BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 36, Number 3, Pages 297–344
S 0273-0979(99)00787-9
Article electronically published on May 24, 1999

ASPECTS OF GLOBAL RIEMANNIAN GEOMETRY

PETER PETERSEN

ABSTRACT. In this article we survey some of the developments in Riemannian geometry. We place special emphasis on explaining the relationship between curvature and topology for Riemannian manifolds with lower curvature bounds.

1. Introduction

We shall in this survey explain the development of the branch of Riemannian geometry called *global Riemannian geometry*. The main goal of this particular type of geometry is to classify topologically, or even metrically, manifolds with certain given geometric conditions on, say, curvature, volume and diameter. The earliest such theorem is the Gauss-Bonnet theorem, which gives the Euler characteristic in terms of the integral of the Gaussian curvature. Since then the subject has developed tremendously, and there is no way we can explain most of the important results. Some exclusions have therefore been necessary. Thus there will be no discussion of submanifold geometry including minimal surfaces. Only a few theorems on manifolds with nonpositive curvature are mentioned, very little on what is called geometric analysis is presented, and we have not discussed Kähler geometry at all. We have thus tried to emphasize results where the main assumption will be a *lower curvature bound* of some sort and the main technique something related to either the *Bochner technique* or the *Rauch estimates*. Even then, it has been hard to include many interesting results. For instance, while many of these ideas have naturally led to very interesting investigations into spaces which are more singular than Riemannian manifolds, and this in turn has given us a better understanding of Riemannian manifolds themselves, we have decided not to talk about these issues. Another related branch we have ignored is that of examples. Some of the references given in the bibliography might help the interested reader to learn more about some of the aspects of geometry not surveyed here.

We have placed particular emphasis on some of the early developments in both global and local Riemannian geometry in the hope that this will make it easier for students and nonexperts to understand some of the later developments. Thus the basic prerequisite for this article is some familiarity with manifold theory including a little knowledge of vector fields and forms. We have also tried to present results in historical rather than logical order, although within each of the artificial periods we have set up, it has been necessary to explain things in parallel rather than linear

Received by the editors November 20, 1997, and in revised form October 20, 1998.
1991 *Mathematics Subject Classification.* Primary 53C20.
Key words and phrases. Riemannian geometry.
Supported in part by the NSF.

order. I have placed particular importance on getting the credits sorted out for many of the classical results. To this end, several discussions with M. Berger have been profoundly helpful.

I would like to thank C. Sprouse and the referee for carefully reading the manuscript and providing me with much constructive criticism.

2. Early global results

The first global geometric results, in the spirit that we wish to consider, go back to Descartes. He considered polyhedral surfaces (usually convex, but here we consider the general case) in $\mathbb{R}^3$. We know that closed surfaces have a topological invariant χ associated to them, called the Euler characteristic. This invariant satisfies the formula $\chi = 2 - 2g$, where g is the genus of the surface. Here $g = 0$ for the sphere, $g = 1$ for the torus, $g = 2$ for the double torus, etc. What Descartes observed is that there is a way of calculating this Euler characteristic by calculating all of the angles of the polygonal faces on the surface. The polyhedral surface consists of a collection of polygons which are glued together along edges in such a way that each edge is met by exactly two polygons. A vertex on the surface can be met by any number of polygons. For each vertex v we add up the angles corresponding to the polygons meeting the vertex. This total angle is denoted T_v. In case $T_v = 2\pi$ the vertex is flat, if $T_v < 2\pi$ it looks like the apex of a pyramid or mountain top, while if $T_v > 2\pi$ it looks like a saddle point (this situation does not occur if the surface is convex). Thus we can construct the angle defect $C_v = 2\pi - T_v$. This is a discrete measure of the curvature at the point v. All other points are flat since they have neighborhoods isometric to balls in the Euclidean plane (yes, even points on edges). The nice discovery is that

$$\sum_v C_v = 2\pi\chi.$$

If we use Euler's theorem,

$$\chi = \#\text{ polygons } - \#\text{ edges } + \#\text{ vertices},$$

together with the classical result that for each polygon with k sides we have that the sum of the interior angles is $(k-2)\,\pi$, then we can easily obtain the above formula.

Moreover, we can refine this result to be more local. Namely, consider a simply connected polyhedral surface with boundary. The boundary is the collection of edges which are only met by one polygon. Now for the vertices on the boundary we need to adjust our angular defect to be $\partial C_v = \pi - T_v$. We then obtain

$$\sum_{v \text{ in interior}} C_v + \sum_{v \text{ on boundary}} \partial C_v = 2\pi.$$

For the simple case of a polygon there are no interior vertices and the angular defect ∂C_v is the exterior angle, so we obtain the classical theorem mentioned above for the angle sum of a planar polygon.

What is interesting about these formulae is that they can be used to measure the curvature of the earth through surveying. It might have been Gauss who first discovered this, or in any case he was the first to observe this in action. He went out and triangulated Lüneburg Heide, thus generating a simply connected polyhedral approximation to the heath. If indeed the heath is planar, the angular defects for the interior vertices should cancel out, but they don't!

Gauss actually generalized the above formula to hold for geodesic polygons on smooth surfaces in Euclidean space. The setup is that we have a smooth surface $S \subset \mathbb{R}^3$ and a simply connected region $F \subset S$ whose boundary consists of a union of geodesic arcs (i.e., curves of shortest length). Where two geodesics meet on the boundary we have as above an exterior angle ∂C_v. Now the interior of this polygon is smooth, so the sum over interior vertices has to be replaced with something else. Gauss discovered that the total curvature $\int_F K$, where $K : S \to \mathbb{R}$ is the Gaussian curvature, does the job. Thus he showed

$$\int_F K + \sum_{v \text{ on boundary}} \partial C_v = 2\pi.$$

For a geodesic triangle this takes the simple form

$$\int_F K = (\alpha + \beta + \gamma) - \pi,$$

where α, β, γ are the interior angles of the triangle and $(\alpha + \beta + \gamma) - \pi$ the angular defect.

It remains to explain what the Gaussian curvature is. At a point p on the surface it measures the defect from the surface being flat at this point, one has that if $l(r)$ is the length of the boundary of a small ball of radius r around p, then

$$l(r) = 2\pi r - \frac{2\pi}{6} K(p) r^3 + O\left(r^4\right).$$

Here $2\pi r$ is exactly what one would expect if the surface were flat at p.

From Gauss's formula for the angle defect of a geodesic triangle one can easily arrive at a global result for closed surfaces. Namely, for a closed surface S, triangulate it by geodesic triangles, and add up the contributions to get

$$\int_S K = 2\pi\chi.$$

It is not clear who first observed this. Even for spheres Gauss doesn't seem to have been aware of the global formula.

In 1848 a variant of Gauss's formula for geodesic polygons was obtained by Bonnet. He considered smooth simply connected surfaces S with smooth boundary ∂S. Now that the boundary is smooth we have instead an infinitesimal angular defect, called the geodesic curvature. If we consider the boundary as a curve parametrized by arclength, the geodesic curvature κ is simply the length of the acceleration vector of the curve in the surface. (The acceleration of a curve on a surface in $\mathbb{R}^3$ is simply the acceleration in $\mathbb{R}^3$ projected down onto the surface.) Bonnet then showed

$$\int_S K + \int_{\partial S} \kappa = 2\pi.$$

This formula can of course also be used to prove the global result, but Bonnet doesn't seem to have observed this either. Surprisingly, this proof seems to first appear in the literature in 1921 with Blaschke's book [18]. In fact, as far as we can discover, the first proof of the global Gauss-Bonnet theorem for embedded surfaces came about in a completely different way. In 1869 Kronecker (see [81]) introduced the degree of a map between surfaces and showed that it could be computed by integrating the Jacobian determinant. He then observed that one can apply this to the Gauss map of a surface. Recall that the Gauss map of an oriented surface in $\mathbb{R}^3$ is the map that takes a point on the surface to the positive normal at the point. Thus the Gauss map is a map

$$G : M^2 \to S^2.$$

Since this map measures how the surface changes from point to point, it is not hard to believe that

$$\det DG = K.$$

Actually, this was Gauss's definition of K. Kronecker now used his formula for the degree to obtain

$$\frac{1}{\mathrm{vol} S^2} \int_M K = \deg(G).$$

This immediately leads to the formula

$$\int_M K = 4\pi \deg(G).$$

In 1888 Dyck then showed that the degree of the Gauss map is related to the Euler characteristic in the following way

$$2 \deg(G) = \chi,$$

thus establishing the global Gauss-Bonnet formula.

Gauss's work is translated in [39] together with explanations and complete references to most of the concepts and results that were developed for surfaces in the nineteenth century.

There are two more important global results for surfaces that we shall have recourse to discuss later. The first is by Bonnet from 1855 (see [20]). He shows that any convex surface with Gauss curvature $\geq k^2 > 0$ has diameter $\leq \pi/k$ (this is the diameter of a sphere of radius $1/k$, and such a sphere has Gauss curvature equal to k^2). The other result is by von Mangoldt and was proved in 1881 (see [89]). He shows that any complete surface of nonpositive curvature has the Euclidean plane as its covering space. This theorem is often incorrectly attributed to Hadamard, yet Hadamard was certainly aware that von Mangoldt first proved it. The mistake seems to occur first in [22], where Cartan proves the same result for abstract Riemannian manifolds.

In the next section we shall explain some of the standard language of Riemannian geometry. After this is done we shall explain why the two results of Bonnet and von Mangoldt are true.

3. Local theory

We shall in this section explain some of the foundational concepts developed by Gauss, Riemann, Levi-Civita, Cartan and others during the period from 1825 to 1925. As promised, we shall also explain how Bonnet's diameter estimate and von Mangoldt's theorem were proved. To save a little space we have decided to develop the abstract theory from the beginning, rather than first considering the surface case as was done historically. While Gauss pretty much developed the theory of surfaces, Riemann was the first to consider higher dimensional manifolds, which moreover don't necessarily lie as submanifolds in Euclidean space. The theory is not much harder for abstract manifolds than for surfaces, but there were some crucial global considerations about completeness of metrics and extendability of geodesics that weren't developed until the late 1920s. Thus many results were on hold, so to speak, until then. In the last subsection on curvature we shall see why it was so hard to generalize even the local links between curvature and the metric itself to higher dimensions. It wasn't until the beginning of the twentieth century that people began to understand some of these relationships.

3.1. **First-order concepts.** We now need to clarify what type of objects we wish to work with. The basic objects are manifolds without boundary. We call them closed if they are compact; otherwise they are said to be open. A Riemannian metric on a manifold is a positive definite symmetric $(0,2)$-tensor g. In other words, each tangent space is endowed with a Euclidean metric that varies smoothly from tangent space to tangent space. Riemannian manifolds are denoted (M,g) if we wish to specify both the manifold M and the metric g.

The most natural type of Riemannian manifold, aside from Euclidean space itself, is a surface in $\mathbb{R}^3$. The Riemannian metric is simply the Euclidean metric restricted to the surface. It is now possible to define all of the natural concepts that we use in Euclidean spaces for surfaces. Namely, compute them in $\mathbb{R}^3$ and project them down to the surface. For instance, if we have a curve on the surface, then its tangent vector field is clearly tangent to the surface, but its acceleration may not be. The acceleration on the surface is then the ambient acceleration projected down to the surface. Thus a curve has zero acceleration on the surface if the acceleration vector is perpendicular to the surface. Such curves are called geodesics. In an abstract Riemannian manifold this construction immediately runs into trouble, as we don't have an ambient space where we know how to compute the acceleration. Riemann, who was the first to work with abstract Riemannian manifolds in 1854, solved this problem in a very ingenious manner (see [121]). He observed that around each point $p \in M$ one can pick a *special* coordinate system $x^1, \ldots, x^n$ such that

$$\begin{aligned} g_{ij}(p) &= g(\partial_i, \partial_j) = g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = \delta_{ij}, \\ \partial_k g_{ij}(p) &= 0. \end{aligned}$$

Thus the metric, at the specified point, in these coordinates looks like the Euclidean metric on $\mathbb{R}^n$ up to first-order. Sometimes such coordinates are said to be normal at p. However, we reserve that term for coordinates that have some further properties (see Section 3.3). Again, it is important to realize that these conditions only hold at p. When passing to different points it is necessary to pick different coordinates. If a curve γ passes through p, say, $\gamma(0) = p$, then the acceleration at 0 is simply defined by first writing the curve out in our special coordinates

$$\gamma(t) = \left(\gamma^1(t), \ldots, \gamma^n(t)\right),$$

and first observing that the tangent field is

$$\dot{\gamma} = \sum_{i=1}^{n} \dot{\gamma}^i(t) \cdot \partial_i.$$

We then define

$$\ddot{\gamma}(0) = \sum_{i=1}^{n} \ddot{\gamma}^i(0) \cdot \partial_i.$$

The real idea behind all this is that we have a *connection*. This concept wasn't developed until much later and probably wasn't completely understood until Levi-Civita's work on parallel transport from the beginning of the twentieth century (see [85]). In essence, it is a way of taking derivatives of vector fields. For a function f on a manifold and a vector v we always have the directional derivative

$$D_v f = \nabla_v f = df\,(v)\,.$$

But there is no natural definition for $\nabla_v X$, where X is a vector field, unless one also has a Riemannian metric. Note that this is really what one wants. Given the tangent field $\dot{\gamma} = \sum_{i=1}^{n} \dot{\gamma}^i\,(t) \cdot \partial_i$, the acceleration can then be computed by using a Leibniz rule on the right-hand side, if we can make sense of the derivative of ∂_i in the direction of $\dot{\gamma}$. This is exactly what the covariant derivative $\nabla_v X$ does for us. If $v \in T_pM$ lies in the tangent space to a point where we have chosen the special coordinates discussed above, then we can simply write $X = \sum \alpha^i \partial_i$ and declare

$$\nabla_v X = \sum \left(D_v \alpha^i\right) \partial_i.$$

Since there are several ways of choosing these coordinates, one must of course check that the definition doesn't depend on the choice, but this is very simple to do. Note that for two vector fields we define $(\nabla_Y X)\,(p) = \nabla_{Y(p)} X$. In the end we get a connection ∇ which satisfies:

(1) $Y \to \nabla_Y X$ is tensorial, i.e. linear and $\nabla_{fY} X = f \nabla_Y X$ for all functions f.

(2) $X \to \nabla_v X$ is linear.

(3) $\nabla_v\,(fX) = (\nabla_v f)\,X\,(p) + f\,(p)\,\nabla_v X$.

(4) $\nabla_X Y - \nabla_Y X = [X, Y]\,.$

(5) $D_v g\,(X, Y) = g\,(\nabla_v X, Y) + g\,(X, \nabla_v Y)\,.$

So, no matter which coordinates we use we can now define the acceleration of a curve in the following way:

$$\begin{aligned}
\gamma\,(t) &= \left(\gamma^1\,(t)\,, \ldots, \gamma^n\,(t)\right), \\
\dot{\gamma}\,(t) &= \sum \dot{\gamma}^i\,(t)\,\partial_i, \\
\ddot{\gamma}\,(t) &= \sum \ddot{\gamma}^i\,(t)\,\partial_i + \dot{\gamma}^i\,(t)\,\nabla_{\dot{\gamma}(t)} \partial_i.
\end{aligned}$$

We call γ a *geodesic* if $\ddot{\gamma} = 0$. This is a second order nonlinear ODE in a fixed coordinate system. Thus we see that given any tangent vector $v \in T_pM$, there is a unique geodesic $\gamma_v\,(t)$ with $\dot{\gamma}_v\,(0) = v$. If the manifold is closed, the geodesic must exist for all time, but in case the manifold is open this might not be so. To see this, simply take as M any open subset of Euclidean space with the induced metric.

Given an arbitrary vector field $E(t)$ along γ, i.e., $E(t) \in T_{\gamma(t)}M$ for all t, we can also define the derivative $\dot{E} = \frac{d}{dt}E$ of E in the direction of $\dot{\gamma}$ by writing

$$\begin{aligned} E(t) &= \sum \alpha^i(t)\,\partial_i, \\ \dot{E}(t) &= \sum \dot{\alpha}^i(t)\,\partial_i + \alpha^i(t)\,\nabla_{\dot{\gamma}(t)}\partial_i. \end{aligned}$$

Note that the derivative of the tangent field $\dot{\gamma}$ is simply the acceleration $\ddot{\gamma}$. The field E is said to be *parallel* provided $\dot{E} = 0$. The equation for a field to be parallel is a first order linear ODE, so we see that for any $v \in T_{\gamma(t_0)}M$ there is a unique parallel field $E(t)$ defined on the entire domain of γ with the property that $E(t_0) = v$. Given two such parallel fields E and F, we have that

$$\begin{aligned} \frac{d}{dt}g(E,F) &= D_{\dot{\gamma}}g(E,F) \\ &= g\left(\dot{E},F\right) + g\left(E,\dot{F}\right) \\ &= 0. \end{aligned}$$

Thus E and F are both of constant length and form constant angles along γ. Hence, "parallel translation" along a curve defines an orthogonal transformation between the tangent spaces to the manifold along the curve. However, in contrast to Euclidean space, this parallel translation will depend on the choice of curve. On the two dimensional unit sphere in $\mathbb{R}^3$ one can, for instance, consider a triangle consisting of going from the North Pole along a longitude down to the Equator, then going along the Equator for a while, and finally going back up to the North Pole along another longitude. In this case all three curves are geodesics and their tangent fields are therefore parallel. If we start at the North Pole and parallel translate the tangent field to the first longitude down to the Equator, then we end up with a vector perpendicular to the Equator. And if we parallel translate this along the Equator, the vector must stay perpendicular. Thus when we reach the second longitude, the vector will be tangent to this longitude and therefore be parallel translated back up to the North Pole tangent to it. In this way we can, by choosing the longitudes suitably, parallel translate any vector at the North Pole to any other vector at the North Pole.

The connection also enables us to define many other classical concepts from calculus in the setting of Riemannian manifolds. Suppose we have a function $f : M \to \mathbb{R}$. If the manifold is not equipped with a Riemannian metric, then we have the differential of f defined by $df(v) = D_v f$. This is a 1-form. The dual concept, the gradient of f, is supposed to be a vector field. But we need a metric to define it. Namely, ∇f is defined by the relationship

$$g(\nabla f, v) = df(v).$$

Having defined the gradient of a function on a Riemannian manifold, we can then use the connection to define the Hessian as the linear map

$$\begin{aligned} \nabla^2 f &: TM \to TM, \\ \nabla^2 f(v) &= \nabla_v \nabla f. \end{aligned}$$

The corresponding bilinear map is then defined as

$$\nabla^2 f(v,w) = g\left(\nabla^2 f(v), w\right).$$

One easily checks that this is a symmetric bilinear form. The Laplacian of f, Δf, is now defined as the trace of the Hessian (as a linear map). This is also called the Laplace-Beltrami operator, since Beltrami first considered this operator on Riemannian manifolds. Note that geometers also use the negative of the trace of the Hessian as a definition for the Laplacian. The "minus" convention has the nice property of making the eigenvalues of the Laplace operator nonnegative. However, as we shall not be discussing eigenvalues, we stick to the above definition.

3.2. **Distances and geodesics.** We are now ready to define the distance between two points on a Riemannian manifold and discuss which curves realize the distance.

For a piecewise smooth (or merely absolutely continuous) curve $\gamma : I \to M$, we define its length on I as

$$L(\gamma, I) = \int_I |\dot{\gamma}| \, dt = \int_I \sqrt{g(\dot{\gamma}, \dot{\gamma})} dt.$$

First, one should note that this length is independent of our parametrization of the curve. Thus the curve can be reparametrized, if we like, in such a way that it has unit velocity. Such curves are said to parametrized by arclength. The distance between two points p and q, $d(p,q)$, can now be defined as the infimum of the lengths of all curves from p to q. This means that the distance measures the shortest way one can travel from p to q. It is easy to check that this indeed defines a metric in the usual sense and that the topology it generates is equivalent to the manifold topology.

If we take a variation $V(s,t) : (-\varepsilon, \varepsilon) \times [0, \ell] \to M$ of a smooth curve $c(t) = V(0,t)$ parametrized by arclength and of length ℓ, then the first derivative of the arclength function

$$L(s) = \int_0^\ell \left| \frac{\partial}{\partial t} V \right| dt$$

is

$$\frac{dL}{ds}(0) = g(\dot{c}, X)|_0^\ell - \int_0^\ell g(\ddot{c}, X) \, dt,$$

where $X(t) = \frac{\partial V}{\partial s}(0,t)$ is the so-called *variational vector field*. This formula is called the *first variation formula*. Given any vector field X along c, one can easily produce a variation whose variational field is X. If the variation fixes the endpoints, $X(a) = X(b) = 0$, then the second term in the formula drops out, and we note that the length of c can always be decreased as long as the acceleration of c is not everywhere zero. Thus the Euler-Lagrange equation for the arclength functional is simply the equation for a curve to be a geodesic.

A curve between two points is called a *segment* if it has length equal to the distance between the points and is parametrized by arclength. Note that a segment when restricted to a subinterval is again a segment. By the first variation formula we therefore see that on each of the intervals where the curve is smooth it must be a geodesic. To see that the curve is smooth everywhere, we can just assume that it is not smooth at just one point. We then pick a variation that forms angles that are strictly less than $\pi/2$ to the two velocity vectors for the segment at this point. We can then add up the two contributions of the first variation formula coming from the variations of each of the two smooth parts and observe that the derivative is negative due to the angle condition.

We now come to the reverse question of when a geodesic is a segment. In Euclidean space this is always the case, but on a unit sphere only geodesics of length $\leq \pi$ can be segments. To get the full answer it is convenient to introduce *exponential coordinates* on a Riemannian manifold. This is done as follows: We first define the exponential map

$$\begin{aligned} \exp_p &: O_p \subset T_pM \to M, \\ \exp_p(v) &= \gamma_v(1). \end{aligned}$$

Thus the line $t \to tv$ in T_pM is mapped to the geodesic whose velocity at p is v. In case (M, g) is not geodesically complete this map might only be defined on an open neighborhood O_p of the origin in T_pM. Since the exponential map takes lines through the origin to geodesics through p and preserves their velocity at p, we see that the differential at the origin is the identity map. Thus the exponential map is a diffeomorphism near the origin and therefore introduces what we call exponential coordinates near $p \in M$. It is a remarkable fact, first observed by Gauss, that for points x which lie in this coordinate chart near p we have

$$d(p, x) = \left|\exp_p^{-1}(x)\right|.$$

This is proved using the *Gauss lemma*, which says that the gradient of the function $x \to \left|\exp_p^{-1}(x)\right|$ is simply the unit radial vector field. In other words, in exponential coordinates the gradient of distance from p on T_pM is the same whether we use the Euclidean metric on T_pM or the pullback of the metric g on M. From this it evidently follows that the distance function $x \to d(p, x)$ is smooth near p, that the integral curves for the gradient are exactly the geodesics emanating from p, and that these curves are segments. The Gauss lemma for surfaces is an immediate consequence of the following four facts. We assume that polar coordinates are chosen on the tangent space and then transported to the manifold via the exponential map.

1. The integral curves for ∂_r are geodesics, i.e., $\nabla_{\partial_r}\partial_r = 0$,
2. $[\partial_r, \partial_\theta] = \nabla_{\partial_r}\partial_\theta - \nabla_{\partial_\theta}\partial_r = 0$,
3. $2\partial_r g(\partial_r, \partial_\theta) = 2g(\partial_r, \nabla_{\partial_r}\partial_\theta) = 2g(\partial_r, \nabla_{\partial_\theta}\partial_r) = \partial_\theta g(\partial_r, \partial_r) = 0$, and
4. $|g(\partial_r, \partial_\theta)| \leq |\partial_\theta| = \phi \to 0$ as $r \to 0$.

Note that 3 and 4 show that the coordinate vector fields are perpendicular.

The *segment domain* $\mathrm{seg}(p) \subset O_p$ is the star-shaped set with the property that $v \in \mathrm{seg}(p)$ if the geodesic $t \to \exp_p(tv)$ is a segment from p to $\exp_p(v)$. The interior of this set is nonempty and contains a neighborhood of the origin. The boundary of the segment domain is also known as the *cut locus*. The exponential map is actually an embedding on the interior of $\mathrm{seg}(p)$. The alternative would be that either the map is not one-to-one or has singular differential somewhere. In the first situation we would then have a segment $c : [0, 1+\varepsilon] \to M$ with the property that there is another segment joining $c(0)$ and $c(1)$. But then we could find a nonsmooth segment from $c(0)$ to $c(1+\varepsilon)$, which we know to be impossible. In the second situation we have a segment $c(t) = \exp_p(tv) : [0, 1+\varepsilon] \to M$ such that $D\exp_p$ is singular at v. We can then find $w \in T_pM$ such that the curve $s \to \exp_p((v + sw))$ has zero velocity at $s = 0$. Then consider the variation $V(s, t) = \exp_p(t(v + sw))$. The curves $t \to \exp_p(t(v + sw))$ are segments for small s. These curves may not reach $c(1) = \exp_p(v)$ as in the above situation, but they will reach $c(1)$ up to first order. Now suppose ε is so small that $c(1)$ lies in a neighborhood around $c(1+\varepsilon)$

on which we have exponential coordinates. Then one can check that the curve which starts out being $t \to \exp_p(t(v+sw))$ on $[0,1-\delta]$ and then ends up being a little segment ending up at $c(1+\varepsilon)$ is indeed shorter than c from $c(0)$ to $c(1+\varepsilon)$ for small ε. Thus we have again arrived at a contradiction.

At this point it is traditional to introduce some more notation. Consider a geodesic $c(t) = \exp_p(tv) : [0,b] \to M$. We say that c has a *conjugate point* at $t_0 \in (0,b]$ if $D\exp_p$ is singular at $t_0 v$. In the next subsection we shall see how conjugate points are controlled by the geometry. In analogy with the segment domain we can define the *interior of the conjugate locus* as the largest star-shaped set $\mathrm{conj}_p \subset O_p$ such that v lies in the interior of conj_p if $\exp_p$ has nonsingular differential at all the points tv, $t \in [0,1]$. Clearly the exponential map is an immersion on the interior of conj_p, and we can therefore pull back the Riemannian metric on M to this open set. In this way we can create a multivalued exponential coordinate system.

Note that the segment domain and the interior of the conjugate locus are not generally the same set. On a flat torus there are no conjugate points as we shall see, but the segment domain has to be bounded.

3.3. **Second order concepts.** We shall now see how curvature enters the picture. The curvature tensor is a rather ominous tensor of type (1,3); i.e., it has three vector variables and its value is a vector as well. However, we shall soon enough see how it appears rather naturally in geometry when we take two derivatives rather than just one.

First recall that the Lie bracket of two vector fields X and Y can be defined implicitly as the vector field $[X,Y]$ which satisfies

$$D_{[X,Y]}f = [D_X, D_Y] f = (D_X D_Y - D_Y D_X) f$$

for all functions f. In other words $D_{[X,Y]} - [D_X, D_Y] = 0$ on functions. If we compute the same thing on vector fields, however, then we get the following quantity:

$$\begin{aligned} R(X,Y)Z &= \left(\nabla_{[X,Y]} - [\nabla_X, \nabla_Y]\right) Z \\ &= \nabla_{[X,Y]} Z - \nabla_X \nabla_Y Z + \nabla_Y \nabla_X Z. \end{aligned}$$

This turns out to be a vector valued (1,3)-tensor in the three variables X, Y, Z. We can then create a (0,4)-tensor with scalar values as follows

$$R(X,Y,Z,W) = g\left(\nabla_{[X,Y]} Z - \nabla_X \nabla_Y Z + \nabla_Y \nabla_X Z, W\right).$$

One can easily see that this tensor is skew-symmetric in X and Y, and also in Z and W. This was already known to Riemann, but there are some further, more subtle properties that were discovered a little later by Bianchi. They play a very important role in geometry, but will not be discussed here, except that they yield the important symmetry condition

$$R(X,Y,Z,W) = R(Z,W,X,Y).$$

Thus the curvature tensor can be thought of as a symmetric operator

$$\mathfrak{R} : \Lambda^2 TM \to \Lambda^2 TM$$

also known as the *curvature operator*.

The *Ricci tensor* is the (1,1)- or (0,2)-tensor defined by

$$\begin{aligned}\operatorname{Ric}(X) &= \sum_{i=1}^{n} R(E_i, X) E_i, \\ \operatorname{Ric}(X,Y) &= \sum_{i=1}^{n} g(R(E_i, X) E_i, Y)\end{aligned}$$

for any orthonormal basis E_i. In other words, the Ricci curvature is simply a trace of the curvature tensor. Similarly one can define the *scalar curvature* as the trace

$$\begin{aligned}\operatorname{scal}(p) &= \operatorname{tr}(\operatorname{Ric}) \\ &= \sum_{i=1}^{n} \operatorname{Ric}(E_i, E_i).\end{aligned}$$

When the Riemannian manifold has dimension 2, all of these curvatures are essentially the same. Since $\dim \Lambda^2 TM = 1$ and is spanned by $X \wedge Y$ where X and Y form an orthonormal basis for T_pM, we see that the curvature tensor depends only on the value

$$K(p) = R(X, Y, X, Y),$$

which also turns out to be the Gauss curvature. The Ricci tensor is a homothety

$$\begin{aligned}\operatorname{Ric}(X) &= K(p) X, \\ \operatorname{Ric}(Y) &= K(p) Y,\end{aligned}$$

and the scalar curvature is twice the Gauss curvature. In dimension 3 there are also some redundancies as $\dim TM = \dim \Lambda^2 TM = 3$. In particular, the Ricci tensor and the curvature tensor contain the same amount of information.

The *sectional curvature* is a kind of generalization of the Gauss curvature whose importance Riemann was already aware of. Given a 2-plane $\pi \subset T_pM$ spanned by an orthonormal basis X, Y it is defined as

$$\sec(\pi) = R(X, Y, X, Y).$$

The remarkable observation by Riemann was that the curvature operator is a homothety, i.e., looks like $\mathfrak{R} = kI$ on $\Lambda^2 T_pM$ iff all sectional curvatures of planes in T_pM are equal to k. This result is not completely trivial, as the sectional curvature is not the entire quadratic form associated to the symmetric operator $\mathfrak{R}$. In fact, it is not true that $\sec \geq 0$ implies that the curvature operator is nonnegative in the sense that all its eigenvalues are nonnegative. What Riemann did was to show that our special coordinates at p can be chosen to be *normal* at p, i.e., satisfy the much stronger condition

$$x^i = \sum_{j=1}^{n} g_{ij} x^j$$

on a neighborhood of p. One can easily show that such coordinates are actually exponential coordinates together with a choice of an orthonormal basis for T_pM so as to identify T_pM with $\mathbb{R}^n$. In these coordinates one can then expand the metric as follows:

$$g_{ij} = \delta_{ij} - \frac{1}{3} \sum_{k,l=1}^{n} R_{ikjl} x^k x^l + O(r^3).$$

Now the equations $\sum_{j=1}^{n} g_{ij}x^j = x^i$ evidently give conditions on the curvatures R_{ijkl} at p. In fact, one can get all of the relevant symmetry conditions needed to prove the constant curvature result just mentioned (including the first Bianchi identity). In dimension two this expansion reduces to a formula known to Gauss:

$$\left(\begin{array}{cc} g_{xx} & g_{xy} \\ g_{xy} & g_{yy} \end{array}\right) = \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right) - \frac{1}{3}K(p)\left(\begin{array}{cc} x^2 & -xy \\ -xy & y^2 \end{array}\right) + O\left(r^3\right).$$

These Taylor formulae for the metric, while telling us something about the infinitesimal geometry and curvature, are not strong enough to give us more global information. Instead we must try to understand how the metric varies over a larger region depending on the behavior of the curvature. We shall, as was done historically, begin by considering 2-dimensional Riemannian manifolds. The higher dimensional case is much more subtle and took a long time to be understood completely.

On $\left(M^2, g\right)$ we fix a point p and with the help of the exponential map $\exp_p : \mathrm{int}(\mathrm{seg}_p) \to M$ introduce polar coordinates. The polar coordinates (r, θ) are constructed on T_pM in the usual fashion and then used on the image of $\mathrm{int}(\mathrm{seg}_p)$ via the exponential map. The Gauss lemma from above asserts that the coordinate vector fields ∂_r and ∂_θ are perpendicular on (M, g) as well as in the Euclidean metric on T_pM; moreover, ∂_r has length 1 on both M and T_pM. The Riemannian metric g can now be written as

$$g = dr^2 + \phi^2(r, \theta)\, d\theta^2$$

on $\exp_p(\mathrm{int}(\mathrm{seg}_p))$. The meaning of this is simply that if we expand $v, w \in T_pM$ in the given coordinate vector fields

$$\begin{aligned} v &= v^r\partial_r + v^\theta\partial_\theta, \\ w &= w^r\partial_r + w^\theta\partial_\theta, \end{aligned}$$

then

$$g(v, w) = v^r w^r + \phi^2(r, \theta)\, v^\theta w^\theta, \text{ for } p = (r, \theta).$$

Notice that even though polar coordinates are defined only away from a half-line, the coordinate vector fields are defined everywhere away from the origin. Thus ∂_r and $\phi^{-1}\partial_\theta$ always form an orthonormal basis. Let us now compute the curvature

$$\begin{aligned} R(\partial_r, \partial_\theta)\partial_r &= \nabla_{[\partial_r, \partial_\theta]}\partial_r - \nabla_{\partial_r}\nabla_{\partial_\theta}\partial_r + \nabla_{\partial_\theta}\nabla_{\partial_r}\partial_r \\ &= -\nabla_{\partial_r}\nabla_{\partial_\theta}\partial_r. \end{aligned}$$

Evidently we must figure out what $\nabla_{\partial_r}\partial_\theta = \nabla_{\partial_\theta}\partial_r$ is. This is done as follows:

$$\begin{aligned} 0 &= \partial_r g(\partial_r, \partial_\theta) = g(\partial_r, \nabla_{\partial_r}\partial_\theta), \\ \phi\partial_r\phi &= \frac{1}{2}\partial_r g(\partial_\theta, \partial_\theta) = g(\partial_\theta, \nabla_{\partial_r}\partial_\theta); \end{aligned}$$

hence

$$\nabla_{\partial_\theta}\partial_r = \nabla_{\partial_r}\partial_\theta = \frac{\partial_r\phi}{\phi}\partial_\theta.$$

Using this in the above expression we obtain

$$\begin{aligned} R(\partial_r,\partial_\theta)\partial_r &= -\nabla_{\partial_r}\nabla_{\partial_\theta}\partial_r \\ &= -\nabla_{\partial_r}\left(\frac{\partial_r\phi}{\phi}\partial_\theta\right) \\ &= -\partial_r\left(\frac{\partial_r\phi}{\phi}\right)\partial_\theta - \frac{\partial_r\phi}{\phi}\nabla_{\partial_r}\partial_\theta \\ &= -\partial_r\left(\frac{\partial_r\phi}{\phi}\right)\partial_\theta - \left(\frac{\partial_r\phi}{\phi}\right)^2\partial_\theta \\ &= -\frac{\partial_r^2\phi}{\phi}\partial_\theta. \end{aligned}$$

After taking inner products with ∂_θ, we obtain the *Jacobi equation*

$$\partial_r^2\phi + \sec(T_pM)\,\phi = 0.$$

This equation was first discovered by Gauss in the same way we just derived it. Later Jacobi extended it to a type of equation that holds for so-called Jacobi fields along geodesics of any length on a surface. We won't need Jacobi's extension or Jacobi fields here, so we shall leave the equation as it is (for a complete discussion see [84]). The important point is that it gives a direct relationship between curvature and the metric in the form of ϕ.

One can even extend the Jacobi equation slightly. Namely, we can via

$$\exp_p : \operatorname{int}(\operatorname{conj}_p) \to M$$

pullback the metric g on the image of $\operatorname{int}(\operatorname{conj}_p)$ to $\operatorname{int}(\operatorname{conj}_p)$ itself. Thus on $\operatorname{int}(\operatorname{conj}_p)$ we can, as above, decompose the metric g in polar coordinates and derive the above Jacobi equation on a domain that might be larger than the manifold itself.

Using this, we can now explain how Bonnet and von Mangoldt proved their results. The important thing to keep in mind is that at the time they were not aware of any precise conditions that ensured the existence of a segment between any two points on abstract Riemannian manifolds. This is a problem that was settled much later and which we shall explain in the next section. For surfaces in $\mathbb{R}^3$, however, they did know that provided the surface was a closed subset of Euclidean space, one would get that any two points could be joined by a segment. The importance of this property lies in the fact that it implies that $\exp_p : \operatorname{seg}_p \to M$ is onto and that the image of $\operatorname{int}(\operatorname{seg}_p)$ is dense in M. To show Bonnet's diameter bound in case the sectional curvature satisfies $\sec \geq k^2 > 0$ for all tangent planes, it therefore suffices to show that $\operatorname{conj}_p \subset \bar{B}(0,\pi/k) \subset T_pM$. (Recall that a geodesic emanating from p cannot minimize after a conjugate point.) Thus we must find a way of detecting conjugate points. If we pull back the metric on M to $\operatorname{int}(\operatorname{conj}_p)$ and write it as $g = dr^2 + \phi^2(r,\theta)\,d\theta^2$, then we observe that for fixed θ we get to a conjugate point along the geodesic $r \to (r,\theta)$ when $\phi(r,\theta) \to 0$ as r approaches the boundary of $\operatorname{int}(\operatorname{conj}_p)$. In other words, the differential of the exponential map becomes singular on the boundary of $\operatorname{int}(\operatorname{conj}_p)$ precisely because the pullback of g to T_pM becomes a degenerate inner product. Now we have that

$$\begin{aligned} \partial_r^2\phi + \sec\cdot\phi &= 0, \\ \phi(0,\theta) &= 0, \\ \sec &\geq k^2 > 0. \end{aligned}$$

Therefore Sturm-Liouville comparison theory tells us that $r \to \phi(r,\theta)$ must become zero before any nontrivial solution to

$$\begin{aligned} y'' + k^2 y &= 0, \\ y(0) &= 0. \end{aligned}$$

Since all nontrivial solutions to this equation look like $y(r) = c\sin(k \cdot r)$, we see that ϕ must become zero before r gets to the value π/k. From all this we can conclude that any segment in (M,g) has length $\leq \pi/k$ and therefore that the diameter of (M,g) must be $\leq \pi/k$.

In order to prove von Mangoldt's theorem, it is necessary to show that there are no conjugate points. In this situation we have on $\mathrm{int}(\mathrm{conj}_p)$ that ϕ satisfies

$$\begin{aligned} \partial_r^2 \phi + \sec \cdot \phi &= 0, \\ \phi(0,\theta) &= 0, \\ \sec &\leq 0. \end{aligned}$$

This time Sturm-Liouville theory then tells us that ϕ cannot become zero before the nontrivial solution to

$$\begin{aligned} y'' &= 0, \\ y(0) &= 0. \end{aligned}$$

But here the solutions are $y(r) = cr$, so they are never zero. Hence ϕ will never become zero either. Thus $\mathrm{int}(\mathrm{conj}_p) = T_pM$, provided we know that geodesics exist for all time. One can then show, and this requires a little argument as well, that the exponential map is a covering map.

Note that in addition to the condition that $\phi(r,\theta) \to 0$ as $r \to 0$, we also know that $\partial_r \phi(r,\theta) \to 1$ as $r \to 0$. To see this just observe that the polar coordinate representation of the Euclidean metric has $\phi = r$ and that near the origin g_{ij} and the Euclidean metric agree up to first order in exponential coordinates as discussed above. In other words, exponential coordinates are normal in the sense discussed above. Refined Sturm-Liouville theory can therefore be used to show that in the case $\sec \geq k^2$ we have $\phi \leq \sin(kr)$, while if $\sec \leq 0$ we have $\phi \geq r$. The latter condition in particular shows that not only is the exponential map nonsingular, but its differential actually increases the length of vectors.

These observations can also be used to demonstrate that spaces of constant sectional curvature k are locally isometric. Namely, ϕ must satisfy

$$\begin{aligned} \partial_r^2 \phi + k \cdot \phi &= 0, \\ \phi(0,\theta) &= 0, \\ \partial_r \phi(0,\theta) &= 1. \end{aligned}$$

Thus it must be given by the single formula

$$\phi(r,\theta) = \mathrm{sn}_k(r),$$

where sn_k is the unique solution to

$$\begin{aligned} y'' + ky &= 0, \\ y(0) &= 0, \\ y'(0) &= 1. \end{aligned}$$

This also tells us that we can construct abstract geometries of any constant curvature. Riemann in 1854 extended this to higher dimensions and showed that for each $k \in \mathbb{R}$ and integer $n \geq 2$ there is a simply connected space S_k^n of dimension n and constant curvature k. This space is Euclidean space when $k = 0$, the Euclidean sphere of radius $1/\sqrt{k}$ when $k > 0$, and hyperbolic space for $k < 0$.

The problem with generalizing the above relationship between curvature and metric to higher dimensions is that we don't have a suitable choice of polar coordinates. This was solved in two different ways. In 1917 Levi-Civita (see [85]) proposed that one should simply fix a geodesic and then select polar coordinates such that the angular coordinates were orthonormal along the geodesic. It is not hard to see that this is possible using the parallel translation developed by Levi-Civita himself. Around the same time, Cartan (see [22]) came up with a more elegant solution. Instead of using angular coordinates, he observed that it suffices to have suitable angular vector fields E_α which are independent of r, perpendicular to the radial coordinate, and orthonormal on the unit sphere in the tangent space. Of course, these fields are only globally defined if the unit sphere is parallelizable, but they always exist locally. This is in contrast to the coordinates by Levi-Civita which change from point to point on the sphere. If we denote the 1-forms dual to E_α by θ^α, we see that the Euclidean metric can be written as

$$dr^2 + r^2\left(\sum_{\alpha=2}^{n} (\theta^\alpha)^2\right).$$

More generally, if we pull back the Riemannian metric on M via the exponential map and write it out in this way, we get a formula which looks like

$$\begin{aligned} g &= dr^2 + \sum_{\alpha,\beta=2}^{n} g_{\alpha\beta}\theta^\alpha\theta^\beta, \\ g_{\alpha\beta} &= g\left(E_\alpha, E_\beta\right), \\ g_{\alpha\beta} &= \begin{pmatrix} r^2 & 0 & \cdots & 0 \\ 0 & & & \vdots \\ \vdots & & & 0 \\ 0 & \cdots & 0 & r^2 \end{pmatrix} + O\left(r^2\right). \end{aligned}$$

The problem now lies in finding a Jacobi-type equation. This time the function ϕ^2 is replaced by the matrix $(g_{\alpha\beta})$. Since the Jacobi equation is really for ϕ, one would expect that the square root of $(g_{\alpha\beta})$ is the important quantity. However, the fact that all of these matrices don't commute for different values of r makes it very hard to work with derivatives of this square root. Several different approaches have therefore been developed. One uses Jacobi fields (these are variational fields of geodesic variations). Another is a very natural direct approach which Cartan used. He developed a much more general theory of how to set up information about the connection and curvature using frames, which are in fact more general than coordinates. When specializing this theory to frames that are naturally associated with polar coordinates and only considering radial derivatives, one obtains the following information: Consider the Hessian $S = \nabla^2 r$ of the distance function $r(x) = d(x, p)$. In invariant notation we then have the two fundamental equations, known respectively as the *first* and *second fundamental equations*,

$$\begin{aligned} (L_{\partial_r} g)(X,Y) &= 2g(S(X),Y), \\ (L_{\partial_r} S)(X) + S^2(X) &= -R(\partial_r, X)\partial_r, \end{aligned}$$

where L_{∂_r} is the Lie derivative and X, Y any vector fields. In the adapted frames we can reduce these equations as follows: First note that $S(\partial_r) = 0$, so we need only look at the matrix for S in terms of the frame E_α. Thus we consider the matrix S^α_β defined by

$$S(E_\beta) = \sum_{\alpha=2}^{n} S^\alpha_\beta E_\alpha.$$

Then the fundamental equations become

$$\begin{aligned} \partial_r(g_{\alpha\beta}) &= 2(S^\gamma_\alpha)(g_{\gamma\beta}), \\ \partial_r(S^\alpha_\beta) + (S^\alpha_\gamma)\left(S^\gamma_\beta\right) &= -\left(R^\alpha_\beta\right), \end{aligned}$$

where the matrix $\left(R^\alpha_\beta\right)$ is the curvature defined by $R(\partial_r, E_\beta)\partial_r = \sum R^\alpha_\beta E_\alpha$. (Note to the expert: These are merely the first and second structural equations of Cartan restricted to radial derivatives.) Thus the metric is related through a linear first order equation to S, and S in turn is related to the curvature through a nonlinear first order equation that is of Riccati type. Through these equations it is now possible, with some extra work, to go from curvature to information about S and finally to information about the metric.

The first test case is that one can use these equations to prove Riemann's theorem from 1854 on the local characterization of manifolds with constant curvature k. Namely, that for fixed k they must all be locally isometric to each other. One can also generalize the Bonnet and von Mangoldt theorems. This will be discussed in the next section.

4. From local to global geometry

The dominant new development from the period 1925-1950 seems now to have been the resolution to the problem of representing the Euler characteristic as an integral of curvature. But there were two other important discoveries. One was the introduction of the second variation formula for the arclength functional. The other was the introduction of a new method of proving many results, now known as the Bochner technique. The techniques and types of theorems developed in this period still form the foundation for many developments today.

4.1. **The second variation formula.**

The modern history of global Riemannian geometry starts with Hopf's classification of space forms in [72] and Cartan's generalization of von Mangoldt's theorem to higher dimensions, together with some other observations about manifolds with nonpositive curvature. At the time some of the foundations of Riemannian geometry were still not completely settled. Bianchi, Levi-Civita, and Ricci had completely developed tensor calculus (what Cartan refers to as the debauch of indices) and had a complete understanding of the relationships between the concepts of connection, parallel translation, and curvature. However, there were still several competing notions of what a suitable global Riemannian manifold should be. Clearly one cannot hope to establish global theorems without also assuming something further about the manifold. For instance, all open subsets of Euclidean space are flat, but their topology can be very complicated. So even for flat manifolds one can't say anything intelligent. Given a Riemannian n-manifold (M^n, g), there were three natural assumptions one could make:

1. that there exists a fixed $r > 0$ such that $\exp_p : B(0,r) \to B(p,r)$ is a diffeomorphism for all $p \in M$,
2. that every bounded infinite set in (M, d) has an accumulation point,
3. that every geodesic exists for all time.

Note that for a closed manifold all three conditions are easily seen to hold, and this was certainly well-known at the time. Thus it is only for open manifolds that one has to worry about these issues. Also one readily sees that $(1) \Rightarrow (2) \Rightarrow (3)$. The first condition was used by Killing in his definition of an appropriate class of constant curvature spaces. Given this condition Killing showed in the late nineteenth century that if (M^n, g) has constant curvature k, then the universal covering is isometric to S^n_k. In today's language condition (1) says that the manifold has a positive lower bound for its *injectivity radius*. Hopf in [72] observed that if an open manifold has finite volume, then it will never satisfy (1), but it might satisfy (3). Thus he settled on this condition as being the correct class to work with. Condition (3) is now referred to as *geodesic completeness*. His result then states that any geodesically complete Riemannian manifold of constant curvature k is covered by S^n_k.

Condition (2) was used by Cartan in [22]. He called such Riemannian manifolds *normal*. Now they are called *complete* or metrically complete. He classified complete constant curvature spaces at the same time that Hopf published his classification assuming "only" geodesic completeness.

In [22] Cartan also proved that complete manifolds with nonpositive curvature have the property that any two points can be joined by a segment. His argument uses that in exponential coordinates the metric is larger that the corresponding Euclidean metric. This is also the fact that implies that the exponential map must have full rank everywhere. He then argued that if one pulls back the metric on M to T_pM via the exponential map, then one gets a new complete metric of nonpositive curvature which is locally isometric to M. In this way we have found the universal covering of M to be Euclidean space. Note that it is important to have that any two points are joined by a geodesic, for otherwise this map can't be a covering map. The proof of all this was explained in detail in section 3.3 for the case of surfaces, and we have also indicated the necessary changes that Cartan made in order to generalize this to higher dimensions. In later editions of Cartan's book, he also includes a proof that in fact on *any* complete Riemannian manifold, any two points can be joined by a segment. After having generalized van Mangoldt's theorem, Cartan also establishes another important property of manifolds with nonpositive curvature. First he observes that all spaces of constant zero curvature have torsion-free fundamental groups. This is because any isometry of finite order on Euclidean space must have a fixed point (the center of mass of any orbit is necessarily a fixed point). Then he notices that one can geometrically describe the L^∞ center of mass of finitely many points $\{p_1, \ldots, p_k\}$ in Euclidean space as the unique minimum for the strictly convex function

$$x \to \max_{i=1,\cdots,k} \frac{1}{2}\left\{(d(p_i, x))^2\right\}.$$

In other words, the center of mass is the center of the ball of smallest radius containing $\{p_1, \ldots, p_k\}$. Now Cartan's observation from above was that the exponential map is expanding and globally distance nondecreasing as a map:

$$(T_pM, \text{ Euclidean metric}) \to (T_pM, \text{ with pullback metric}).$$

Thus distance functions are convex in nonpositive curvature as well as in Euclidean space. Hence the above argument can in fact be used to conclude that any Riemannian manifold of nonpositive curvature must also have torsion free fundamental group.

The next important development, which occurred almost simultaneously but was virtually ignored at the time, was Synge's paper [130]. In this paper he develops the first and second variation formulae for arclength. Various versions of these formulae were already known, but Synge gave the version of the second variation formula that we now know and use (there are several, and not all of them are equally transparent). Moreover, he for the first time finds estimates for conjugate points on abstract Riemannian manifolds, thus generalizing Bonnet's work on surfaces. However, he seems not to have pointed out that for complete manifolds this gives diameter bounds. Later in [131], after Myers in [98] published his generalization of Bonnet's diameter bound, Synge then points out that this is indeed a trivial conclusion and also berates Schoenberg for his paper [125], where Schoenberg obtains conjugate point estimates similar to those in Synge's 1925 paper. What is even more interesting is that Synge correctly points out that it is in fact like shooting flies with cannon balls to get these diameter bounds from conjugate point estimates, for one can simply show directly from the second variation formula that geodesics that are too long can't minimize length.

In order to get a better picture of what is happening here, let us set up the *second variation formula* and explain how it is used. We have already seen the first variation formula and how it can be used to characterize geodesics. Now suppose that we have a unit speed geodesic $c(t)$ parametrized on $[0, \ell]$ and consider a variation $V(s,t)$, where $V(0,t) = c(t)$. Synge then shows that

$$\frac{d^2L}{ds^2}(0) = \int_0^\ell \left\{ g\left(\dot{X}, \dot{X}\right) - \left(g\left(\dot{X}, \dot{c}\right)\right)^2 - g\left(R(X, \dot{c})X, \dot{c}\right) \right\} dt + g(\dot{c}, A)|_0^\ell,$$

where $X(t) = \frac{\partial V}{\partial s}(0,t)$ is the variational vector field, $\dot{X} = \nabla_{\dot{c}} X$, and $A(t) = \nabla_{\frac{\partial V}{\partial s}} X$. In the special case where the variation fixes the endpoints, i.e., $s \to V(s,a)$ and $s \to V(s,b)$ are constant, the term with A in it falls out. We can also assume that the variation is perpendicular to the geodesic and then drop the term $g\left(\dot{X}, \dot{c}\right)$. Thus, we arrive at the following simple form:

$$\begin{aligned} \frac{d^2L}{ds^2}(0) &= \int_0^\ell \left\{ g\left(\dot{X}, \dot{X}\right) - g\left(R(X, \dot{c})X, \dot{c}\right) \right\} dt \\ &= \int_0^\ell \left\{ \left|\dot{X}\right|^2 - \sec(\dot{c}, X)\,|X|^2 \right\} dt. \end{aligned}$$

Therefore, if the sectional curvature is nonpositive, we immediately observe that any geodesic locally minimizes length (that is, among close-by curves), even if it doesn't minimize globally (for instance c could be a closed geodesic). On the other hand, in positive curvature we can see that if a geodesic is too long, then it cannot minimize even locally. The motivation for this result comes from the unit sphere, where we can consider geodesics of length $> \pi$. Globally, we of course know that it would be shorter to go in the opposite direction. However, if we consider a variation of c where the variational field looks like $X = \sin\left(t \cdot \frac{\pi}{\ell}\right) E$ and E is a unit length parallel field along c which is also perpendicular to c, then we get

$$\begin{aligned}
\frac{d^2 L}{ds^2}(0) &= \int_0^\ell \left\{ \left|\dot{X}\right|^2 - \sec(\dot{c}, X)\left|X\right|^2 \right\} dt \\
&= \int_0^\ell \left\{ \left(\frac{\pi}{\ell}\right)^2 \cdot \cos^2\left(t \cdot \frac{\pi}{\ell}\right) - \sec(\dot{c}, X)\sin^2\left(t \cdot \frac{\pi}{\ell}\right) \right\} dt \\
&= \int_0^\ell \left(\left(\frac{\pi}{\ell}\right)^2 \cdot \cos^2\left(t \cdot \frac{\pi}{\ell}\right) - \sin^2\left(t \cdot \frac{\pi}{\ell}\right) \right) dt \\
&= -\frac{1}{2\ell}\left(\ell^2 - \pi^2\right),
\end{aligned}$$

which of course is negative if the length ℓ of the geodesic is greater than π. Therefore, the variation gives a family of curves that are both close to and shorter than c. In the general case, we can then observe that if $\sec \geq 1$, then for the same type of variation we obtain

$$\frac{d^2 L}{ds^2}(0) \leq -\frac{1}{2\ell}\left(\ell^2 - \pi^2\right).$$

Thus we can conclude that, if the space is complete, then the diameter must be $\leq \pi$ because in this case any two points are joined by a segment, which can't minimize if it has length $> \pi$. With some minor modifications one can now conclude that any complete Riemannian manifold (M, g) with $\sec \geq k^2 > 0$ must satisfy $\operatorname{diam}(M, g) \leq \pi \cdot k^{-1}$. In particular, M must be compact. Since the universal covering of M satisfies the same curvature hypothesis, the conclusion must also hold for this space; hence M must have compact universal covering space and finite fundamental group.

In the above form, these observations were first made explicitly by Hopf and Rinow in [74], but only for abstract surfaces. In this foundational paper the authors demonstrate the Hopf-Rinow theorem, which says that the conditions of metric and geodesic completeness from above are equivalent. Thus Cartan's normal spaces are no more restrictive that Hopf's geodesically complete spaces.

The next result, which even by today's jaded standards seems truly amazing, is another one of Synge's from 1936 (see [132]). It simply states that any orientable, closed, even dimensional Riemannian manifold of positive sectional curvature must be simply connected. The motivation here again comes from constant curvature. Namely, we know from Hopf's classification that any space of constant positive curvature is the quotient of a sphere. Moreover, the deck transformations are isometries on this sphere and therefore simply orthogonal transformations. Now on an even dimensional sphere (i.e., in an odd dimensional Euclidean space) any orientation preserving orthogonal transformation must have a fixed point and can therefore not belong to the deck group. Thus the only possible element in the deck group is the antipodal map. With this we conclude that in even dimensions only the sphere and the real projective spaces have constant positive curvature, and the latter is not orientable. Note that the same arguments show that in odd dimensions all spaces of constant positive curvature must be orientable, as orientation reversing orthogonal transformation on odd dimensional spheres have fixed points. This can now be generalized to manifolds of varying positive curvature. Synge did it in the following way: Suppose M is not simply connected (or not orientable), and use this to find a shortest closed geodesic in a free homotopy class of curves (that reverses orientation). Now consider parallel translation around this geodesic. As the tangent field to the geodesic is itself a parallel field, we see that parallel translation preserves the orthogonal complement to the geodesic. This complement is now odd dimensional (even dimensional), and by assumption parallel translation preserves (reverses) the orientation; thus it must have a fixed point. In other words, there must exist a closed parallel field X perpendicular to the closed geodesic c. We can now use the above second variation formula

$$\begin{aligned}\frac{d^2L}{ds^2}(0) &= \int_0^\ell \left\{ \left|\dot{X}\right|^2 - |X|^2 \sec(\dot{c}, X) \right\} dt + g(\dot{c}, A)\big|_0^\ell \\ &= -\int_0^\ell |X|^2 \sec(\dot{c}, X)\, dt.\end{aligned}$$

Here the boundary term drops out because the variation closes up at the endpoints, and $\dot{X} = 0$ since we used a parallel field. In case the sectional curvature is always positive we then see that the above quantity is negative. But this means that the closed geodesic has nearby closed curves which are shorter. This is, however, in contradiction with the fact that the geodesic was constructed as a length minimizing curve in a free homotopy class.

Meanwhile, for surfaces another interesting result was discovered by Cohn-Vossen in the mid 1930's (see [35]). Using some detailed analyses of the Gauss-Bonnet theorem for polygonal regions, he shows that any complete open surface of nonnegative curvature must either be isometric to a cylinder or diffeomorphic to Euclidean space. In particular, any complete open surface of positive curvature must be diffeomorphic to Euclidean space. We shall see how this was generalized to higher dimensions in the next section.

In 1941 Myers (see [99]) generalized the diameter bound to the situation where one only has a lower bound for the Ricci curvature. The idea is simply that $\mathrm{Ric}(\dot{c},\dot{c}) = \sum_{i=1}^{n-1} \sec(E_i, \dot{c})$ for any set of vector fields E_i along c such that $\dot{c}, E_1, \ldots, E_{n-1}$ forms an orthonormal frame. Now assume that the fields are parallel and consider the $n-1$ variations coming from the variational vector fields $\sin\left(t \cdot \frac{\pi}{\ell}\right) E_i$. Adding up the contributions from the variational formula applied to these fields then yields

$$\begin{aligned}\sum_{i=1}^{n-1} \frac{d^2 L_i}{ds^2}(0) &= \sum_{i=1}^{n-1} \int_0^\ell \left\{ \left(\frac{\pi}{\ell}\right)^2 \cdot \cos^2\left(t \cdot \frac{\pi}{\ell}\right) - \sec(\dot{c}, E_i) \sin^2\left(t \cdot \frac{\pi}{\ell}\right) \right\} dt \\ &= \int_0^\ell \left\{ (n-1)\left(\frac{\pi}{\ell}\right)^2 \cdot \cos^2\left(t \cdot \frac{\pi}{\ell}\right) - \mathrm{Ric}(\dot{c},\dot{c}) \sin^2\left(t \cdot \frac{\pi}{\ell}\right) \right\} dt.\end{aligned}$$

Therefore, if $\mathrm{Ric}(\dot{c},\dot{c}) \geq (n-1)k^2$ (this is the Ricci curvature of S_k^n), then we obtain

$$\begin{aligned}\sum_{i=1}^{n-1} \frac{d^2 L_i}{ds^2}(0) &\leq (n-1) \int_0^\ell \left\{ \left(\frac{\pi}{\ell}\right)^2 \cdot \cos^2\left(t \cdot \frac{\pi}{\ell}\right) - k^2 \sin^2\left(t \cdot \frac{\pi}{\ell}\right) \right\} dt \\ &= -(n-1)\frac{1}{2\ell}\left(\ell^2 k^2 - \pi^2\right),\end{aligned}$$

which is negative when $\ell > \pi \cdot k^{-1}$ (the diameter of S_k^n). Thus at least one of the contributions $\frac{d^2 L_i}{ds^2}(0)$ must be negative as well, implying that the geodesic can't be a segment in this situation.

The era of the second variation formula was ended with Preissmann's paper [118]. In this paper he gave rigorous proofs for all of the above results and added a very nice result of his own. Preissmann's observation was that the convexity of distance functions in nonpositive curvature shows that geodesic triangles have angle sum $\leq \pi$. (For surfaces this is of course a consequence of the Gauss-Bonnet theorem.) Geodesic quadrilaterals must therefore have angle sum $\leq 2\pi$. If we now assume that the sectional curvature is negative, then these inequalities become strict. Using this very geometric description of curvature, he then shows that all Abelian subgroups of the fundamental group of a closed Riemannian manifold with negative curvature must be cyclic. The idea is simply that two commuting isometries can be used to create a quadrilateral whose angles add up to 2π. In particular, we see that no closed product manifold $M \times N$ admits a metric with negative sectional curvature. The same question for positive curvature was asked by Hopf and is still completely open even for the simple manifold $S^2 \times S^2$.

4.2. **Morse theory.** At the same time the variational formulae were discovered, a related technique, called Morse theory, was introduced into Riemannian geometry. This theory was developed by Morse, first for functions on manifolds in 1925 (see [96]), and then in 1934 (see [97]) for the loop space. The latter theory, as we shall see, sets up a very nice connection between the first and second variation formulae from the previous section and the topology of M. It is this relationship that we shall explore at a general level here. In section 5 we shall then see how this theory was applied in various specific settings.

If we have a proper function $f : M \to \mathbb{R}$, then its Hessian (as a quadratic form) is in fact well defined at its critical points without specifying an underlying Riemannian metric. The *nullity* of f at a critical point is defined as the dimension of the kernel of $\nabla^2 f$, while the *index* is the number of negative eigenvalues counted with multiplicity. A function is said to be a *Morse function* if the nullity at any of its critical points is zero. Note that this guarantees in particular that all critical points are isolated. The first fundamental theorem of Morse theory is that one can determine the topological structure of a manifold from a Morse function. More specifically, if one can order the critical points $x_1, \ldots, x_k$ so that $f(x_1) < \cdots < f(x_k)$ and the index of x_i is denoted λ_i, then M has the structure of a CW complex with a cell of dimension λ_i for each i. Note that in case M is closed then x_1 must be a minimum and so $\lambda_1 = 0$, while x_k is a maximum and $\lambda_k = n$. The classical example of this theorem in action is a torus in 3-space and f the height function (see [92]).

We are now left with the problem of trying to find appropriate Morse functions. While there are always plenty of such functions, there doesn't seem to be a natural way of finding one. However, there are natural choices for Morse functions on the loop space to a Riemannian manifold. This is, somewhat inconveniently, infinite dimensional. Still, one can develop Morse theory as above for suitable functions, and moreover the loop space of a manifold determines the topology of the underlying manifold.

If $p, q \in M$, then we denote by Ω_{pq} the space of all smooth, piecewise smooth, or absolutely continuous paths from p to q. The first observation about this space is that

$$\pi_{i+1}(M) = \pi_i(\Omega_{pq}).$$

To see this, just fix a path from p to q and then join this path to every curve in Ω_{pq}. In this way Ω_{pq} is identified with Ω_p, the space of loops fixed at p. For this space the above relationship between the homotopy groups is almost self-evident.

On the space Ω_{pq} we have two naturally defined functions, the arclength and energy functionals:

$$\begin{aligned} L(\gamma) &= \int |\dot{\gamma}|, \\ E(\gamma) &= \frac{1}{2}\int |\dot{\gamma}|^2. \end{aligned}$$

While the energy functional is easier to work with, it is of course the arclength functional that we are really interested in. In order to make things work out nicely for the arclength functional, it is convenient to parametrize all curves on $[0,1]$ and proportionally to arclength. We shall think of Ω_{pq} as an infinite dimensional manifold. For each curve γ in Ω_{pq} the natural choice for the tangent space consists of the vector fields along γ which vanish at the endpoints of γ. This is because these vector fields are exactly the variational fields for curves through γ in Ω_{pq}, i.e. fixed endpoint variations of γ. An inner product on the tangent space is then naturally defined by

$$(X,Y) = \int_0^1 g(X,Y)\,dt.$$

Now the first variation formula for arclength tells us that the gradient for L at γ is $-\nabla_{\dot{\gamma}}\dot{\gamma}$. Actually this can't be quite right, as $-\nabla_{\dot{\gamma}}\dot{\gamma}$ doesn't vanish at the endpoints. The real gradient is gotten in the same way we find the gradient for a function on a surface in space, namely, by projecting it down into the correct tangent space. In any case we note that the critical points for L are exactly the geodesics from p to q. The second variation formula tells us that the Hessian of L at these critical points is given by

$$\nabla^2 L(X) = \ddot{X} + R(X,\dot{\gamma})\dot{\gamma},$$

at least for vector fields X which are perpendicular to γ. Again we ignore the fact that we have the same trouble with endpoint conditions as above. We now need to impose the Morse condition that this Hessian isn't allowed to have any kernel. The vector fields J for which $\ddot{J} + R(J,\dot{\gamma})\dot{\gamma} = 0$ are called *Jacobi fields*. Thus we have to figure out whether there are any Jacobi fields which vanish at the endpoints of γ. The first observation is that Jacobi fields must always come from geodesic variations. The Jacobi fields which vanish at p can therefore be found using the exponential map $\exp_p$. If the Jacobi field also has to vanish at q, then q must be a critical value for $\exp_p$. Now Sard's theorem asserts that the set of critical values has measure zero. For given $p \in M$ it will therefore be true that the arclength functional on Ω_{pq} is a Morse function for almost all $q \in M$. Note that it may not be possible to choose $q = p$, the simplest example being the standard sphere. We are now left with trying to decide what the *index* should be. This is of course the dimension of the largest subspace on which the Hessian is negative definite. It turns out that this index can also be computed using Jacobi fields and is in fact always finite. Thus one can compute the topology of Ω_{pq}, and hence M, by finding all the geodesics from p to q and then computing their index.

In geometric situations it is often unrealistic to suppose that one can compute the index precisely, but as we shall see it is often possible to given lower bounds for the index. As an example, note that if M is not simply connected, then Ω_{pq} is not connected. Each curve of minimal length in the path components is a geodesic from p to q which is a local minimum for the arclength functional. Such geodesics evidently have index zero. In particular, if one can show that all geodesics, except for the minimal ones from p to q, have index > 0, then the manifold must be simply connected.

4.3. **Gauss-Bonnet revisited.** In 1926 Hopf proved (see [73]) that in fact there is a Gauss-Bonnet formula for all even dimensional hypersurfaces $H^{2n} \subset \mathbb{R}^{2n+1}$. The idea is simply that the determinant of the differential of the Gauss map $G : H^{2n} \to S^{2n}$ is the Gaussian curvature of the hypersurface. Moreover, this is an intrinsically computable quantity. If we integrate this over the hypersurface, we obtain, as Kronecker did for surfaces,

$$\frac{1}{\mathrm{vol} S^{2n}} \int_H \det (DG) = \deg (G) ,$$

where $\deg (G)$ is the Brouwer degree of the Gauss map. Note that this can also be done for odd dimensional surfaces, in particular curves, but in this case the degree of the Gauss map will depend on the embedding or immersion of the hypersurface. Instead one gets the so-called winding number. Hopf then showed, as Dyck had earlier done for surfaces, that $\deg (G)$ is always half the Euler characteristic of H, thus yielding

$$\frac{2}{\mathrm{vol} S^{2n}} \int_H \det (DG) = \chi (H) .$$

Since the left-hand side of this formula is in fact intrinsic, it is natural to conjecture that such a formula should hold for all manifolds. The problem was to find the correct type of curvature to integrate. Namely, just because $\det (DG)$ is an intrinsic quantity doesn't mean that it has an expression that makes sense on manifolds which are not hypersurfaces. The simplest manifold for which the above formula doesn't necessarily make sense is complex projective space. Given the work of Kronecker and Dyck the above formula seems less than surprising, but the other important new discovery in Hopf's paper was the new formula for the Euler characteristic in terms of the indices of a vector field.

In the 1930s Allendoerfer and Fenchel independently of each other established a Gauss-Bonnet formula for all closed submanifolds of Euclidean space, thus generalizing Hopf's result to arbitrary codimension. In 1943 Allendoerfer and Weil in [2] then established a completely general Gauss-Bonnet formula for all closed Riemannian manifolds. Their proof used a local polyhedral version of the Gauss-Bonnet formula and also a result of Janet-Burstin-Cartan that real analytic metrics can locally be isometrically imbedded in Euclidean space. Thus from Whitney's result that any manifold admits an analytic structure and the fact that any Riemannian metric locally can be approximated by an analytic metric, one can get the whole result. Almost simultaneously, Chern in a very important paper in 1944 gave an intrinsic proof of the Gauss-Bonnet theorem, and therefore completely bypassed the idea that one needs to (locally) embed the space into Euclidean space. Chern's proof is also important in many other respects, for it gave him the idea to compute characteristic classes using curvature, something that eventually became very important in connection with the index theory developed by Atiyah and Singer.

4.4. **The Bochner technique.** In the 1946 paper [19], Bochner develops an entirely new way of working with curvature and analysis on manifolds. The idea goes back at least to Bernstein around 1900. He used that if $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is a harmonic function, then

$$\Delta \frac{1}{2} |\nabla u|^2 = \left|\nabla^2 u\right|^2 \geq 0.$$

Thus the energy $\frac{1}{2} |\nabla u|^2$ of u is always a subharmonic function. Bochner realized that this game could also be played on Riemannian manifolds and moreover with several different kinds of quantities. However, an interesting thing happens. Namely, various curvature quantities also enter into these formulae. Let X be a vector field on a closed Riemannian manifold (M, g). We think of ∇X as a $(1,1)$-tensor $v \to \nabla_v X$. When this tensor is skew-symmetric, the vector field is called a Killing field. This condition is equivalent to assuming that the flow of X acts by isometries. When ∇X is symmetric, X is locally the gradient for a function, or in other words the 1-form $\omega(v) = g(v, X)$ is closed. To this we wish to add the condition that $\mathrm{div} X = 0$, or in other words that X is locally the gradient of a harmonic function. Such vector fields are called harmonic. In these cases of Killing and harmonic vector fields, Bochner established the formulae

$$\Delta \frac{1}{2} |X|^2 = |\nabla X|^2 \pm \operatorname{Ric}(X, X),$$

where the $-$ occurs when X is a Killing field. Integrating over M and using Stokes' theorem then imply

$$\begin{aligned} 0 &= \int_M \Delta \frac{1}{2} |X|^2 \\ &= \int_M |\nabla X|^2 \pm \int_M \operatorname{Ric}(X, X). \end{aligned}$$

Therefore, if $\operatorname{Ric} \leq 0$ and X is a Killing field, we see that $\nabla X = 0$; i.e., X is a parallel field. Moreover, we must have that $\operatorname{Ric}(X, X) = 0$. Thus $X = 0$ in case the Ricci curvature is negative. By a result of Myers and Steenrod (see [100]) from 1939, it was known at the time that the isometry group of a Riemannian manifold is a Lie group whose Lie algebra is the set of Killing fields. Hence closed manifolds of negative Ricci curvature must have finite isometry groups.

In the other case, where $\operatorname{Ric} \geq 0$ and X is harmonic, we see again that $\nabla X = 0$ and that $\operatorname{Ric}(X, X) = 0$. Thus we again conclude that $X = 0$ when $\operatorname{Ric} > 0$. This time the importance of the result lies in the fact that the space of harmonic vector fields is the same as the first cohomology with real coefficients. More generally it was established by Hodge in 1936 (see [71]) that the space of harmonic p-forms is exactly the pth cohomology group. Thus we get the conclusion that a closed Riemannian manifold with positive Ricci curvature has vanishing first Betti number. This is much weaker than Myers' result which says that the fundamental group must be finite. The added benefit is that we can also show that if the Ricci curvature is merely nonnegative, then the first Betti number is less than or equal to the dimension. Moreover, equality occurs only for flat tori. In particular, any metric on the torus with nonnegative Ricci curvature must be flat. We shall see in the next sections how this type of result has been generalized in many different directions.

Some people also refer to the formulae

$$\Delta\frac{1}{2}|X|^2 = |\nabla X|^2 \pm \operatorname{Ric}(X,X)$$

as *Weitzenböck formulae.* In 1923 in his book on invariant theory (see [139]), Weitzenböck actually did establish a similar formula for p-forms. More precisely, he considered the exterior derivative d and its dual δ, which is defined on a Riemannian manifold (this operator is up to sign the divergence of the form). He then computed $d\delta + \delta d$ and showed that it could be decomposed as follows:

$$d\delta + \delta d = \nabla^*\nabla + C(R).$$

Here ∇^* is the dual to the connection considered as a map $\nabla : \Gamma(TM) \to \operatorname{End}(TM, TM)$ (i.e., ∇ takes vector fields X to the $(1,1)$-tensor ∇X), and $C(R)$ a contraction of the curvature tensor on forms. We should warn the reader here that both $d\delta + \delta d$ and $\nabla^*\nabla$ have positive eigenvalues and therefore look like the negative of a trace.

From this formula one can get the above Bochner identity at least for 1-forms after working out that in this case $C(R)$ is the dual of the Ricci tensor. Weitzenböck, however, did not go that far. More generally one often has a natural type of Laplace operator $\triangle$ on sections of a bundle $E \to M$, where E has an inner product structure and M is a Riemannian manifold. Given a natural type of connection on E, i.e., a map $\nabla : \Gamma(E) \to \operatorname{End}(TM, E)$ one can then usually find formulae of Weitzenböck type

$$\triangle = \nabla^*\nabla + C(R),$$

where $C(R)$ is a contraction in the curvature of the connection ∇. A particularly interesting case, as we shall see in the next sections, is when E is a spinor bundle and $\triangle$ the square of the Dirac operator. But there are many other examples as well.

5. Structure results

In this section, which covers the period 1950-1975, we shall see how the second variation and Bochner techniques were substantially improved to the point where they yielded several deep connections between the topology and geometry of Riemannian manifolds. We shall see that Morse theory was instrumental in the developments of many of the new results. Morse theory was explained briefly in section 4.2. For a more detailed description, the reader is referred to [92] and [80].

5.1. **Rauch comparison.** First we should point out that the methods of proof we use here are adopted from the two fundamental equations of Riemannian geometry from Section 3.3. Originally people used variational methods and Jacobi fields for this. However, we feel that the approach developed here is quite natural and perhaps easier for the uninitiated to comprehend.

In his 1926 paper Synge actually proves that if a closed manifold (M,g) satisfies $k \le \sec \le K$, then we have that $B\left(0, \pi/\sqrt{K}\right) \subset \operatorname{int}(\operatorname{conj}_p) \subset B\left(0, \pi/\sqrt{k}\right)$, where $\pi/\sqrt{k}$ is interpreted as ∞ in case $k \le 0$. Thus the exponential map is nonsingular on a given large region. Rauch in 1951 (see [119]) refined this statement to actually get bounds for the differential of the exponential map. The method is the same as that used by Cartan to get that the exponential map is distance nondecreasing in nonpositive curvature. In the language of polar coordinates on $\operatorname{int}(\operatorname{conj}_p)$ as explained in 3.3, we obtain that the metric as a bilinear form satisfies

$$\operatorname{sn}_K^2(r)(\delta_{\alpha\beta}) \le (g_{\alpha\beta}(r)) \le \operatorname{sn}_k^2(r)(\delta_{\alpha\beta}),$$

where sn_k is the function

$$\operatorname{sn}_k(r) = \begin{cases} \frac{\sin(\sqrt{k}r)}{\sqrt{k}} & \text{if} \quad k > 0, \\ r & \text{if} \quad k = 0, \\ \frac{\sinh(\sqrt{-k}r)}{\sqrt{-k}} & \text{if} \quad k < 0. \end{cases}$$

This means that the exponential map is increasing when compared to constant curvature K, and decreasing when compared to constant curvature k. The above inequalities are referred to as the *Rauch comparison estimates*. With these estimates Rauch started a whole new trend by proving the first *sphere theorem*. Specifically, he shows that if the curvatures are *pinched* so that $0.74 \cdot K \le \sec \le K$, then the universal covering is homeomorphic to a sphere. It is very likely that H. Hopf suggested this type of pinching problem to Rauch. Certainly Hopf thought about trying to generalize his own result on manifolds with constant curvature to manifolds which almost have constant curvature.

The proof is quite complicated. The idea is to try to cover M by two metric balls of radius $\pi/\sqrt{K}$ and then, using the exponential maps centered at these two points, to glue together the corresponding balls in the tangent spaces in such a way that one obtains a covering map. When a closed manifold is exhibited as a union of two discs, it follows from the Schoenflies theorem that it is a *twisted sphere* and in particular homeomorphic to a sphere. Work of Milnor, however, shows that it might not be diffeomorphic to a standard sphere.

In 1958 Klingenberg announced an improvement to Rauch's result. The full version appeared in 1959 in [78]. Klingenberg's idea was that one should be able to control the injectivity radius. In an ingenious argument that builds on Synge's paper of 1936, he shows that in even dimensions the exponential map will in addition be one-to-one on balls of size $\pi/\sqrt{K}$. That is, the injectivity radius is $\ge \pi/\sqrt{K}$, provided that the manifold is simply connected and $0 < \sec \le K$. Therefore, as long as one can exhibit M as the union of two balls of size $\pi/\sqrt{K}$, one will immediately have shown that M is the union of two topological balls and thus that M is a sphere. Using this, Klingenberg was able to improve the pinching constant of Rauch from 0.74 to 0.55.

Shortly afterwards in 1958, Berger discovered a new method for showing that a space becomes a homotopy sphere using Morse theory (see [11]). His idea is that if $\sec \ge k > 0$, then geodesics of length $\ell > \pi/\sqrt{k}$ have index $\ge (\dim M - 1)$. This follows immediately from Synge's estimates, because for any parallel field E perpendicular to the geodesic, the variational field $\sin(t\pi/\ell)E$ has negative second variation. Thus there is an $n-1$ dimensional subspace on which the Hessian is negative definite. Therefore, if one can find $p, q \in M$ such that all but the shortest geodesics from p to q have length $> \pi/\sqrt{k}$, then Morse theory would guarantee that the homotopy groups in dimensions $1, \ldots, \dim M - 1$ vanish, and hence the space must be a homotopy sphere. Berger then proceeded to show that if the manifold is simply connected and strictly quarter pinched, i.e., for some

$\delta > 1/4$ we have $\delta \cdot K \leq \sec \leq K$, then in fact all geodesic loops must have length $\geq \pi/\sqrt{\delta \cdot K} > \pi/\sqrt{K}$. Using Sard's theorem as described above, we can then find q and p close together so that they are not conjugate to each other and have the desired index property for the geodesics joining these points.

Using Klingenberg's injectivity radius estimate, this result was soon improved by Berger in 1960 (see [13] and [12]). In this work the idea is that if $\sec \geq k$, then the manifold can be covered by two balls of radius $\pi/2\sqrt{k}$. In the quarter-pinched situation this means the space will be covered by two balls of radius $> \pi/\sqrt{K}$. But by Klingenberg's argument such balls are discs, at least in even dimensions; thus the manifold must be homeomorphic to a sphere. This same argument also seems to have been noticed by Toponogov around the same time (see [135]). The proof of why such manifolds are covered by two such discs and also why geodesics between appropriate points are long will be discussed below. In fact, one can show that if $\sec \geq k$ and $\operatorname{diam} > \pi/2\sqrt{k}$, then it is possible to find points such that all but the shortest geodesics between these points have length $> \pi/\sqrt{k}$ and therefore index $> \dim M - 1$. Thus such manifolds have to be homotopy spheres. While the main ingredients for this result were available already in Berger's 1958 paper, this result doesn't seem to have attracted much attention at the time. Usually people refer to [14] for this result. In the literature this result is strangely enough often stated only for simply connected manifolds and often without attribution to Berger. This result is known as the *diameter sphere theorem* and will also be discussed below in Section 6.1.

In 1961 Klingenberg in [79], using very delicate arguments based on the Morse theory for the loop space of M, showed that his injectivity radius estimate also holds in odd dimensions, provided the manifold is simply connected and quarter pinched. For a completely different proof of this see also [43]. Thus the quarter-pinched sphere theorem was shown to hold in all dimensions. In even dimensions this result is optimal, as the Fubini-Study metric on the projective space satisfies $1/4 \leq \sec \leq 1$. In odd dimensions the story is still not completely settled. However, Berger in 1962 discovered some very important examples (see [14]). By shrinking the Hopf fibers on S^3 he was able to get a sequence of metrics g_ε on this space such that $0 < \sec \leq 1$ and with volume going to zero as $\varepsilon \to 0$. Thus some sort of strong pinching condition is needed in order to get an injectivity radius estimate in odd dimensions. As late as 1994, Abresch and Meyer showed that an injectivity radius estimate can be gotten provided the metric is almost quarter pinched. Abresch and Meyer have a very nice survey article in [62] that explains in detail all of the above material on sphere theorems and injectivity radius estimates, together with many other related results that we are not going to mention here.

In the above-mentioned results related to the sphere theorem a new result by Toponogov (see [134]) turned out to be very convenient and later became one of the cornerstones of global Riemannian geometry. It is essentially just a generalization of the Rauch estimates to all of the segment domain. To explain the idea in full, as it is understood these days, we shall first mention another important result from this period that Toponogov was not aware of. The story starts with a paper from 1958 by Calabi (see [21]), where he extends the maximum principle of E. Hopf for C^2 functions to functions which are merely continuous and weakly subharmonic. In this paper Calabi also shows how this can be used in Riemannian geometry to get global weak estimates for the Laplacian of a distance function. Furthermore, the same technique yields weak estimates on the Hessian, but this wasn't observed until much later (see Karcher's article in [33]). Given a distance function $f(x) = d(x, p)$ on a complete Riemannian manifold, Calabi observed that on the segment domain the Laplacian Δf satisfies the differential equation

$$\partial_r (\Delta f) + \operatorname{tr} \left(\nabla^2 f\right)^2 = -\operatorname{Ric}(\partial_r, \partial_r).$$

Note that this comes from taking traces in the second fundamental equation from section 3.3. The Cauchy-Schwarz inequality then yields

$$\partial_r (\Delta f) + \frac{(\Delta f)^2}{n-1} \leq -\operatorname{Ric}(\partial_r, \partial_r).$$

Therefore, if $\operatorname{Ric} \geq (n-1)k$, then one observes that the Laplacian is less than what it is in constant curvature k. Specifically,

$$\Delta f \leq (n-1) \frac{\operatorname{sn}_k' \circ f}{\operatorname{sn}_k \circ f}.$$

The important discovery is now that in fact this estimate holds on all of M in the *weak sense*. More precisely, we say that $(\Delta f)(x) \leq a$ in the weak sense if for each $\varepsilon > 0$ there is a smooth function g_ε defined in a neighborhood of x such that

$$\begin{aligned} f(x) &= g_\varepsilon(x), \\ f &\leq g_\varepsilon, \\ \Delta g_\varepsilon &\leq a + \varepsilon. \end{aligned}$$

In case f is smooth at x this evidently says $(\Delta f)(x) \leq a$.

To see that the above estimate holds in the weak sense, one simply notes that at points q where f is not smooth one can choose a unit speed segment σ from p to q and then use that the functions

$$f_\varepsilon(x) = \varepsilon + d(\sigma(\varepsilon), x)$$

are smooth. By the triangle inequality these functions are support functions from above for f at q; that is,

$$\begin{aligned} f(x) &\leq f_\varepsilon(x), \\ f(q) &= f_\varepsilon(q). \end{aligned}$$

Moreover, by the above estimate on the Laplacian we also have

$$\Delta f_\varepsilon \leq (n-1) \frac{\operatorname{sn}_k' \circ f_\varepsilon}{\operatorname{sn}_k \circ f_\varepsilon} \underset{\varepsilon \to 0}{\to} (n-1) \frac{\operatorname{sn}_k' \circ f}{\operatorname{sn}_k \circ f},$$

thus yielding the desired weak Laplacian estimate.

However, our interest is in understanding the Hessian of f. Using the second fundamental equation

$$L_{\partial_r} \nabla^2 f + \nabla^2 f \circ \nabla^2 f = -R(\partial_r, \cdot)\, \partial_r,$$

one can as before obtain estimates where f is smooth:

$$\nabla^2 f \leq \frac{\operatorname{sn}_k' \circ f}{\operatorname{sn}_k \circ f} (I - dr \partial_r).$$

Here, $I - dr\partial_r$ is the tensor that orthogonally projects onto the complement of ∂_r. One slight problem, however, is that the Hessian is always zero in the direction of ∂_r, and so it is not really possible to extend the estimate to nonsmooth points. Assuming that $\sec \geq k$, we can replace f by a *modified distance function* f_k. This function is obtained from f by composing it with a function ϕ_k, which has the property that in constant curvature k the modified distance function has a Hessian which is a homothety, thus eliminating the zero eigenvalue. In the case where $k = 0$, we simply change f to $\frac{1}{2}f^2$, and if $k = 1$, then $f_1 = 1 - \cos \circ f$. As in the Ricci curvature case one then has that

$$\nabla^2 f_k \leq (\mathrm{sn}_k' \circ f)\, I$$

on all of M. Here the right-hand side is the Hessian one has in constant curvature k. One can then use the generalized maximum principle to show *Toponogov's comparison theorem*: Any triangle on a Riemannian manifold with $\sec \geq k$ has larger interior angles than the triangle on S_k^2 with the same side lengths. Here a triangle is simply three points joined by segments.

Toponogov first proved this for $k > 0$ in [134] and then for $k = 0$ in [136]. However, both results had been announced without proofs already in 1957. The case where $k = 0$ turns out to be the easiest to establish, and $k < 0$ is also not too hard. Toponogov's interest in these results seems to have been motivated by two rigidity results he also proves in these papers. The first is the *maximal diameter theorem*: If $\sec \geq k > 0$ and the diameter has the maximal possible value of $\pi/\sqrt{k}$, then (M, g) is isometric to S_k^n. The second is the so-called *splitting theorem*. Instead of having positive curvature, one simply has $\sec \geq 0$. To replace the maximal diameter assumption one then supposes the manifold contains a *line*, i.e., a unit speed geodesic $\sigma : \mathbb{R} \to M$ with $d(\sigma(s), \sigma(t)) = |s - t|$ for all $s, t \in \mathbb{R}$. The conclusion is then that (M, g) is isometric to $\left(N \times \mathbb{R}, g_{|H} + dt^2\right)$, where N is a totally geodesic hypersurface with the induced metric from M. Note that, in particular, such manifolds do not have any points where all curvatures are positive. The proofs of these two results are quite simple given the above Hessian estimates.

In the first case simply consider two points p, q at maximal distance, and let f and h be the distance functions from these points. Then we have from the triangle inequality that

$$f(x) + h(x) \geq \pi/\sqrt{k}$$

for all $x \in M$, with equality holding for x on any segment joining p and q. The modified distance functions will then satisfy

$$f_k(x) + h_k(x) \geq 2,$$

with equality for x on any segment joining p and q. The Hessian estimates on the other hand show that this sum is concave:

$$\nabla^2 (f_k(x) + h_k(x)) \leq 0.$$

Thus the sum is constant. From this it follows that both functions are smooth away from $\{p, q\}$ and that the Hessians are what one would expect in constant curvature. Now recall that the first fundamental equation from 3.3 tells us what the metric must be if we know the Hessian of a distance function. Consequently, if the Hessian is the same as it would be in constant curvature k, then it follows that the manifold must have constant curvature k.

With this we can now also see how Berger established the important results that lead to the quarter-pinched sphere theorem and the diameter sphere theorem. Given the injectivity radius estimate, we have a manifold which satisfies $\sec \geq k > 0$ and $\mathrm{inj} > \pi/2\sqrt{k}$. For simplicity let us suppose that $k = 1$. The important point is that if p and q realize the diameter ($> \pi/2$), then the two metric balls $B(p, \mathrm{inj}_p)$ and

$B(q, \mathrm{inj}_q)$ cover M. To see this, fix $x \notin B(p, \mathrm{inj}_p)$. The modified distance function $f_1 = 1 - \cos \circ f$ for the distance from p satisfies

$$\nabla^2 f_1 \leq \cos \circ f = (1 - f_1) I.$$

The Hessian is therefore negative-definite on the complement of $\bar{B}(p, \pi/2) \subset B(p, \mathrm{inj}_p)$. Thus q must be the *unique* maximum for f_1 (and f). Therefore, if σ is a segment joining q and x, the composition $\phi = f_1 \circ \sigma$ will satisfy

$$\begin{aligned} \phi'' &\leq 1 - \phi, \\ \phi(0) &= 1 - \cos d, d = \mathrm{diam} M, \\ \phi'(0) &= 0. \end{aligned}$$

It is then fairly easy to check that ϕ must be smaller than the solution to

$$\begin{aligned} y'' &= 1 - y, \\ y(0) &= 1 - \cos d, \\ y'(0) &= 0. \end{aligned}$$

The solution to this equation is $y(t) = 1 - \cos d \cos(t)$. As $\phi(t) \leq y(t) = 1 - \cos d \cos(t)$ we see that $\phi \leq 1$ for $t \geq \pi/2$. Or, in other words, that $d(p, \sigma(t)) \leq \pi/2$ if $t \geq \pi/2$. But this means exactly that x must lie in $B(p, \mathrm{inj}_p)$ if it doesn't lie in $B(q, \mathrm{inj}_q)$.

A similar argument will now also establish the other result by Berger. Instead, the assumption is that $\sec \geq 1$ and $\mathrm{diam} > \pi/2$. If we pick $p, r \in M$ such that $d(p, r) = \mathrm{diam} M = d > \pi/2$, then we can argue as above that for q sufficiently close to p, any geodesic from p to q is either minimal or has length $> \pi$ (and therefore index $\geq n - 1$).

In the case where $\sec \geq 0$ and we have a line σ, a construction similar to the one used to prove the maximal diameter theorem can be devised. Instead of using modified distance functions we simply use that

$$\nabla^2 f \leq \frac{1}{f}(I - dr\partial_r) \leq \frac{1}{f} I$$

to get the estimate

$$\nabla^2 f \leq \frac{1}{f} I$$

everywhere. Given a *ray* $\gamma : [0, \infty] \to M$, i.e., $d(\gamma(s), \gamma(t)) = |s - t|$ for all $s, t \in [0, \infty]$, we can construct a Busemann function

$$f_\gamma(x) = \lim_{t \to \infty} (d(x, \gamma(t)) - t).$$

This is simply the renormalized distance from infinity in the direction of the ray. On Euclidean space Busemann functions look like $f_\gamma(x) = \dot{\gamma}(0) \cdot (\gamma(0) - x)$; that is, they measure the length of the projection of x down on the line γ. The above Hessian estimate now translates into the fact that these Busemann functions are concave:

$$\nabla^2 f_\gamma \leq 0,$$

since the actual distance from any point x to $\gamma(\infty)$ is infinite. Given the line, we can construct two Busemann functions $f_\pm$ by going off in both directions. The triangle inequality now implies that

$$f_+ + f_- \geq 0,$$

with equality holding on the line. However, as the sum is also concave it must be constant. Then each of the Busemann functions are linear ($\nabla^2 f_\pm = 0$), and the hypersurface we seek is $N = f_\pm^{-1}(0)$. Rigidity of the metrics follows as before.

Given these results and the fact that Calabi developed the necessary analytic tools to work with Laplacians rather than just Hessians, one would expect that these two results would immediately have been generalized to the case where one merely has a lower Ricci curvature bound. But it actually took some time for this to happen. In the case of the splitting theorem for $\mathrm{Ric} \geq 0$, this was done by Cheeger and Gromoll in 1971 (see [30] and [44]). The same proof works, except this time we only have that the Busemann functions are superharmonic:

$$\Delta f_\pm \leq 0.$$

However, the maximum principle according to Calabi still works. So the fact that the sum has a global minimum implies that the sum is constant. Thus the functions are harmonic. Now the equation

$$\partial_r (\Delta f_\pm) + \operatorname{tr} \left(\nabla^2 f_\pm\right)^2 = -\mathrm{Ric}\,(\partial_r, \partial_r)$$

shows that, in fact, the Busemann functions are linear, thus putting us in the same situation as before. The maximal diameter result for manifolds with $\mathrm{Ric} \geq (n-1)\,k > 0$ and diameter $\pi/\sqrt{k}$ was established by S.-Y. Cheng only in 1976. His proof used eigenvalue techniques. However, the above proof for sectional curvature again works without using modified distance functions. Thus, modulo a little analysis, the proof is actually simpler than in the sectional curvature case.

In the paper [30] some interesting consequences for closed manifolds with $\mathrm{Ric} \geq 0$ are also observed. Notably, any such manifold has the property that the universal covering is isometric to $N \times \mathbb{R}^k$, for some closed Riemannian manifold N with $\mathrm{Ric} \geq 0$. From this we can see that the structure of closed manifolds with $\mathrm{Ric} \geq 0$ is essentially reduced to that of flat manifolds and closed simply connected manifolds with $\mathrm{Ric} \geq 0$. A particularly interesting consequence of this is that any closed $K(\pi, 1)$ (i.e., the universal covering is contractible) with $\mathrm{Ric} \geq 0$ must be flat (see also the weaker but earlier result of Eells-Sampson in section 5.3 and compare to Bochner's rigidity result for such metrics on tori).

In the two papers [55] and [30], from 1969 and 1972 respectively, Cheeger, Gromoll, and W. Meyer extended Cohn-Vossen's results for surfaces to all dimensions. As the Gauss-Bonnet theorem is no longer useful in this more general context, entirely new methods were necessary. In the first paper it is shown that any open complete manifold of positive sectional curvature is diffeomorphic to Euclidean space. In the second paper, open complete manifolds of nonnegative sectional curvature are studied. The main result there is the so-called soul theorem, which asserts that such manifolds are diffeomorphic to the normal bundle of a closed totally geodesic submanifold, called the *soul*. Thus the first result merely says that in positive curvature the soul must be a point. The first step in proving these results is to use that all Busemann functions are concave. We fix $x \in M$ and consider all rays γ emanating from x and with those the minimum f of all of the Busemann functions f_γ corresponding to these rays. This function must again be concave. Moreover, the superlevel sets are convex, and hence they must also be compact, for otherwise some superlevel set would contain a ray γ_0 emanating from x, thus contradicting that $f \leq f_{\gamma_0}$. In particular, the maximal superlevel is compact. In case the manifold has positive curvature this maximal level set is a point. Otherwise it is just a compact convex subset C. If C is a submanifold, we have found the soul. If not, it is necessary to study the structure of C and show that it contains a closed convex submanifold. This is done by considering the distance function to the boundary of C and showing that this function is concave on C. To this end it is of course necessary to show, in analogy with convex subsets for Euclidean space, that if C is not a submanifold, it is at least topologically a submanifold with boundary. The final step after having identified the appropriate soul is then to use that the concave functions, while not smooth, are sufficiently regular that one can use them to nicely contract the manifold onto the soul. All of the steps in the proof are quite geometric, but also very technical. A more complete analysis with a full set of references is available in Greene's article in [62].

In the context of Morse theory note that if a set $S \subset M$ is totally convex (e.g., a sublevel set of a convex function), then any geodesic which begins and ends in S must lie entirely in S. Thus all the critical points for the length functional are curves which lie in S. This means, in particular, that M and S have the same homotopy type. While this argument doesn't give the exact statement of the soul theorem, it at least gives one the correct result in homotopy theory relatively easily.

5.2. **Cheeger's thesis.** All of the above results are structure results, in the sense that from very few geometric assumptions one gains a rather complete picture of what the underlying manifold looks like. With Cheeger's thesis in 1967, a completely new type of result was introduced into global Riemannian geometry (see [24] and [25]). About the same time Weinstein proved a similar but much weaker and more specialized result in [138]. The idea is to see what one can say with much weaker assumptions on curvature. Now keep in mind that any closed Riemannian manifold can be scaled (blown up) until it has bounded curvature, say, $|\sec| \leq 1$. Such a condition therefore says very little about the manifold. However, one can in addition also try to make assumptions on diameter and/or volume. The Gauss-Bonnet theorem, for instance, says that only finitely many surfaces admit metrics with $\sec \geq k$ and $\mathrm{vol} \leq V$ for any $k \in \mathbb{R}$ and $V \in (0, \infty)$. In higher dimensions there is more to topology than just the Euler characteristic. In dimension 3, for example, we have infinitely many space forms of constant curvature 1. They certainly have both bounded diameter and volume, but as the fundamental group increases the volume decreases. The optimal situation that might give us a sort of *finiteness result* is therefore to consider the class of closed Riemannian n-manifolds with

$$\begin{aligned} |\sec| &\leq K, \\ \mathrm{diam} &\leq D, \\ \mathrm{vol} &\geq v, \end{aligned}$$

for fixed but arbitrary positive constants K, D, v. Cheeger's *finiteness theorem* states that each such class contains only finitely many diffeomorphism types.

The first step in his proof is to show that each manifold in such a class has a lower bound for the injectivity radius which depends only on n, K, D, v (see also [70] for a different proof of this). Thus exponential coordinates always exist on balls of an a priori size. The curvature bound then tells us that the differential of the exponential map is bounded. Moreover, as the manifold is bounded in size it takes only an a priori number of such charts to cover the manifold. Suppose now we have an infinite collection (M_i, g_i) of Riemannian manifolds in this class. Then cover each of them by charts as just described. By passing to a subsequence we can even assume that the same number of charts are used on each manifold in this sequence. Let us denote them by $\phi_{ij} : B(0,r) \to M_i$. The transition functions $\phi_{ij}^{-1} \circ \phi_{ik}$ for fixed j, k are uniformly bi-Lipschitz and can therefore be assumed to converge as well, after possibly passing to a further subsequence, by the Arzela-Ascoli Theorem. For large i the manifolds therefore have charts whose transition functions are close. This indicates that the manifolds should be diffeomorphic. However, all one can show is that they are homeomorphic, and even this required using some nontrivial results from geometric topology. Using some general results from topology (see, e.g., [77]) it is then possible to show that in dimensions $n \geq 5$ any class of manifolds which contains finitely many homeomorphism types must also contain only finitely many diffeomorphism types. Much more work is of course needed to make all this rigorous, but subsequent developments have made statements and proofs sufficiently simple that the above outline really can be justified to the point of establishing finiteness of diffeomorphism types directly. This will be discussed further in section 6.2. The first proof in the literature of finiteness of diffeomorphism types occurs in [56, Section 3], [59], and with more details in [110].

5.3. **The Bochner technique comes of age.** The next two interesting results using the Bochner technique are due to Lichnerowicz. In the mid 1950s he developed a Bochner formula for the curvature tensor (see [86]). In the special case where the curvature is harmonic (i.e., $\operatorname{div}\nabla R = 0$, which is weaker than having constant Ricci curvature), the formula looks like

$$\Delta \frac{1}{2} |R|^2 = |\nabla R|^2 + 2K(R).$$

Lichnerowicz at the time used this for various things related to mathematical physics (see [87]), but we shall in a minute see how it can be used in a more geometric setting. More importantly, in 1963 Lichnerowicz (see [88]) discovered a Bochner formula for *spinors*. In case the spinor is harmonic it says

$$\Delta \frac{1}{2} |\sigma|^2 = |\nabla \sigma|^2 + \frac{\mathrm{scal}}{4} |\sigma|^2.$$

He was thus able to conclude that any spin manifold with positive scalar curvature has vanishing $\hat{A}$ genus. This formula has been very important in various general forms for almost all developments relating to scalar curvature including, in particular, Seiberg-Witten theory (see [95]). Given that the Bochner technique works for more than just forms and vector fields, one is naturally led to conjecture that there is some general abstract setting where one can apply this technique. The idea is that it works for any type of section that satisfies appropriate symmetry conditions on its (covariant) derivative (see e.g. [15]).

Eells and Sampson in 1964 (see also the survey [42]) showed how a Bochner formula for functions between manifolds can be used to show various rigidity results. Their idea is to first generalize Hodge theory to the point where one can show that maps between closed manifolds such that the target has nonpositive sectional curvature are always homotopic to a harmonic map. For harmonic functions one then has a suitable Bochner formula which asserts that in case the domain has nonnegative Ricci curvature and the target nonpositive sectional curvature, then the map is totally geodesic. Using these ideas, one can prove a weak type of von Mangoldt-Cartan theorem for closed manifolds of nonpositive curvature. Namely, all of their higher homotopy groups must vanish, as spheres of dimension greater than 1 have positive curvature. They also realized that one could generalize Bochner's maximal Betti number result in a very interesting direction. Namely, any metric with nonnegative Ricci curvature on a closed manifolds which also admits a metric with nonpositive sectional curvature must be flat. To see this find a harmonic map homotopic to the identity, where the domain metric has nonnegative Ricci curvature and the target metric nonpositive sectional curvature. As noted in section 5.1, this result was generalized by Cheeger and Gromoll in 1971.

The next development came in 1971 with D. Meyer's proof that any manifold with positive curvature operator must have trivial Hodge cohomology. This had been conjectured for a long time, and Bochner, Yano and others had many partial results in this direction. Berger established the conjecture for two-forms in 1961. Finally in 1975 D. Meyer and Gallot in [48] gave an almost complete classification of all closed manifolds with nonnegative curvature operator. The classification states that aside from taking products of such spaces, only spheres, complex projective spaces, and locally symmetric spaces occur (in the first two cases the metric is not necessarily the standard one). Their proof uses a nice result of Tachibana from 1974 to the effect that any manifold with nonnegative curvature operator and harmonic curvature is locally symmetric. The proof of this uses the above-mentioned Bochner formula for the curvature tensor.

6. Spaces of Riemannian manifolds

In the last period (1975-1998) several new techniques were developed that led to a much better understanding of the topological properties of Riemannian manifolds. The main developments were critical point theory for nonsmooth distance functions, Ricci flow, and various types of convergence for Riemannian manifolds.

6.1. **Critical point theory.** Critical point theory should really be called regular point theory, for it is the lack of critical points which is used throughout. Thus one should be reminded of Lusternik-Schnirelmann theory rather than Morse theory. The best comprehensive account of this theory in the setting we wish to concentrate on is the article by Grove in [53, Vol. 3]. In Greene's article in [62] it is also explained how this theory works for convex functions and in connection with the soul theorem. The idea began with a strengthening by Grove and Shiohama in 1977 (see [65]) of Berger's diameter theorem: Any closed Riemannian manifold with $\sec \geq k > 0$ and $\operatorname{diam} > \pi/2\sqrt{k}$ is a homotopy sphere. Not only did they show directly that the manifold must be a union of two discs and therefore homeomorphic to a sphere, but they also introduced a new way of proving this. Without upper curvature bounds one cannot hope to get injectivity radius estimates, but this doesn't prevent appropriate metric balls from being discs. Their idea was to consider two points p, q at maximal distance. We have already seen how the distance functions from these points are concave near the opposite points, and thus that these distance functions have a unique global maximum. The important realization is that, in fact, these functions in a suitable sense are regular on the whole region $M - \{p, q\}$. The way to see this is via Toponogov's comparison theorem. Take any $x \in M - \{p, q\}$ and join the three points x, p, q by segments. As the base segment from p to q has length $> \pi/2\sqrt{k}$ we see that a comparison triangle with the same side lengths in S^2_k has angle $> \pi/2$ at $\bar{x}$. Thus the angle at x must also be $> \pi/2$. But this means that *all* segments from p to x will form an angle $> \pi/2$ with a given segment from x to q. A first variation argument then shows that if we go in the direction of the fixed segment from x to q, then the distance from p must increase. Thus x is in a suitable sense a regular point for the distance to p. A partition of unity argument can now be used to get a vector field X on $M - \{p, q\}$ with the property that the distance from p increases along the integral curves for X. In this way we can exhibit $M - \{p, q\}$ as an open cylinder $I \times S^{n-1}$. From this we conclude that M is homeomorphic to a sphere.

More generally we say that x is *critical* for p if for any direction from x we can find a segment from p to x that forms angle $\leq \pi/2$ with the chosen direction at x. Thus x is critical if one can't find a direction in which the distance function increases. Already in Berger's work on the sphere theorem the idea of critical points was important. Specifically, he showed that a point were the distance function is maximal must be a critical point. His method of proof was simply to use a first variation argument to show that at any point which is not critical one can find a direction in which the distance function increases.

The next major development using this type of regular point theory for the distance function came with Gromov's Betti number estimate from 1981 (see [58]): There is a constant $C(n, k, D)$ so that any closed Riemannian n-manifold, (M, g), with $\sec \geq k \in \mathbb{R}$ and $\operatorname{diam} \leq D$ has Betti numbers with any field coefficients, F, bounded by

$$\sum_{i=0}^{n} b_i(M, F) \leq C\left(n, k \cdot D^2\right).$$

Various extensions of the Bochner technique can also be used to get similar bounds for the Betti numbers. But one can of course only bound the dimension of the de Rham cohomology this way. Also much stronger curvature hypotheses are needed in order to control the curvature quantities that appear in the Bochner identities for forms. The best condition seems to be a lower bound for the curvature operator and an upper bound on the diameter (see [9]).

With Cheeger's finiteness theorem for the class: $k \leq \sec \leq K$, $\operatorname{diam} \leq D$, and $\operatorname{vol} \geq v$ and Gromov's Betti number estimate for the much larger class: $k \leq \sec$ and $\operatorname{diam} \leq D$, it is natural to conjecture that with the intermediate hypotheses: $k \leq \sec$, $\operatorname{diam} \leq D$, and $\operatorname{vol} \geq v$ one should be able to obtain an intermediate finiteness result, say, a bound on the number of homotopy types. Again using critical point theory, but now for the distance function to the diagonal in $M \times M$, Grove and Petersen in 1988 were able to prove this (see [63]). The main hurdle to proving this result is the fact that one cannot get a lower bound for the radius at which metric balls are contractible. This is why one has to go to the Cartesian product $M \times M$ and study how neighborhoods of the diagonal can be contracted onto the diagonal. Soon after, it was realized that techniques from geometric topology could be used to improve the conclusion of this result so that one essentially has finiteness of diffeomorphism types, at least in dimensions > 4 (see [64]). Perel'man refined all of the above results that use critical point theory to hold for a larger class of metric spaces (see [108]). Given that it has not been possible to prove these results with more analytic methods, it is perhaps interesting that in fact the results of [64] can be generalized to a setting where one has only integral curvature bounds (see [113]). Not only that, but the critical point theory still works in this setting, despite the fact that a substitute for Toponogov's theorem is no longer available.

In 1990 Abresch and Gromoll discovered a new way of estimating critical points using an excess estimate (see [1]). The *excess function* for two points p, q measures the deviation of the triangle inequality to be an equality:

$$\mathrm{exc}_{p,q}(x) = d(p,x) + d(x,q) - d(p,q).$$

Thus $\mathrm{exc}_{p,q}(x) = 0$ iff x lies on a segment joining p and q. The *height,* h, of a triangle p, x, q is the distance from x to the chosen segment from p to q. In a metric space we have that these quantities are related by

$$\mathrm{exc}_{p,q}(x) \leq 2h.$$

In general this cannot be improved. However, if we are on a Riemannian manifold with $\mathrm{Ric} \geq (n-1)k$, then

$$\mathrm{exc}_{p,q}(x) \leq E_{k \cdot R^2, n}(h/s) \cdot h,$$

where $s = \min\{d(p,x), d(x,q)\}$ and $E_{k \cdot R^2, n}$ is a function which depends only on the lower Ricci curvature bound and dimension, and R, where the triangle pxq is contained in a ball of size R. In addition we have that $E_{k \cdot R^2, n}(t) \to 0$ as $t \to 0$. The lower Ricci curvature bound therefore leads to an improved excess estimate for triangles where x isn't too close to p or q. In the special case where $k = 0$ we have $E_{0,n}(t) = 8t^{1/(n-1)}$. Note that the dependence on $k \cdot R^2$ is very important in case k isn't zero. Namely, we could apply the excess estimate to the case where we have a line and two Busemann functions as in Section 5.1. In this case

$$(f_- + f_+)(x) = \lim_{t \to \infty} \mathrm{exc}_{\ell(t), \ell(-t)}(x)$$

is simply the excess of a triangle with fixed height and base going to infinity. Thus we should obtain $f_- + f_+ = 0$ from the excess estimate. However, due to the dependence of R this only works in case we have $k = 0$. Note that, in particular, the excess estimate gives us a new proof of the splitting theorem in nonnegative Ricci curvature. While this is a roundabout way of proving the splitting theorem, it has the important consequence that one can prove it for appropriate limit spaces (see [26] and also Section 6.2 below).

The way to use the excess estimate in connection with critical point theory is to observe that if x is critical for p, then triangles p, x, q can be chosen to be acute at x. Acute triangles, however, have a tendency to have rather large excess, although one can only get a specific lower bound from using Toponogov's comparison theorem. Using this, Abresch and Gromoll showed that any complete Riemannian manifold with $\mathrm{Ric} \geq 0$, $\sec \geq k$ and bounded diameter growth, i.e., the diameter of the distance spheres $S(p,r)$ remain bounded as $r \to \infty$, has the property that the distance function to p has no critical points outside some compact set. In particular, such manifolds must be homeomorphic to the interior of a compact manifold with boundary. There seems to be no way of getting rid of the lower sectional curvature bound. In 1994 Perel'man discovered a new way of using this excess estimate (see [109] and Zhu's article in [62]). His idea was that instead of assuming that distance spheres are small relative to their radius, one could get information out of assuming that they have large volume. Specifically, we have that if $\mathrm{Ric} \geq (n-1)k$, then $\mathrm{vol}B(p,r) \leq \mathrm{vol}(B(\bar{p},r) \subset S^n_k)$. If the volume of $B(p,r)$ attains this maximal value, then it is easy to see that it is isometric to $B(\bar{p},r) \subset S^n_k$. Perel'man's idea is that if $\mathrm{vol}B(p,r)$ is close to the maximal possible value, then there are lots of long segments emanating from p that go all the way out to the boundary of this ball. The excess estimate can then be used to conclude that this ball is contractible, at least inside a slightly larger ball. This conclusion is similar to the work in [63], and indeed one obtains a similar finiteness result: There is an $\varepsilon(n,k,r) > 0$ such that the class of closed Riemannian n-manifolds with

$$\begin{aligned} \mathrm{Ric} &\geq (n-1)k, \\ \mathrm{diam} &\leq D, \\ \mathrm{vol}B(p,r) &\geq (1-\varepsilon)\,\mathrm{vol}\left(B(\bar{p},r) \subset S^n_k\right) \end{aligned}$$

contains only finitely many homotopy types. The methods from [64] still work in this context, and so the conclusion can be strengthened to finiteness of diffeomorphism types in dimensions > 4. Below we shall see how Cheeger and Colding have improved this substantiality to conclude that the balls $B(p,r)$ are both metrically and topologically close to $B(\bar{p},r) \subset S^n_k$. From this one can get finiteness of diffeomorphism types in all dimensions.

Finally we should mention a few other results which don't really fall under this heading, but are related to some of the classical results already mentioned.

First, related to the soul theorem, there is the paper of Özaydın and Walschap (see [105]) where it is shown that not all vector bundles over closed manifolds with nonnegative curvature admit complete metrics with nonnegative curvature. In particular, nontrivial bundles over tori do not admit such metrics.

Then there is Perel'man's improvement on the soul theorem (see [107]), which says that there is a Riemannian submersion from the entire space onto the soul. This result, in particular, shows that if such a manifold has positive curvature somewhere, then the soul must be a point and hence the space diffeomorphic to Euclidean space.

Wilhelm in [140] extends Synge's theorem. Synge's theorem does not hold if one merely has positive Ricci curvature, but something can still be done if the closed geodesic under investigation is long. To make this precise define the *first systole*, sys_1, as being the lower bound on the lengths of noncontractible loops (simply connected spaces thus have infinite first systole). The result can now be stated as: If $\mathrm{Ric} \geq (n-1)k > 0$ and $\mathrm{sys}_1 > \sqrt{\frac{n-2}{n-1}} \cdot \pi/\sqrt{k}$, then (1) if the manifold is even dimensional and orientable, it must be simply connected; while (2) if it is odd dimensional, it must be orientable. The technique used is a very ingenious extension of Synge's original argument. The systole bound is, moreover, easily seen to be a necessary condition for this result to hold.

Finally Myers' diameter bound was generalized to situations where one has integral curvature bounds in [114].

6.2. **Convergence.** Recall Berger's example of metrics on the 3-sphere with the property that the curvature stays bounded while the volume goes to zero. More generally, if we have a Riemannian manifold (M,g) and a Killing field X without zeros, then we can shrink the metric g in the direction of X so as to obtain a family of metrics with bounded curvature, bounded diameter and volume going to zero. In the case of the Berger sphere we used the Killing field associated to the Hopf fibration; thus the space looks as if it is *collapsing* to the base of this fibration, namely the 2-sphere. The Killing field can, however, be positioned in such a way that the space collapses even further. Note for instance that on the standard 3-sphere we have an isometric torus action. If we quotient out by this action, then we obtain an interval. Now simply choose the Killing field to correspond to an irrational flow on this torus to get a sequence of metrics on the 3-sphere collapsing to an interval. This construction can easily be generalized to a situation where one has a commuting set of Killing fields, although no further examples will be obtained, as one can always use the trick of irrational flows to reduce the construction to one Killing field. The next step in trying to generalize this is to consider a nilpotent Lie algebra of Killing fields, i.e., a collection of Killing fields that can be spanned by Killing fields without zeroes, $\{X_1, \ldots, X_k\}$, subject to the condition that $[X_i, X_j] = \sum_{k>\max\{i,j\}} c_{ij}^k X_k$. The interesting idea is now that one can still shrink the manifold in the direction of these vector fields and obtain a collapsing sequence with bounded curvature and diameter. This time, however, one must shrink the metrics more in some directions than in others. A very interesting example is gotten from this. In dimension three there are many closed manifolds which admit a nilpotent three dimensional Lie algebra of Killing fields, namely, the closed quotients of a three dimensional nilpotent Lie group. Thus one obtains a sequence of metrics on such manifolds with bounded curvatures and such that it collapses to a point; i.e., the diameter goes to zero. By rescaling, we can arrange matters so that the diameter is bounded and the curvature goes to zero. In this way we have examples of manifolds which are almost flat but do not admit flat metrics. Note that, in fact, any metric can be scaled to have almost zero curvature, but then the diameter will necessarily be very large. Thus some extra condition is necessary in order to say something intelligent. In 1978 Gromov (see [56]) proved that if the scale invariant quantity

$$\mathrm{diam}^2 \cdot \max|\sec|$$

is sufficiently small depending only on dimension, then the universal covering can in a natural way be identified with a nilpotent Lie group. Subsequently Ruh gave a complete characterization of such manifolds (see [124]). Better yet, in [32] Cheeger, Fukaya, and Gromov gave a complete picture of what a manifold with bounded curvature must look like if it is collapsed in certain places. Note how this complements Cheeger's finiteness theorem: When no collapse occurs we have bounds on the geometry and topology of the space.

Simultaneously with the idea of collapse the idea of *compactness* was also developed by Gromov in [59]. There are several ways in which metric spaces and Riemannian manifolds can converge. The weakest form is *Gromov-Hausdorff convergence*. This is merely an abstraction of the classical Hausdorff distance between sets in a given metric space. Thus two metric spaces have Gromov-Hausdorff distance $\leq \varepsilon$ if they can be isometrically embedded into some metric space where they have Hausdorff distance $\leq \varepsilon$. (Note that an isometric embedding means a distance preserving map; thus Euclidean spheres are not isometrically embedded into Euclidean space.) The useful result here is that the class of Riemannian n-manifolds with $\mathrm{Ric} \geq (n-1)\,k$ and $\mathrm{diam} \leq D$ is precompact in the Gromov-Hausdorff topology for any $k \in \mathbb{R}$ and $D \in (0, \infty)$. The idea of the proof is to use *the relative volume comparison estimate*. This estimate was first observed by Bishop around 1960 for small balls (see [17]), and then extended to hold for all balls by Gromov around 1980. Gromov was also the first to observe the importance of this result in many different contexts. The relative volume comparison result is a type of monotonicity formula that asserts that if $\operatorname{Ric}(M^n, g) \geq (n-1)\,k$, then the volume ratio with constant curvature

$$\frac{\operatorname{vol}(B(p,r) \subset M)}{\operatorname{vol}(B(\bar{p},r) \subset S^n_k)}$$

is decreasing. Given that M also has bounded diameter, one can from this inequality conclude that M can be covered by $N(\varepsilon)$ balls of radius ε, where the function $N(\varepsilon)$ depends only on n, k, D. This is the idea behind establishing the precompactness theorem mentioned above. With the help of the precompactness result, one can under very minimal hypotheses always assume that a sequence of closed Riemannian manifolds converges to a compact metric space in the Gromov-Hausdorff topology. From this weak convergence result, one can then with more stringent hypotheses try to get better information about the limit space and then use this to study the sequence or class under investigation.

One of the most useful stronger topologies on closed Riemannian manifolds is that of the $C^{k,\alpha}$ *topology*. We say that $(M_i, g_i) \to (M, g)$ in the $C^{k,\alpha}$ topology if for large i there are diffeomorphisms $f_i : M \to M_i$, such that the pullback metrics $f_i^* g_i \to g$ in the $C^{k,\alpha}$ topology on M. Note that even for a sequence of metrics on a given manifold one might not have convergence without moving the metrics in the sequence by a gauge transformation. Thus, our convergence concept is gauge invariant. Note that in each of these stronger topologies a given class of Riemannian manifolds can only be precompact if it contains finitely many diffeomorphism types. In order to see when one obtains precompactness in these topologies, we introduce a more modern concept: The $C^{k,\alpha}$*-norm on the scale of* r of a closed Riemannian manifold (M^n, g), $\|(M,g)\|_{r,C^{k,\alpha}}$ measures (in the $C^{k,\alpha}$ sense) how the metric deviates from the Euclidean metric (δ_{ij}) in coordinates $\phi : B(0,r) \subset \mathbb{R}^n \to M$. The idea is that Euclidean space has norm zero on all scales, flat manifolds have zero norm on small scales, and all Riemannian manifolds have small norm on small scales. The content of the developments Cheeger started with his thesis is then that for fixed but arbitrary r, Q, the set of closed Riemannian manifolds with $C^{k,\alpha}$-norm $\leq Q$ on the scale of r is precompact in the $C^{k,\beta}$ topology for each $\beta < \alpha$. This, of course, makes the theory look like Hölder's extension of the Arzela-Ascoli lemma.

From a geometric point of view we are now left with the question of how these norms are bounded in terms of the geometry. Under the conditions

$$\begin{aligned} |\sec| &\leq K, \\ \operatorname{diam} &\leq D, \\ \operatorname{vol} &\geq v, \end{aligned}$$

we saw that the exponential map made the C^0 norm bounded on the scale of the injectivity radius. The fact that one only has a C^0 bound of course complicates matters a bit, and this was something Cheeger had to worry about. However, using distance functions as coordinates gives us C^1 bounds on the scale of the injectivity radius. This follows from the Hessian estimates for distance functions. But this doesn't seem to have been used until 1980 (see [59]). Even better bounds can be obtained if one uses *harmonic coordinates* as in [75]. Harmonic coordinates have the property that the gradient of the coordinate functions is harmonic. As early as 1922 harmonic coordinates (see [83] and [38]) were used in relativity theory to simplify the description of gravitational waves. They were also used in [37] to show that the metric has the best regularity properties in such coordinates. If we polarize the Bochner identity to a formula for the inner product of harmonic vector fields X, Y, then we have

$$\frac{1}{2}\Delta g\left(X,Y\right) = g\left(\nabla X, \nabla Y\right) + \text{ Ric}\left(X,Y\right).$$

If we use $X = \nabla x^i$ and $Y = \nabla x^j$ for a harmonic coordinate system, then we arrive at

$$\frac{1}{2}\Delta g\left(\nabla x^i, \nabla x^j\right) = g\left(\nabla^2 x^i, \nabla^2 x^j\right) + \text{ Ric}\left(\nabla x^i, \nabla x^j\right).$$

Writing ∇x^k and $\nabla^2 x^k$ out in terms of the metric g then yields a formula of the type

$$\frac{1}{2}\Delta g_{ij} = Q\left(g, \partial g\right) + \text{ Ric}_{ij},$$

where Q is a universal function in the metric components and its derivatives. From this equation one sees that if harmonic coordinates exist on a certain size region, where one has already obtained C^1 bounds on the metric and has bounds on the Ricci curvature, then one gets $C^{1,\alpha}$ bounds on the metric as well from standard elliptic theory (see, e.g., [51]). The problem, of course, is to show the existence of such harmonic coordinates. Jost and Karcher in the above-mentioned paper managed to do this for the class of manifolds that Cheeger considered. Thus for each α one can bound the $C^{1,\alpha}$-norm on a given scale for this class. Consequently one also gets $C^{1,\alpha}$ precompactness for this class and, in particular, finiteness of diffeomorphism types. This result is known as the *convergence theorem of Riemannian geometry*.

At this point it is worthwhile to explain a little bit about who did what, aside from what has already been mentioned. Suppose we have a sequence of Riemannian manifolds satisfying Cheeger's conditions which converges in the Gromov-Hausdorff topology to a metric space. Around 1980 Nikolaev was already aware that this metric space has bounded curvature in the comparison sense, and from this he was able to conclude that it must be a smooth manifold with a $C^{1,\alpha}$ Riemannian metric (see [102]). Using the harmonic coordinates of Jost and Karcher, Peters and Greene-Wu around 1987 showed that the convergence actually happens in the Lipschitz topology (see [52] and [111]). This topology is, however, weaker than even the C^0 topology, and thus the above statement is much stronger. Probably Kasue was the first to state the result as it is stated above (see [76]).

In 1990 Anderson discovered a new way of obtaining bounds on the metric. The idea was to use a contradiction type argument in connection with harmonic coordinates (see [5]), the main point being that one gets a boost in the regularity of the metric from using these coordinates, as we just explained. His idea then centers around assuming that one can't find these coordinates on certain size balls and extract a contradiction from this, rather than attacking the problem directly, as Jost and Karcher did. With this, Anderson obtained a $C^{1,\alpha}$ precompactness result for the class satisfying

$$\begin{aligned} |\text{Ric}| &\leq K, \\ \text{diam} &\leq D, \\ \text{inj} &\geq i. \end{aligned}$$

In this case a lower volume bound no longer suffices, as Cheeger's lemma for estimating the injectivity radius fails when one only has bounded Ricci curvature (see, however, [6]).

For more on compactness and collapsing results (and how this fits into the language of norms) we refer the reader to Petersen's article in [62], and also to [112], [117]. Fukaya in [45] has an extensive survey of Gromov-Hausdorff convergence and its uses, including a proof of Gromov's almost flat manifold theorem.

The latest development in convergence theory came with Colding's work and its elaborations by Cheeger and Colding. This work has significantly enhanced our understanding of manifolds with a lower Ricci curvature bound. We explain very briefly some of the main points here. It will be convenient to generalize the norm concept from above. The *Reifenberg norm on the scale of* r of an n-dimensional metric space (X, d) is simply

$$\|(X,d)\|_r = r^{-1} \cdot \max_{p \in X} d_{G-H}\left(B\left(p,r\right), B\left(0,r\right)\right),$$

where d_{G-H} denotes the Gromov-Hausdorff distance and $B(0,r) \subset \mathbb{R}^n$. A Riemannian manifold has the property that this norm goes to zero as the scale goes to zero. Conversely if this norm is sufficiently small for all small scales, then the metric space is in a weak sense a Riemannian manifold (see [27]). This extends an older result by Reifenberg for subspaces of Euclidean space. The other important information Colding obtains in [36] is that for a manifold with a given lower Ricci curvature bound, the Reifenberg norm on the scale of r is small iff the volume of r-balls is close to the volume of a Euclidean r-ball. From relative volume comparison one then has that in the presence of a lower Ricci curvature bound the Reifenberg norm is small on all scales provided it is small on just one scale. This gives one an amazing control over the metric, and it can be used to show some very nice results for manifolds with almost maximal volume that extend Perel'man's results mentioned above. In particular, one obtains differentiable stability rather than just topological stability for a sequence of manifolds with these properties. One of the main ideas in this new development of Colding is an integral version of Toponogov's comparison result for thin triangles as in the excess estimate. This, in turn, comes from a new L^2 bound on the Hessian for distance functions. Recall that Toponogov's result depended on an upper bound for the Hessian. The fact that one can get an L^2 bound from lower bounds on Ricci curvature might seem strange, but in the places where the Hessian is very negative the metric develops conjugate points, and this means that the volume form in these places is very small as well. Thus, when integrating, these nasty spots disappear. In [116] much of this work has been generalized to the situation where one has integral Ricci curvature bounds.

In a completely different direction Rong has gained an understanding of convergence when one has positive curvature. Some of his results can be paraphrased as follows: If the universal covering space of a positively curved manifold admits a nontrivial Killing field, then the fundamental group contains a cyclic subgroup whose index is bounded from above by the dimension. This result is, of course, only interesting in odd dimensions, given Synge's result, but the important point is that manifolds with small volume and curvature bounds of the type $0 < \delta \leq \sec \leq K$ have nearby metrics with sufficient symmetry that one almost gets the desired

Killing field. In any case, the result still holds for such manifolds (see [122]). Another interesting result by Rong can be found in [123], where he extends Bochner's theorem of the nonexistence Killing fields in negative Ricci curvature.

6.3. **More on the Bochner technique and other analytic methods.** In a very interesting development, Gromov and Lawson managed to almost classify simply connected manifolds with positive scalar curvature and also established some new restrictions on manifolds with nonnegative scalar curvature. They did this work in 1980. Similar but weaker results were also obtained by Schoen and Yau just prior to this. The classification (completed by Stolz in [129]) simply says that either the manifold is not spin, in which case it admits positive scalar curvature, or if it is spin it only has positive scalar curvature when $\hat{A} = 0$ (thus yielding a converse to Lichnerowicz's theorem). The new restrictions they obtained, in particular, generalize the rigidity result of Eells and Sampson mentioned above. Namely, they show that on a closed spin manifold which admits nonpositive sectional curvature, any metric with nonnegative scalar curvature must be flat. Their result is actually much more general and also works for infrasolvemanifolds (spaces covered by solvable Lie groups). The story is explained nicely in [82].

Another important technique was introduced by Hamilton in 1982. He studied the *Ricci flow* on a Riemannian manifold

$$\partial_t g_t = -2\mathrm{Ric}_{g_t}.$$

This flow deforms the metric in the direction of the Ricci tensor. The equation, while not quite parabolic, is nice enough that one can show local existence for this flow. The problem then is to study to what extent one gets long time existence. Note that if one starts with positive curvature, then the metric will shrink in finite time, as the curvature will always increase. This can be partially averted by suitably rescaling the time variable. Hamilton obtains two results that so far have not been obtained with other methods: (1) Any closed three manifold with positive Ricci curvature admits constant positive curvature. (2) Any closed four manifold with positive curvature operator admits a metric with constant curvature. These results and many other things are explained quite well in the article [67], and for some other interesting results that use the Ricci flow see the articles by Min-Oo and Ruh and also Nishikawa in [126]. The short time Ricci flow can also be used to give a different proof of the Gallot-Meyer classification of manifolds with nonnegative curvature operator (see [34] and Chen's article in [53, Vol. 3]). This work relies heavily on other important developments by Mok (see [94]) and Micallef-Moore (see the next paragraph).

In a different direction, Micallef and Moore in the mid 1980s (see, in particular, [90]) were able to generalize Synge's result on closed geodesics in positive curvature to a similar result for harmonic two-spheres. First they obtain a second variation formula for such harmonic two-spheres and rewrite it in such a way that one gets the curvature term to depend on what is known as the *isotropic curvature,* which is the sectional curvature on isotropic two planes when one complexifies the tangent bundle. Then they use this formula to show that any closed simply connected n-manifold $n \geq 4$ with positive isotropic curvature is a homotopy sphere. This result is interesting, not only in its own right, but also because both pointwise quarter pinching and positivity of the curvature operator imply positivity of the isotropic curvature. Thus they generalized simultaneously the hypothesis of the classical sphere theorem from section 5.1 and the conclusion of D. Meyer's theorem from section 5.3. It is an interesting problem to study manifolds with positive (or nonnegative) isotropic curvature that are not simply connected, as this class is closed under taking connected sums (see [91]). Hence the fundamental group of these spaces can be quite large, something which is not possible when one assumes positive Ricci curvature. It is possible that while the Ricci curvature is known to control one forms, the isotropic curvature controls forms of higher degree and perhaps better yet the integral homology interpreted as $H^p\left(M, S^1\right) = H^{p+1}(M, \mathbb{Z})$, $p = 1, \ldots, n-3$, $n \geq 4$. In [91] the authors show that in even dimensions isotropic curvature controls two-forms, and also that this is not true for odd dimensional manifolds. Finally in [68] Hamilton uses the Ricci flow to classify four manifolds with positive isotropic curvature.

7. Bibliography

The book that most extensively covers the material discussed here in greater detail is [112], but there are other books which also cover other aspects of Riemannian geometry. See, for instance, [80] (gives the most comprehensive treatment of the variational calculus of geodesics), [47] (does a little bit of everything including eigenvalues), [23] (also covers curvature free geometry and eigenvalues), and [103] (also covers general relativity). For the Bochner technique we have [112], [141], and [42] (which emphasizes maps), and with more on how the Bochner technique can be used to estimate topology there is [9]. Finally, on how spin geometry unifies some of the Bochner techniques see [82]. Finiteness and convergence theorems are covered in [112] and [29], and for more general and recent material there is [62]. This last book contains surveys on sphere theorems, manifolds of nonnegative curvature, Ricci curvature, convergence theorems and much more. Most of the articles in this book give at least a good indication of proof techniques as well. M. Berger has written a much more extensive survey of Riemannian geometry (see [10]). There are two other nice survey articles that are particularly appropriate in the context of this article, namely, the articles by Abresch and Kreck in [49].

The bibliography contains several articles which have not been mentioned in the text. The reader who wishes to learn more about nonpositive curvature is referred to the comprehensive text [41]; two other goods texts on this subject are [7] and [8]. For more on examples and counterexamples there are [4], [3], [16], [15], the articles by Abresch-Meyer and Perel'man in [62], [104], and the article by Sha and Yang in [53].

References

[1] U. Abresch and D. Gromoll, *On complete manifolds with nonnegative Ricci curvature*, J. Amer. Math. Soc. 3 (1990) 355-374. MR **91a:**53071

[2] C.B. Allendoerfer and A. Weil, *The Gauss-Bonnet theorem for Riemannian polyhedra*, Trans. Amer. Math. Soc. 53 (1943) 101-129. MR **4:**169e

[3] M.T. Anderson, *Short geodesics and gravitational instantons*, J. Differential Geom. 31 (1990) 265-275. MR **91b:**53040

[4] M.T. Anderson, *Metrics of positive Ricci curvature with large diameter*, Manu. Math. 68 (1990) 405-415. MR **91g:**53045

[5] M.T. Anderson, *Convergence and rigidity of manifolds under Ricci curvature bounds*, Invent. Math. 102 (1990) 429-445. MR **92c:**53024

[6] M.T. Anderson and J. Cheeger, *Diffeomorphism finiteness for manifolds with Ricci curvature and $L^{n/2}$ curvature bounded,* Geom. and Funct. Anal. 1 (1991), 231-251. MR **92h:**53052

[7] W. Ballmann, *Lectures on spaces of non-positive curvature*, Basel: Birkhäuser, 1995. MR **97a:**53053

[8] W. Ballmann, V. Schroeder and M. Gromov, *Manifolds of non-positive curvature*, Boston: Birkhäuser, 1985. MR **87h:**53050

[9] P. Bérard, *From vanishing theorems to estimating theorems: the Bochner technique revisited*, Bull. Amer. Math. Soc. (N.S.) 19 (1988), 371-406. MR **89i:**58152

[10] M. Berger, *Riemannian geometry during the second half of the twentieth century*, Jahresber. Deutsch. Math.-Verein. 100 (1998), no. 2, 45–208. CMP 98:16

[11] M. Berger, *Les variétés riemanniennes à courbure positive*, Bull. Soc. Math. Belg. 10 (1958) 88-104. See also Bull. Soc. Math. France 87 (1959), 285–292; C. R. Acad. Sci Paris 247 (1958), 1165–1168. MR **23:**A1338; MR **20:**2754

[12] M. Berger, *Le variétés riemanniennes 1/4-pincées*, C. R. Acad. Sci. Paris 250 (1960) 442-444. MR **22:**1865

[13] M. Berger, *Le variétés riemanniennes 1/4-pincées*, Ann. Scuole Norm. Sup. Pisa 14 (1960), 161-170. MR **25:**3478

[14] M. Berger, *On the diameter of some Riemannian manifolds*, preprint, Berkeley, 1962.

[15] A.L. Besse, *Einstein Manifolds*, Berlin-Heidelberg: Springer-Verlag, 1987. MR **88f:**53087

[16] A.L. Besse, *Manifolds all of whose geodesics are closed*, Berlin-Heidelberg: Springer-Verlag, 1978. MR **80c:**53044

[17] R.L. Bishop and R.J. Crittenden, *Geometry of Manifolds*, New York: Academic Press, 1964. MR **29:**6401

[18] W. Blaschke, *Vorlesungen über Differentialgeometrie*, Berlin-Heidelberg: Springer Verlag, first edition in 1921. MR **7:**391g

[19] S. Bochner, *Vector fields and Ricci curvature*, Bull. Amer. Math. Soc. 52 (1946), 776-797. MR **8:**230a

[20] O. Bonnet, *Sur quelques propriétés des lignes géodésiques*, C. R. Acad. Sci. Paris 40 (1855), 1311-1313.

[21] E. Calabi, *An extension of E. Hopf's maximum principle with an application to Riemannian geometry*, Duke Math. J. 25 (1958), 45-56. MR 19,1056e

[22] E. Cartan, *Geometry of Riemannian spaces*, Boston: Math. Sci. Press, 1983. First French edition in 1925 and second French edition 1951. MR **85m:**53001

[23] I. Chavel, *Riemannian Geometry—A Modern Introduction*, New York: Cambridge University Press, 1993. MR **95j:**53001

[24] J. Cheeger, *Comparison and finiteness theorems for Riemannian manifolds*, Ph.D. thesis, Princeton University, 1967.

[25] J. Cheeger, *Finiteness theorems for Riemannian manifolds,* Amer. J. Math. 92 (1970), 61-74. MR **41:**7697

[26] J. Cheeger and T.H. Colding, *Lower bounds on Ricci curvature and almost rigidity of warped products*, Ann. Math. 144 (1996), 189-237. MR **97h:**53038

[27] J. Cheeger and T.H. Colding, *On the structure of spaces with Ricci curvature bounded below I*, J. Differential Geom. 46 (1997), 406–480. MR **98k:**53044

[28] J. Cheeger and D. Ebin, *Comparison theorems in Riemannian geometry*, North-Holland, 1975. MR **56:**16538

[29] J. Cheeger, Critical points of distance functions and applications to geometry, *Geometric Topology: Recent developments*, Lecture Notes in Math., 1504, Berlin-Heidelberg: Springer-Verlag, 1991. MR **94a:**53075

[30] J. Cheeger and D. Gromoll, *The splitting theorem for manifolds of nonnegative Ricci curvature,* J. Differential Geometry 6 (1971) 119-128. MR **46:**2597

[31] J. Cheeger and D. Gromoll, *On the structure of complete manifolds of nonnegative curvature*, Ann. Math. 96 (1972), 413-443. MR **46:**8121

[32] J. Cheeger, K. Fukaya, and M. Gromov, *Nilpotent structures and invariant metrics on collapsed manifolds,* J. Amer. Math. Soc. 5 (1992), 372-372. MR **93a:**53036

[33] S.S. Chern, ed., *Global Differential Geometry*, 2nd edition, MAA Studies 27, Washington: Mathematical Association of America, 1989. MR **90d:**53003

[34] B. Chow and D. Yang, *Rigidity of nonnegatively curved compact quaternionic-Kähler manifolds*, J. Differential Geom. 29 (1989), 361-372. MR **91f:**53065

[35] S. Cohn-Vossen, *Kützeste Wege und Totalkrümming auf Flächen*, Comp. Math. 2 (1935), 69-133.

[36] T.H. Colding, *Ricci curvature and volume convergence*, Ann. Math. 145 (1997), 477-501. MR **98d:**53050

[37] D. DeTurk and J. Kazdan, *Some regularity theorems in Riemannian geometry*, Ann. Sci. École Norm. Sup. (4) 14 (1981) 249-260. MR **83f:**53018

[38] P.A.M. Dirac, *General Theory of Relativity*, First published by Wiley in 1975, reprinted by Princeton University Press 1996. MR **97a:**83002

[39] P. Dombrowski, *150 years after Gauss' "Disquisitiones generales circa superficies curvas"*, Asterique 62, 1979. MR **80g:**01013

[40] W. Dyck, *Beitrage zur Analysis Situs I. Ein- und zwei-dimensionale Mannifaltigkeiten*, Math. Ann. 32 (1888), 457-512.

[41] P.B. Eberlein, *Geometry of nonpositively curved manifolds*, Chicago Lectures in Mathematics, Chicago University Press, 1996. MR **98h:**53002

[42] J. Eells and L. Lemaire, *A report on harmonic maps*, Bull. London Math. Soc. 10 (1978), 1-68. MR **82b:**58033

[43] J.-H. Eschenburg, *Local convexity and nonnegative curvature—Gromov's proof of the sphere theorem*, Invent. Math. 84 (1986) 507-522. MR **87j:**53080

[44] J.-H. Eschenburg and E. Heintze, *An elementary proof of the Cheeger-Gromoll splitting theorem*, Ann. Global Anal. and Geom. 2 (1984) 141-151. MR **86h:**53042

[45] K. Fukaya, *Hausdorff convergence of Riemannian manifolds and its applications*, Advanced Studies in Pure Math. 18-I (1990) Recent topics in Differential and Analytic Geometry pp. 143-238. MR **92k:**53076

[46] S. Gallot, *Isoperimetric inequalities based on integral norms of Ricci curvature*, Astérisque, 157-158 (1988), 191-216. MR **90a:**58179

[47] S. Gallot, D. Hulin and J. Lafontaine, *Riemannian Geometry*, Berlin-Heidelberg: Springer-Verlag, 1987. MR **88k:**53001

[48] S. Gallot and D. Meyer, *Opérateur de courbure et laplacien des formes différentielles d'une variété riemannienne*, J. Math. Pures Appl. 54 (1975), 259-284. MR **56:**13128

[49] W.-D. Geyer, ed., *Jahresbericht der Deutschen Mathematiker-Vereinigung: Jubiläumstagung, 100 Jahre DMV, Bremen 1990*, Stuttgart : B.G. Teubner, 1992.

[50] G. Gibbons and S. Hawking, *Gravitational multi-instantons*, Phys. Lett. B 78 (1978) 430-432.

[51] D. Gilbarg and N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 2nd edition, Berlin-Heidelberg: Springer-Verlag, 1983. MR **86c:**35035

[52] R. Greene and H.-H. Wu, *Lipschitz convergence of Riemannian manifolds*, Pac. J. Math. 131 (1988), 119-141. MR **90j:**53061

[53] R. Greene and S.T. Yau, eds., *Differential Geometry*, Proc. Symp. Pure Math. 54, vol. 3, Amer. Math. Soc. (1993). MR **94a:**00012

[54] D. Gromoll, W. Klingenberg, and W. Meyer, *Riemansche Geometrie im Grossen*, Lecture Notes in Mathematics, 55, Berlin-Heidelberg: Springer-Verlag, 1968 and 1975. MR **51:**1651

[55] D. Gromoll and W. Meyer, *On complete open manifolds of positive curvature*, Ann. Math. 90 (1969), 74-90. MR **40:**854

[56] M. Gromov, *Almost flat manifolds,* J. Differential Geom. 13 (1978), 231-241. MR **80h:**53041

[57] M. Gromov, *Manifolds of negative curvature*, J. Differential Geom. 13 (1978), 223-230. MR **80h:**53040

[58] M. Gromov, *Curvature, diameter, and Betti numbers*, Comment. Math. Helv. 56 (1981), 179-195. MR **82k:**53062

[59] M. Gromov, *Structures métriques pour les variétés riemanniennes* (J. Lafontaine and P. Pansu, eds.) Paris: Cedic/Fernand Nathan, 1981. MR **85e:**53051

[60] M. Gromov and W. Thurston, *Pinching constants for hyperbolic manifolds*, Invent. Math. 89 (1987) 1-12. MR **88e:**53058

[61] K. Grove, H. Karcher, and E. Ruh, *Group actions and curvature*, Invent. Math. 23 (1974), 31-48. MR **52:**6609

[62] K. Grove and P. Petersen, eds., *Comparison Geometry*, MSRI publications vol. 30, New York: Cambridge University Press, 1997. MR **97m:**53001

[63] K. Grove and P. Petersen, *Bounding homotopy types by geometry*, Ann. Math. (1988), 195-206. MR **90a:**53044

[64] K. Grove, P. Petersen, and J.-Y. Wu, *Geometric finiteness theorems via controlled topology*, Invent. Math. 99 (1990) 205-213, *Erratum* Invent. Math. 104 (1991) 221-222. MR **90k:**53075; MR **92b:**53065

[65] K. Grove and K. Shiohama, *A generalized sphere theorem*, Ann. Math. (1977), 201-211. MR **58:**18268

[66] J. Hadamard, *Les surfaces à courbures opposées*, J. Math. Pures Appl. 4 (1889), 27-73.

[67] R.S. Hamilton, *The formation of singularities in the Ricci flow*, Surveys in Differential Geometry, Vol. II, International Press (1995), 7-136. MR **97e:**53075

[68] R.S. Hamilton, *Four-manifolds with positive isotropic curvature*, Comm. Anal. Geom. 5 (1997), 1-92. MR **99e:**53049

[69] S.W. Hawking and G.F.R Ellis, *The large scale structure of space-time*, Cambridge University Press, 1973. MR **54:**12154

[70] E. Heintze and H. Karcher, *A general comparison theorem with applications to volume estimates for submanifolds*, Ann. Sci. École Norm. Sup. 11 (1978), 451-470. MR **80i:**53026

[71] W.V.D. Hodge, *The existence theorem for harmonic integrals*, Proc. London Math. Soc. 41 (1936), 483-496.

[72] H. Hopf, *Zum Clifford-Kleinschen Raumproblem*, Math. Ann. 95 (1926), 313-339.

[73] H. Hopf, *Über die Curvatura integra geschlossener Hyperflächen*, Math. Ann. 95 (1926), 340-367.

[74] H. Hopf and W. Rinow, *Über den Begriff der vollständigen differential-geometrischen Flächen*, Comment. Math. Helv. 3 (1931), 209-225.

[75] J. Jost and H. Karcher, *Geometrische Methoden zur Gewinnung von a-priori-Schranken für harmonische Abbildungen*, Manuscripta Math. 40 (1982) 27-77. MR **84e:**58023

[76] A. Kasue, *A convergence theorem for Riemannian manifolds and some applications*, Nagoya Math. J. 114 (1989), 21-51. MR **90g:**53053

[77] P. Kirby and L. Siebenmann, *Foundational essays on topological manifolds, smoothings, and triangulations*, Ann. Math. Stud. 88, Princeton Univ. Press (1977). MR **58:**31082

[78] W. Klingenberg, *Contributions to Riemannian geometry in the large*, Ann. Math. 69 (1959), 654-666. MR **21:**4445

[79] W. Klingenberg, *Über Riemannsche Mannigfaltigheiten mit positiver Krümmung*, Comment. Math. Helv. 35 (1961), 47-54. MR **25:**2559

[80] W. Klingenberg, *Riemannian Geometry*, de Gruyter Studies in Mathematics, 1, Berlin, 1982. MR **84j:**53001

[81] L. Kronecker, *Über Systeme von Funktionen mehrer Variablen*, Monatsberichte Berlin Akad. (1869), 159-193 and 688-698.

[82] H.B. Lawson Jr. and M.-L. Michelsohn, *Spin Geometry*, Princeton: Princeton University Press, 1989. MR **91g:**53001

[83] C. Lanczos, *Ein vereinfachendes Koordinatensystem für die Einsteinschen Gravitationsgleichungen*, Phys. Z. 23 (1922) 537-539.

[84] J.M. Lee, *Riemannian Manifolds*, Graduate Texts in Mathematics, 176, Berlin-Heidelberg: Springer Verlag (1997). MR **98d:**53001

[85] T. Levi-Civita, *The absolute differential calculus*, London and Glasgow: Blackie and Son, first Italian ed. 1923, English ed. in 1926, and revised ed. in 1961.

[86] A. Lichnerowicz, *Géométrie des groupes de transformations*, Travaux et Recherches Mathématiques, III. Dunod, Paris, 1958. MR **23:**A1329

[87] A. Lichnerowicz, *Propagateurs et Commutateurs en relativité générale*, Inst. Hautes Études Sci. Publ. Math. No. 10 (1961) 293–343. MR **28:**967

[88] A. Lichnerowicz, *Spineurs harmoniques*, C. R. Acad. Sci. Paris 257 (1963), 7-9. MR **27:**6218

[89] H. v. Mangoldt, *Über diejenigen Punkte aur positiv gekrummten Flächen, welche die Eigenschaft haben, dass die von ihnen ausgehenden geodätischen Linien nie aufhören, kürzeste Linien zu sien*, J. Reine Angew. Math. 91 (1881), 23-52.

[90] M.J. Micallef and J.D. Moore, *Minimal 2-spheres and the topology of manifolds with positive curvature on totally isotropic 2-planes*, Ann. of Math. 127 (1988), 199-227. MR **89e:**53088

[91] M.J. Micallef and M.Y. Wang, *Metrics with nonnegative isotropic curvature*, Duke Math. J. 72 (1993), 649-672. MR **94k:**53052

[92] J.W. Milnor, *Morse Theory*, Princeton University Press, Princeton, N.J., 1963. MR **29:**634

[93] C.W. Misner, K.S. Thorne and J.A. Wheeler, *Gravitation*, W. H. Freeman and Co., San Francisco, Calif., 1973. MR **54:**6869

[94] N. Mok, *The uniformization theorem for compact Kähler manifolds of non-negative holomorphic bi-sectional curvature*, J. Diff. Geom. 27 (1988), 179-214. MR **89d:**53115

[95] J.W. Morgan, *The Seiberg-Witten equations and applications to the topology of smooth four-manifolds*, Princeton Univ. Press, Princeton, N.J., 1996. MR **97d:**57042

[96] M. Morse, *Relations between the critical points of a real analytic function of n independent variables*, Trans. Amer. Math. Soc. 27 (1925) 345-396.

[97] M. Morse, *The calculus of variations in the large*, AMS Coll. Publ. 18, New York, 1934.

[98] S.B. Myers, *Riemannian manifolds in the large*, Duke Math. J. 1 (1935), 39-49.

[99] S.B. Myers, *Riemannian manifolds with positive mean curvature*, Duke Math. J. 8 (1941) 401-404. MR 3,18f

[100] S.B. Myers and N. Steenrod, *The group of isometries of a Riemannian manifold*, Ann. Math. 40 (1939), 400-416.

[101] P. Nabonnand, *Contribution à l'histoire de la théorie des géodésiques au XIX^e siècle,* Rev. Histoire Math. 1 (1995) 159-200. MR **96j:**01027

[102] I.G. Nikolaev, *Parallel translation and smoothness of the metric of spaces of bounded curvature*, Dokl. Akad. Nauk SSSR 250 (1980), 1056-1058. MR **81d:**53052

[103] B. O'Neill, *Semi-Riemannian Geometry*, Pure and Applied Mathematics, 103, Academic Press, New York-London, 1983. MR **85f:**53002

[104] Y. Otsu, *On manifolds of positive Ricci curvature with large diameters*, Math. Z. 206 (1991) 255-264. MR **91m:**53033

[105] M. Özaydın and G. Walschap, *Vector bundles with no soul*, Proc. Amer. Math. Soc 120 (1994) 565-567. MR **94d:**53057

[106] R. Penrose, *Techniques of differential topology in relativity*, CBMS-NSF Regional Conference Series in Applied Mathematics 7, 1973. MR **57:**8942

[107] G. Perel'man, *Proof of the soul conjecture of Cheeger and Gromoll*, J. Diff. Geom. 40 (1994), 209-212. MR **95d:**53037

[108] G. Perel'man, *Alexandrov's spaces with curvatures bounded from below II*, preprint.

[109] G. Perel'man, *Manifolds of positive Ricci curvature with almost maximal volume*, J. Amer. Math. Soc. 7 (1994) 299-305. MR **94f:**53077

[110] S. Peters, *Cheeger's finiteness theorem for diffeomorphism classes of manifolds*, J. Reine Angew. Math. 349 (1984) 77-82. MR **85j:**53046

[111] S. Peters, *Convergence of Riemannian manifolds*, Compositio. Math. 62 (1987), 3-16. MR **88i:**53076

[112] P. Petersen, *Riemannian geometry*, Graduate Texts in Mathematics, 171, Springer-Verlag, New York, 1998. MR **98m:**53001

[113] P. Petersen, S. Shteingold and G. Wei, *Comparison geometry with integral curvature bounds*, Geom. Funct. Anal. 7 (1997), 1011-1030. MR **99c:**53022

[114] P. Petersen and C. Sprouse, *Integral curvature bounds, distance estimates and applications*, J. Diff. Geom. 50 (1998), 269–298.

[115] P. Petersen and G. Wei, *Relative volume comparison with integral Ricci curvature bounds*, Geom. Funct. Anal. 7 (1997) 1031-1045. MR **99c:**53023

[116] P. Petersen and G. Wei, *Geometry and analysis on manifolds with integral Ricci curvature bounds*, preprint from UCLA. To appear in Trans. Amer. Math. Soc.

[117] P. Petersen, G. Wei, and R. Ye, Controlled geometry via smoothing, to appear in Comm. Math. Helv.

[118] A. Preissmann, *Quelques propriétés globales des espaces de Riemann*, Comment. Math. Helv. 15 (1943), 175-216. MR **6:**20g

[119] H.E. Rauch, *A contribution to differential geometry in the large*, Ann. Math. 54 (1951), 38-55. MR **13:**159b

[120] G. de Rham, *Differentiable Manifolds*, Berlin-Heidelberg: Springer-Verlag, 1984. MR **85m:**58005

[121] B. Riemann, *Über die Hypothesen welche der Geometrie zu Grunde liegen*, Habilitationschrift 1854; Gött. Abh. 13, 1868; Gesamm. Werke, 2nd ed. Leipzig, 1892. Edited with comments by H. Weyl, Berlin, 1919 and revised in 1923.

[122] X. Rong, *The almost cyclicity of the fundamental groups of positively curved manifolds*, Invent. Math. 126 (1996) 47-64. And *Positive curvature, local and global symmetry, and fundamental groups*, preprint, Rutgers, New Brunswick. MR **97g:**53045

[123] X. Rong, *A Bochner Theorem and Applications*, Duke Math. J. 91 (1998), 381–392. MR **99e:**53043

[124] E. Ruh, *Almost flat manifolds*, J. Diff. Geom. 17 (1982), 1-14. MR **84a:**53047

[125] I.J. Schoenberg, *Some applications of the calculus of variations to Riemannian geometry*, Ann. Math. 33 (1932) 485-495.

[126] K. Shiohama, T. Sakai and T. Sunada, Eds., *Curvature and Topology of Riemannian Manifolds*, Lecture Notes in Mathematics 1201, Berlin-Heidelberg: Springer-Verlag, 1986. MR **87i:**53001

[127] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, vols. I-V, Wilmington: Publish or Perish, 1979. MR **82g:**53003a; MR **82g:**53003b; MR **82g:**53003c; MR **82g:**53003d; MR **82g:**53003e

[128] J.J. Stoker, *Differential Geometry*, New York: Wiley-Interscience, 1989. MR **90f:**53002

[129] S. Stolz, *Simply connected manifolds with positive scalar curvature*, Ann. Math. 136 (1992), 511-540. MR **93i:**57033

[130] J.L. Synge, *The first and second variations of the length-integral in Riemann spaces*, Proc. London Math. Soc. 25 (1926), 247-264.

[131] J.L. Synge, *On the neighborhood of a geodesic in Riemannian space*, Duke Math. J. 1 (1935), 527-537.

[132] J.L. Synge, *On the connectivity of spaces of positive curvature*, Quart. J. Math. 7 (1936), 316-320.

[133] S. Tachibana, *A theorem on Riemannian manifolds of positive curvature operator*, Proc. Japan Acad. 50 (1974), 301-302. MR **51:**1667

[134] V.A. Toponogov, *Riemannian spaces having their curvature bounded below by a positive number*, Uspehi Mat. Nauk 14 (1959), 87-130; Dokl. Akad. Nauk SSSR 120 (1958), 719–721. MR **21:**2278; MR **20:**6139

[135] V.A. Toponogov, *Dependence between curvature and topological structure of Riemann spaces of even dimension*, Soviet Math. Dokl. 1 (1961), 943–945. MR **25:**3476

[136] V.A. Toponogov, *The metric structure of Riemannian spaces of nonnegative curvature which contain straight lines*, Sibirsk. Mat. Z. 5 (1964), 1358-1369; Dokl. Akad. Nauk. SSSR 127 (1959), 977–979. MR **32:**3017; MR **21:**7520

[137] F.W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, New York: Springer-Verlag, 1983. MR **84k:**58001

[138] A. Weinstein, *On the homotopy type of positively pinched manifolds*, Archiv Math. 18 (1967), 523-524. MR **36:**3376

[139] R. Weitzenbock, *Invariantentheorie*, Groningen: Nordhoff, 1923.

[140] F.H. Wilhelm, *On intermediate Ricci curvature and the fundamental group*, Illinois J. of Math. 41 (1997), 488–494. MR **98h:**53065

[141] H.H. Wu, *The Bochner Technique in Differential Geometry*, Mathematical Reports vol. 3, part 2, London: Harwood Academic Publishers, 1988. MR **91h:**58031

Department of Mathematics, University of California, Los Angeles, CA 90095-1555
E-mail address: petersen@math.ucla.edu