deck transformations on the covering space

$$\tilde{M}/[\pi_1(M),\pi_1(M)]$$

with quotient M. Since $H_1(M,\mathbb{Z})$ is a finitely generated Abelian group, we know that the set of torsion elements T is a finite normal subgroup. We can then consider $\Gamma=H_1(M,\mathbb{Z})/T$ as acting by deck transformations on

$$\bar{M}=\tilde{M}/[\pi_1(M),\pi_1(M)]/T.$$

Thus, we have a covering $\pi:\bar{M}\to M$ with a torsion free and Abelian Galois group of deck transformations. The rank of the torsion-free group Γ is clearly equal to

$$b_1(M)=\dim H_1(M,\mathbb{R}).$$

Next recall that any finite-index subgroup of Γ has the same rank as Γ. So if we can find a finite-index subgroup that is generated by elements that can be geometrically controlled, then we might be able to bound b_1. To this end we have a very interesting result.

LEMMA 37. (M. Gromov, 1980) *For fixed* $x \in \bar{M}$ *there exists a finite-index subgroup* $\Gamma' \subset \Gamma$ *that is generated by elements* $\gamma_1, \ldots, \gamma_m$ *such that*

$$d\left(x, \gamma_i\left(x\right)\right) \leq 2 \cdot \operatorname{diam}\left(M\right).$$

Furthermore, for all $\gamma \in \Gamma' - \{1\}$ *we have*

$$d\left(x, \gamma\left(x\right)\right) > \operatorname{diam}\left(M\right).$$

PROOF. First we find a finite-index subgroup that can be generated by elements satisfying the first condition. Then we modify this group so that it also satisfies the second condition.

For each $\varepsilon > 0$ let Γ_ε be the group generated by

$$\left\{\gamma \in \Gamma : d\left(x, \gamma\left(x\right)\right) < 2\operatorname{diam}\left(M\right) + \varepsilon\right\},$$

and let $\pi_\varepsilon : \bar{M} \to \bar{M}/\Gamma_\varepsilon$ denote the covering projection. We claim that for each $z \in \bar{M}$ we have

$$d\left(\pi_\varepsilon\left(z\right), \pi_\varepsilon\left(x\right)\right) < \operatorname{diam}\left(M\right) + \varepsilon.$$

Otherwise, we could find $z \in \bar{M}$ such that

$$d\left(x, z\right) = d\left(\pi_\varepsilon\left(z\right), \pi_\varepsilon\left(x\right)\right) = \operatorname{diam}\left(M\right) + \varepsilon.$$

Now, we can find $\gamma \in \Gamma$ such that $d\left(\gamma\left(x\right), z\right) \leq \operatorname{diam}\left(M\right)$, but then we would have

$$\begin{aligned} d\left(\pi_\varepsilon\left(\gamma\left(x\right)\right), \pi_\varepsilon\left(x\right)\right) &\geq d\left(\pi_\varepsilon\left(z\right), \pi_\varepsilon\left(x\right)\right) - d\left(\pi_\varepsilon\left(z\right), \pi_\varepsilon\left(\gamma\left(x\right)\right)\right) \geq \varepsilon, \\ d\left(x, \gamma\left(x\right)\right) &\leq d\left(x, z\right) + d\left(z, \gamma\left(x\right)\right) \leq 2\operatorname{diam}\left(M\right) + \varepsilon. \end{aligned}$$

Here we have a contradiction, as the first line says that $\gamma \notin \Gamma_\varepsilon$, while the second line says $\gamma \in \Gamma_\varepsilon$.

Note that compactness of $\bar{M}/\Gamma_\varepsilon$ shows that $\Gamma_\varepsilon \subset \Gamma$ has finite index.

Now observe that there are at most finitely many elements in the set

$$\left\{\gamma \in \Gamma : d\left(x, \gamma\left(x\right)\right) < 3\operatorname{diam}\left(M\right)\right\},$$

as Γ acts discretely on $\bar{M}$. Hence, there must be a sufficiently small $\varepsilon > 0$ such that

$$\left\{\gamma \in \Gamma : d\left(x, \gamma\left(x\right)\right) < 2\operatorname{diam}\left(M\right) + \varepsilon\right\} = \left\{\gamma \in \Gamma : d\left(x, \gamma\left(x\right)\right) \leq 2\operatorname{diam}\left(M\right)\right\}.$$

Then we have a finite-index subgroup Γ_ε of Γ generated by

$$\left\{\gamma \in \Gamma : d\left(x, \gamma\left(x\right)\right) \leq 2\operatorname{diam}\left(M\right)\right\} = \left\{\gamma_1, \ldots, \gamma_m\right\}.$$

We shall now modify these generators until we get the desired group Γ'.

First, observe that as the rank of Γ_ε is b_1, we can assume that $\left\{\gamma_1, \ldots, \gamma_{b_1}\right\}$ are linearly independent and generate a subgroup $\Gamma'' \subset \Gamma_\varepsilon$ of finite index. Next, we recall that only finitely many elements γ in Γ'' lie in

$$\left\{\gamma \in \Gamma : d\left(x, \gamma\left(x\right)\right) \leq 2\operatorname{diam}\left(M\right)\right\}.$$

We can therefore choose

$$\left\{\tilde{\gamma}_1, \ldots, \tilde{\gamma}_{b_1}\right\} \subset \left\{\gamma \in \Gamma : d\left(x, \gamma\left(x\right)\right) \leq 2\operatorname{diam}\left(M\right)\right\}$$

with the following properties (we use additive notation here, as it is easier to read):

(1) $\operatorname{span}\left\{\tilde{\gamma}_1, \ldots, \tilde{\gamma}_k\right\} \subset \operatorname{span}\left\{\gamma_1, \ldots, \gamma_k\right\}$ has finite index for all $k = 1, \ldots, b_1$.

(2) $\tilde{\gamma}_k = l_{1k} \cdot \gamma_1 + \cdots + l_{kk} \cdot \gamma_k$ is chosen such that l_{kk} is maximal in absolute value among all elements in

$$\Gamma'' \cap \{\gamma \in \Gamma : d(x, \gamma(x)) \leq 2\mathrm{diam}(M)\}.$$

The group Γ' generated by $\{\tilde{\gamma}_1, \ldots, \tilde{\gamma}_{b_1}\}$ clearly has finite index in Γ'' and hence also in Γ. The generators lie in

$$\{\gamma \in \Gamma : d(x, \gamma(x)) \leq 2\mathrm{diam}(M)\},$$

as demanded by the first property. It only remains to show that the second property is also satisfied. The see this, let

$$\gamma = m_1 \cdot \tilde{\gamma}_1 + \cdots + m_k \cdot \tilde{\gamma}_k$$

be chosen such that $m_k \neq 0$. If $d(x, \gamma(x)) \leq \mathrm{diam}(M)$, then we also have that

$$\begin{aligned} d\left(x, \gamma^2(x)\right) &\leq d(x, \gamma(x)) + d\left(\gamma(x), \gamma^2(x)\right) \\ &= 2d(x, \gamma(x)) \\ &\leq 2\mathrm{diam}(M). \end{aligned}$$

Thus,

$$\gamma^2 \in \Gamma'' \cap \{\gamma \in \Gamma : d(x, \gamma(x)) \leq 2\mathrm{diam}(M)\},$$

and also,

$$\begin{aligned} \gamma^2 &= 2m_1 \cdot \tilde{\gamma}_1 + \cdots + 2m_k \cdot \tilde{\gamma}_k \\ &= \sum_{i=1}^{k-1} n_i \cdot \gamma_i + 2m_k \cdot l_{kk} \cdot \gamma_k. \end{aligned}$$

But this violates the maximality of l_{kk}, as we assumed $m_k \neq 0$. □

With this lemma we can now give Gromov's proof of

THEOREM 63. (S. Gallot and M. Gromov, 1980) *If M is a Riemannian manifold of dimension n such that* Ric $\geq (n-1)k$ *and* $\mathrm{diam}(M) \leq D$, *then there is a function $C\left(n, k \cdot D^2\right)$ such that*

$$b_1(M) \leq C\left(n, k \cdot D^2\right).$$

Moreover, $\lim_{\varepsilon \to 0} C(n, \varepsilon) = n$.In particular, there is $\varepsilon(n) > 0$ such that if $k \cdot D^2 \geq -\varepsilon(n)$, then $b_1(M) \leq n$.

PROOF. First observe that for $k > 0$ there is nothing to prove, as we know that $b_1 = 0$ from Myers' theorem.

Suppose we have chosen a covering $\bar{M}$ of M with torsion-free Abelian Galois group of deck transformations $\Gamma = \left\langle \gamma_1, \ldots, \gamma_{b_1} \right\rangle$ such that for some $x \in \bar{M}$ we have

$$\begin{aligned} d(x, \gamma_i(x)) &\leq 2\mathrm{diam}(M), \\ d(x, \gamma(x)) &> \mathrm{diam}(M), \gamma \neq 1. \end{aligned}$$

Then we clearly have that all of the balls $B\left(\gamma(x), \frac{\mathrm{diam}(M)}{2}\right)$ are disjoint. Now set

$$I_r = \left\{\gamma \in \Gamma : \gamma = l_1 \cdot \gamma_1 + \cdots + l_{b_1} \cdot \gamma_{b_1}, |l_1| + \cdots + |l_{b_1}| \leq r\right\}.$$

Note that for $\gamma \in I_r$ we have

$$B\left(\gamma(x), \frac{\mathrm{diam}(M)}{2}\right) \subset B\left(x, r \cdot 2\mathrm{diam}(M) + \frac{\mathrm{diam}(M)}{2}\right).$$

All of these balls are disjoint and have the same volume, as γ acts isometrically. We can therefore use the relative volume comparison theorem to conclude that the cardinality of I_r is bounded from above by

$$\frac{\mathrm{vol}B\left(x, r\cdot 2\mathrm{diam}\left(M\right)+\frac{\mathrm{diam}(M)}{2}\right)}{\mathrm{vol}B\left(x,\frac{\mathrm{diam}(M)}{2}\right)}\leq\frac{v\left(n,k,r\cdot 2\mathrm{diam}\left(M\right)+\frac{\mathrm{diam}(M)}{2}\right)}{v\left(n,k,\frac{\mathrm{diam}(M)}{2}\right)}.$$

This shows that

$$\begin{aligned} b_1 &\leq |I_1| \\ &\leq \frac{v\left(n,k,2\mathrm{diam}\left(M\right)+\frac{\mathrm{diam}(M)}{2}\right)}{v\left(n,k,\frac{\mathrm{diam}(M)}{2}\right)}, \end{aligned}$$

which gives us a general bound for b_1. To get a more refined bound we have to use I_r for larger r. If r is an integer, then

$$|I_r| = (2r+1)^{b_1}.$$

The upper bound for $|I_r|$ can be reduced to

$$\begin{aligned} \frac{v\left(n,k,r\cdot 2\mathrm{diam}\left(M\right)+\frac{\mathrm{diam}(M)}{2}\right)}{v\left(n,k,\frac{\mathrm{diam}(M)}{2}\right)} &\leq \frac{v\left(n,k,\left(r\cdot 2+\frac{1}{2}\right)D\right)}{v\left(n,k,\frac{D}{2}\right)} \\ &= \frac{\int_0^{\left(r\cdot 2+\frac{1}{2}\right)D}\left(\frac{\sinh\left(\sqrt{-k}t\right)}{\sqrt{-k}}\right)^{n-1}dt}{\int_0^{\frac{1}{2}D}\left(\frac{\sinh\left(\sqrt{-k}t\right)}{\sqrt{-k}}\right)^{n-1}dt} \\ &= \frac{\int_0^{\left(r\cdot 2+\frac{1}{2}\right)D\sqrt{-k}}\sinh^{n-1}(t)\,dt}{\int_0^{\frac{1}{2}D\sqrt{-k}}\sinh^{n-1}(t)\,dt} \\ &= 2^n\left(r\cdot 2+\frac{1}{2}\right)^n+\cdots\leq 5^n\cdot r^n, \end{aligned}$$

where in the last step we assume that $D\sqrt{-k}$ is very small relative to r. If $b_1\geq n+1$, this cannot be larger than $|I_r|=(2r+1)^{b_1}$ when $r=5^n$. Thus select $r=5^n$ and the assume $D\sqrt{-k}$ is small enough that

$$\frac{\int_0^{\left(r\cdot 2+\frac{1}{2}\right)D\sqrt{-k}}\sinh^{n-1}(t)\,dt}{\int_0^{\frac{1}{2}D\sqrt{-k}}\sinh^{n-1}(t)\,dt}\leq 5^n\cdot r^n$$

in order to force $b_1\leq n$. □

Gallot's proof of the above theorem uses techniques that are sophisticated generalizations of the Bochner technique.

2.2. Finiteness of Fundamental Groups. One can get even more information from these volume comparison techniques. Instead of considering just the first homology group, we can actually get some information about fundamental groups as well.

For our next result we need a different kind of understanding of how fundamental groups can be represented.

Lemma 38. (M. Gromov, 1980) *For a Riemannian manifold M and $\tilde{x} \in \tilde{M}$, we can always find generators $\{\gamma_1, \dots \gamma_m\}$ for the fundamental group $\Gamma = \pi_1(M)$ such that $d(x, \gamma_i(x)) \leq 2\text{diam}(M)$ and such that all relations for Γ in these generators are of the form $\gamma_i \cdot \gamma_j \cdot \gamma_k^{-1} = 1$.*

Proof. For any $\varepsilon \in (0, \text{inj}(M))$ choose a triangulation of M such that adjacent vertices in this triangulation are joined by a curve of length less that ε. Let $\{x_1, \dots, x_k\}$ denote the set of vertices and $\{e_{ij}\}$ the edges joining adjacent vertices (thus, e_{ij} is not necessarily defined for all i, j). If x is the projection of $\tilde{x} \in \tilde{M}$, then join x and x_i by a segment σ_i for all $i = 1, \dots, k$ and construct the loops

$$\sigma_{ij} = \sigma_i e_{ij} \sigma_j^{-1}$$

for adjacent vertices. Now, any loop in M based at x is homotopic to a loop in the 1-skeleton of the triangulation, i.e., a loop that is constructed out of juxtaposing edges e_{ij}. Since $e_{ij}e_{jk} = e_{ij}\sigma_j^{-1}\sigma_j e_{jk}$ such loops are the product of loops of the form σ_{ij}. Therefore Γ is generated by σ_{ij}.

Now observe that if three vertices x_i, x_j, x_k are adjacent to each other, then they span a 2-simplex $\triangle_{ijk}$. Thus, we have that the loop $\sigma_{ij}\sigma_{jk}\sigma_{ki} = \sigma_{ij}\sigma_{jk}\sigma_{ik}^{-1}$ is homotopically trivial. We claim that these are the only relations needed to describe Γ. To see this, let σ be any loop in the 1-skeleton that is homotopically trivial. Now use that σ in fact contracts in the 2-skeleton. Thus, a homotopy corresponds to a collection of 2-simplices $\triangle_{ijk}$. In this way we can represent the relation $\sigma = 1$ as a product of elementary relations of the form $\sigma_{ij}\sigma_{jk}\sigma_{ik}^{-1} = 1$.

Finally, use discreteness of Γ to get rid of ε as in the above case. □

A simple example might be instructive here.

Example 48. *Consider $M_k = S^3/\mathbb{Z}_k$; the constant-curvature 3-sphere divided out by the cyclic group of order k. As $k \to \infty$ the volume of these manifolds goes to zero, while the curvature is 1 and the diameter $\frac{\pi}{2}$. Thus, the fundamental groups can only get bigger at the expense of having small volume. If we insist on writing the cyclic group $\mathbb{Z}_k$ in the above manner, then the number of generators needed goes to infinity as $k \to \infty$. This is also justified by the next theorem.*

For numbers $n \in \mathbb{N}$, $k \in \mathbb{R}$, and $v, D \in (0, \infty)$, let $\mathfrak{M}(n, k, v, D)$ denote the class of compact Riemannian n-manifolds with

$$\begin{aligned} \text{Ric} &\geq (n-1)k, \\ \text{vol} &\geq v, \\ \text{diam} &\leq D. \end{aligned}$$

We can now prove:

Theorem 64. (M. Anderson, 1990) *There are only finitely many fundamental groups among the manifolds in $\mathfrak{M}(n, k, v, D)$ for fixed n, k, v, D.*

Proof. Choose generators $\{\gamma_1, \dots, \gamma_m\}$ as in the lemma. Since the number of possible relations is bounded by 2^{m^3}, we have reduced the problem to showing that m is bounded. We have that $d(x, \gamma_i(x)) \leq 2D$. Fix a fundamental domain $F \subset \tilde{M}$

that contains x, i.e., a closed set such that $\pi : F \to M$ is onto and $\text{vol}F = \text{vol}M$. One could, for example, choose the Dirichlet domain

$$F = \left\{ z \in \tilde{M} : d(x, z) \le d(\gamma(x), z) \text{ for all } \gamma \in \pi_1(M) \right\}.$$

Then we have that the sets $\gamma_i(F)$ are disjoint up to sets of measure 0, all have the same volume, and all lie in the ball $B(x, 4D)$. Thus,

$$m \le \frac{\text{vol}B(x, 4D)}{\text{vol}F} \le \frac{v(n, k, 4D)}{v}.$$

In other words, we have bounded the number of generators in terms of n, D, v, k alone. □

Another related result shows that groups generated by short loops must in fact be finite.

LEMMA 39. (M. Anderson, 1990) *For fixed numbers $n \in \mathbb{N}$, $k \in \mathbb{R}$, and $v, D \in (0, \infty)$ we can find $L = L(n, k, v, D)$ and $N = N(n, k, v, D)$ such that if $M \in \mathfrak{M}(n, k, v, D)$, then any subgroup of $\pi_1(M)$ that is generated by loops of length $\le L$ must have order $\le N$.*

PROOF. Let $\Gamma \subset \pi_1(M)$ be a group generated by loops $\{\gamma_1, \ldots, \gamma_k\}$ of length $\le L$. Consider the universal covering $\pi : \tilde{M} \to M$ and let $x \in \tilde{M}$ be chosen such that the loops are based at $\pi(x)$. Then select a fundamental domain $F \subset \tilde{M}$ as above with $x \in F$. Thus for any $\gamma_1, \gamma_2 \in \pi_1(M)$, either $\gamma_1 = \gamma_2$ or $\gamma_1(F) \cap \gamma_2(F)$ has measure 0.

Now define $U(r)$ as the set of $\gamma \in \Gamma$ such that γ can be written as a product of at most r elements from $\{\gamma_1, \ldots, \gamma_k\}$. We assumed that $d(x, \gamma_i(x)) \le L$ for all i, and thus $d(x, \gamma(x)) \le r \cdot L$ for all $\gamma \in U(r)$. This means that $\gamma(F) \subset B(x, r \cdot L + D)$. As the sets $\gamma(F)$ are disjoint up to sets of measure zero, we obtain

$$\begin{aligned} |U(r)| &\le \frac{\text{vol}B(x, r \cdot L + D)}{\text{vol}F} \\ &\le \frac{v(n, k, r \cdot L + D)}{v}. \end{aligned}$$

Now define

$$\begin{aligned} N &= \frac{v(n, k, 2D)}{v} + 1, \\ L &= \frac{D}{N}. \end{aligned}$$

If Γ has more than N elements we get a contradiction by using $r = N$ as we would have

$$\begin{aligned} \frac{v(n, k, 2D)}{v} + 1 &= N \\ &\le |U(N)| \\ &\le \frac{v(n, k, 2D)}{v}. \end{aligned}$$

□

3. Manifolds of Nonnegative Ricci Curvature

In this section we shall prove the splitting theorem of Cheeger-Gromoll. This theorem is analogous to the maximal diameter theorem in many ways. It also has far-reaching consequences for compact manifolds with nonnegative Ricci curvature. For instance, we shall see that $S^3 \times S^1$ does not admit any complete metrics with zero Ricci curvature. One of the critical ingredients in the proof of the splitting theorem is the maximum principle for continuous functions. These analytical matters will be taken care of in the first subsection.

3.1. The Maximum Principle.

We shall try to understand how one can assign second derivatives to (distance) functions at points where the function is not smooth. In chapter 11 we shall also discuss generalized gradients, but this theory is completely different and works only for Lipschitz functions.

The key observation for our development of generalized Hessians and Laplacians is

LEMMA 40. *If $f, h : (M, g) \to \mathbb{R}$ are C^2 functions such that $f(p) = h(p)$ and $f(x) \geq h(x)$ for all x near p, then*

$$\begin{aligned} \nabla f(p) &= \nabla h(p), \\ \mathrm{Hess} f|_p &\geq \mathrm{Hess} h|_p, \\ \Delta f(p) &\geq \Delta h(p). \end{aligned}$$

PROOF. If $(M, g) \subset (\mathbb{R}, \mathrm{can})$, then the theorem is simple calculus. In general, We can take $\gamma : (-\varepsilon, \varepsilon) \to M$ to be a geodesic with $\gamma(0) = p$, then use this observation on $f \circ \gamma$, $h \circ \gamma$ to see that

$$\begin{aligned} df(\dot\gamma(0)) &= dh(\dot\gamma(0)), \\ \mathrm{Hess} f(\dot\gamma(0), \dot\gamma(0)) &\geq \mathrm{Hess} h(\dot\gamma(0), \dot\gamma(0)). \end{aligned}$$

This clearly implies the lemma if we let $v = \dot\gamma(0)$ run over all $v \in T_pM$. □

This lemma implies that a C^2 function $f : M \to \mathbb{R}$ has $\mathrm{Hess} f|_p \geq B$, where B is a symmetric bilinear map on T_pM (or $\Delta f(p) \geq a \in \mathbb{R}$), iff for every $\varepsilon > 0$ there exists a function $f_\varepsilon(x)$ defined in a neighborhood of p such that

(1) $f_\varepsilon(p) = f(p)$.
(2) $f(x) \geq f_\varepsilon(x)$ in some neighborhood of p.
(3) $\mathrm{Hess} f_\varepsilon|_p \geq B - \varepsilon \cdot g|_p$ (or $\Delta f_\varepsilon(p) \geq a - \varepsilon$).

Such functions f_ε are called *support functions from below*. One can analogously use *support functions from above* to find upper bounds for $\mathrm{Hess} f$ and Δf. Support functions are also known as barrier functions in PDE theory.

For a continuous function $f : (M, g) \to \mathbb{R}$ we say that: $\mathrm{Hess} f|_p \geq B$ (or $\Delta f(p) \geq a$) iff there exist smooth support functions f_ε satisfying (1)-(3). One also says that $\mathrm{Hess} f|_p \geq B$ (or $\Delta f(p) \geq a$) hold in the support or barrier sense. In PDE theory there are other important ways of defining weak derivatives. The notion used here is guided by what we can obtain from geometry.

One can easily check that if $(M, g) \subset (\mathbb{R}, \mathrm{can})$, then f is convex if $\mathrm{Hess} f \geq 0$ everywhere. Thus, $f : (M, g) \to \mathbb{R}$ is convex if $\mathrm{Hess} f \geq 0$ everywhere. Using this, one can easily prove

THEOREM 65. *If $f : (M, g) \to \mathbb{R}$ is continuous with* $\text{Hess} f \geq 0$ *everywhere, then f is constant near any local maximum. In particular, f cannot have a global maximum unless f is constant.*

We shall need a more general version of this theorem called the maximum principle. As stated below, it was first proved for smooth functions by E. Hopf in 1927 and then later for continuous functions by Calabi in 1958 using the idea of support functions. A continuous function $f : (M, g) \to \mathbb{R}$ with $\Delta f \geq 0$ everywhere is said to be *subharmonic*. If $\Delta f \leq 0$, then f is *superharmonic*.

THEOREM 66. (The Strong Maximum Principle) *If $f : (M, g) \to \mathbb{R}$ is continuous and subharmonic, then f is constant in a neighborhood of every local maximum. In particular, if f has a global maximum, then f is constant.*

PROOF. First, suppose that $\Delta f > 0$ everywhere. Then f can't have any local maxima at all. For if f has a local maximum at $p \in M$, then there would exist a smooth support function $f_\varepsilon(x)$ with

(1) $f_\varepsilon(p) = f(p)$,
(2) $f_\varepsilon(x) \leq f(x)$ for all x near p,
(3) $\Delta f_\varepsilon(p) > 0$.

Here (1) and (2) imply that f_ε must also have a local maximum at p. But this implies that $\text{Hess} f_\varepsilon(p) \leq 0$, which contradicts (3).

Next just assume that $\Delta f \geq 0$ and let $p \in M$ be a local maximum for f. For sufficiently small $r < \text{inj}(p)$ we therefore have a function $f : (B(p, r), g) \to \mathbb{R}$ with $\Delta f \geq 0$ and a global maximum at p. If f is constant on $B(p, r)$, then we are done, otherwise, we can assume (by possibly decreasing r) that $f(x) \neq f(p)$ for some

$$x \in S(p, r) = \{x \in M : d(p, x) = r\}.$$

Then define

$$V = \{x \in S(p, r) : f(x) = f(p)\}.$$

Our goal is to construct a smooth function $h = e^{\alpha\varphi} - 1$ such that

$$\begin{aligned} h &< 0 \text{ on } V, \\ h(p) &= 0, \\ \Delta h &> 0 \text{ on } \bar{B}(p, r). \end{aligned}$$

This function is found by first selecting an open disc $U \subset S(p, r)$ that contains V. We can then find φ such that

$$\begin{aligned} \varphi(p) &= 0, \\ \varphi &< 0 \text{ on } U, \\ \nabla\varphi &\neq 0 \text{ on } \bar{B}(p, r). \end{aligned}$$

In an appropriate coordinate system $\left(x^1, \ldots, x^n\right)$ we can simply assume that U lies in the lower half-plane: $x^1 < 0$ and then let $\varphi = x^1$ (see also Figure 9.2). Lastly, choose α so big that

$$\Delta h = \alpha e^{\alpha\varphi}(\alpha|\nabla\varphi|^2 + \Delta\varphi) > 0 \text{ on } \overline{B}(p, r).$$

Now consider the function $\bar{f} = f + \delta h$ on $\overline{B}(p, r)$. Provided δ is very small, this function has a local maximum in the interior $B(p, r)$, since

$$\begin{aligned} \bar{f}(p) &= f(p) \\ &> \max\left\{f(x) + \delta h(x) = \bar{f}(x) : x \in \partial B(p, r)\right\}. \end{aligned}$$

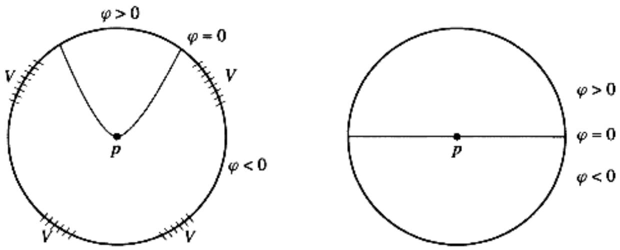

Figure 9.2

On the other hand, we can also show that $\bar{f}$ has positive Laplacian, thus giving a contradiction with the first part of the proof. To see that the Laplacian is positive, select f_ε as a support function from below for f at $q \in B(p, r)$. Then $f_\varepsilon + \delta h$ is a support function from below for $\bar{f}$ at q. The Laplacian of this support function is estimated by

$$\Delta(f_\varepsilon + \delta h)(p) \geq -\varepsilon + \delta \Delta h(p),$$

which for given δ must become positive as $\varepsilon \to 0$. □

A continuous function $f : (M, g) \to \mathbb{R}$ is said to be *linear* if $\mathrm{Hess} f \equiv 0$ (i.e., both of the inequalities $\mathrm{Hess} f \geq 0$, $\mathrm{Hess} f \leq 0$ hold everywhere). One can easily prove that this implies that

$$(f \circ \gamma)(t) = f(\gamma(0)) + \alpha t$$

for each geodesic γ. This implies that

$$f \circ \exp_p(x) = f(p) + g(v_p, x)$$

for each $p \in M$ and some $v_p \in T_p M$. In particular f is C^∞ with $\nabla f|_p = v_p$.

More generally, we have the concept of a harmonic function. This is a continuous function $f : (M, g) \to \mathbb{R}$ with $\Delta f = 0$. The maximum principle shows that if M is closed, then all harmonic functions are constant. On incomplete or complete open manifolds, however, there are often many harmonic functions. This is in contrast to the existence of linear functions, where ∇f is necessary parallel and therefore splits the manifold locally into a product where one factor is an interval. It is an important fact that any harmonic function is C^∞ if the metric is C^∞. Using the above maximum principle we can reduce this to a standard result in PDE theory (see also chapter 10).

THEOREM 67. (Regularity of harmonic functions) *If $f : (M, g) \to \mathbb{R}$ is continuous and harmonic in the weak sense, then f is smooth.*

PROOF. We fix $p \in M$ and a neighborhood Ω around p with smooth boundary. We can in addition assume that Ω is contained in a coordinate neighborhood. It is now a standard fact from PDE theory that the following Dirichlet boundary value problem has a solution:

$$\begin{aligned} \Delta u &= 0, \\ u|_{\partial\Omega} &= f|_{\partial\Omega}. \end{aligned}$$

Moreover, such a solution u is smooth on the interior of Ω. Now consider the two functions $u - f$ and $f - u$ on Ω. If they are both nonpositive, then they must vanish and hence $f = u$ is smooth near p. Otherwise one of these functions must be

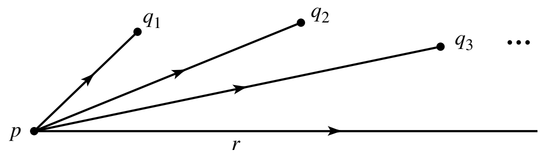

Figure 9.3

positive somewhere. However, as it vanishes on the boundary and is subharmonic this implies that it has an interior global maximum. The maximum principle then shows that the function is constant, but this is only possible if it vanishes. □

3.2. Rays and Lines. We will work only with complete and noncompact manifolds in this section. A *ray* $r(t) : [0, \infty) \to (M, g)$ is a unit speed geodesic such that

$$d\,(r(t),\ r(s)) = |t - s| \text{ for all } t, s \geq 0.$$

One can think of a ray as a semi-infinite segment or as a segment from $r(0)$ to infinity. A *line* $\ell(t) : \mathbb{R} \to (M, g)$ is a unit speed geodesic such that

$$d(\gamma(t), \gamma(s)) = |t - s| \text{ for all } t, s \in \mathbb{R}.$$

Lemma 41. *If $p \in (M, g)$, then there is always a ray emanating from p. If M is disconnected at infinity then (M, g) contains a line.*

Proof. Let $p \in M$ and consider a sequence $q_i \to \infty$. Find unit vectors $v_i \in T_pM$ such that:

$$\sigma_i(t) = \exp_p(tv_i),\ t \in [0, d(p, q_i)]$$

is a segment from p to q_i. By possibly passing to a subsequence, we can assume that $v_i \to v \in T_pM$ (see Figure 9.3). Now

$$\sigma(t) = \exp_p(tv),\ t \in [0, \infty),$$

becomes a segment. This is because σ_i converges pointwise to σ by continuity of $\exp_p$, and thus

$$d(\sigma(s),\ \sigma(t)) = \lim d(\sigma_i(s),\ \sigma_i(t)) = |s - t|.$$

A complete manifold is *connected at infinity* if for every compact set $K \subset M$ there is a compact set $C \supset K$ such that any two points in $M - C$ can be joined by a curve in $M - K$. If M is not connected at infinity, we say that M is *disconnected at infinity*.

If M is disconnected at infinity, we can obviously find a compact set K and sequences of points $p_i \to \infty$, $q_i \to \infty$ such that any curve from p_i to q_i must pass through K. If we join these points by segments $\sigma_i : (-a_i, b_i) \to M$ such that $a_i, b_i \to \infty$, $\sigma_i(0) \in K$, then the sequence will subconverge to a line (see Figure 9.4). □

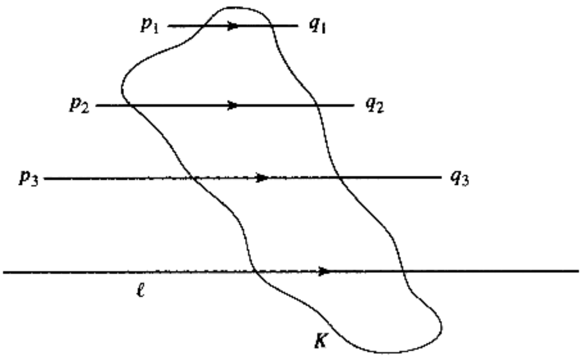

Figure 9.4

EXAMPLE 49. *Surfaces of revolution* $dr^2 + \varphi^2(r)ds^2_{n-1}$, *where* $\varphi : [0, \infty) \to [0, \infty)$ *and* $\dot{\varphi}(t) < 1$, $\ddot{\varphi}(t) < 0$, $t > 0$, *cannot contain any lines. These manifolds look like paraboloids.*

EXAMPLE 50. *Any complete metric on* $S^{n-1} \times \mathbb{R}$ *must contain a line, since the manifold is disconnected at infinity.*

EXAMPLE 51. *The Schwarzschild metric on* $S^2 \times \mathbb{R}^2$ *does not contain any lines. This will also follow from our main result in this section.*

THEOREM 68. (The Splitting Theorem, Cheeger-Gromoll, 1971)*: If* (M, g) *contains a line and has* $\mathrm{Ric} \geq 0$*, then* (M, g) *is isometric to a product* $(H \times \mathbb{R}, g_0 + dt^2)$*.*

The proof is quite involved and will require several constructions. The main idea is to find a distance function $r : M \to \mathbb{R}$ (i.e. $|\nabla r| \equiv 1$) that is linear (i.e. $\mathrm{Hess} r \equiv 0$). Having found such a function, one can easily see that $M = U_0 \times \mathbb{R}$, where $U_0 = \{r = 0\}$ and $g = dt^2 + g_0$. The maximum principle will play a key role in showing that r, when it has been constructed, is both smooth and linear. Recall that in the proof of the maximal diameter theorem we used two distance functions $r, \tilde{r}$ placed at maximal distance from each other and then proceeded to show that $r + \tilde{r} =$ constant. This implied that $r, \tilde{r}$ were smooth, except at the two chosen points, and that Δr is exactly what it is in constant curvature. We then used the rigidity part of the Cauchy-Schwarz inequality to compute $\mathrm{Hess} r$. In the construction of our linear distance function we shall do something similar. In this situation the two ends of the line play the role of the points at maximal distance. Using this line we will construct two distance functions $b_\pm$ from infinity that are continuous, satisfy $b_+ + b_- \geq 0$ (from the triangle inequality), $\Delta b_\pm \leq 0$, and $b_+ + b_- = 0$ on the line. Thus, $b_+ + b_-$ is superharmonic and has a global minimum. The minimum principle will therefore show that $b_+ + b_- \equiv 0$. Thus, $b_+ = -b_-$ and

$$0 \geq \Delta b_+ = -\Delta b_- \geq 0,$$

which shows that both of $b_\pm$ are harmonic and therefore C^∞. We then show that they are actually distance functions (i.e., $|\nabla b_\pm| \equiv 1$). Finally we can conclude that

$$\begin{aligned} 0 &= \nabla b_\pm(\Delta b_\pm) + \frac{(\Delta b_\pm)^2}{n-1} \\ &\leq \nabla b_\pm(\Delta b_\pm) + |\mathrm{Hess} b_\pm|^2 \\ &= |\mathrm{Hess} b_\pm|^2 \\ &\leq -\mathrm{Ric}(\nabla b_\pm, \nabla b_\pm) \leq 0. \end{aligned}$$

This establishes that $|\mathrm{Hess} b_\pm|^2 = 0$, so that we have two linear distance functions $b_\pm$ as desired.

The proof proceeds through several results some of which we will need later.

3.3. Laplacian Comparison.

LEMMA 42. (E. Calabi, 1958) *Let* $r(x) = d(x,p),\ p \in (M,g)$. *If* $\mathrm{Ric}(M,g) \geq 0$, *then*

$$\Delta r(x) \leq \frac{n-1}{r(x)} \text{ for all } x \in M.$$

PROOF. We know that the result is true whenever r is smooth. For any other $q \in M$, choose a unit speed segment $\sigma : [0,\ell] \to M$ with $\sigma(0) = p,\ \sigma(\ell) = q$. Then the triangle inequality implies that $r_\varepsilon(x) = \varepsilon + d(\sigma(\varepsilon), x)$ is a support function from above for r at q. If all these support functions are smooth at q, then

$$\begin{aligned} \Delta r_\varepsilon(q) &\leq \frac{n-1}{r_\varepsilon(q) - \varepsilon} \\ &= \frac{n-1}{r(q) - \varepsilon} \\ &\leq \frac{n-1}{r(q)} + \varepsilon \cdot \frac{2(n-1)}{(r(q))^2} \end{aligned}$$

for small ε, and hence $\Delta r(q) \leq \frac{n-1}{r(q)}$ in the support sense.

Now for the smoothness. Fix $\varepsilon > 0$ and suppose r_ε is not smooth at q. Then we know that either

(1) there are two segments from $\sigma(\varepsilon)$ to q, or
(2) q is a critical value for $\exp_{\sigma(\varepsilon)} : \mathrm{seg}(\sigma(\varepsilon)) \to M$.

Case (1) would give us a nonsmooth curve of length ℓ from p to q, which we know is impossible. Thus, case (2) must hold. To get a contradiction out of this, we show that this implies that $\exp_q$ has $\sigma(\varepsilon)$ as a critical value.

Using that q is critical for $\exp_{\sigma(\varepsilon)}$, we find a Jacobi field $J(t) : [\varepsilon, \ell] \to TM$ along $\sigma|_{[\varepsilon,\ell]}$ such that $J(\varepsilon) = 0$, $\dot{J}(\varepsilon) \neq 0$ and $J(\ell) = 0$ (see chapter 6). Then also $\dot{J}(\ell) \neq 0$ as it solves a linear second order equation. Running backwards from q to $\sigma(\varepsilon)$ then shows that $\exp_q$ is critical at $\sigma(\varepsilon)$. This however contradicts that $\sigma : [0,\ell] \to M$ is a segment. □

By a similar analysis, we can prove

LEMMA 43. *If* (M,g) *is complete and* $\mathrm{Ric}(M,g) \geq (n-1)k$, *then any distance function* $r(x) = d(x,p)$ *satisfies:*

$$\Delta r(x) \leq (n-1)\frac{\mathrm{sn}_k'(r(x))}{\mathrm{sn}_k(r(x))}.$$

This lemma together with the maximum principle allows us to eliminate the use of relative volume comparison in the proof of Cheng's diameter theorem.

As in the other proof, consider $\tilde{r}(x) = d(x,q)$, $r(x) = d(x,p)$, where $d(p,q) = \pi/\sqrt{k}$. Then we have $r + \tilde{r} \geq \pi/\sqrt{k}$, and equality will hold for any $x \in M - \{p,q\}$ that lies on a segment joining p and q. On the other hand, the above lemma tells us that

$$
\begin{aligned}
\Delta(r+\tilde{r}) &\leq \Delta r + \Delta\tilde{r} \\
&\leq (n-1)\sqrt{k}\cot(\sqrt{k}r(x)) + (n-1)\sqrt{k}\cdot(\sqrt{k}\tilde{r}(x)) \\
&\leq (n-1)\sqrt{k}\cot(\sqrt{k}r(x)) + (n-1)\sqrt{k}\cot\left(\sqrt{k}\left(\frac{\pi}{\sqrt{k}} - r(x)\right)\right) \\
&= (n-1)\sqrt{k}(\cot(\sqrt{k}r(x)) + \cot(\pi - \sqrt{k}r(x))) = 0.
\end{aligned}
$$

So $r + \tilde{r}$ is superharmonic on $M - \{p,q\}$ and has a global minimum on this set. Thus, the minimum principle tells us that $r + \tilde{r} = \pi/\sqrt{k}$ on M. The proof can now be completed as before.

3.4. Busemann Functions. For the rest of this section we fix a complete noncompact Riemannian manifold (M,g) with nonnegative Ricci curvature. Let $\gamma : [0,\infty) \to (M,g)$ be a unit speed ray, and define

$$b_t(x) = d(x,\gamma(t)) - t.$$

PROPOSITION 40. *(1) For fixed x, the function $t \to b_t(x)$ is decreasing and bounded in absolute value by $d(x,\gamma(0))$.*
(2) $|b_t(x) - b_t(y)| \leq d(x,y)$.
(3) $\Delta b_t(x) \leq \frac{n-1}{b_t+t}$ everywhere.

PROOF. (2) and (3) are obvious, since $b_t(x)+t$ is a distance function from $\gamma(t)$. For (1), first observe that the triangle inequality implies

$$|b_t(x)| = |d(x,\gamma(t) - t| = |d(x,\gamma(t)) - d(\gamma(0),\gamma(t))| \leq d(x,\gamma(0)).$$

Second, if $s < t$ then

$$
\begin{aligned}
b_t(x) - b_s(x) &= d(x,\gamma(t)) - t - d(x,\gamma(s)) + s \\
&= d(x,\gamma(t)) - d(x,\gamma(s)) - d(\gamma(t),\gamma(s)) \\
&\leq d(\gamma(t),\gamma(s)) - d(\gamma(t),\gamma(s)) = 0.
\end{aligned}
$$

□

This proposition shows that the family of functions $\{b_t\}_{t\geq 0}$ forms a pointwise bounded equicontinuous family that is also pointwise decreasing. Thus, b_t must converge to a distance-decreasing function b_γ satisfying

$$
\begin{aligned}
|b_\gamma(x) - b_\gamma(y)| &\leq d(x,y), \\
|b_\gamma(x)| &\leq d(x,\gamma(0)),
\end{aligned}
$$

and

$$b_\gamma(\gamma(r)) = \lim b_t(\gamma(r)) = \lim(d(\gamma(r),\gamma(t)) - t) = -r.$$

This function b_γ is called the *Busemann function* for γ and should be interpreted as a distance function from "$\gamma(\infty)$."

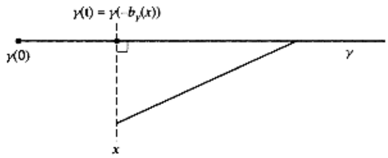

Figure 9.5

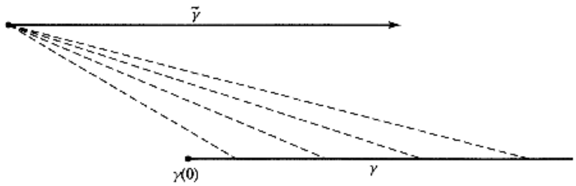

Figure 9.6

Example 52. *If $M = (\mathbb{R}^n, \text{can})$, then all Busemann functions are of the form*

$$b_\gamma(x) = \gamma(0) - \dot{\gamma}(0) \cdot x$$

(see Figure 9.5).

The level sets $b_\gamma^{-1}(t)$ are called *horospheres*. In $(\mathbb{R}^n, \text{can})$ these are obviously hyperplanes.

Given our ray γ, as before, and $p \in M$, consider a family of unit speed segments $\sigma_t : [0, \ell_t] \to (M, g)$ from p to $\gamma(t)$. As when we constructed rays, this family must subconverge to some ray $\tilde{\gamma} : [0, \infty) \to M$, with $\tilde{\gamma}(0) = p$. A ray coming from such a construction is called an *asymptote* for γ from p (see Figure 9.6). Such asymptotes from p need not be unique.

Proposition 41. *(1) $b_\gamma(x) \leq b_\gamma(p) + b_{\tilde{\gamma}}(x)$.*
(2) $b_\gamma(\tilde{\gamma}(t)) = b_\gamma(p) + b_{\tilde{\gamma}}(\tilde{\gamma}(t)) = b_\gamma(p) - t$.

Proof. Let $\sigma_i : [0, \ell_i] \to (M, g)$ be the segments converging to $\tilde{\gamma}$. To check (1), observe that

$$\begin{aligned} d(x, \gamma(s)) - s &\leq d(x, \tilde{\gamma}(t)) + d(\tilde{\gamma}(t), \gamma(s)) - s \\ &= d(x, \tilde{\gamma}(t)) - t + d(p, \tilde{\gamma}(t)) + d(\tilde{\gamma}(t), \gamma(s)) - s \\ &\to d(x, \tilde{\gamma}(t)) - t + d(p, \tilde{\gamma}(t)) + b_\gamma(\tilde{\gamma}(t)) \quad \text{as} \quad s \to \infty. \end{aligned}$$

Thus, we see that (1) is true provided that (2) is true. To establish (2), we notice that

$$d(p, \gamma(t_i)) = d(p, \sigma_i(s)) + d(\sigma_i(s), \gamma(t_i))$$

for some sequence $t_i \to \infty$. Now, $\sigma_i(s) \to \tilde{\gamma}(s)$, so we obtain

$$\begin{aligned} b_\gamma(p) &= \lim(d(p, \gamma(t_i)) - t_i) \\ &= \lim(d(p, \tilde{\gamma}(s)) + d(\tilde{\gamma}(s), \gamma(t_i)) - t_i) \\ &= d(p, \tilde{\gamma}(s)) + \lim(d(\tilde{\gamma}(s), \gamma(t_i)) - t_i) \\ &= s + b_\gamma(\tilde{\gamma}(s)) \\ &= -b_{\tilde{\gamma}}(\tilde{\gamma}(s)) + b_\gamma(\tilde{\gamma}(s)). \end{aligned}$$

□

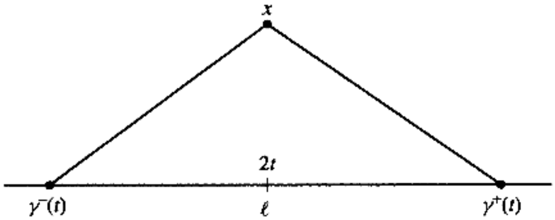

Figure 9.7

We have now shown that b_γ has $b_\gamma(p) + b_{\tilde{\gamma}}$ as support function from above at $p \in M$.

LEMMA 44. *If* $\mathrm{Ric}(M, g) \geq 0$, *then* $\Delta b_\gamma \leq 0$ *everywhere.*

PROOF. Since $b_\gamma(p) + b_{\tilde{\gamma}}$ is a support function from above at p, we only need to check that $\Delta b_{\tilde{\gamma}}(p) \leq 0$. To see this, observe that the functions

$$b_t(x) = d(x, \tilde{\gamma}(t)) - t$$

are actually support functions from above for $b_{\tilde{\gamma}}$ at p. Furthermore, these functions are smooth at p with

$$\Delta b_t(p) \leq \frac{n-1}{t} \to 0 \text{ as } t \to \infty.$$

□

Now suppose (M, g) has Ric ≥ 0 and contains a line $\gamma(t) : \mathbb{R} \to M$. Let b^+ be the Busemann function for $\gamma : [0, \infty) \to M$, and b^- the Busemann function for $\gamma : (-\infty, 0] \to M$. Thus,

$$\begin{aligned} b^+(x) &= \lim_{t \to +\infty} (d(x, \gamma(t)) - t), \\ b^-(x) &= \lim_{t \to +\infty} (d(x, \gamma(-t)) - t). \end{aligned}$$

Clearly,

$$b^+(x) + b^-(x) = \lim_{t \to +\infty} (d(x, \gamma(t)) + d(x, \gamma(-t)) - 2t),$$

so by the triangle inequality

$$\left(b^+ + b^-\right)(x) \geq 0 \text{ for all } x.$$

Moreover,

$$\left(b^+ + b^-\right)(\gamma(t)) = 0$$

since γ is a line (see Figure 9.7).

This gives us a function $b^+ + b^-$ with $\Delta(b^+ + b^-) \leq 0$ and a global minimum at $\gamma(t)$. The minimum principle then shows that $b^+ + b^- = 0$ everywhere. In particular, $b^+ = -b^-$ and $\Delta b^+ = \Delta b^- = 0$ everywhere.

To finish the proof of the splitting theorem, we still need to show that $b^\pm$ are distance functions, i.e. $|\nabla b^\pm| \equiv 1$. To see this, let $p \in M$ and construct asymptotes $\tilde{\gamma}^\pm$ for $\gamma^\pm$ from p. Then consider

$$b_t^\pm(x) = d(x, \tilde{\gamma}^\pm(t)) - t,$$

and observe:

$$b_t^+(x) \geq b^+(x) - b^+(p) = -b^-(x) + b^-(p) \geq -b_t^-(x)$$

with equality holding for $x = p$. Since both $b_t^{\pm}$ are smooth at p with unit gradient, we must therefore have that $\nabla b_t^+(p) = -\nabla b_t^-(p)$. Then also, $b^{\pm}$ must be differentiable at p with unit gradient. We have therefore shown (without using that $b^{\pm}$ are smooth from $\Delta b^{\pm} = 0$) that $b^{\pm}$ are everywhere differentiable with unit gradient. The result that harmonic functions are smooth can now be invoked and the proof is finished as explained in the beginning of the section.

3.5. Structure Results in Nonnegative Ricci Curvature. The splitting theorem gives several nice structure results for compact manifolds with nonnegative Ricci curvature.

COROLLARY 26. *$S^p \times S^1$ does not admit any Ricci flat metrics when $p = 2, 3$.*

PROOF. The universal covering is $S^p \times \mathbb{R}$, As this space is disconnected at infinity any metric with nonnegative Ricci curvature must split. If the original metric is Ricci flat, then after the splitting, we will get a Ricci flat metric on S^p. If $p \leq 3$, such a metric must also be flat. But we know that $S^p,\ p = 2, 3$ do not admit any flat metrics. □

When $p \geq 4$ it is not known whether S^p admits a Ricci flat metric.

THEOREM 69. (Structure Theorem for Nonnegative Ricci Curvature, Cheeger-Gromoll, 1971) *Suppose (M, g) is a compact Riemannian manifold with* $\operatorname{Ric} \geq 0$. *Then the universal cover $(\widetilde{M}, \tilde{g})$ splits isometrically as a product $N \times \mathbb{R}^p$, where N is a compact manifold.*

PROOF. By the splitting theorem, we can write $\widetilde{M} = N \times \mathbb{R}^p$, where N does not contain any lines. Observe that if

$$\gamma(t) = (\gamma_1(t), \gamma_2(t)) \in N \times \mathbb{R}^p$$

is a geodesic, then both γ_i are geodesics, and if γ is a line, then both γ_i are also lines unless they are constant. Thus, all lines in $\widetilde{M}$ must be of the form $\gamma(t) = (x, \sigma(t))$, where $x \in N$ and σ is a line in $\mathbb{R}^p$.

If N is not compact, then it must contain a ray $\gamma(t) : [0, \infty) \to N$. If $\pi : \widetilde{M} \to M$ is the covering map, then we can consider $c(t) = \pi \circ (\gamma(t), 0)$ in M. This is of course a geodesic in M, and since M is compact, there must be a sequence $t_i \to \infty$ such that $\dot{c}(t_i) \to v \in T_xM$ for some $x \in M$, $v \in T_xM$. Choose $\tilde{x} \in \widetilde{M}$ such that $\pi(\tilde{x}) = x$, and consider lifts $\gamma_i(t) : [-t_i, \infty) \to \widetilde{M}$ of $c(t + t_i)$, where $D\pi(\dot{\gamma}_i(0)) = \dot{c}(t_i)$ and $\gamma_i(0) \to \tilde{x}$. On the one hand, these geodesics converge to a geodesic $\hat{\gamma} : (-\infty, \infty) \to \widetilde{M}$ with $\hat{\gamma}(0) = \tilde{x}$. On the other hand, since $D\pi(\dot{\gamma}(t_i)) = \dot{c}(t_i)$, there must be deck transformations $g_i \in \pi_1(M)$ such that $g_i \circ \gamma(t + t_i) = \gamma_i(t)$. Thus, the γ_is are rays and must converge to a line. From our earlier observations, this line must be in $\mathbb{R}^p$. The deck transformations g_i therefore map $\dot{\gamma}(t + t_i)$, which are tangent to N, to vectors that are almost perpendicular to N. This, however, contradicts the following property for isometries on $\widetilde{M}$.

Let $F : \widetilde{M} \to \widetilde{M}$ be an isometry, e.g., $F = g_i$. If $\ell(t)$ is a line in $\widetilde{M}$, then $F \circ \ell$ must also be a line in $\widetilde{M}$. Since all lines in $\widetilde{M}$ lie in $\mathbb{R}^p$ and every vector tangent to $\mathbb{R}^p$ is the velocity of some line, we see that for each $c \in N$ we can find $F_1(c) \in N$ such that

$$F : \{c\} \times \mathbb{R}^p \to \{F_1(c)\} \times \mathbb{R}^p.$$

This implies that F must be of the form $F = (F_1, F_2)$, where $F_1 : N \to N$ is an isometry and $F_2 : N \times \mathbb{R}^p \to \mathbb{R}^p$. In particular, the tangent bundles TN and $T\mathbb{R}^p$ are preserved by DF. □

This theorem also gives a strong structure for $\pi_1(M)$. Consider the group G of isometries on N that are split off by the action of $\pi_1(M)$ on $\widetilde{M} = N \times \mathbb{R}^p$. Since N is compact and G acts discretely on N we see that it is finite. The kernel of the homomorphism $\pi_1(M) \to G$ is then a finite index subgroup that acts discretely and cocompactly on $\mathbb{R}^p$. Such groups are known as *crystallographic groups* and are fairly well understood. It is a theorem of Bieberbach that any group of isometries $\Gamma \subset \mathrm{Iso}(\mathbb{R}^p)$ that is discrete and cocompact must contain a rank p Abelian group $\mathbb{Z}^p$ of finite index. This structure comes from the exact sequence

$$1 \to \mathbb{R}^p \to \mathrm{Iso}(\mathbb{R}^p) \to O(p) \to 1,$$

where the map $\mathrm{Iso}(\mathbb{R}^p) \to O(p)$ is the assignment that takes the isometry $Ox + v$ to O. If we restrict this short exact sequence to Γ we see that the kernel is an Abelian subgroup of $\mathbb{R}^p$ which acts discretely. This shows that it must be of the form $\mathbb{Z}^q$ where $q \leq p$. If $q < p$, then the action of $\mathbb{Z}^q$ leaves the q-dimensional subspace $V = \mathrm{span}\{\mathbb{Z}^q\}$ invariant and therefore fixes the orthogonal complement. This shows that the action can't be cocompact. Finally we also note that the image in $O(p)$ is discrete and hence finite. (For more details see also [**34**], [**96**]). Note that there are non-discrete actions of $\mathbb{Z}^n$ on $\mathbb{R}$ for any $n \geq 1$. To see this simply take real numbers $\alpha_1, ..., \alpha_n$ that are linearly independent over $\mathbb{Q}$ and use these as a basis for the action. Note, however, that all orbits of this action are dense so it is not a discrete action.

We can now prove some further results about the structure of compact manifolds with nonnegative Ricci curvature.

COROLLARY 27. *Suppose (M, g) is a complete, compact Riemannian manifold with* Ric ≥ 0. *If M is $K(\pi, 1)$, i.e., the universal cover is contractible, then the universal covering is Euclidean space and (M, g) is a flat manifold.*

PROOF. We know that $\widetilde{M} = \mathbb{R}^p \times C$, where C is compact. The only way in which this space can be contractible is if C is contractible. But the only compact manifold that is contractible is the one-point space. □

COROLLARY 28. *If (M, g) is compact with* Ric ≥ 0 *and has* Ric > 0 *on some tangent space T_pM, then $\pi_1(M)$ is finite.*

PROOF. Since Ric > 0 on an entire tangent space, the universal cover cannot split into a product $\mathbb{R}^p \times C$, where $p \geq 1$. Thus, the universal covering is compact. □

COROLLARY 29. *If (M, g) is compact and has* Ric ≥ 0*, then $b_1(M) \leq \dim M = n$, with equality holding iff (M, g) is a flat torus.*

PROOF. We always have a surjection

$$h : \pi_1(M) \to H_1(M, \mathbb{Z}),$$

that maps loops to cycles. The above mentioned structure result for the fundamental groups shows that we have a finite index subgroup $\mathbb{Z}^p \subset \pi_1(M)$ with $p \leq n$. The image $h(\mathbb{Z}^p) \subset H_1(M, \mathbb{Z})$ is therefore also of finite index. This shows that the rank of the torsion free part of $H_1(M, \mathbb{Z})$ must be $\leq p$. In case $b_1 = n$, we must have that $p = n$ as $h(\mathbb{Z}^p)$ otherwise couldn't have finite index. This shows in addition that

$$h|_{\mathbb{Z}^n} : \mathbb{Z}^n \to H_1(M, \mathbb{Z})$$

has trivial kernel as the image otherwise couldn't have finite index. Thus $\widetilde{M} = \mathbb{R}^n$ as p was the dimension of the Euclidean factor. Consequently M is flat. We now observe that the kernel of

$$h : \pi_1(M) \to H_1(M, \mathbb{Z})$$

has to be a finite subgroup as it does not intersect the finite index subgroup $\mathbb{Z}^n \subset \pi_1(M)$. Since all isometries on $\mathbb{R}^n$ of finite order have a fixed point we have shown that the inclusion $\mathbb{Z}^n \subset \pi_1(M)$ is an isomorphism. This shows that M is a torus. □

The penultimate result is a bit stronger than simply showing that $H^1(M, \mathbb{R}) = 0$ as we did using the Bochner technique. The last result is equivalent to Bochner's theorem, but the proof is quite a bit different.

4. Further Study

The adventurous reader could consult [**47**] for further discussions. Anderson's article [**2**] contains the finiteness results for fundamental groups mentioned here and also some interesting examples of manifolds with nonnegative Ricci curvature. For the examples with almost maximal diameter we refer the reader to [**3**] and [**74**]. It is also worthwhile to consult the original paper on the splitting theorem [**27**] and the elementary proof of it in [**37**]. We already mentioned in chapter 7 Gallot's contributions to Betti number bounds, and the reference [**40**] works here as well. The reader should also consult the articles by Colding, Perel'man, and Zhu in [**50**] to get an idea of how rapidly this subject has grown in the past few years.

5. Exercises

(1) With notation as in the first section:

$$d\text{vol} = \lambda dr \wedge d\text{vol}_{n-1}.$$

Show that $\mu = \lambda^{\frac{1}{n-1}}$ satisfies

$$\begin{aligned} \partial_r^2 \mu &\leq -\frac{\mu}{n-1} \operatorname{Ric}(\partial_r, \partial_r), \\ \mu(0, \theta) &= 0, \\ \lim_{r \to 0} \partial_r \mu(r, \theta) &= 1. \end{aligned}$$

This can be used to show the desired estimates for the volume form as well.

(2) Assume the distance function $r = d(\cdot, p)$ is smooth on $B(p, R)$. If in our usual polar coordinates

$$\text{Hess} r = \frac{\text{sn}_k'(r)}{\text{sn}_k(r)} g_r,$$

then all sectional curvatures on $B(p, R)$ are equal to k.

(3) Show that if (M, g) has $\operatorname{Ric} \geq (n-1)k$ and for some $p \in M$ we have $\text{vol} B(p, R) = v(n, k, R)$, then the metric has constant curvature k on $B(p, R)$.

(4) Let X be a vector field on a Riemannian manifold and consider $F_t(p) = \exp_p(tX|_p)$.
   (a) For $v \in T_pM$ show that $J(t) = DF_t(v)$ is a Jacobi field along $t \to \gamma(t) = \exp(tX)$ with the initial conditions $J(0) = v$, $\dot{J}(0) = \nabla_v X$.
   (b) Select an orthonormal basis e_i for T_pM and let $J_i(t) = DF_t(e_i)$. Show that
$$(\det(DF_t))^2 = \det(g(J_i(t), J_j(t))).$$
   (c) Show that as long as det $(DF_t) \neq 0$ it satisfies
$$\frac{d^2(\det(DF_t))^{\frac{1}{n}}}{dt^2} \leq -\frac{(\det(DF_t))^{\frac{1}{n}}}{n}\operatorname{Ric}(\dot{\gamma}, \dot{\gamma}).$$
   Hint: Use that any $n \times n$ matrix satisfies $(\operatorname{tr}(A))^2 \leq n\operatorname{tr}(A^*A)$.

(5) Show that a complete manifold (M, g) with the property that
$$\begin{aligned} \operatorname{Ric} &\geq 0, \\ \lim_{r\to\infty} \frac{\operatorname{vol}B(p,r)}{\omega_n r^n} &= 1, \end{aligned}$$
for some $p \in M$, must be isometric to Euclidean space.

(6) *(Cheeger)* The relative volume comparison estimate can be generalized as follows: Suppose (M, g) has $\operatorname{Ric} \geq (n-1)k$ and dimension n.
   (a) Select points $p_1, \ldots, p_k \in M$. Then the function
$$r \to \frac{\operatorname{vol}\left(\bigcup_{i=1}^k B(p_i, r)\right)}{v(n, k, r)}$$
   is nonincreasing and converges to k as $r \to 0$.
   (b) If $A \subset M$, then
$$r \to \frac{\operatorname{vol}\left(\bigcup_{p\in A} B(p, r)\right)}{v(n, k, r)}$$
   is nonincreasing. To prove this, use the above with the finite collection of points taken to be very dense in A.

(7) The absolute volume comparison can also be slightly generalized. Namely, for $p \in M$ and a subset $\Gamma \subset T_pM$ of unit vectors, consider the cones defined in polar coordinates:
$$B^\Gamma(p, r) = \{(t, \theta) \in M : t \leq r \text{ and } \theta \in \Gamma\}.$$
If $\operatorname{Ric}M \geq (n-1)k$, show that
$$\operatorname{vol}B^\Gamma(p, r) \leq \operatorname{vol}\Gamma \cdot \int_0^r (\operatorname{sn}_k(t))^{n-1} dt.$$

(8) Let G be a compact connected Lie group with a bi-invariant metric. Use the results from this chapter to prove
   (a) If G has finite center, then G has finite fundamental group.
   (b) A finite covering of G looks like $G' \times T^k$, where G' is compact simply connected, and T^k is a torus.
   (c) If G has finite fundamental group, then the center is finite.

(9) Show that a compact Riemannian manifold with irreducible restricted holonomy and $\operatorname{Ric} \geq 0$ has finite fundamental group.

(10) Let (M, g) be an n-dimensional Riemannian manifold that is isometric to Euclidean space outside some compact subset $K \subset M$, i.e., $M - K$ is isometric to $\mathbb{R}^n - C$ for some compact set $C \subset \mathbb{R}^n$. If $\text{Ric}_g \geq 0$, show that $M = \mathbb{R}^n$. (In chapter 7 we gave two different hints for this problem, here is a third. Use the splitting theorem.)

(11) Show that if $\text{Ric} \geq n - 1$, then $\text{diam} \leq \pi$, by showing that if $d(p, q) > \pi$, then

$$e_{p,q}(x) = d(p, x) + d(x, q) - d(p, q)$$

has negative Laplacian at a local minimum.