universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

"Probability estimates for manifold learning from random samples"

Verfasser

Martin Heuschober

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer. nat.)

Wien, im April 2013

Studienkennzahl lt. Studienblatt:	A 405
Studienrichtung lt. Studienblatt:	Mathematik
Betreuer:	Ass. Prof. Mag. Dr. Privatdoz. Stefan Haller

Abstract

The paper *"Finding the Homology of Submanifolds with High Confidence from Random Samples"*([NSW08]) by Partha Niyogi, Stephen Smale and Shmuel Weinberger shows that a compact submanifold M of $\mathbb{R}^n$ with condition number τ is homotopy equivalent to the union of ε-balls around some sample points $x_1, \dots, x_N \in \mathcal{B}_r(M)$ with probability greater $1 - \delta$, if the samples are taken identically and independently distributed according to a probability measure μ, which has a lower bound $k_s > 0$ for all $\mu(\mathcal{B}_s(x))$, and the parameters N,r,τ,s,ε obey three geometric relations.

This will be extended as follows. Let M be a complete Riemannian manifold with bounded sectional curvature $\kappa \leq \sec \leq \mathcal{K}$ and $S \subseteq M$ a closed submanifold with condition number τ. And let $x_1, \dots, x_N \in \mathcal{B}_r(M)$ such that $S \subseteq \bigcup_{i=1}^N \mathcal{B}_s(x_i)$. Then the union of ε-balls in M around these points, is diffeomorphic to the normal bundle $\mathcal{B}^\perp(S)$, moreover $\exp^{-1}(\mathcal{U})$ is an open and fibrewise star-shaped subset of the normal bundle, if the parameters r,τ,s,ε and the injectivity and convexity radius obey some geometric relations. This will be done with the use of the Rauch Comparison Estimate and Toponogov's Theorem. At last we will combine this result with a probabilistic estimate for the number of sample points to get a similar high confidence result as in [NSW08].

Zusammenfassung

Im Artikel *"Finding the Homology of Submanifolds with High Confidence from Random Samples"*([NSW08]) von Partha Niyogi, Stephen Smale and Shmuel Weinberger wurde gezeigt, dass eine kompakte Teilmannigfaltigkeit M des $\mathbb{R}^n$ mit Konditionszahl τ, homotopieäquivalent zur Vereinigung von ε-Bällen um Samplepunkte $x_1, \dots, x_N \in \mathcal{B}_r(M)$, mit Wahrscheinlichkeit größer als $1 - \delta$ ist, wenn die Punkte unabhängig und identisch verteilt bezüglich eines Wahrscheinlichkeitsmaßes μ gewählt wurden, und für alle x das Maß von $\mu(\mathcal{B}_s(x))$ eine untere Schranke $k_s > 0$ besitzt, und wenn die Parameter N,r,τ,s,ε drei geometrische Bedingungen erfüllen.

Dieses Resultat wird folgendermaßen verallgemeinert. Sei M eine vollständige Riemann'sche Mannigfaltigkeit mit beschränkter Schnittkrümmung $\kappa \leq \sec \leq \mathcal{K}$ und $S \subseteq M$ eine geschlossene Teilmannigfaltigkeit mit Konditionszahl τ. Und seien $x_1, \dots, x_N \in \mathcal{B}_r(M)$ mit der Eigenschaft $S \subseteq \bigcup_{i=1}^N \mathcal{B}_s(x_i)$. Dann ist die Vereinigung von ε-Bällen in M um diese Punkte diffeomorph zum Normalenbündel $\mathcal{B}^\perp(S)$. Wir werden sogar zeigen dass $\exp^{-1}(\mathcal{U})$ eine offene faserweise sternförmige Teilmenge des Normalenbündels ist. Vorausgesetzt dass die Parameter r,τ,s,ε sowie der Injektivitätsradius und Konvexitätsradius einige geometrische Bedingungen erfüllen. Dabei werden wir als zentrales Werkzeug das Vergleichstheorem von Rauch sowie den Satz von Toponogov verwenden. Zum Schluss werden wir dieses Ergebnis mit Wahrscheinlichkeitsabschätzungen kombinieren, um ein analoges "high-confidence" Resultat wie in [NSW08] zu erhalten.

Danksagung

Zu allererst möchte ich meinem Betreuer Stefan Haller danken, für das Thema der Arbeit, das mich nach Jahren des Auseinandersetzens immer noch interessiert. Dafür, dass ich viel von ihm gelernt habe und dass er mir Zusammenhänge schnell anhand eines Beispiels oder Gegenbeispiels klar machen konnte. Dass er mich in unzähligen Gesprächen immer wieder aus mathematischen Sackgassen geleitet hat und mir den nötigen Rat und Hilfe bereitgestellt hat, um diese Arbeit anzufangen und zu beenden. Ich möchte ihm für die Art und Weise mich zu kritisieren und auf Fehler hinzuweisen danken, was mich bestärkt hat weiter zu machen und mich zu verbessern.

Außerdem möchte ich meinen Eltern danken, dass sie mir mein Studium ermöglicht haben und mich die lange Zeit über, sowohl seelisch als auch finanziell unterstützt haben, für ihre Geduld und die Tatsache, dass sie bis zum Schluss an mich geglaubt haben.

Weiters möchte ich Christina danken, die mich im letzten Jahr durch alle Höhen und Tiefen begleitet und mich immer wieder bestärkt hat.

Zum Schluss möchte ich noch Herrn Spraider danken, der manchmal der Einzige war, der mich davon abgehalten konnte, das Handtuch zu werfen.

Contents

Intro

What is manifold learning?

Manifold learning or non-linear dimensionality reduction is an algorithm or mapping to embed a set of high dimensional data into a low dimensional manifold. In some cases this can lead to a visualization of the data. One of the most prominent examples in manifold learning is face/feature recognition. See [ZCPR00] for an overview over the development these years or more recently [LJ11] for further reading.

High dimensional data - and possibly low dimensional substructure

I want to start with an example of face recognition. Let us assume we have a set of pictures of faces, and assume that all of those pictures have the same amount of horizontal and vertical pixels, or at least can be converted in such a format. If we start with the set of 64×64-pixel images as in [LZ08], then we can think of every image as a point in a 4096-dimensional space. It is reasonable, as all human faces have about the same shape, that all of these images have a common substructure, an underlying *"ideal face"*. And given an appropriate metric one should be able to distinguish a picture of a tree, from that of a face. However the task of biometry is still more complicated: One would like to find out whether a given face in one picture is the same as that in another picture, which may be hard due to the fact that the second picture is taken from a different angle, or at a different lighting environment, or that the person is not the same on both pictures. Where in the first two cases an algorithm should confirm the identity with high confidence, in the latter case it should falsify the equality of the pictures. Coming back of the idea of the *"ideal face"*, it should be possible to extract a number of features in all facial pictures, that have a common substructure.

It is unlikely for this substructure to be a linear subspace of this high dimensional base space, so we will try to model it with the tools of smooth manifolds. More precisely in this thesis I will assume that this substructure is a closed submanifold of a Riemannian manifold.

One setting - Riemannian manifold learning

In the setting of Riemannian geometry a lot of algorithms for *"manifold learning"* or *"non-linear dimension reduction"* have been developed, the English Wikipedia entry of *Nonlinear Dimensionality Reduction* alone lists 24 different manifold learning algorithms [Wik12]. For a short comparison see [LZL06] or [LZ08].

another prominent choice - persistent homology

Persistent homology is a framework for computing the homology of a *simplicial complex*. Usually one constructs a filtered complex depending on a parameter ε, which defines whether $(n+1)$-many points are connected to an n-simplex, and a set of sample points. Then one tries to argue why the resulting simplicial complex has the same homology type as the underlying manifold given the current configuration of sample points and ε. Persistent Homology takes a different approach, one calculates many homologies by varying the parameter ε and then filters out the *"short lived"* cycles in homology by embedding each complex with smaller parameter into the ones with greater ε. And thus getting a homology that persists over almost all ε, hence the name.

Why manifold learning?

Data is everywhere

The 21^{st} century could be called the age of data, if by any means we are judged by our primary tools, and the waste we leave behind ourselves. Everyday more people are filling in their profiles at social media sites, posting pictures and tagging them. Hormones and Proteins are being analysed by biologists and put into databases. In Geneva the European Organization for Nuclear Research, is running the LHC and producing approximately 25 Petabytes per year [Gri12]. All of this data contains useful information, which has to be extracted. And there is even data, where one wouldn't expect it, as we will see [LPM03] had a closer look at 3×3 pixel patches taken from natural grey-scale images provided by [vHvdS98] and did find, quite unexpectedly in my opinion, a quite complex structure within it.

Data is hard to manage

Most of this data is provided by measurement, and thus has inherently a few problems. One is that there is a certain inaccuracy due to measurement errors or other human errors. Another problem is the "curse of dimensionality" mentioned before, i.e. that data is often high dimensional but can be embedded in a low dimensional manifold. And after finding a substructure one has to interpret the result and apply heuristics to provide feedback to the problem at hand.

Data has structure - obvious and hidden

At last I would like to show an example given in *"Topology & Data"* by Gunnar Carlsson (see [Car09] for a short summary or [LPM03] for more details). Though the methods used there differ from the ones used in this thesis, I want to spend some time with it as it was the first paper that introduced me to the idea of manifold learning and made me want to explore this subject more thoroughly. Following [LPM03] I will briefly describe the procedure that lead from the grey-scale pictures to data within the manifold of the 7-dimensional sphere.

- Each high-contrast picture was partitioned in 3×3 pixel-squares, which can be viewed as points in $\mathbb{R}^9$, in addition they switched to logarithmic values, to compensate for the variety of intensity found in the optical pictures.
- From the resulting 4.2 million pixel patches Lee, Pedersen and Mumford selected 5000 at random.
- Subtracted mean value and normalizing the contrast for each patch.
- Then they filtered out the lower 80% by contrast, which is measured by the so called *"D-norm"*, and kept 1000 image patches for further processing.
- The next step was to mean centring the *"points"*, which is equivalent to factor out a plane
- at last we normalize the *"D-norm"*.

So we obtain a data set in a 7-sphere. Starting with this set Carlsson shows that in first approximation step one finds a substructure of a circle, in the second step two more circles appear that intersect the first one but not with each other, and in the third and last step one can reason that these three circles happen all to be on a Klein's bottle. All these observations were made with the tool of *"Persistent Homology"*. Which seems to be quite powerful, though I couldn't find any estimates how big – or small the likelihood of an error, that the persistent homology differed from the homology of the underlying manifold. Which lead me to the paper by Partha Niyogi, Stephen Smale and Shmuel Weinberger [NSW08].

Finding the Homology of Submanifolds with High Confidence from Random Samples

The authors discover that provided a *"well behaved"* submanifold M of $\mathbb{R}^n$ and sufficiently dense sample points taken from M guarantee, that the Čech complex obtained from the sample points has the same homology as the underlying manifold. Furthermore they give an estimate, given that the number of sample points, identically independently drawn according to a probability measure, is larger than a certain number, then the ε-Čech complex of the points is homotopy equivalent to the manifold with high confidence. This number can be expressed in terms of the condition number τ, which measures how the manifold is embedded in $\mathbb{R}^n$, the noise r, the $\frac{\varepsilon}{2}$-covering number, the probability measure μ and the error δ one is bound to accept. In addition the noise has to be relatively small compared to the condition number, and the construction number ε for the Čech complex has to be in bounds given in terms of τ and r.

Generalization of [NSW08]

In this thesis I attempt to put the aforementioned result of P. Niyogi, S. Smale and S. Weinberger in the context of Riemannian geometry in the following sense. Replace $\mathbb{R}^n$ with a complete Riemannian manifold M with bounded sectional curvature, injectivity radius $r_{inj} > 0$ and convexity radius $r_{cvx} > 0$. Then I will prove that for a closed submanifold S that has condition number τ then the union $\mathcal{U}$ of ε-balls around r-noisy sample points is homotopy equivalent to S, moreover $\exp^{-1}(\mathcal{U})$ is an open, fibrewise star-shaped subset of the normal bundle if the following conditions hold

$$r < \min\{r_{inj}, \tau - \varepsilon\} \tag{3.1}$$

$$\varepsilon \leq r_{cvx} \tag{3.2}$$

$$2\varepsilon + r \leq r_{inj} \tag{3.3}$$

$$2\varepsilon + r \leq \frac{\pi}{\sqrt{\mathcal{K}}} \text{ in the case of } \mathcal{K} > 0 \tag{3.4}$$

$$\operatorname{tn}_{\mathcal{K}}(\varepsilon + r) \leq \operatorname{tn}_{\kappa}(\tau) \text{ and} \tag{3.5}$$

$$\frac{\operatorname{tn}_{\mathcal{K}}(\varepsilon)}{\operatorname{sn}_{\kappa}(\varepsilon)} \cdot \left(\frac{\operatorname{md}_{\kappa}(\varepsilon) - \operatorname{md}_{\kappa}(r)}{\operatorname{tn}_{\kappa}(\tau)} + \operatorname{sn}_{\kappa}(r) \right) \leq \operatorname{tn}_{\mathcal{K}} \left(\frac{\varepsilon - s}{2} \right) \tag{3.6}$$

In particular ε-Čech complex constructed from the samples is homotopy equivalent to S.

In order to do that we will start to reprove the deformation retraction lemma from [NSW08, Lemma 4.1], with slightly different methods. Which will be used quite similarly in our main argument; most prominently will be the Law of Cosines and the monotonicity of the solution function. Then we will introduce the modifying function for the distance on manifolds md_* and prove some of its basic properties. And relate manifolds of constant sectional curvature with the surrounding manifold, more precisely the distance, by using the famous result by Harry Rauch, to get a generalized version of the Law of Cosines, employing the methods introduced by Hermann Karcher. All this methods lead to the conclusion that the union of balls of radius ε is retractible along geodesics to S. And applying the nerve lemma we get our desired result. Afterwards we spend some time with the special cases of manifolds with constant curvature. In the last section we transform the restriction of the sample points being s-dense into a probabilistic estimate on the number of sample points that asserts that the ε-Čech complex constructed from them is homotopy equivalent to the submanifold S with high confidence.

Though through all of the paper it is assumed that the reader is familiar with basic differential geometry I provide some definitions in Section 2.1, otherwise the interested reader might consult some textbook on differential geometry like [Pet06] or [Jos08]. And as some understanding of homology is needed I provide the necessary tools to understand the nerve lemma, though the prove would need some more knowledge, for an introduction in the subject matter I would recommend to read [Hat02] or [Die08].

I. Setting and short recapitulation

When I introduce the topic of my thesis to laymen I always tell them the following story: *"Imagine a sphere and you have obtained some points on that sphere by a measuring procedure, that all are near the equator of this sphere, then the aim of my work is to give an estimate on how likely it is, that all of these points connected are an approximation of the equator, and on what parameters this probability depends. Obviously two or three points lead to a very bad approximation of the equator, thus more sample points should lead to a higher probability. Another parameter that has to affect the probability is the radius of the sphere, the bigger it is compared to the sampling error, the more likely you should get a good approximation."*

1.1. Finding the Homology of Submanifolds with High Confidence from Random Samples

The original idea that lead to this thesis can be found in the paper *"Finding the Homology of Submanifolds with high Confidence from Random Samples"* by Partha Niyogi, Stephen Smale and Shmuel Weinberger. So I will summarize their work and point out the difference compared to the results in this thesis.

The setting in [NSW08] is divided in two parts. One without noise, where sample points are taken from a compact submanifold and the other part where the samples are noisy, i.e. taken from the tubular neighbourhood of radius r around M. Then the aim then is to learn the homology of the manifold by studying the sample points.

Without noise:

For the first part this is done in three steps. A geometric one that shows that M is a deformation retract of the union of ε-balls around $x_1, \dots, x_N$, provided that ε is small in relation to the condition number and that the sample points are $\frac{\varepsilon}{2}$-dense. The condition number τ is a parameter, which measures how M is embedded in $\mathbb{R}^n$ and bounds its curvature. It is defined as the maximal number τ such that $\mathcal{B}_r^{\perp}(M)$, the open normal bundle of radius $r < \tau$ is embedded in $\mathbb{R}^n$ by the exponential mapping.

Then there is a probabilistic step that asserts sample points are sufficiently dense with probability higher $1-\delta$ if $\mathrm{cov}_{\frac{\varepsilon}{4}}(M)(1-\alpha)^N < \delta$. Where α is a lower bound for the volume of $\frac{\varepsilon}{4}$-balls intersected with M.

The last step is to provide a lower bound for the volume of $\mathcal{B}_\varepsilon^n(p) \cap M$

$$\mathrm{vol}(\mathcal{B}_\varepsilon(p) \cap M) \geq \cos^k(\theta)\,\mathrm{vol}(\mathcal{B}_\varepsilon^k(0)),$$

where $p \in M$ and $\mathcal{B}_\varepsilon^k(x)$ denotes the Euclidean ball in $\mathbb{R}^k$ with centre x, and $\theta := \arcsin(\frac{\varepsilon}{2\tau})$. And estimate the covering number by the packing number, which again is bounded by

$$\mathrm{pck}_\varepsilon(M) \leq \frac{\mathrm{vol}(M)}{\cos^k(\theta)\,\mathrm{vol}(\mathcal{B}_\varepsilon^k(0))}.$$

And relate the condition number τ to the curvature and the norm of the second fundamental form.[1]

Combined this gives the following theorem [NSW08, Theorem 3.1].

I.1 Proposition: *Let M be a compact k-submanifold of $\mathbb{R}^n$ with condition number τ. And let $x_1, \ldots, x_N$ be sample points in M, drawn identically and independently distributed according to the uniform probability measure on M. For $0 < \varepsilon < \frac{\tau}{2}$ we define $\mathcal{U} := \bigcup_{i=1}^N$. Then for all*

$$N > \beta_1\big(\log(\beta_2) + \log(\frac{1}{\delta})\big)$$

the homology of $\mathcal{U}$ equals the homology of M with high confidence (probability $> 1-\delta$), where $\beta_1 := \frac{\mathrm{vol}(M)}{\cos^k(\arcsin(\frac{\varepsilon}{8\tau}))\,\mathrm{vol}(\mathcal{B}_{\varepsilon/4}^k(0))}$ and $\beta_2 := \frac{\mathrm{vol}(M)}{\cos^k(\arcsin(\frac{\varepsilon}{16\tau}))\,\mathrm{vol}(\mathcal{B}_{\varepsilon/8}^k(0))}$.

With noise:

The second part is introducing the notion of noise. This is done by taking a probability measure μ that satisfies the following conditions

$$\mathrm{supp}(\mu) \subseteq \mathcal{B}_r(M) \text{ and}$$
$$\forall s > 0 : \exists k_s > 0 \text{ such that } k_s \leq \inf\{\mu(\mathcal{B}_s(p)) | p \in M\}.$$

[1] Note that in [NSW08] both τ and $(1/\tau)$ are referred to as the condition number.

Then there are two geometric results that provide sufficient conditions that assert $\mathcal{U} := \bigcup_{i=1}^{N} \mathcal{B}_\varepsilon(x_i)$ is homotopy equivalent to M, where $x_1, \ldots, x_N$ are sample points drawn identically and independently distributed according to μ. The latter is the most general and has the following form.

For a compact submanifold M of $\mathbb{R}^n$ with condition number τ, we have sample points $x_1, \ldots, x_N \in \mathcal{B}_r(M)$ such that $M \subseteq \bigcup_{i=1}^{N} \mathcal{B}_s(x_i)$. Then for $\varepsilon > 0$ there are three geometric conditions such that the manifold M is a deformation retract of $\mathcal{U} = \bigcup_{i=1}^{N} \mathcal{B}_\varepsilon(x_i)$

$$(\tau - w)^2 < (\tau - r)^2 - \varepsilon^2 \tag{1.1}$$

$$(\tau - r)^2 - (\tau - \beta)^2 = s^2 - \beta^2 \tag{1.2}$$

$$s^2 - \beta^2 + (\beta + w)^2 = \varepsilon^2, \tag{1.3}$$

where the lengths β and w are as denoted in Figure I.1.

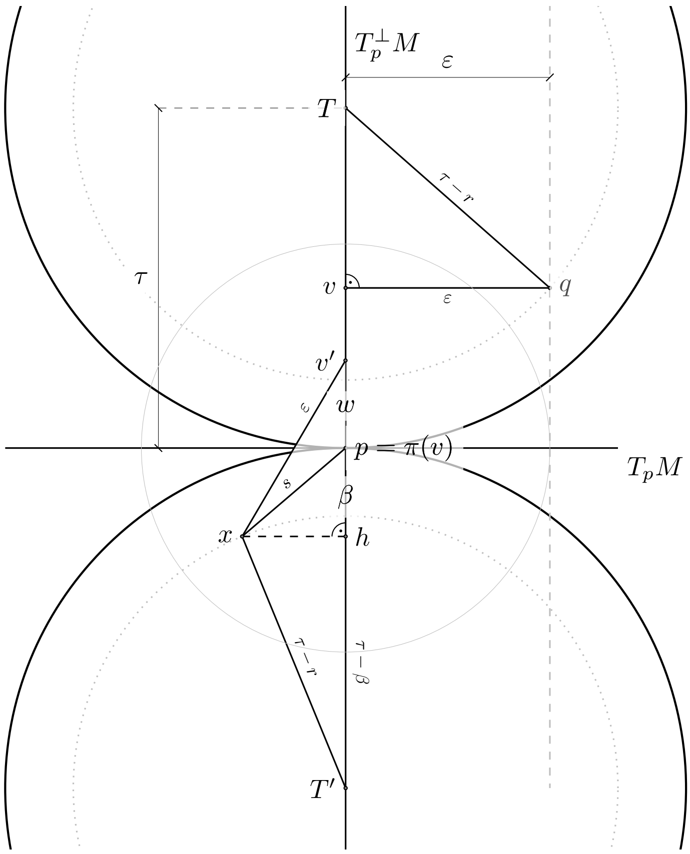


Figure I.1.: the triangles discussed in [NSW08, Lemma 7.1]: The points p,q and v again span a plane, and p,x and v do too, without loss of generality we may assume that these planes coincide. We can retract the line $\bar{pv}_{max}$ if the lines $\mathcal{B}_\varepsilon(x_i) \cap T_p^\perp M$ and $\mathcal{B}_\varepsilon(q) \cap T_p^\perp M$ intersect.

Niyogi, Smale and Weinberger argue that the situation depicted in Figure I.1 is the worst case, for a $q \in \bar{x}$ and for any $v \in \mathcal{B}_\varepsilon(q)$ the line $\bar{v}p$ can be retracted if the distance $\mathrm{d}(v,p) < w$, where w is the maximal length of the segment of $\{t \cdot \vec{pv} | t > 0\} \cap \mathcal{B}_\varepsilon(x)$ and $x \in \bar{x}$ is the farthest point from v such that $\mathrm{d}(p,x) < s$. And β is the projection of the segment $\bar{xp}$ to $T_p^\perp S$. Then the first inequality (1.1) is a consequence of the Law of Cosines. The second condition (1.2) follows from the Theorem of Pythagoras applied to the triangle (x,h,T') and the third condition (1.3) uses the Theorem of Pythagoras applied to the triangles (v',x,h) and (p,x,h), which have the segment $\bar{xh}$ in common.

1.2. Reproving the deformation retraction argument without noise

As a first step we want to reprove two geometric lemmata, presented in [NSW08] and introduce the core idea for generalizing this proof to the setting of submanifolds within manifold. But before we need to define the condition number.

I.2 Definition: condition number
Let M be a submanifold in $\mathbb{R}^n$, then the condition number *is defined analogous to [NSW08], as follows*

$$\tau := \sup\{r \geq 0 | \mathcal{B}_r^\perp(M) \text{ is embedded in } \mathbb{R}^n \text{ by the exponential mapping}\},$$

where $\mathcal{B}_r^\perp(M)$ is the open normal bundle of radius r around M.

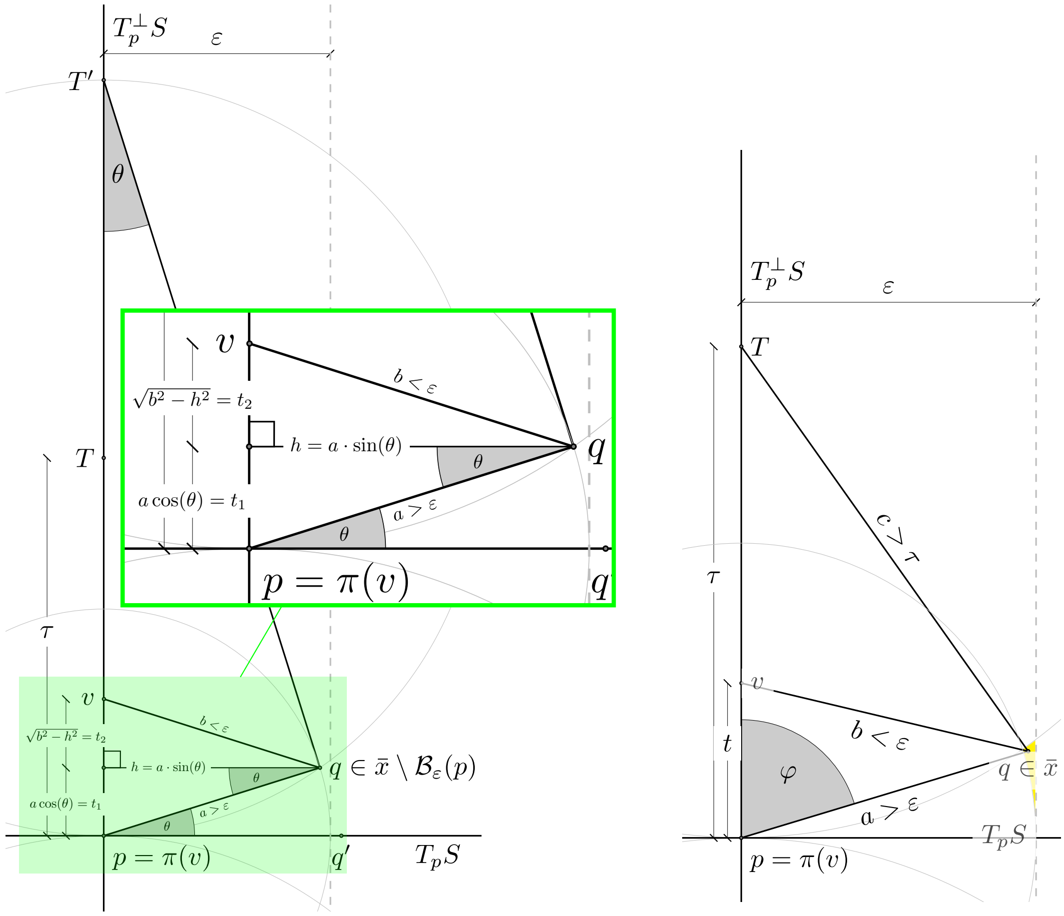
Note that later on we will define a generalization of this condition number for submanifolds of Riemannian manifolds. See Definition III.1 for more details.

I.3 Proposition: an alternative proof of [NSW08] Lemma 4.1
Let M be a manifold embedded in $\mathbb{R}^n$ with condition number τ and $\bar{x} := x_1, \ldots, x_N$ a finite number of points in M, and let v be a point in $\mathcal{B}_\varepsilon(q) \cap T_p^\perp \cap \mathcal{B}_\varepsilon(p) \subseteq \mathcal{B}_\varepsilon(\bar{x})$, where $0 < \varepsilon < \tau$, and p is the unique nearest point of v in M, and $q \in \bar{x} \setminus B_\varepsilon(p)$. Then the distance

$$\mathrm{d}(v,p) < \frac{\varepsilon^2}{\tau}. \tag{1.4}$$

Proof. At first take a look at the situation described in [NSW08]. The three points

p, q and v span a plane in $\mathbb{R}^n$. And the points T and T' are given by $T := p + \tau \frac{\vec{pv}}{\|\vec{pv}\|}$ and $T' := p + 2\tau \frac{\vec{pv}}{\|\vec{pv}\|}$



(a)

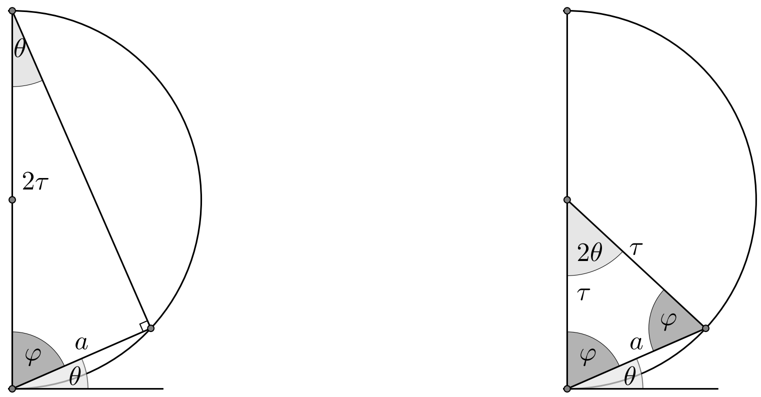
A detailed picture of the original version. Note any manifold M with condition number τ has to lie outside the circle with centre T and radius τ. The dashed line denotes all points that have distance ε from $T_x^\perp(M)$. Then the distance $\mathrm{d}(p, v)$ can be decomposed in the two segments $t_1 := a \cdot \cos(\theta)$ and $t_2 := \sqrt{b^2 - h^2}$, where the auxiliary length is given by $h := a \cdot \sin(\theta)$, the angle θ at the bottom is equal to the angle at T' as the triangles (T', p, q) and (p, q', q) are similar.

(b)

alternative view on the situation, where the region $(M \cap \mathcal{B}_\varepsilon(v)) \setminus \mathcal{B}_\varepsilon(p)$ is indicated by the yellow area. a denotes the distance $\mathrm{d}(p, q) > \varepsilon$, $b := \mathrm{d}(v, q) < \varepsilon$ and $c := \mathrm{d}(T, q) > \tau$.

Figure I.2.: Comparison of Lemma 4.1 found in [NSW08]

I would like to start with comparing the approach taken in [NSW08] with the one here. Niyogi, Smale and Weinberger start with a the larger triangle in Figure **??** and use the angle θ. We will now calculate the length of the side a which have



(a) fig:anglecomp1 (b) fig:anglecomp2

Figure I.3.: Equivalence of the angles φ and θ.

both triangles in common. As the triangle in the right picture is equilateral, we know that the angles at the segment a have to be equal (φ), if we do a simple calculation, we see that the angle at the centre of the circle is given by 2θ. The triangle on the left hand side is a right-angled one by construction. So the length of a is given in terms of the condition number τ and the angles θ/φ:

$$
\begin{aligned}
a &= 2\tau \cdot \sin(\theta) \\
a &= 2\tau \cdot \cos(\varphi)
\end{aligned}
$$

In our approach we will always consider the smaller triangle of Figure **??** and the angle φ. With a short calculation we see that the length of a is equal to the one derived before.

$$
\begin{aligned}
a^2 &= 2\tau^2 - 2\tau^2 \cdot \cos(2\theta) \\
a^2 &= 2\tau^2(1 - \cos^2(\theta) + \sin^2(\theta)) \\
a^2 &= 4\tau^2 \sin^2(\theta) \\
a &= 2\tau \sin(\theta) \\
a &= 2\tau \sin(\frac{\pi}{2} - \varphi) \\
a &= 2\tau \cos(\varphi)
\end{aligned}
$$

We will resume all calculations with φ being the angle of choice. If we want to calculate the distance $\mathrm{d}(v,p)$ in the original case we first need to express φ in terms of a and c, which is given by:

$$\varphi(a,c) = \arccos\left(\frac{a^2+\tau^2-c^2}{2a\tau}\right) \tag{1.5}$$

Now we denote $\mathrm{d}(v,p) = t(a,b,c) = t_1(a,b,c) + t_2(a,b,c)$ where

$$t_1(a,b,c) = t_1(a,c) = a\cos\varphi(d,a) = a\cdot\frac{a^2+\tau^2-c^2}{2a\tau} = \frac{a^2+\tau^2-c^2}{2\tau}$$

and

$$\begin{aligned} t_2 &= \sqrt{b^2-h(a,c)} = \sqrt{b^2-a^2\cdot\sin^2\varphi(a,c)} = \sqrt{b^2-a^2(1-\cos^2\varphi(a,c))} \\ &= \sqrt{b^2-a^2+a^2\cdot\frac{(a^2+\tau^2-c^2)^2}{4a^2\tau^2}} = \sqrt{\frac{(a^2+\tau^2-c^2)^2}{4\tau^2}+b^2-a^2}. \end{aligned}$$

Or more easily with the law of cosines:

$$\begin{aligned} b^2 &= t^2+a^2-2\cdot t\cdot a\cdot\cos\varphi \\ 0 &= t^2-2\cdot t\cdot a\cdot\cos\varphi-(b^2-a^2) \\ t_{1,2} &= a\cdot\cos\varphi\pm\sqrt{a^2\cdot\cos^2\varphi+b^2-a^2} \end{aligned}$$

$$t(a,b,c) = \frac{a^2+\tau^2-c^2}{2\tau}+\sqrt{\frac{(a^2+\tau^2-c^2)^2}{4\tau^2}+b^2-a^2}$$

Now to conclude what happens in the *"worst case"* we take a closer look at the partial derivatives of t, and want to prove that

(i) $t(a,b,c)$ is increasing with respect to b

$$\frac{\partial t}{\partial b} = \frac{b}{\sqrt{\frac{(a^2+\tau^2-c^2)^2}{4\tau^2}+b^2-a^2}} > 0$$

(ii) $t(a,b,c)$ is decreasing with respect to a

$$\frac{\partial t}{\partial a} = \frac{a}{\tau}+\frac{\frac{2(a^2+\tau^2-c^2)2a}{4\tau^2}-2a}{2\sqrt{\frac{(a^2+\tau^2-c^2)^2}{4\tau^2}+b^2-a^2}}$$

$$= \frac{a}{\tau} + \frac{a(a^2+\tau^2-c^2-4\tau^2 2)}{4\tau^2\sqrt{\frac{(a^2+\tau^2-c^2)^2}{4\tau^2}+b^2-a^2}} < 0$$

which is true, for we have the following chain of equivalences:

$$\begin{aligned}
\Leftrightarrow \quad & a^2-3\tau^2-c^2+2\tau\sqrt{\frac{(a^2+\tau^2-c^2)^2}{4\tau^2}+b^2-a^2} < 0 \\
\Leftrightarrow \quad & 2\tau\sqrt{\frac{(a^2+\tau^2-c^2)^2}{4\tau^2}+b^2-a^2} < 3\tau^2+c^2-a^2 \\
\Leftrightarrow \quad & 4\tau^2\frac{(a^2+\tau^2-c^2)^2}{4\tau^2}+b^2-a^2 < (3\tau^2+c^2-a^2)^2 \\
\Leftrightarrow \quad & (a^2+\tau^2-c^2)^2+4\tau^2(b^2-a^2) < (3\tau^2+c^2-a^2)^2
\end{aligned}$$

as we know that $b < \varepsilon < a$ it remains to show

$$\begin{aligned}
0 &< (3\tau^2+c^2-a^2)^2-(a^2+\tau^2-c^2)^2 = \\
&= 9\tau^4+6\tau^2c^2-6\tau^2a^2+c^4-2c^2a^2+a^4 \\
&\qquad -(\tau^4-2\tau^2c^2+2\tau^2a^2+c^4-2c^2a^2+a^4) = \\
&= 9\tau^4+4\tau^2c^2-4\tau^2a^2
\end{aligned}$$

which is equivalent to

$$0 < 2\tau^2+c^2-a^2$$

and this is true as we have $\tau > a$.

(iii) $t(a,b,c)$ is decreasing with respect to c

$$\frac{\partial t}{\partial c} = -\frac{c}{\tau} + \frac{-\frac{(a^2+\tau^2-c^2)c}{\tau^2}}{2\sqrt{\frac{(a^2+\tau^2-c^2)^2}{4\tau^2}+b^2-a^2}} < 0$$

so the "worst case scenario" is as follows. We replace $a \to \varepsilon$, $b \to \varepsilon$ and $c \to \tau$ and get

$$t(\tau,\varepsilon,\varepsilon) = \frac{\varepsilon^2+\tau^2-\tau^2}{2\tau} + \sqrt{\frac{(\varepsilon^2+\tau^2-\tau^2)^2}{4\tau^2}+\varepsilon^2-\varepsilon^2} = \frac{\varepsilon^2}{2\tau} + \sqrt{\frac{\varepsilon^4}{4\tau^2}} = \frac{\varepsilon^2}{\tau}$$

□

We saw that in the proof of the last theorem argumentation was a bit more easily when using the Law of Cosines, which is a bit more complicated than the

Theorem of Pythagoras but it allows us to generalize the previous theorem (see Proposition II.5). Another thing we will see later on, there is a way to avoid deriving an expression like that above. Of course one cannot hope to get an explicit formula for the distance t if the setting is on a manifold.

Next I want to give an example that if the maximal distance is $\frac{\varepsilon^2}{\tau}$ one needs the condition that the sample points are $\frac{\varepsilon}{2}$-dense, and in addition ε has to be small, to be precise one can prove that S is a deformation retraction of $\mathcal{B}_\varepsilon(\bar{x})$, if $\bar{x}$ is $\frac{\varepsilon}{2}$-dense and $0 \leq \varepsilon < \sqrt{\frac{3}{5}}\tau$. See [NSW08, Lemma 4.2].

I.4 Example: *Let* $M = \mathbb{R}^2$ *and* $S = \mathcal{S}^1$, *let the sample points* $\bar{\varphi} := \varphi_{-4}, \ldots, \varphi_4 \in \mathcal{S}^1$ *given by* $\varphi_0 := \arccos(\frac{7}{10})$ *for* $i = -4, \ldots, 4$ *we define* $\varphi_i := \varphi_0 - 2i \arccos(\frac{37}{40})$ *Then for* $\varepsilon = \sqrt{\frac{3}{5}}$ *we have* $\mathcal{S}^1 \subseteq \bigcup_{i=-4}^{4} \bar{\mathcal{B}}_{\frac{\varepsilon}{2}}(\varphi_i)$ *and* $v = (\frac{2}{5}, 0)$ $\pi(v) = p = (1,0)$. *Note that* $\tau = 1$ *and the distance* $\mathrm{d}(v,p) = \frac{\varepsilon^2}{\tau} = \frac{3}{5}$ *and the whole line* $\overline{vp}$ *is contained in* $\bar{\mathcal{B}}_\varepsilon(\varphi_0)$.

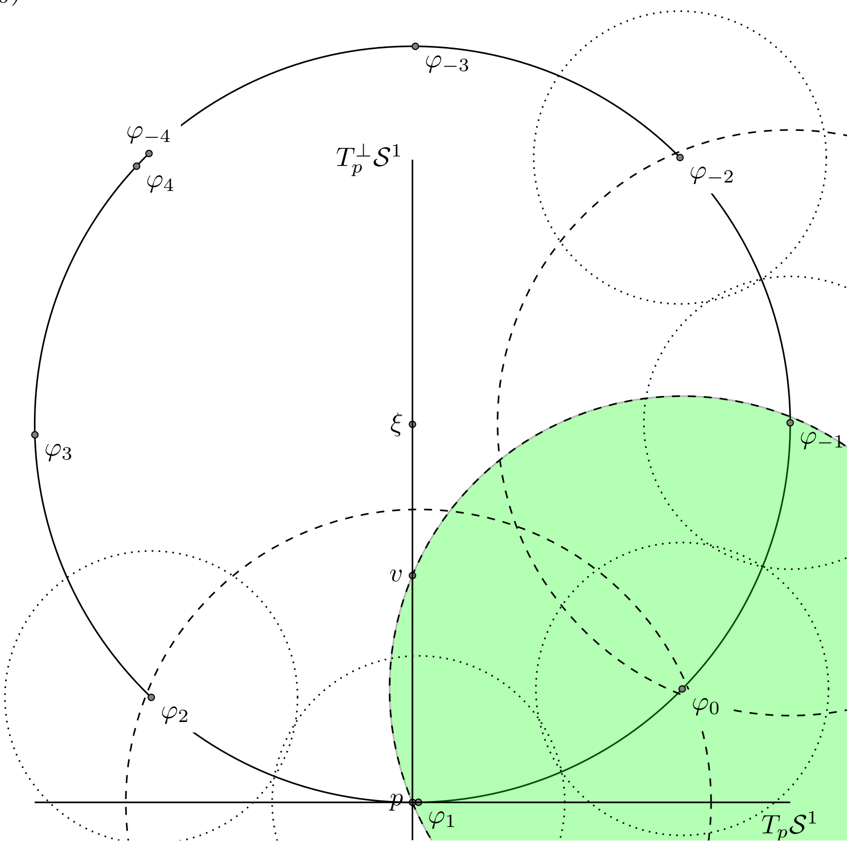


Figure I.4.: Example 4.1: The dashed circles are the ε-balls around the sample points, and the smaller dotted ones indicate that the samples are $\varepsilon/2$-dense. Note that the picture is rotated by 90 degrees.

1.3. Reproving the deformation retraction argument with noise

One thing to note is, that one does not need to have both points, v and p in one ball, but rather that it is only necessary to have the *"lowest point"* in the q-ball covered by a ball containing p. Which is the underlying idea of the second approach, taken in [NSW08]. In addition we take noise $r > 0$ into account, in the following graphic Figure I.5 this will be marked as the dotted circles. And we replace the condition of $\bar{x}$ being $\frac{\varepsilon}{2}$-dense by it being s-dense, which is slightly more general.

I.5 Proposition: improved distance estimate in [NSW08]
Let M be a complete Riemannian manifold with condition number τ. Let $\bar{x} := x_1, \ldots, x_N \in \mathcal{B}_r(M)$ be s-dense, then $\mathcal{B}_\varepsilon(\bar{x})$ is contractible to M if the following conditions hold:

$$t_q < t_{x_i} \tag{1.6}$$

$$(\tau - t_q)^2 < (\tau - r)^2 - \varepsilon^2 \tag{1.7}$$

$$t_{x_i}^2 + \frac{r^2 + 2\tau r + s^2}{\tau} t_{x_i} + s^2 - \varepsilon^2 < 0 \tag{1.8}$$

Let q be a sample point, and $v_q \in \bar{\mathcal{B}}_\varepsilon(q)$, then we denote its unique nearest point in M by p and we define $t_q := \mathrm{d}(v_q, p)$. As the sample points $x_1, \ldots, x_N$ are s-dense there exists an i such that x_i contains p, the farthest position for x_i to be is shown in Figure I.5 by v_{x_i} we denote the farthest point to p in $\{t \cdot p\vec{v}_q | t > 0\} \cap \bar{\mathcal{B}}_\varepsilon(x_i)$ and $t_{x_i} := \mathrm{d}(v_{x_i}, p)$.

Proof. As before we begin with a sketch of the current situation see Fig.I.5. If we look at the triangle (T, v, q) we note that the law of cosines gives

$$(\tau - r)^2 < c^2 = b^2 + (\tau - t_q)^2 - 2b(\tau - t_q)\cos(\psi) \tag{1.9}$$

in addition we know the angle ψ is always acute, as v_q is the unique closest point to p within the closed ball of radius ε around q. So we have $\cos(\psi) > 0$ and get

$$(\tau - r)^2 < b^2 + (\tau - t_q)^2 < \varepsilon^2 + (\tau - t_q)^2. \tag{1.10}$$

So the next step is to estimate t_{x_i}, which we do by looking at the second triangle

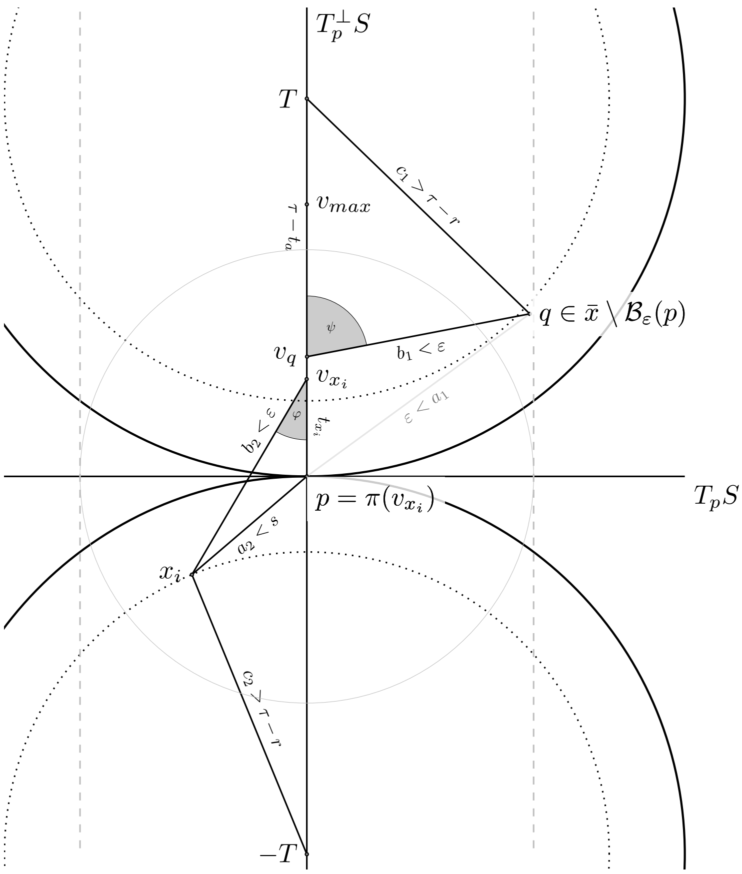


Figure I.5.: the triangles discussed in [NSW08, Lemma 7.1]: The points p,q and v_q again span a plane, and p,x_i and v_{x_i} do too, without loss of generality we may assume that these planes coincide. We can retract the line $p\bar{v}_{max}$ if the lines $\mathcal{B}_\varepsilon(x_i) \cap T_p^\perp M$ and $\mathcal{B}_\varepsilon(q) \cap T_p^\perp M$ intersect.

and apply the same logic as in Proposition I.3.

$$s^2 > a_2^2 = b_2^2 + t_{x_i}^2 - 2b_2 t_{x_i} \cos(\varphi) \tag{1.11}$$

$$(\tau - r)^2 > c_2^2 = b_2^2 + (\tau + t_{x_i})^2 - 2b_2(\tau + t_{x_i}) \cos(\varphi). \tag{1.12}$$

For reasons of simplicity we omit the subscripts during our calculation and get

$$\frac{b^2 + (\tau + t)^2 - (\tau - r)^2}{(\tau + t)} < \frac{b^2 + t^2 - s}{t}, \tag{1.13}$$

which leads to

$$t^2 + \frac{r^2 + 2\tau r + s}{\tau} t - b^2 + s^2 < 0 \tag{1.14}$$

and keeping $b < \varepsilon$ in mind and reattaching the indices we now have

$$t_{x_i}^2 + \frac{r^2 + 2\tau r + s}{\tau} t_{x_i} - \varepsilon^2 + s^2 < 0; \tag{1.15}$$

the desired implicit inequality for t_{x_i}. □

See [NSW08, Section 7] for a discussion of the special cases, where the noise r is zero; and the samples are r-dense.

Getting the distance estimate was the hardest step in proving a probability estimate on how many points ought to collect in order to retrieve the homotopy type of a manifold. Now the last step is to find a constraint for the sample points being s-dense, which can easily be done in terms of the covering number a probability measure and the maximal error δ, and then one can estimate the packing number, which itself gives control over the covering number.

II. Comparison Results in Riemannian Geometry

In the following I will assume that the reader is familiar with knowledge provided in a basic course on differential geometry or in some standard text on that subject like [Lee03], [Lee09] and [Lan96]. Nevertheless we will recapitulate some facts and definitions from Riemannian Geometry. Further on we will define the auxiliary functions $\mathrm{sn}, \mathrm{cs}, \mathrm{tn}$, which are generalized versions of sine, cosine and tangent function. As sn and cs have origin in a parametrized ordinary differential equation, which has sine and cosine as a special case, we have similar sum and difference identities for these functions. We will also define md, a modifying function for the distance functional, which allows us to formulate the *"Law of Cosines"* for triangles in the setting of manifolds with constant sectional curvature.

The next section is devoted to collect the tools to compare distances and volumes of manifolds with bounded sectional curvature to the volumes of manifolds of constant sectional curvature. The starting point is the famous *Rauch Comparison Theorem*, a local statement which compares Jacobi-fields along geodesics between two manifolds. This leads to comparing distances between a manifold with bounded sectional curvature and one of constant sectional curvature. With the use of the "Law of Cosines" we are then able to provide local upper and lower bounds for the inner angles of a triangle with given side-lengths. Then we will take a look at *Toponogov's Theorem*, which is a global distance estimate for manifolds with sectional curvature bounded from below. Combining this with the "Law of Cosines" we get a global upper bound for the inner angles of a triangle in such a manifold.

Then we will study the volume form on manifolds, and see that for Riemannian manifolds with sectional curvature bounded from above, we have the result of the *"Volume Comparison theorem of Günther"*, which gives a lower bound for the volume of sufficiently small balls. Which itself is again a consequence of Rauch's Comparison Theorem.

At last we will take a closer look on the tubular neighbourhood of a submanifold S of M and prove that for a point $x \in M$ but not in S a geodesic from x to S is always perpendicular to S.

2.1. Notation and some Results from Riemannian Geometry

A *Riemannian manifold* M is a smooth manifold equipped with the *Riemannian metric* g, which will also be denoted $\langle \_, \_ \rangle$, which is an inner product in every tangent space, that depends smoothly on the base point. This means, for two smooth vector fields $\xi, \eta \in \mathfrak{X}(M)$ the map $x \mapsto g(\xi(x), \eta(x))$ is smooth. For a submanifold $S \subseteq M$ this inner product allows us to decompose the tangent space at $y \in S$ in $T_y S \oplus T_y^\perp S$, so we can define the normal bundle as $T^\perp S := \bigsqcup_{y \in S} T_y^\perp S$.

With the Riemannian metric we are also able to define the length of a piecewise smooth curve $c : [a, b] \to M$ by

$$\ell(c) := \int_a^b \sqrt{\langle c'(t), c'(t) \rangle} dt.$$

And thus we can endow a Riemannian manifold with a metric by defining the distance between two points $x, y \in M$ as

$$\mathrm{d}(x, y) := \inf\{\ell(c) | c[0, a] \to M \text{ piecewise smooth, with } c(0) = x \text{ and } c(a) = y\}.$$

Note that in general one is not able to connect two points $x, y \in M$ by a curve, if for example they lie in different connected components. But there is a class of special curves connecting two points in M, that generalize the concept of straight lines to non-Euclidean geometry. These so called *geodesics* γ are characterised by the property of vanishing acceleration $\frac{d\gamma}{d^2t} = 0$, and with the theory of ordinary differential equations one can show that geodesics do exist locally and are unique for given starting point and initial direction. Note that no acceleration implies that the speed of a geodesic is constant, i.e. $\frac{d\gamma}{dt} = \text{const}$, and one can show that there always exists a reparametrization $\tilde{\gamma}$, such that $\|\frac{d\tilde{\gamma}}{dt}\| = 1$, which is called *parametrized by arclength*. In addition we define a *segment* σ between points $x, y \in M$ to be a geodesic parametrized by arclength such that $\mathrm{d}(x, y) = \ell(\sigma)$ and x and y are starting and endpoint of σ.

Another thing in which geodesics differ from lines in Euclidean space, is that geodesics may not exist for all times, see [Pet06, Example 28, p. 119]. So we say a manifold M is *geodesically complete*, if every geodesic in M exists for all times $t > 0$. By the Hopf-Rinow Theorem, this new notion of completeness is not different from M being complete as a metric space, with the distance defined above. For a proof see [Pet06, Theorem 16, p. 137].

At last let us recall that we can define the *exponential map* as

$$\begin{aligned} T_xM \supseteq U &\longrightarrow M \\ \exp_x(v) &:= \gamma_v(1), \end{aligned}$$

where γ_v is the unique geodesic with $\gamma_v(0) = x$ and $\gamma_v'(0) = v$, and U is the set of all $v \in T_xM$, such that 1 is in the domain of γ_v. Note that we have $\exp_x(tv) = \gamma_v(t)$, $\ell(\exp_x(tv)) = \|v\|$ and that $\exp_x$ is a diffeomorphism between $T_xM \supseteq \mathcal{B}_\varepsilon(0) \cong \mathcal{B}_\varepsilon(x) \subseteq M$ for sufficiently small $\varepsilon > 0$. The largest ε, such that $\exp_x$ is a diffeomorphism of ε-balls is called the *injectivity radius* and will be denoted $r_{inj}(x)$, and

$$r_{inj} := \inf\{r_{inj}(x) | x \in M\}$$

is the global variant of it. Another number closely related to the injectivity radius is the convexity radius. The *convexity radius*(see [Pet06, p. 177]) $r_{cvx}(x)$ is the largest R such that the radius function $r_x(y) := \mathrm{d}(x, y)$ is convex on $\mathcal{B}_R(x)$ and for any two points in $y_1, y_2 \in \mathcal{B}_R(x)$ we have a unique segment $\sigma : [0, \alpha] \to \mathcal{B}_R(x)$ with $\sigma(0) = y_1$ and $\sigma(\alpha) = y_2$. We will denote the global infimum of all such radii by

$$r_{cvx} := \inf\{r_{cvx}(x) | x \in M\}.$$

To connect this with the injectivity radius we note the following proposition see [Pet06, Theorem 29, p. 177] for a proof.

II.1 Proposition: *For a Riemannian manifold and a point $x \in M$ if we have for $\mathcal{B} = \mathcal{B}_R(x)$*

$$R \leq \frac{1}{2} \min\{r_{inj}(\mathcal{B}), \frac{\pi}{\sqrt{\sec(\mathcal{B})}}\}. \tag{2.1}$$

then $R \leq r_{inj}(x)$. Here $\sec(\mathcal{B}) := \sup\{\sec_y | y \in \mathcal{B}\}$ is the least upper bound for the sectional curvature (see Definition II.2) on $\mathcal{B}$.

2.2. Special functions and the Law of Cosines

Both the study of the exp-mapping and geodesic-variations, i.e. variational curves, where every such variation is a geodesic, naturally lead to the idea of Jacobi-fields. We call $J \in \Gamma(\gamma^* TM)$ a vector-field along a geodesic γ a *Jacobi-field*, if it solves the Jacobi-equation

$$\nabla_{\partial_t} \nabla_{\partial_t} J + R(J, \gamma')\gamma' = 0. \tag{2.2}$$

Here ∇_{∂_t} denotes the *Levi-Civita connexion*, which is the unique torsion-free linear connexion compatible with the Riemannian metric g, i.e. it is characterised by the equations

$$\nabla_X Y - \nabla_Y X = [X, Y] \tag{2.3}$$
$$X \cdot g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z). \tag{2.4}$$

And R is the Riemannian curvature tensor defined as

$$R(X, Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z. \tag{2.5}$$

If we have a submanifold $S \subseteq M$ we can use the Levi-Civita connexion and a decomposition of the tangent bundle $TM|_S = TS \oplus T^\perp S$ and corresponding projections $\mathrm{pr}_{TS} : TM|_S \to TS$ and $\mathrm{pr}_{T^\perp S} : TM|_S \to T^\perp S$, to define

$$\nabla^{\|}_X Y := \mathrm{pr}_{TS}(\nabla_X Y) \qquad X, Y \in \mathfrak{X}(S) \tag{2.6}$$
$$\mathrm{II}(X, Y) := \mathrm{pr}_{T^\perp S}(\nabla_X Y) \qquad X, Y \in \mathfrak{X}(S) \tag{2.7}$$
$$\nabla^{\|}_X \xi := \mathrm{pr}_{T^\perp S}(\nabla_X \xi) \qquad X \in \mathfrak{X}(S), \xi \in \Gamma(T^\perp S) \tag{2.8}$$
$$B_X \xi := -\mathrm{pr}_{TS}(\nabla_X \xi) \qquad X \in \mathfrak{X}(S), \xi \in \Gamma(T^\perp S). \tag{2.9}$$

Where II is called the *second fundamental form* and B is the *Weingarten map*. We will write $\mathrm{II}_y(X, Y)$ for the second fundamental form in a point $y \in S$, where $X, Y \in T_y S$. We can also define a norm for the second fundamental form by

$$\| \mathrm{II}_y \| := \sup \left\{ \| \mathrm{II}_y(X, Y) \| \middle| X, Y \in T_y S \text{ with } \|X\| = \|Y\| = 1 \right\}. \tag{2.10}$$

Note that in the following $\gamma : [0, \alpha] \to M$ will always denote a geodesic with starting point $\gamma(0) = x$. If we have such a geodesic, then we want to analyse variations of such a curve, i.e. a function $\bar{\gamma} : (-\varepsilon, \varepsilon) \times I \to M$, where $\bar{\gamma}(s, 0) = x$, for fixed s the curve $\gamma_s(t) := \bar{\gamma}(s, t)$ is a geodesic and $\bar{\gamma}(0, t) = \gamma(t)$. Then the *variational field* along γ defined as $J(t) = \frac{\partial \bar{\gamma}}{\partial s}(0, t)$ is a Jacobi-field. Another important fact is that, given $\xi \in T_x M$ and $\eta \in T_\xi T_x M \cong T_x M$ we have the unique Jacobi-field J, along γ with $J(0) = \xi$ and $\nabla_{\partial_t} J(0) = \eta$. On the other hand, given a Jacobi field along a geodesic with $J(0) = 0$ and fixed $\nabla_{\partial_t} J(0)$, we can define a geodesic variation of γ by

$$\bar{\gamma}(s, t) = \exp_x(t(\gamma'(0) + s\nabla_{\partial_t} J(0))). \tag{2.11}$$

The equation (2.11) can be used to connect the derivative of the exp-mapping with Jacobi-fields, a short calculation shows that

$$J(t) = T_{t\gamma'(0)} \exp(x) \cdot (t\eta) \tag{2.12}$$

is a unique Jacobi-field with $J(0) = 0$ and $\nabla_{\partial_t} J(0) = \eta$.

Another link between Jacobi-fields and the exponential mapping are conjugate points. We call two points $x, y \in M$ *conjugated along* $\gamma : [0, \alpha] \to M$, if $\gamma(0) = x$ and $\gamma(\alpha) = y$ and there exists a non-vanishing Jacobi field along γ with $\gamma(0) = 0$ and $\gamma(\alpha) = 0$. This can be equivalently characterized by the fact that $\exp_x : T_x M \to M$ is not a local diffeomorphism at $\alpha\gamma'(0) \in T_x M$.

At last we note that an arbitrary Jacobi-field J along a geodesic γ can be decomposed $J = J^\perp + J^\parallel$ in an orthogonal and a parallel Jacobi-field. Where $J^\perp$ and $J^\parallel$ are defined as

$$X^\parallel := \frac{\langle X, \gamma' \rangle}{\langle \gamma', \gamma' \rangle}, \qquad X^\perp := X - X^\parallel. \tag{2.13}$$

We now take a look at Jacobi-fields in the setting of constant sectional curvature, and see that there exists a concrete decomposition in the form of equation (2.23).

II.2 Definition: sectional curvature
For a two dimensional subspace $E \subseteq T_x M$ we define the sectional curvature *at a point $x \in M$*

$$\sec_x(E) := -g_x(R(X, Y)X, Y), \tag{2.14}$$

where $X, Y \in E$ are an orthonormal base for this subspace. Note that this definition does not depend on the choice of this orthonormal base, see [Pet06, 3.3

Sectional Curvature, p. 36] for more details.

In most cases we will restrict ourselves to manifolds with bounded sectional curvature, so we have to introduce a bit of notation. We say a manifold has sectional curvature bounded from above resp. below, if for all $x \in M$ and all planes $E \subseteq T_xM$ we have $sec_x(E) \leq \mathcal{K}$ for some $\mathcal{K} \in \mathbb{R}$ resp. $\sec_x(E) \geq \kappa$ for some $\kappa \in \mathbb{R}$, and we will denote this by $\sec(M) \leq \mathcal{K}$ and $\sec(M) \geq \kappa$ respectively.

Note that for a submanifold $S \subseteq M$ we have an upper bound for the sectional curvature given by the sectional curvature of M and its second fundamental form. Let R^S be the Riemannian curvature tensor of (S, g_S) then we have the Gauss-Equation [Pet06, Theorem 3, p. 44] for $X, Y, Z, W \in \mathfrak{X}(S)$

$$\begin{aligned} g_y(R(X,Y)Z,W) = g_y^S(&R^S(X,Y)Z,W) \\ &+ g_y^{\perp}(\mathrm{II}(X,Z),\mathrm{II}(Y,W)) - g_y^{\perp}(\mathrm{II}(Y,Z),\mathrm{II}(X,W)). \end{aligned} \tag{2.15}$$

Here $g^{\perp}$ denotes the orthogonal part of g. And using the definitions of sec and $\|\,\mathrm{II}\,\|$ we obtain

$$\sec^S \leq \sec^M + 2\|\,\mathrm{II}\,\|^2. \tag{2.16}$$

If we have bounds on the sectional curvature of a manifold M we can compare the distances and volumes in M with the ones of the space form $\mathbb{M}^n_\kappa$, which we define as follows.

II.3 Definition: *For $\kappa \in \mathbb{R}$ and $n \in \mathbb{N}$ we say a manifold is a* model space, *if it is n-dimensional and has constant sectional curvature equal to κ. If M is simply connected and $\kappa = -1$,$\kappa = 0$ and $\kappa = 1$ we have $\mathbb{H}^n$ the hyperbolic space, $\mathbb{R}^n$ and $\mathcal{S}^n$, the n-dimensional sphere, as the most important examples. We define the space forms to be*

$$(\mathbb{M}^n_\kappa, g_\kappa) := \begin{cases} \textit{sphere with sectional curvature } \sec = \kappa & \textit{if } \kappa > 0 \\ \textit{Euclidean space with standard metric} & \textit{if } \kappa = 0 \\ \textit{hyperbolic space with sectional curvature } \sec = \kappa & \textit{if } \kappa < 0 \end{cases}$$

Note that the *Hopf-Killing Theorem* says that, if M is a simply connected n-dimensional model space then M is isometric to either $\mathbb{H}^n$, $\mathbb{R}^n$ or $\mathcal{S}^n$, if the sectional curvature is -1, 0 or $+1$. For more information see [Cha06, Note II.6, p. 104; Theorem IV.2.1, p. 198].

Another important result to note is that the Riemannian curvature tensor is

$$R(X,Y)Z = -\sec(M)g(X,Z)Y + \sec(M)g(Y,Z)X \tag{2.17}$$

for a manifold M with constant sectional curvature $\sec(M)$.

When studying Jacobi-fields on model spaces, we have the functions sn_κ and cs_κ as coefficients for a decomposition in the sum of two parallel vector fields. In addition these functions as well as tn_κ and md_κ will be used in Theorem III.2 and Proposition II.5 extensively. So let us define.

II.4 Definition: sine, cosine, tangent and a modified distance
We define the sn-*function, called the* κ-sine function

$$\mathrm{sn}_\kappa(t) := \begin{cases} \frac{1}{\sqrt{\kappa}}\sin(\sqrt{\kappa}t) & \text{for } \kappa > 0 \\ t & \text{for } \kappa = 0 \\ \frac{1}{\sqrt{|\kappa|}}\sinh(\sqrt{|\kappa|}t) & \text{for } \kappa < 0 \end{cases} \tag{2.18}$$

and the cs-*function, called the* κ-cosine function, *which is the derivative of* sn_κ,

$$\mathrm{cs}_\kappa(t) := \begin{cases} \cos(\sqrt{\kappa}t) & \text{for } \kappa > 0 \\ 1 & \text{for } \kappa = 0 \\ \cosh(\sqrt{|\kappa|}t) & \text{for } \kappa < 0 \end{cases} \tag{2.19}$$

and at last the tn-*function, called the* κ-tangent function, *which is the quotient* $\mathrm{sn}_\kappa / \mathrm{cs}_\kappa$,

$$\mathrm{tn}_\kappa(t) := \begin{cases} \frac{1}{\sqrt{\kappa}}\tan(\sqrt{\kappa}t) & \text{for } \kappa > 0 \\ t & \text{for } \kappa = 0 \\ \frac{1}{\sqrt{|\kappa|}}\tanh(\sqrt{|\kappa|}t) & \text{for } \kappa < 0 \end{cases} \tag{2.20}$$

At last we define the modifying function md_κ *for the radial distance*

$$\mathrm{md}_\kappa(r) := \int_0^r \mathrm{sn}_\kappa(t)dt = \begin{cases} \frac{1}{\kappa}(1 - \mathrm{cs}_\kappa(r)) \text{ for } \kappa \neq 0 \\ \frac{1}{2}r^2 \text{ for } \kappa = 0 \end{cases} \tag{2.21}$$

Going back to the Jacobi-equation (2.2) we see that sn_κ and cs_κ are the simplest cases, where we can solve it.

So we see that the functions above (sn_κ) and (cs_κ) are solutions to the Jacobi-equation with respective initial condition.

$$f'' + \kappa f = 0$$

sn_κ	cs_κ
$\mathrm{sn}_\kappa(0) = 0$	$\mathrm{cs}_\kappa(0) = 1$
$\mathrm{sn}'_\kappa(0) = 1$	$\mathrm{cs}'_\kappa(0) = 0$

In addition we have the derivatives

$$\mathrm{sn}'_\kappa = \mathrm{cs}_\kappa, \qquad \mathrm{cs}'_\kappa = -\kappa\,\mathrm{sn}_\kappa\,. \tag{2.22}$$

With the help of (2.17) one can derive that on a space form of sectional curvature κ a Jacobi-field along a geodesic γ parametrized by arclength with initial condition $J(0) = v \perp \gamma'(0)$ and $\nabla_{\partial_t} J = w \perp \gamma'(0)$, can be written as

$$J(t) = \mathrm{cs}_\kappa(t)V(t) + \mathrm{sn}_\kappa(t)W(t). \tag{2.23}$$

Where $V(t)$ and $W(t)$ are the unique parallel vector fields with $V(0) = v$ and $W(0) = w$.

For the modifying function we have analogous to sn_κ and cs_κ that md_κ satisfies the inhomogeneous Jacobi equation.

$$\mathrm{md}''_\kappa + \kappa\,\mathrm{md}_\kappa = 1. \tag{2.24}$$

Which is can be seen as follows: for $\kappa = 0$ this is a simple calculation; for $\kappa \neq 0$ it is an application of the fundamental theorem of calculus:

$$\mathrm{md}'_\kappa(r) = \frac{d}{dr}\int_0^r \mathrm{sn}_\kappa(t)dt = \mathrm{sn}_\kappa(r) \tag{2.25}$$

$$\mathrm{md}''_\kappa(r) = \mathrm{sn}'_\kappa(r) = \mathrm{cs}_\kappa(r) \tag{2.26}$$

so from the definition of md_κ we get

$$\mathrm{md}''_\kappa(r) + \kappa\,\mathrm{md}_\kappa(r) = \mathrm{cs}_\kappa + \kappa \cdot \frac{1}{\kappa}(1 - \mathrm{cs}_\kappa(r)) = 1. \tag{2.27}$$

We also have the following sum and difference identities for sn_κ and cs_κ and a very useful identity that links cs_κ with md_κ.

$$1 = \mathrm{cs}^2_\kappa + \kappa\,\mathrm{sn}^2_\kappa \tag{2.28}$$

$$\mathrm{sn}_\kappa(a \pm b) = \mathrm{sn}_\kappa(a)\,\mathrm{cs}_\kappa(b) \pm \mathrm{cs}_\kappa(a)\,\mathrm{sn}_\kappa(b) \tag{2.29}$$

$$\mathrm{cs}_\kappa(a \pm b) = \mathrm{cs}_\kappa(a)\,\mathrm{cs}_\kappa(b) \mp \kappa\,\mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b) \tag{2.30}$$

$$1 = \mathrm{cs}_\kappa + \kappa\,\mathrm{md}_\kappa\,. \tag{2.31}$$

In addition the functions md_κ and cs_κ are even functions and sn_κ is odd.

From [Mey89, Section 1.5] we have a generalized version of the *"Law of Cosines"*.

II.5 Proposition: Generalized Law of Cosines

Let $\mathbb{M}_\kappa$ be a complete Riemannian manifold with constant sectional curvature κ and let x, y and z be points in $\mathbb{M}_\kappa$. Let us denote the distances $a := \mathrm{d}(y,z)$, $b := \mathrm{d}(x,z)$ and $c := \mathrm{d}(x,y)$, and let φ be the angle at z on the inside of the triangle. As depicted in Figure II.1.

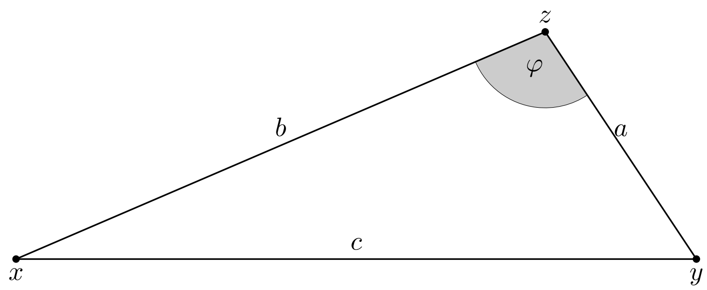


Figure II.1.: Law of Cosines

then the following equation is true

$$\mathrm{md}_\kappa(c) = \mathrm{md}_\kappa(a-b) + \mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b)(1-\cos(\varphi)). \tag{2.32}$$

This formula combines the three "classical" versions of the Law of Cosines for the cases $\kappa > 0$, $\kappa = 0$ and $\kappa < 0$ respectively:

$$\cos(\sqrt{\kappa}c) = \cos(\sqrt{\kappa}a)\cos(\sqrt{\kappa}b) + \sin(\sqrt{\kappa}a)\sin(\sqrt{\kappa}b)\cos(\varphi) \tag{2.33}$$

$$c^2 = a^2 + b^2 - 2ab\cos(\varphi) \tag{2.34}$$

$$\cosh(\sqrt{|\kappa|}c) = \cosh(\sqrt{|\kappa|}a)\cosh(\sqrt{|\kappa|}b) - \sinh(\sqrt{|\kappa|}a)\sinh(\sqrt{|\kappa|}b)\cos(\varphi). \tag{2.35}$$

Proof. Let us denote the radial distance from the corner point x to any $q \in \mathbb{M}_\kappa$ as $r(q) := \mathrm{d}(x,q)$. Composing this with the modifying function md_κ we get a smooth function $f(q) := \mathrm{md}_\kappa(r(q))$. And for its derivative we have

$$df = d(\mathrm{md}_\kappa \circ r) = (\mathrm{md}_\kappa' \circ r)dr = (\mathrm{sn}_\kappa \circ r)dr,$$

and thus we can calculate the Hessian of f as follows:

$$\begin{aligned}\mathrm{Hess}(f) = \nabla df = \nabla((\mathrm{sn}_\kappa \circ r)dr) &= (\mathrm{sn}_\kappa \circ r)\nabla dr + (\mathrm{cs}_\kappa \circ r)dr \otimes dr\\ &= (\mathrm{sn}_\kappa \circ r)\,\mathrm{Hess}(r) + (\mathrm{cs}_\kappa \circ r)dr \otimes dr\end{aligned}$$

Knowing that an orthogonal Jacobi-field along a given curve γ, parametrized by arclength, can be written in the form $J(t) = \mathrm{cs}_\kappa(t)J_1(t) + \mathrm{sn}_\kappa(t)J_2(t)$ (see equation (2.23)), where J_1 and J_2 are parallel vector fields along γ. Combined with $\mathrm{Hess}(\frac{1}{2}r^2)(X,X) = \langle \nabla_{\partial_t} X, X\rangle$ we get

$$\mathrm{Hess}(r) = \frac{\mathrm{cs}_\kappa \circ r}{\mathrm{sn}_\kappa \circ r}(g - dr \otimes dr). \tag{2.36}$$

Now using this and (2.31) we obtain

$$\mathrm{Hess}(f) = (1 - \kappa f)g. \tag{2.37}$$

For a geodesic γ with $\gamma(0) = z$, $\gamma(a) = y$ and $\|\gamma'\| = 1$ we define the composed function $\varphi(t) := f(\gamma(t))$. We show that this function is solution to the inhomogeneous Jacobi equation

$$\varphi'' + \kappa\varphi = 1, \tag{2.38}$$

by calculating the first derivative $\varphi' = df(\gamma')$ and with $\nabla_{\partial_t}\gamma' = 0$ the second derivative is

$$\begin{aligned}\varphi'' &= (\nabla df)(\gamma', \gamma') - df(\nabla_{\partial_t}\gamma') = \mathrm{Hess}(f)(\gamma', \gamma')\\ &= (1 - \kappa(f \circ \gamma))g(\gamma', \gamma') = 1 - \kappa\varphi.\end{aligned}$$

From (2.24) we know md_κ is also a solution to (2.38). Furthermore sn_κ and cs_κ are solutions to the Jacobi-equations, which gives combined

$$\varphi(t) = \mathrm{md}_\kappa(t) + C_0\,\mathrm{cs}_\kappa(t) + C_1\,\mathrm{sn}_\kappa(t). \tag{2.39}$$

Where

$$C_0 = \varphi(0) = \mathrm{md}_\kappa(r(\gamma(0))) = \mathrm{md}_\kappa(r(z)) = \mathrm{md}_\kappa(\mathrm{d}(x,z)) = \mathrm{md}_\kappa(b) \qquad (2.40)$$

and

$$\begin{aligned} C_1 = \varphi'(0) &= df(\gamma'(0))\,\mathrm{sn}_\kappa(r(\gamma(0)))dr(\gamma'(0)) = \mathrm{sn}_\kappa(b)dr(\gamma'(0)) \\ &= \mathrm{sn}_\kappa(b)\langle r, \gamma'(0)\rangle = \mathrm{sn}_\kappa(b)\|\nabla r\| \cdot \|\gamma'(0)\| \cos(\pi - \varphi) = -\mathrm{sn}_\kappa(b)\cos(\varphi). \end{aligned} \qquad (2.41)$$

At last we evaluate (2.39) at $t = a$ to see

$$\mathrm{md}_\kappa(c) = \varphi(a) = \mathrm{md}_\kappa(a) + \mathrm{cs}_\kappa(a)\,\mathrm{md}_\kappa(b) - \mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b)\cos(\varphi). \qquad (2.42)$$

Which concludes the proof. □

For a detailed proof of the “Law of Cosines“ on $\mathbb{M}_\kappa$ in the special cases $\kappa = 0, \kappa = 1$ and $\kappa = -1$ see [Pet06, Sec. 11, Prop. 48, p. 340]. Note that the general case of $\kappa < 0$ and $\kappa > 0$ follow, with a rescaling of the metric.

After this study of geometry on space forms we want to take a closer look on the basic properties of the modifying function as they will be used extensively in Lemma III.4.

II.6 Remark: sum and difference identities of md_*
If we have a closer look at the md_κ*-function defined above, we can prove the following properties:*

$$\begin{aligned} \mathrm{md}_\kappa(a+b) &= \mathrm{md}_\kappa(a) + \mathrm{cs}_\kappa(a)\,\mathrm{md}_\kappa(b) + \mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b) && (2.43) \\ &= \mathrm{md}_\kappa(b) + \mathrm{cs}_\kappa(b)\,\mathrm{md}_\kappa(a) + \mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b) && (2.44) \\ \mathrm{md}_\kappa(a-b) &= \mathrm{md}_\kappa(a) + \mathrm{cs}_\kappa(a)\,\mathrm{md}_\kappa(b) - \mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b) && (2.45) \\ &= \mathrm{md}_\kappa(b) + \mathrm{cs}_\kappa(b)\,\mathrm{md}_\kappa(a) - \mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b) && (2.46) \end{aligned}$$

Proof. For $\kappa = 0$ we have $\mathrm{md}_\kappa(a+b) = \frac{(a+b)^2}{2} = \frac{a^2}{2} + 1 \cdot \frac{b^2}{2} + \frac{2\cdot a\cdot b}{2} = \mathrm{md}_\kappa(a) + \mathrm{cs}_\kappa(a)\,\mathrm{md}_\kappa(b) + \mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b)$.

Let $\kappa \neq 0$ then note that md_κ has the property $\mathrm{cs}_\kappa + \kappa\,\mathrm{md}_\kappa = 1$ (see (2.31)). If we evaluate this equation at $a + b$ we get the following

$$\mathrm{cs}_\kappa(a+b) + \kappa\,\mathrm{md}_\kappa(a+b) = 1 \qquad (2.47)$$

And using the sum identities of cs_κ from equation (2.30) we get

$$\mathrm{cs}_\kappa(a)\,\mathrm{cs}_\kappa(b) - \kappa\,\mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b) + \kappa\,\mathrm{md}_\kappa(a+b) = 1, \tag{2.48}$$

which is equivalent to

$$(1-\kappa\,\mathrm{md}_\kappa(a))(1-\kappa\,\mathrm{md}_\kappa(b)) - \kappa\,\mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b) + \kappa\,\mathrm{md}_\kappa(a+b) = 1. \tag{2.49}$$

This can be simplified to

$$\mathrm{md}_\kappa(a+b) = \mathrm{md}_\kappa(a) + \mathrm{md}_\kappa(b)\,\mathrm{cs}_\kappa(a) + \mathrm{sn}_\kappa(a)\,\mathrm{sn}_\kappa(b) \tag{2.50}$$

Now the other equations (2.44)-(2.46) follow easily. □

Another very helpful identity is the following

II.7 Lemma: *The function* md_κ *can be expressed in terms of* sn_κ *and* tn_κ

$$\mathrm{md}_\kappa(r) = \mathrm{sn}_\kappa(r)\,\mathrm{tn}_\kappa(\frac{r}{2}) \tag{2.51}$$

Proof. A simple calculation using the basic identities from before shows

$$\begin{aligned}
\mathrm{md}_\kappa(r) &= \frac{1}{\kappa}\big(1-\mathrm{cs}_\kappa\big(\frac{r}{2}+\frac{r}{2}\big)\big) = \frac{1}{\kappa}\big[1-\Big(\mathrm{cs}_\kappa^2\big(\frac{r}{2}\big) - \kappa\,\mathrm{sn}_\kappa^2(\frac{r}{2})\Big)\big] \\
&= \frac{1}{\kappa}\big[1-\mathrm{cs}_\kappa^2(\frac{r}{2}) + \kappa\,\mathrm{sn}_\kappa^2(\frac{r}{2})\big] = 2\,\mathrm{sn}_\kappa^2(\frac{r}{2}) \\
&= \frac{2\,\mathrm{sn}_\kappa(\frac{r}{2})\,\mathrm{cs}_\kappa(\frac{r}{2})\,\mathrm{sn}_\kappa(\frac{r}{2})}{\mathrm{cs}_\kappa(\frac{r}{2})} = \mathrm{sn}_\kappa\big(\frac{r}{2}+\frac{r}{2}\big)\frac{\mathrm{sn}_\kappa(\frac{r}{2})}{\mathrm{cs}_\kappa(\frac{r}{2})} \\
&= \mathrm{sn}_\kappa(r)\,\mathrm{tn}_\kappa(\frac{r}{2})
\end{aligned}$$

In the case of $\kappa = 0$ this is simply proven by the observation $\frac{r^2}{2} = r\cdot\frac{r}{2}$ □

Note that one can use this formula to derive the half angle formula for the tangent and hyperbolic tangent function.

2.3. Comparison Estimates

Distance Comparison

The first part will be focussed on the distance comparison estimate by Harry Ernest Rauch, which allows us to reduce the situation of bounded sectional curvature to the case of constant curvature. We will formulate it in terms of Jacobi-fields and derive a distance version using the previously defined function md to treat both tangential and normal Jacobi fields at the same time. Another key theorem is by Victor Andreevich Toponogov a global extension of the Rauch Comparison Theorem. Later on we will use both to provide bounds on the inner angle.

In this section we will derive a few important corollaries, which have their origin in the classical theorem by Rauch.

The Rauch Comparison Theorem has found its way into mathematical literature in numerous versions see [Jos08], [Kar87] or [Cha06] or [dC92], which is the source for the following formulation. For a complete prove of this estimate see [dC92, Theorem 2.3, p. 215].

II.8 Theorem: (Rauch) Comparison Theorem
Let $\gamma : [0, \alpha] \to M^n$ and $\tilde{\gamma} : [0, \alpha] \to \tilde{M}^{n+k}$ $(k \geq 0)$, be geodesics with $\|\gamma'\| = \|\tilde{\gamma}'\|$. And let J and $\tilde{J}$ be Jacobi-fields along γ and $\tilde{\gamma}$, such that:

$$J(0) = 0 = \tilde{J}(0) \tag{2.52}$$

$$\langle J'(0), \gamma'(0)\rangle = \langle \tilde{J}'(0), \tilde{\gamma}'(0)\rangle \tag{2.53}$$

$$\|J'(0)\| = \|\tilde{J}'(0)\| \tag{2.54}$$

Assume that $\tilde{\gamma}$ does not have conjugate points on $(0, \alpha]$, and for all t and all $X \in T_{\gamma(t)}M$, $\tilde{X} \in T_{\tilde{\gamma}(t)}\tilde{M}$, for which X, γ' and $\tilde{X}$, $\tilde{\gamma}'$ are linearly independent, we have

$$\sec_{\gamma(t)}\langle X, \gamma'(t)\rangle_{vs} \leq \tilde{\sec}_{\tilde{\gamma}(t)}\langle \tilde{X}, \tilde{\gamma}'(t)\rangle_{vs},$$

where $\langle v_1, \ldots, v_n\rangle_{vs}$ denotes the vector space generated by the vectors $v_1, \ldots, v_n$. Then for all t we have

$$\|\tilde{J}(t)\| \leq \|J(t)\|.$$

In addition, if for some $t_0 \in (0, l]$, we have $\|\tilde{J}(t_0)\| = \|J(t_0)\|$, then for all $t \in [0, t_0]$ we have

$$\tilde{\sec}_{\gamma(t)}(\tilde{J}(t), \tilde{\gamma}'(t)) = \sec_{\tilde{\gamma}(t)}(J(t), \gamma'(t)).$$

Note that the idea of the following proof and also the notation md_* were coined by Hermann Karcher in *"Riemannian comparison constructions"* [Kar87]. A rather lengthy proof of the following statement can be found in [Pet06, Theorem 27, p. 175].

II.9 Corollary: *Let M be a Riemannian manifold with bounded sectional curvature $\kappa \le \sec(M) \le \mathcal{K}$. And for a geodesic $\gamma : [0, \alpha] \to M$ let X be a vector field orthogonal to γ. Then the following is true*

$$\frac{\mathrm{sn}'_{\mathcal{K}}(r)}{\mathrm{sn}_{\mathcal{K}}(r)} g(X, X) \le \mathrm{Hess}(r)(X, X) \le \frac{\mathrm{sn}'_{\kappa}(r)}{\mathrm{sn}_{\kappa}(r)} g(X, X), \tag{2.55}$$

as long as there is no conjugate point to $\gamma(0)$ on γ for all $t \in (0, \alpha]$.

Combining this corollary with the idea of md we see that the Riemannian metric is bounded. We will follow the proof given in [Hal11] as the other proof I found in [Pet06, Lemma 53, p. 342] is unfortunately rather short. Another proof can be found in [Mey89, Section 1.6, p. 13].

II.10 Proposition: (Karcher) Hessian Comparison
Let M be a Riemannian manifold with bounded sectional curvature $\kappa \le \sec \le \mathcal{K}$ and r the radial distance function induced by the Riemannian metric g at a point $x \in M$ defined as $r(y) := \mathrm{d}(x, y)$ and $f_\kappa(y) := (\mathrm{md}_\kappa \circ r)(y)$, $f_{\mathcal{K}}(y) := (\mathrm{md}_{\mathcal{K}} \circ r)(y)$ then its Hessian satisfies

$$\mathrm{Hess}(f_\kappa) + \kappa f_\kappa g \le g \le \mathrm{Hess}(f_{\mathcal{K}}) + \mathcal{K} f_{\mathcal{K}} g \tag{2.56}$$

on $\mathcal{B}_R(x)$ for $R < r_{inj}$.

Proof. For reasons of simplicity we just prove the estimate using the lower bound on the sectional curvature, as the one for the upper bound can be proven in the same manner. To get an estimate on the Hessian of the modified radial distance function we start with

$$df_\kappa = d(\mathrm{md}_\kappa \circ r) = (\mathrm{md}'_\kappa \circ r) dr = (\mathrm{sn}_\kappa \circ r) dr,$$

which leads to

$$\begin{aligned} \mathrm{Hess}(f_\kappa) &= (\mathrm{sn}'_\kappa \circ r) dr \otimes dr + (\mathrm{sn}_\kappa \circ r) \nabla dr \\ &= (\mathrm{sn}'_\kappa \circ r) dr^2 + (\mathrm{sn}_\kappa \circ r) \,\mathrm{Hess}(r) \\ &= (\mathrm{cs}_\kappa \circ r) dr^2 + (\mathrm{sn}_\kappa \circ r) \,\mathrm{Hess}(r). \end{aligned}$$

Now we want to calculate $\mathrm{Hess}(r)(Y, Y)$ for an arbitrary vector field $Y \in \mathfrak{X}(M)$ along a minimal geodesic γ, which is parametrized by arclength. As this vector field

can be decomposed in the sum of two vector fields $X^\perp$ a perpendicular component, and $X^\parallel$ a parallel to γ, we only have to consider the two cases:
Let $X^\perp \in \mathfrak{X}(M)$ perpendicular to γ in every point, then $dr(X^\perp) = 0$, and we get

$$\operatorname{Hess}(f_\kappa)(X^\perp, X^\perp) = 0 + (\operatorname{sn}_\kappa \circ r) \operatorname{Hess}(r)(X^\perp, X^\perp) \tag{2.57}$$
$$\leq (\operatorname{sn}_\kappa' \circ r) g(X^\perp, X^\perp) = (\operatorname{cs}_\kappa \circ r) g(X^\perp, X^\perp). \tag{2.58}$$

In the second case let $X^\parallel$ be a parallel vector field, i.e. it has the form $X^\parallel = \lambda(t) \cdot \gamma'(t)$, and let Y be an arbitrary vector field along γ. Then $\operatorname{Hess}(r)(X^\parallel, Y) = (\nabla_{X^\parallel} dr)Y = (\nabla_{\lambda(t)\cdot\gamma'(t)} dr)Y$. From $(\nabla_{\gamma'(t)} dr)Y = \langle \operatorname{grad}(r), Y\rangle = \langle \nabla_{\partial_t}\gamma', Y\rangle = 0$ we get $\operatorname{Hess}(X^\parallel, Y) = 0$. So now easily follows that

$$\operatorname{Hess}(f_\kappa)(X^\parallel, Y) = (\operatorname{cs}_\kappa \circ r) dr^2(X^\parallel, Y) + 0. \tag{2.59}$$

Combining both results we obtain for an arbitrary vector field $Y \in \mathfrak{X}(M)$:

$$\begin{aligned}\operatorname{Hess}(f_\kappa)(Y, Y) &= \operatorname{Hess}(Y^\perp, Y^\perp) + \operatorname{Hess}(Y^\parallel, Y^\perp) + \operatorname{Hess}(Y, Y^\parallel)\\ &\leq (\operatorname{cs}_\kappa \circ r)(g(Y^\perp, Y^\perp) + g(Y^\parallel, Y^\perp) + g(Y, Y^\parallel))\\ &= (\operatorname{cs}_\kappa \circ r) g(Y, Y)\\ &= (1 - \kappa f_\kappa) g(Y, Y).\end{aligned}$$

We see that for the upper bound of the Riemannian metric, one could use the same line of arguments.

□

For the interested reader I want to note that the original idea of the modifying function and additional information can be found in [Kar87], but I would rather advise you to look into the collection [CC89] since it includes the paper *"Conjugate and Cut loci"* by Shoshichi Kobayashi which Karcher frequently refers to.

Now as we have this auxiliary statement proven we will now compose the result with Rauch's Theorem.

II.11 Corollary: (Rauch) Comparison Theorem cosine version *Let M be a complete Riemannian manifold with bounded sectional curvature $\kappa \leq \sec \leq \mathcal{K}$. Suppose x, y and z be three points in M. Then we define $a := \mathrm{d}(y, z)$, $b := \mathrm{d}(x, z)$ and $c := \mathrm{d}(x, y)$ and let φ be the inner angle at the point z. Then*

$$\operatorname{md}_\mathcal{K}(b - a) + \operatorname{sn}_\mathcal{K}(b) \operatorname{sn}_\mathcal{K}(a)\big(1 - \cos(\varphi)\big) \leq \operatorname{md}_\mathcal{K}(c) \tag{2.60}$$

$$\mathrm{md}_\kappa(c) \leq \mathrm{md}_\kappa(b-a) + \mathrm{sn}_\kappa(b)\,\mathrm{sn}_\kappa(a)\big(1-\cos(\varphi)\big) \tag{2.61}$$

holds as long as the whole geodesic, γ connecting y with z, i.e. $\gamma(0) = y$ and $\gamma(a) = z$, is contained in $\mathcal{B}_R(x)$ with $0 \leq R < r_{inj}(x)$. In case of $\mathcal{K} > 0$, we additionally have to assume a, $R < \frac{\pi}{\sqrt{\mathcal{K}}}$.

Proof. First note that the function $f(w) := \mathrm{md}_\mathcal{K}(\mathrm{d}(x,w))$ is smooth on $\mathcal{B}_R(x)$, since we have $R < r_{inj}$. And the distance function for fixed x is the same as the radial distance on the tangent space at x, $r(w) := \mathrm{d}(x,w)$. So within $\mathcal{B}_R(x)$ the Hessian of f satisfies

$$\mathrm{Hess}(f) + \mathcal{K}fg \geq g \tag{2.62}$$

by Proposition II.10. Also note $(f\circ\gamma)(0) = \mathrm{md}_\mathcal{K}(b)$ and

$$(f\circ\gamma)'(0) = \mathrm{sn}_\mathcal{K}(b)\langle\nabla r, \gamma'\rangle = -\,\mathrm{sn}_\mathcal{K}(b)\cos(\varphi). \tag{2.63}$$

And with equation (2.2) we may conclude that the function

$$h(t) := f(\gamma(t)) - \big[\,\mathrm{md}_\mathcal{K}(t-b) + \mathrm{sn}_\mathcal{K}(t)\,\mathrm{sn}_\mathcal{K}(b)(1-\cos(\varphi))\big] \tag{2.64}$$

satisfies the following initial value problem

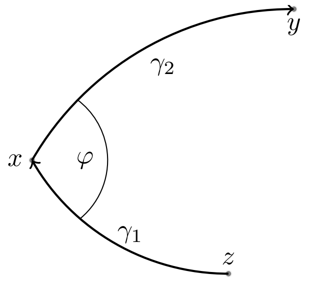
$$h'' + \mathcal{K}h \geq 0 \qquad\qquad h(0) = 0 \qquad\qquad h'(0) = 0. \tag{2.65}$$

From this differential inequality we obtain $(h'\,\mathrm{sn}_\mathcal{K} - h\,\mathrm{sn}_\mathcal{K}')' = (h''\,\mathrm{sn}_\mathcal{K} - h\,\mathrm{sn}_\mathcal{K}'') = (h'' + \mathcal{K}h)\,\mathrm{sn}_\mathcal{K} \geq 0$, hence the function $h'\,\mathrm{sn}_\mathcal{K} - h\,\mathrm{sn}_\mathcal{K}'$ is increasing. With $h(0) = 0$ this leads to $h'\,\mathrm{sn}_\mathcal{K} - h\,\mathrm{sn}_\mathcal{K}' \geq 0$. Which yields $\big(\frac{h}{\mathrm{sn}_\mathcal{K}}\big)' = \frac{h'\,\mathrm{sn}_\mathcal{K} - h\,\mathrm{sn}_\mathcal{K}'}{\mathrm{sn}_\mathcal{K}^2} \geq 0$, which implies $\frac{h}{\mathrm{sn}_\mathcal{K}}$ is increasing. From the calculation $\frac{h}{\mathrm{sn}_\mathcal{K}}(0) = \lim_{x\to 0}\frac{h(x)}{\mathrm{sn}_\mathcal{K}(x)} = \frac{h'(0)}{\mathcal{K}\,\mathrm{cs}_\mathcal{K}(0)} = h'(0) = 0$ we conclude $\frac{h}{\mathrm{sn}_\mathcal{K}} \geq 0$, and hence the desired inequality $h \geq 0$ follows. □

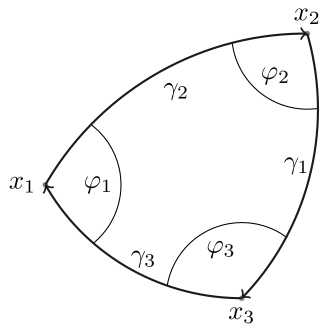
The next theorem is an important generalization of the Rauch Comparison Theorem. One can show that for Riemannian manifolds with sectional curvature bounded from below there exists a global comparison of lengths in a model space and the lengths measured in the manifold. This theorem is due to Victor Andreevich Toponogov and has two equivalent formulations, the triangle version and the hinge version. A proof for this theorem can be found in many textbooks on Riemannian Geometry like [Cha06, Theorem IX.5.1/2 p. 400], [CES75, 2.2 Theorem, p. 42], [Kli82, Theorem 2.7.12, p. 226] or [Pet06, Theorem 79, p. 339] The proof requires quite some preparation and is rather lengthy so we will just cite the theorem. But before we need the following definition.

II.12 Definition: geodesic triangle and hinge

A hinge *or* geodesic hinge *is a triple $(\gamma_1, \gamma_2, \varphi)$, where $\gamma_1 : [0, \alpha_1] \to M$ is a unit speed geodesic from z to x and $\gamma_2 : [0, \alpha_2] \to M$ a segment from x to y. And $\varphi := \measuredangle(-\gamma_1'(\alpha_1), \gamma_2'(0))$ is the angle on the inside of the triangle. A* geodesic triangle *is a triple $(\gamma_1, \gamma_2, \gamma_3)$ of three unit speed geodesics $\gamma_i : [0, \alpha_i] \to M$, called the sides, such that $\gamma_i(\alpha_i) = \gamma_{i+1}(0)$ and $\alpha_i + \alpha_{i+1} \geq \alpha_{i+2}$. The points $x_i = \gamma_{i+2}(0)$ are called the vertices of the triangle and $\alpha_i = \measuredangle(-\gamma_{i+1}'(\alpha_{i+1}), \gamma_{i+2}'(0))$ the corresponding angles. Note that indices are taken modulo 3.*



(a) Sketch of a hinge



(b) Sketch of a geodesic triangle

II.13 Theorem: (Toponogov)

Let M be a complete Riemannian manifold with sectional curvature bounded from below $\sec(M) \geq \kappa$.

***(hinge version)** Let $(\gamma_1, \gamma_2, \varphi)$ be a geodesic hinge in M, such that γ_1 is minimal, and if $\kappa > 0$, then suppose $\alpha_2 \leq \frac{\pi}{\sqrt{\kappa}}$. Let $(\tilde{\gamma}_1, \tilde{\gamma}_2, \varphi)$ be a geodesic hinge in $\mathbb{M}_\kappa^2$ with $\ell(\tilde{\gamma}_1) = \alpha_1$ and $\ell(\tilde{\gamma}_2) = \alpha_2$. Then*

$$\mathrm{d}(\gamma_1(0), \gamma_2(\alpha_2)) \leq \mathrm{d}(\tilde{\gamma}_1(0), \tilde{\gamma}_2(\alpha_2)). \tag{2.66}$$

***(triangle version)** Let $(\gamma_1, \gamma_2, \gamma_3)$ be a geodesic triangle in M. Suppose γ_1 and γ_3 are minimal, and if $\kappa > 0$, we assume $\alpha_2 \leq \frac{\pi}{\kappa}$. Then there exists a geodesic triangle $(\tilde{\gamma}_1, \tilde{\gamma}_2, \tilde{\gamma}_3)$ in $\mathbb{M}_\kappa^2$ such that for $i = 1, 2, 3$*

$$\ell(\tilde{\gamma}_i) = \alpha_i, \quad \tilde{\alpha}_1 \leq \alpha_1, \quad \tilde{\alpha}_3 \leq \alpha_3. \tag{2.67}$$

This triangle $(\tilde\gamma_1, \tilde\gamma_2, \tilde\gamma_3)$ *is uniquely determined, unless* $\kappa > 0$ *and* $\alpha_i = \frac{\pi}{\sqrt{\kappa}}$.

We will use a more convenient form of the Toponogov theorem by combining it with the “Law of Cosines” II.5.

II.14 Corollary: (Toponogov) cosine version *Let* M *be a complete Riemannian manifold with sectional curvature bounded from below* $\kappa \leq \sec$. *Suppose* $x, y \in M$ *and* σ *a geodesic parametrized by arclength, with starting point* $x = \sigma(0)$. *Now define* $a := \mathrm{d}(x, y)$ *and furthermore let* φ *be the inner angle at* x *between* σ *and a minimal geodesic from* x *to* y. *Then for all* t *we have*

$$\mathrm{md}_\kappa(\mathrm{d}(y, \sigma(t))) \leq \mathrm{md}_\kappa(t - a) + \mathrm{sn}_\kappa(t)\,\mathrm{sn}_\kappa(a)\big(1 - \cos(\varphi)\big). \tag{2.68}$$

In case of $\kappa > 0$, *we additionally have to assume* $|t| < \frac{\pi}{\sqrt{\kappa}}$.

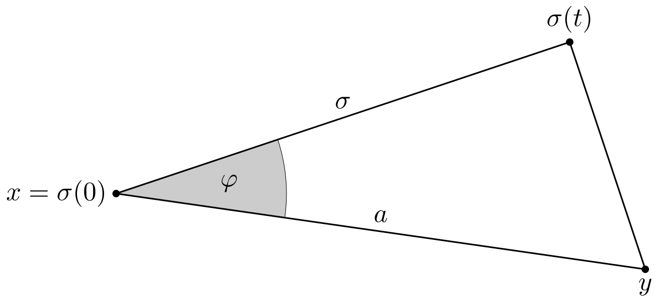


Figure II.2.: Triangle for Toponogov's Theorem

At last we cite the *diameter estimate* by Sumner Byron Myers, which gives an upper bound for the maximal distance two points in a manifold with Ricci curvature bounded from below can have. See [Pet06, Theorem 25, p. 171] for a complete proof. But before we start we have to define the *Ricci-curvature*.

The Ricci curvature is the trace of the Riemannian curvature tensor R, i.e. $\mathrm{ric}(X, Y) := \mathrm{tr}(\xi \to R(\xi, X)Y)$. If we choose an orthonormal basis $(E_1, \dots, E_n)$ of the n-dimensional tangent space, we can write this as

$$\mathrm{ric}(X, Y) = \sum_{i=1}^{n} \langle R(X, E_i)E_i, Y\rangle.$$

We say $\mathrm{ric} \geq k$, if and only if for all $X \in T_xM$ we have $\mathrm{ric}(X, X) \geq k\langle X, X\rangle$. Note that we can write ric in terms of sec. For if we have a unit vector X we can complete it to an orthonormal basis $\{X, E_2, \ldots, E_n\}$ for T_xM then

$$\operatorname{ric}(X,X) = \langle R(X,X)X,X\rangle + \sum_{i=2}^{n}\langle R(E_i,X)X,E_i\rangle = \sum_{i=2}^{n}\sec(X,E_i),$$

and thus if sec is bounded from below by κ, we also have $\operatorname{ric} \geq (n-1)\kappa$.

II.15 Theorem: Myers' diameter estimate
Let M be a complete n-dimensional Riemannian manifold with Ricci curvature bounded from below by $\operatorname{ric} \geq (n-1)k > 0$. Then $\operatorname{diam}(M) \leq \frac{\pi}{\sqrt{k}}$. Where the diameter is defined as $\operatorname{diam}(M) := \sup\{\mathrm{d}(x,y)|x,y \in M\}$. In particular the condition holds if $\sec \geq \kappa > 0$.

Note that if we combine Myers' Theorem II.15 and Lemma II.19 we see that the condition number as defined below in Definition III.1 $\tau \leq \operatorname{diam}(M) \leq \frac{\pi}{\sqrt{\mathcal{K}}}$.

Volume Comparison

Another consequence of the Rauch Comparison Theorem is the following. We can estimate the volume of balls in M, by the volume of balls in the model spaces, which is due to Paul Günther [Gün60]. A proof for this theorem can be found in [Cha06, Theorem III.4.2, p. 129].

II.16 Lemma: metric lemma
Let $g \leq \tilde{g}$ be two Riemannian metrics on an oriented n-manifold M. Then the associated volume forms can be compared $\mathrm{vol}_g \leq \mathrm{vol}_{\tilde{g}}$

Proof. Let $x \in M$ and A a linear, orientation preserving automorphism of T_xM such that for all $X, Y \in T_xM : g(X,Y) = \tilde{g}(AX, AY)$. Which leads to $\|AX\|^2 = \tilde{g}(AX,AX) = g(X,X) \leq \tilde{g}(X,X) = \|X\|^2$, this yields $\|A\|_{\tilde{g}} \leq 1$ and therefore for all eigenvalues λ of A we have $|\lambda| \leq 1$ and therefore $\det(A) \leq 1$. If we have $(E_1, \ldots, E_n)$, a positively oriented orthonormal base with respect to g, then $1 = \mathrm{vol}_g(E_1,\ldots,E_n) = \mathrm{vol}_{\tilde{g}}(A(E_1),\ldots,A(E_n)) = \det(A)\,\mathrm{vol}_{\tilde{g}}(E_1,\ldots,E_n) \leq \mathrm{vol}_{\tilde{g}}(E_1,\ldots,E_n)$. □

Using this we can prove the next theorem.

II.17 Theorem: (Günther) volume comparison
Let M be a Riemannian manifold of dimension n, with sectional curvature $\sec < \mathcal{K}$

bounded from above. And let the injectivity radius of M be denoted by r_{inj}. Then we have for $0 \leq r < r_{inj}$ and a point $x \in M$

$$\mathrm{vol}(\mathcal{B}_r^{\mathbb{M}^n_{\mathcal{K}}}(0)) \leq \mathrm{vol}(\mathcal{B}_r^M(x)). \tag{2.69}$$

Proof. Let g be the Riemannian metric on M, then define a Riemannian metric on $\mathcal{S}_r(0) := \{X \in T_xM | g(X,X) = r^2\}$ by $\tilde{g} := (\exp_x |_{\mathcal{S}_r(0)})^* g$ and a comparison metric $\tilde{g}_{\mathcal{K}} := \frac{\mathrm{sn}^2_{\mathcal{K}}(r)}{r^2} g_0$ on the model space with constant curvature $\mathcal{K}$, where g_0 is the flat Riemannian metric induced by g on the tangent space. Then we apply Rauch's Comparison Theorem II.8 to see

$$\tilde{g} \geq \tilde{g}_{\mathcal{K}}, \tag{2.70}$$

Combined with Lemma II.16 this implies:

$$\mathrm{vol}_{\tilde{g}} \geq \mathrm{vol}_{\tilde{g}_{\mathcal{K}}}$$

And with $(\mathcal{S}_r(0), \tilde{g}) \cong (\mathcal{S}_r(x), g)$ by construction and exp being a diffeomorphism on $\mathcal{S}_r(x)$.

$$\begin{tikzcd}
(\mathcal{S}_r(0), \exp^* g) \arrow[r, leftrightarrow, "\mathrm{id}"] \arrow[d, "\exp"] & (\mathcal{S}_r(0), \tilde{g}_{\mathcal{K}}) \arrow[d, "\psi"] \\
(\mathcal{S}_r(x), g) & (\mathcal{S}_r^{\mathbb{M}^n_{\mathcal{K}}}(0), g_{\mathcal{K}})
\end{tikzcd}$$

where ψ is just the isometry given by $x \mapsto \frac{\mathrm{sn}_{\mathcal{K}}(r)}{r} x$ and $\psi^* g_0(X,Y) = g_0(\psi X, \psi Y) = g_0(\frac{\mathrm{sn}_{\mathcal{K}}(r)}{r} X, \frac{\mathrm{sn}_{\mathcal{K}}(r)}{r} Y) = \frac{\mathrm{sn}^2_{\mathcal{K}}(r)}{r^2} g_0(X,Y)$, so we get

$$\mathrm{vol}(\mathcal{S}_r(x)) = \mathrm{vol}_{\tilde{g}}(\mathcal{S}_r(0)) \geq \mathrm{vol}_{\tilde{g}_{\mathcal{K}}}(\mathcal{S}_r(0)) = \mathrm{vol}(\mathcal{S}_r^{\mathbb{M}^n_{\mathcal{K}}}(0)).$$

Now we get $\mathrm{vol}(\mathcal{B}_r) \geq \mathrm{vol}(\mathcal{B}_r^{\mathbb{M}^n_{\mathcal{K}}}(0))$ by $\mathrm{vol}(\mathcal{B}_r(0)) = \int_0^r \mathrm{vol}(\mathcal{S}_t(0)) dt$, since we know the radial vector field ∂_r is normed and perpendicular to the sphere. □

Note that we can explicitly calculate the volume of an s-ball in the model space of constant sectional curvature in terms of Γ-functions

$$\mathrm{vol}(B_s^{\mathbb{M}^n_\kappa}(0)) = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} \int_0^s \mathrm{sn}_\kappa^{n-1}(t) dt. \tag{2.71}$$

see [?, III.7.1 Bemerkung, p. 189]quation III.89.

2.4. On the tubular neighbourhood

As a last part in this chapter we will prove that for a Riemannian manifold M, a closed submanifold S and a point $x \in M$ every geodesic realizing the distance $\mathrm{d}(x, S)$ is orthogonal to S.

II.18 Lemma: minimal Geodesics are orthogonal
Let M be a complete manifold, S a closed submanifold, i.e. compact without boundary, and $x \in M$ then every distance realizing geodesic $\sigma : [0, \alpha] \to M$ from S to x is orthogonal to S.

Proof. By [Pet06, Proposition 17, p. 126] we have, that the minima of the length function are equal to the minima of the *energy functional*

$$E(\gamma) := \frac{1}{2} \int_0^1 \| \frac{d\gamma}{dt} \|^2. \tag{2.72}$$

So we take a geodesic variation $\bar{\sigma}$ of σ with$\bar{\sigma}(s, \alpha) = x$ and $\bar{\sigma}(s, 0) \in S$ for all s. So we can calculate the following

$$\begin{aligned} 0 &= \frac{d}{ds}|_{s=0} E(\sigma_s) = \frac{d}{ds}|_{s=0} \frac{1}{2} \int_0^\alpha \langle \frac{\partial}{\partial t}\bar{\sigma}, \frac{\partial}{\partial t}\bar{\sigma} \rangle dt = \frac{1}{2} \int_0^\alpha \frac{d}{ds}|_{s=0} \langle \frac{\partial}{\partial t}\bar{\sigma}, \frac{\partial}{\partial t}\bar{\sigma} \rangle dt \\ &= \int_0^\alpha \langle \frac{\partial^2}{\partial s \partial t}\bar{\sigma}, \frac{\partial}{\partial t}\bar{\sigma} \rangle dt|_{s=0} = \int_0^\alpha \frac{\partial}{\partial t} \langle \frac{\partial}{\partial s}\bar{\sigma}, \frac{\partial}{\partial t}\bar{\sigma} \rangle dt|_{s=0} - \int_0^\alpha \langle \frac{\partial}{\partial t}\bar{\sigma}, \underbrace{\frac{\partial^2}{\partial t^2}\bar{\sigma}}_{=0} \rangle dt|_{s=0}. \end{aligned}$$

The underbraced term here is zero because all σ_s are geodesics, which are characterised by $\frac{d}{dt^2}\sigma_s = 0$. With this we have

$$0 = \langle \frac{\partial}{\partial s}\bar{\sigma}, \frac{\partial}{\partial t}\bar{\sigma} \rangle dt|_{(0,0)}^{(0,\alpha)} = \langle \underbrace{\frac{\partial}{\partial s}\bar{\sigma}(0, \alpha)}_{=0}, \frac{\partial}{\partial t}\bar{\sigma}(0, \alpha) \rangle dt - \langle \frac{\partial}{\partial s}\bar{\sigma}(0, 0), \frac{\partial}{\partial t}\bar{\sigma}(0, 0) \rangle dt$$

The underbraced term is zero because $\bar{\sigma}(s, \alpha)$ is constant. Finally note that for any curve $c : (-\varepsilon, \varepsilon) \to S$ with $c(0) = y$, there exists a variation of σ such that $\frac{\partial}{\partial s}\sigma(s, 0) = c'(0) \in T_y S$. □

The previous lemma is based an on "Exercise 1" given in [Jos08, Chapter 4, p. 232].

II.19 Lemma: *Let M be a complete Riemannian manifold, and $S \subseteq M$ a closed submanifold. If the exponential mapping of $\mathcal{B}_r^{\perp}(y) \xrightarrow[\cong]{\exp} \mathcal{B}_r^M(S)$ is a diffeomorphism onto its image. Then for all $X \in \exp_y^{-1}(\mathcal{B}_r(S))$ the curve $\gamma_X(t) := \exp(tX)$ is the unique minimal geodesic connecting $\exp_y(X)$ with S.*

Proof. Suppose $x \in \mathcal{B}_r(S)$, since S is compact there exist a point $y \in S$ realizing, the distance of x to S. Using the Hopf-Rinow Theorem (see [Pet06, Theorem 16, p. 137]) we see that there exists a geodesic of minimal length connecting x with S. By the previous Lemma II.18 this geodesic is perpendicular to the submanifold S, with unit direction $\nu \in T^{\perp}S$. So for every point in $x \in \mathcal{B}_r(S)$ we have a geodesic connecting this point x with S. So for $X := \mathrm{d}(x, S)\nu$ the geodesic $\gamma := \exp_y(tX)$ is the unique geodesic connecting x with S, otherwise the exponential mapping would not be injective and thus no diffeomorphism. $\square$

III. Manifold Learning from *s*-dense Samples

We are now finally able to proof the central theorem. We will show that for a Riemannian manifold and noisy sample points around a submanifold that the union of ε-balls around the sample points is diffeomorphic to $B_\tau^\perp(S)$ by a fibre-invariant diffeomorphism.

3.1. Learning the structure of a submanifold by *s*-dense samples

The proof of this central theorem will be split up in three parts. At first we will see that we have an increasing series of inclusions of neighbourhoods of S in M. The next part shows that if the sample points are taken from a small neighbourhood of S then the set $\mathcal{U}$ constructed by the union of ε-balls around the sample points is open, contains S and is fibrewise star-shaped with centre 0, i.e. for $x \in \mathcal{U}$ and for its nearest point $y \in S$, the minimal geodesic connecting x and y is fully contained in $\mathcal{U}$. In order to do that we need an auxiliary lemma, that reflects the derivation step in Proposition I.3. At last we will show that open star-shaped neighbourhoods of the zero-section of a vector bundle are diffeomorphic to the whole vector bundle, by a diffeomorphism that is fibre invariant. In particular S is a strong deformation retract of $\mathcal{B}_\varepsilon(\bar{x})$.

At last I want to note that the proof of the next theorem is based on notes of my supervisor Stefan Haller, for which I am very thankful that he provided me with. We begin with defining the condition number τ for submanifold analogous to Definition I.2 from [NSW08]. This number is a measure for the curvature of S in M, and gives a bound for S to intersect with itself.

III.1 Definition: condition number
Let M a complete Riemannian manifold, $S \subseteq M$ a closed submanifold, i.e. compact

without boundary. So we define the condition number

$$\tau := \sup\{r \geq 0 | T^{\perp}S \supseteq \mathcal{B}_r(0) \xrightarrow{\exp^M} U \subseteq M \textit{is a diffeomorphism onto its image}\}.$$

Note that both Lemma II.18 and Lemma II.19 apply and we have that exp is a diffeomorphism of $\mathcal{B}_\tau^\perp(S) \cong \mathcal{B}_\tau^M(S)$ and within $\mathcal{B}_\tau^M(S)$ geodesics are unique.

With all the preparation done we can now prove our main theorem.

III.2 Theorem: *Let M be a complete Riemannian manifold with bounded sectional curvature $\kappa \leq \sec \leq \mathcal{K}$, injectivity radius $r_{inj} > 0$, convexity radius $r_{cvx} > 0$. Let S be a closed submanifold, with condition number τ. Furthermore for $r \geq 0$ we have points $\bar{x} := (x_1, \ldots, x_N) \in \mathcal{B}_r(S)$ such that $S \subseteq \mathcal{B}_s(\bar{x})$, where $0 < s < \varepsilon$. Now suppose the following conditions hold*

$$r + \varepsilon < \tau \tag{3.1}$$

$$\varepsilon \leq r_{cvx} \tag{3.2}$$

$$2\varepsilon + r \leq r_{inj} \tag{3.3}$$

$$2\varepsilon + r \leq \frac{\pi}{\sqrt{\mathcal{K}}} \textit{ in the case of } \mathcal{K} > 0 \tag{3.4}$$

$$\operatorname{tn}_{\mathcal{K}}(\varepsilon + r) \leq \operatorname{tn}_\kappa(\tau) \textit{ and} \tag{3.5}$$

$$\frac{\operatorname{tn}_{\mathcal{K}}(\varepsilon)}{\operatorname{sn}_\kappa(\varepsilon)} \cdot \left(\frac{\operatorname{md}_\kappa(\varepsilon) - \operatorname{md}_\kappa(r)}{\operatorname{tn}_\kappa(\tau)} + \operatorname{sn}_\kappa(r) \right) \leq \operatorname{tn}_{\mathcal{K}} \left(\frac{\varepsilon - s}{2} \right). \tag{3.6}$$

If we denote $\mathcal{U} := \bigcup_{i=1}^N \mathcal{B}_\varepsilon(x_i)$ then $\exp^{-1}(\mathcal{U}) \cap \mathcal{B}_\tau(S)$ is an open subset of $T^\perp S$ containing the zero section and fibrewise star shaped with centre 0. In particular $\mathcal{U}$ is a tubular neighbourhood of S, hence a diffeomorphism to the total space of $T^\perp S$. Moreover this diffeomorphism can be chosen to intertwine vector bundle projection π with the retraction $r : \mathcal{U} \to S$.

Proof. At first note that the condition number τ is strictly positive, by compactness of S and the *Inverse Function Theorem.* The proof will be split in 3 parts.

Lemma III.3

Proposition III.4 and Sublemma III.5

Proposition III.6

So we start with the chain of neighbourhoods of S.

III.3 Lemma: *Let M be a Riemannian manifold and S a submanifold with condition number τ. For $0 < s < \varepsilon$, $r > 0$ and $\varepsilon + r < \tau$ let $\bar{x} := x_1, \ldots, x_N$ be points in $\mathcal{B}_r(S)$ such that $S \subseteq \mathcal{B}_s(\bar{x})$. Then we have*

$$S \subseteq \mathcal{B}_{\varepsilon-s}(S) \subseteq \mathcal{B}_\varepsilon(\bar{x}) \subseteq \mathcal{B}_{\varepsilon+r}(S) \subseteq \mathcal{B}_\tau(S). \tag{3.7}$$

Proof. The first inclusion is trivial, since $\mathrm{d}(y, S) = 0 < \varepsilon - s$ for $y \in S$. To see the second inclusion let $x \in \mathcal{B}_{\varepsilon-s}(S)$ be arbitrary then the distance $\mathrm{d}(x, \bar{x}) < \varepsilon$ as there exists a point $y \in S$ with $\mathrm{d}(x, y) < \varepsilon - s$ and as there has to be an i such that $\mathrm{d}(x_i, y) < s$ and so $\mathrm{d}(x, x_i) < \mathrm{d}(x, y) + \mathrm{d}(y, x_i) < \varepsilon$. The third inclusion is true, because for an arbitrary $x \in \mathcal{B}_\varepsilon(\bar{x})$ there exists an i such that $x \in \mathcal{B}_\varepsilon(x_i)$ and as $\mathrm{d}(x_i, S) < r$ we get $\mathrm{d}(x, S) < \mathrm{d}(x, x_i) + \mathrm{d}(x_i, S) < \varepsilon + r$. The last inclusion is true, as $\varepsilon + r < \tau$. □

III.4 Proposition: *Let M, S and $\mathcal{U}$ be as defined in Theorem III.2. Furthermore let the inequalities* (3.1)-(3.6) *hold. Then for each $x \in \mathcal{U}$ we have the segment σ connecting x with the base-point $y := \pi^\perp(x)$, fully contained in $\mathcal{U}$. Where $\pi^\perp$ denotes the projection of x to its unique nearest point in S.*

Proof. Now suppose $x \in \mathcal{U}$ and let $\sigma : [0, t] \to M$ be the segment in M connecting $y := \sigma(0) \in S$ with $\sigma(t) = x$. Since we have $x \in \mathcal{U} = \mathcal{B}_\varepsilon \bar{x}$, there exists an i such that $x \in \mathcal{B}_\varepsilon(x_i)$. If $y \in \mathcal{B}_\varepsilon(x_i)$, then $\sigma([0, t]) \subseteq \mathcal{B}_\varepsilon(x_i)$ by geodesic convexity of the ε-ball, since $\varepsilon < r_{cvx}$.

Otherwise there exits an $q \in \sigma([0, t])$, such that $q \notin \mathcal{B}_\varepsilon(x_i)$. So it suffices to show $t < \varepsilon - s$, then $\sigma \subseteq \mathcal{B}_{\varepsilon-s}(S) \subseteq \mathcal{B}_\varepsilon(\bar{x})$ follows by Lemma III.3. In order

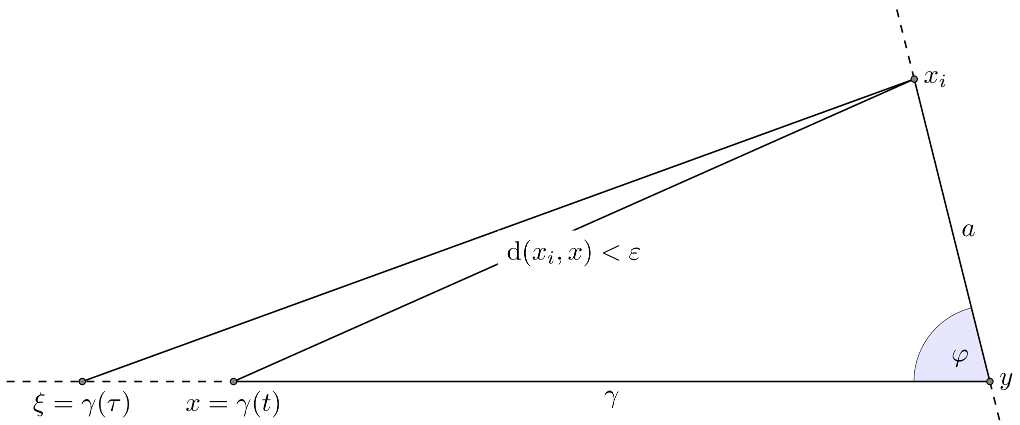


Figure III.1.: The geodesic triangle spanned by the points x, x_i and y, and the auxiliary geodesic triangle spanned by ξ, x_i and y.

to do that, we study the geodesic triangles given by the vertices (x, x_i, y) and (ξ, x_i, y). For a minimal geodesic γ connecting x_i with $y : [0, \alpha] \to M$ we define $\varphi := \langle -\gamma(\alpha), \sigma(0) \rangle$ to be the inner angle at y as denoted in the picture, and we define $a := \mathrm{d}(x_i, y)$, $\xi := \sigma(\tau)$. Note that we have $a > \varepsilon$, $\mathrm{d}(\xi, y) = \tau$, $\mathrm{d}(x, y) = t$, $\mathrm{d}(x, x_i) < \varepsilon$, $\mathrm{d}(x_i, S) < r$ and $\mathrm{d}(\xi, x_i) > \tau - r$, where the last estimate follows from

$$\tau = \mathrm{d}(\xi, S) \leq \mathrm{d}(\xi, x_i) + \mathrm{d}(x_i, S) < \mathrm{d}(\xi, x_i) + r. \tag{3.8}$$

And both τ and $\mathrm{d}(\xi, x_i)$ are less than $\frac{\pi}{\sqrt{\mathcal{K}}}$ by Myers' Diameter Theorem II.15. We first analyse the triangle with vertices (ξ, x_i, y), and use the lower bound on the sectional curvature for M and apply Toponogov's Theorem (Corollary II.14) to see

$$\mathrm{md}_\kappa(\mathrm{d}(\xi, x_i)) \leq \mathrm{md}_\kappa(\tau - a) + \mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(a)(1 - \cos(\varphi)). \tag{3.9}$$

Next note that we have $\mathrm{md}_\kappa'(r) = \mathrm{sn}_\kappa(r)$ and thus $\mathrm{md}_\kappa(r)$ is increasing for r, as we have $r < \frac{\pi}{\sqrt{\mathcal{K}}} \leq \frac{\pi}{\sqrt{\kappa}}$. So we have an implicit lower bound for φ

$$\frac{\mathrm{md}_\kappa(\tau - r) - \mathrm{md}_\kappa(\tau - a)}{\mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(a)} < 1 - \cos(\varphi). \tag{3.10}$$

For the triangle with vertices x,x_i and y we use the upper bound on the sectional curvature and Rauch's theorem (Corollary II.9) to get

$$\mathrm{md}_{\mathcal{K}}(t - a) + \mathrm{sn}_{\mathcal{K}}(t)\,\mathrm{sn}_{\mathcal{K}}(a)(1 - \cos(\varphi)) \leq \mathrm{md}_{\mathcal{K}}(d(x, x_i)). \tag{3.11}$$

Using $x \in \mathcal{B}_\varepsilon(x_i)$ we get $\mathrm{md}_{\mathcal{K}}(\mathrm{d}(x_i, x)) \leq \mathrm{md}_{\mathcal{K}}(\varepsilon)$, and combining this with the inequality before we have the implicit lower bound for the angle φ

$$1 - \cos(\varphi) \leq \frac{\mathrm{md}_{\mathcal{K}}(\varepsilon) - \mathrm{md}_{\mathcal{K}}(t - a)}{\mathrm{sn}_{\mathcal{K}}(t)\,\mathrm{sn}_{\mathcal{K}}(a)}. \tag{3.12}$$

Combining both inequalities (3.10) and (3.12) we obtain

$$\frac{\mathrm{md}_\kappa(\tau - r) - \mathrm{md}_\kappa(\tau - a)}{\mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(a)} \leq \frac{\mathrm{md}_{\mathcal{K}}(\varepsilon) - \mathrm{md}_{\mathcal{K}}(t - a)}{\mathrm{sn}_{\mathcal{K}}(t)\,\mathrm{sn}_{\mathcal{K}}(a)}. \tag{3.13}$$

Using the addition property of md_* (see Remark II.6) to expand the previous equation (3.13) we have

$$\frac{\mathrm{md}_\kappa(\tau) + \mathrm{cs}_\kappa(\tau)\,\mathrm{md}_\kappa(r) - \mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(r) - [\mathrm{md}_\kappa(\tau) + \mathrm{cs}_\kappa(\tau)\,\mathrm{md}_\kappa(a) - \mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(a)]}{\mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(a)}$$
$$\leq \frac{\mathrm{md}_\mathcal{K}(\varepsilon) - [\mathrm{md}_\mathcal{K}(a) + \mathrm{cs}_\mathcal{K}(a)\,\mathrm{md}_\mathcal{K}(t) - \mathrm{sn}_\mathcal{K}(t)\,\mathrm{sn}_\mathcal{K}(a)]}{\mathrm{sn}_\mathcal{K}(t)\,\mathrm{sn}_\mathcal{K}(a)}. \quad (3.14)$$

Which is equivalent to

$$\frac{\mathrm{cs}_\kappa(\tau)\,\mathrm{md}_\kappa(r) - \mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(r) - \mathrm{cs}_\kappa(\tau)\,\mathrm{md}_\kappa(a)}{\mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(a)} + 1$$
$$\leq \frac{\mathrm{md}_\mathcal{K}(\varepsilon) - \mathrm{md}_\mathcal{K}(a) - \mathrm{cs}_\mathcal{K}(a)\,\mathrm{md}_\mathcal{K}(t)}{\mathrm{sn}_\mathcal{K}(t)\,\mathrm{sn}_\mathcal{K}(a)} + 1. \quad (3.15)$$

We apply Sublemma III.5 below, to see that the inequality above is still valid, if we replace the parameter a by ε, for $0 < \varepsilon \leq a$ to get

$$\frac{\mathrm{cs}_\kappa(\tau)\,\mathrm{md}_\kappa(r) - \mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(r) - \mathrm{cs}_\kappa(\tau)\,\mathrm{md}_\kappa(\varepsilon)}{\mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(\varepsilon)} \leq -\frac{\mathrm{cs}_\mathcal{K}(\varepsilon)\,\mathrm{md}_\mathcal{K}(t)}{\mathrm{sn}_\mathcal{K}(t)\,\mathrm{sn}_\mathcal{K}(\varepsilon)}. \quad (3.16)$$

Which is equivalent to

$$\frac{\mathrm{md}_\kappa(r) - \mathrm{md}_\kappa(\varepsilon)}{\mathrm{tn}_\kappa(\tau)\,\mathrm{sn}_\kappa(\varepsilon)} - \frac{\mathrm{sn}_\kappa(r)}{\mathrm{sn}_\kappa(\varepsilon)} \leq -\frac{\mathrm{md}_\mathcal{K}(t)}{\mathrm{sn}_\mathcal{K}(t)\,\mathrm{tn}_\mathcal{K}(\varepsilon)} \quad (3.17)$$

and by the identity given in (2.46) $\mathrm{tn}_\mathcal{K}(\frac{t}{2}) = \frac{\mathrm{md}_\mathcal{K}(t)}{\mathrm{sn}_\mathcal{K}(t)}$ this leads to

$$\mathrm{tn}_\mathcal{K}(\frac{t}{2}) \leq \frac{\mathrm{tn}_\mathcal{K}(\varepsilon)}{\mathrm{sn}_\kappa(\varepsilon)} \cdot \left(\frac{\mathrm{md}_\kappa(\varepsilon) - \mathrm{md}_\kappa(r)}{\mathrm{tn}_\kappa(\tau)} + \mathrm{sn}_\kappa(r)\right). \quad (3.18)$$

Combining this with the requirement (3.6) we finally conclude $\mathrm{tn}_\mathcal{K}(\frac{t}{2}) \leq \mathrm{tn}_\mathcal{K}(\frac{\varepsilon - s}{2})$, and by monotonicity of tn_* we have $t \leq \varepsilon - s$. □

III.5 Sublemma: Monotonicity of the md-inequality
The inequality (3.13) *is still valid if we replace a with ε, where $0 < \varepsilon \leq a$, as long as inequality* (3.5) *holds.*

Proof. We see with the same kind of reasoning as before and by

$$\mathrm{md}(a) + \mathrm{md}(b)\,\mathrm{cs}(a) = \mathrm{md}(b) + \mathrm{md}(a)\,\mathrm{cs}_\kappa(b),$$

that (3.13) is equivalent to

$$\frac{\mathrm{md}_\kappa(\tau - r) - \mathrm{md}_\kappa(\tau) - \mathrm{cs}_\kappa(\tau)\,\mathrm{md}_\kappa(a)}{\mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(a)} \leq \frac{\mathrm{md}_\mathcal{K}(\varepsilon) - \mathrm{md}_\mathcal{K}(t) - \mathrm{cs}_\mathcal{K}(t)\,\mathrm{md}_\mathcal{K}(a)}{\mathrm{sn}_\mathcal{K}(t)\,\mathrm{sn}_\mathcal{K}(a)} \tag{3.19}$$

this transforms to the following function being increasing with respect to a

$$f(a) = \frac{\mathrm{sn}_\mathcal{K}(t)\,\mathrm{sn}_\mathcal{K}(a)}{\mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(a)}\left(\mathrm{md}_\kappa(\tau - r) - \mathrm{md}_\kappa(\tau) - \mathrm{cs}_\kappa(\tau)\,\mathrm{md}_\kappa(a)\right) + \mathrm{cs}_\mathcal{K}(t)\,\mathrm{md}_\mathcal{K}(a). \tag{3.20}$$

Which can be checked by looking at its derivative:

$$\begin{aligned}\frac{d}{da}f'(a) = \frac{\mathrm{sn}_\mathcal{K}(t)}{\mathrm{sn}_\kappa(\tau)}\underbrace{\frac{\mathrm{cs}_\mathcal{K}(a)\,\mathrm{sn}_\kappa(a) - \mathrm{sn}_\mathcal{K}(a)\,\mathrm{cs}_\kappa(a)}{\mathrm{sn}^2_\kappa(a)}}_{\leq 0} \cdot \underbrace{[\mathrm{md}_\kappa(\tau - r) - \mathrm{md}_\kappa(\tau) - \mathrm{cs}_\kappa(\tau)\,\mathrm{md}_\kappa(a)]}_{\leq 0} \\ + \underbrace{\frac{\mathrm{sn}_\mathcal{K}(t)\,\mathrm{sn}_\mathcal{K}(a)}{\mathrm{sn}_\kappa(\tau)\,\mathrm{sn}_\kappa(a)}\big(- \mathrm{cs}_\kappa(\tau)\,\mathrm{sn}_\kappa(a)\big) + \mathrm{cs}_\mathcal{K}(t)\,\mathrm{sn}_\mathcal{K}(a)}_{\geq 0} \geq 0\end{aligned} \tag{3.21}$$

Here the first underbraced term is negative as we have the equivalent term

$$\frac{\mathrm{cs}_\mathcal{K}(a)\,\mathrm{cs}_\kappa(a)}{\mathrm{sn}^2_\kappa(a)}(\mathrm{tn}_\kappa(a) - \mathrm{tn}_\mathcal{K}(a)) \leq 0,$$

It is negative since tn_k is monotonic increasing with respect to k. The second underbraced term is negative by monotonicity of the modifying function. And the third term being positive is equivalent to

$$\frac{\mathrm{sn}_\mathcal{K}(t)\,\mathrm{sn}_\mathcal{K}(a)}{\mathrm{tn}_\mathcal{K}(t)\,\mathrm{tn}_\kappa(\tau)}\big(\,\mathrm{tn}_\kappa(\tau) - \mathrm{tn}_\mathcal{K}(t)\big) \geq 0$$

which is true as we have $t \leq \varepsilon + r$ and inequality (3.5). $\square$

III.6 Proposition: *Let $E \xrightarrow{\pi} B$ a vector bundle and $\mathcal{U}$ an open, fibrewise star-shaped subset of E containing the zero section as centre. Then there exists a fibre-invariant diffeomorphism $\Phi : \mathcal{U} \to E$, which is the identity on a neighbourhood of the zero-section. Where fibre-invariant means that for $X \in \mathcal{U}$ we have $\Phi(X) \in \mathbb{R} \cdot X$*

Proof. We start with defining auxiliary the functions $\mathcal{R} : \mathcal{U} \to (1, \infty]$

$$\mathcal{R}(x) := \sup\{t > 0 | t \cdot x \in \mathcal{U}\}$$

and $r : \mathcal{U} \to (0;1]$

$$r(x) := \min\{\mathcal{R}(x) - 1, 1\}.$$

At first we show that the function r is semi-continuous from below, i.e. for every point $x \in \mathcal{U}$ and for all $\varepsilon > 0$ we have a neighbourhood V_ε of x such that $r(V_\varepsilon) > r(x) - \varepsilon$. So for $s > 0$ we define the open set $U_s := \{y \in \mathcal{U} | (1+s)y \in \mathcal{U}\}$. We have two cases:

If $r(x) - \varepsilon \leq 0$, then put $V_\varepsilon = \mathcal{U}$.

If $r(x) - \varepsilon > 0$, then there exists an $s_\varepsilon \in (r(x) - \varepsilon, r(x))$, then put $V_\varepsilon := U_{s_\varepsilon}$.

Which proves the semi-continuity of r. By partition of unity we get a smooth function $\chi : \mathcal{U} \to (0,1]$ such that $\chi(x) \leq r(x)$ and $\chi|_V \equiv 1$ for an open set V, such that $V \subseteq U_1$. Next we define the smooth scaling function $\lambda : \mathcal{U} \longrightarrow [1, \infty]$ by

$$\lambda(x) := \int_0^1 \frac{1}{\chi(\tau x)} d\tau$$

and at last the smooth map $\Phi : \mathcal{U} \longrightarrow E$

$$\Phi(x) := \lambda(x) \cdot x$$

A short calculation shows $\lambda(tx) = \frac{1}{t} \int_0^t \frac{1}{\chi(\tau x)} d\tau$, which implies $\Phi(tx) = \int_0^t \frac{1}{\chi(\tau x)} d\tau x$. Now note that the mapping $t \mapsto \int_0^t \frac{1}{\chi(\tau x)} d\tau$ is strictly monotonic increasing, since $\chi > 0$ which shows that Φ is injective. To see that Φ is surjective we have to look at two cases:

(i) If $\mathcal{R}(x) = \infty$ we have

$$\int_0^{\mathcal{R}(x)} \frac{1}{\chi(\tau x)} d\tau = \int_0^{\infty} 1 d\tau = \infty.$$

(ii) If $\mathcal{R}(x) < \infty$ we have $\mathcal{R}(tx) = \frac{\mathcal{R}(x)}{t}$, and $\chi(tx) \leq \frac{\mathcal{R}(x)}{t} - 1$. So the integral may be estimated by

$$\int_0^{\mathcal{R}(x)} \frac{1}{\chi(\tau x)} d\tau \geq \int_0^{\mathcal{R}(x)} \frac{\tau}{\mathcal{R}(x) - \tau} d\tau = \infty.$$

It remains to show that the derivative of Φ is a linear isomorphism, then we apply the *Inverse Function Theorem* see [AE06, Theorem 7.3, p. 215] to see that Φ is a diffeomorphism.

The derivative of Φ at a point x is a local concept, so it can be calculated in $E|_V$, a small neighbourhood of x. For this neighbourhood $E|_V$ we can use a local trivialization $V \times \mathbb{R}^d$. Then we write $(x, \xi) \overset{\Phi}{\longmapsto} (x, \lambda(x;\xi)\xi)$ and the derivative is the block matrix

$$D_{(x,\xi)}\Phi = \left(\begin{array}{c|c} \mathrm{id} & 0 \\ \hline * & A_{(x,\xi)} \end{array}\right)$$

where the sub-matrix $A_{(x,\xi)}$ is the derivative of Φ restricted to the fibre over x. This mapping Φ_x is an endomorphism on the fibre $\Phi_x : E_x \to E_x$ and given by $\xi \mapsto \lambda_x(\xi)\xi$, thus its derivative is $(D_\xi \Phi_x)\eta = d_\xi\lambda_x(\eta)\xi + \lambda_x(\xi)\eta$. If we calculate the kernel of $D\Phi_x$ we see the following equivalent equations are true

$$\begin{aligned} (D_\xi \Phi_x)\eta &= 0 \\ d_\xi\lambda_x(\eta)\xi + \lambda_x(\xi)\eta &= 0 \\ \eta &= -\frac{d_\xi\lambda_x(\eta)}{\lambda_x(\xi)}\xi. \end{aligned}$$

So η is a radial vector, i.e. it is of the form $\mu \cdot \xi$ for some constant μ. If η is the zero vector, then Φ_x is the identity mapping for a small neighbourhood, thus the derivative is invertible. Otherwise equation (3.1) implies that

$$d\lambda_x(\mu\xi)\xi + \lambda_x(\xi)\mu\xi = (d\lambda_x(\xi) + \lambda_x(\xi))\mu\xi = 0,$$

and using that $t \mapsto \int_0^t \frac{1}{\chi(\tau x)} d\tau$ is strictly monotonic increasing we see $(d_\xi\lambda_x(\xi) + \lambda_x(\xi)) > 0$, so $\mu = 0$. Thus $D\Phi_x$ is a linear isomorphism and $\Phi|_V \equiv \mathrm{id}$ by construction. $\square$

This concludes the proof of Theorem III.2. We have shown that for small distances $t < (\varepsilon - s)$ the tubular neighbourhood $\mathcal{B}_t(S)$ of S is contained in $\mathcal{U}$. Then we have shown that $\exp^{-1}(\mathcal{U}) \cap \mathcal{B}_\tau(S)$ is open and fibrewise star-shaped with centre 0. And in the last step we saw that a set like this is diffeomorphic to the full normal bundle. $\square$

III.7 Corollary: *The submanifold S is a strong deformation retract of $\mathcal{U}$.*

Proof. As we have $\exp^{-1}(\mathcal{U}) \cap \mathcal{B}_\tau(S)$ being diffeomorphic to the normal bundle $T^\perp S$ by a fibre-invariant diffeomorphism Φ, i.e. $\Phi(x,\xi) \in \{(x,t\xi)|t \in \mathbb{R}\}$. It suffices to show that $T^\perp S$ is a strong deformation - which can be seen with the homotopy relative the zero-section given by $H_t(x,\xi) := (x,t\xi)$. □

3.2. Building a homotopy equivalent simplicial complex

In this section we will derive an important corollary of the previous theorem, namely that the ε-Čech complex built from the sample points, has the same homotopy type as the submanifold S. Where the Čech-complex is a simplicial complex, whose vertices are the sample points $x_1, \dots, x_N$ and we add a k-cell for each non-empty intersection $\bigcap_{i=1}^k \mathcal{B}_\varepsilon(x_{j_i})$. Moreover this finite complex makes it feasible to calculate the homology groups as well as the Betti numbers. I will assume that the reader is familiar with basic concepts of algebraic topology see [Koz08], [Die08], [Spa94] or [Hat02] for further reference. Nevertheless I will provide the most basic definitions to understand the Nerve Lemma III.8, which asserts that for a paracompact Hausdorff space X and an open covering $\mathcal{U}$, such that every finite, non-empty intersections of cover sets is contractible, the nerve of this covering is homotopy equivalent to X. Combining this lemma with the result of Theorem III.2 we get a simplicial complex to calculate the homology and Betti numbers from, which are the same as the homology and Betti numbers of the submanifold S.

We need to introduce two concepts from algebraic topology one is homotopy equivalence, which is an equivalence relation on spaces and the other is simplicial complexes.

Homotopy and Homotopy Equivalence

For X, Y topological spaces and two continuous functions $h_0, h_1 : X \to Y$ a *homotopy* is a function $H : [0,1] \to \mathcal{C}(X,Y)$ with the two properties:

(i) $\hat{H} : [0,1] \times X \to Y$ is continuous as a mapping from the product (with product topology) to Y

(ii) $H(0) = h_0$ and $H(1) = h_1$.

The functions h_0, h_1 are then called *homotopic*, denoted by $h_0 \sim h_1$. This is an equivalence relation with equivalence classes $[f] := \{h \in \mathcal{C}(X,Y)|h \sim f\}$.

Let $f, g : X \to Y$ be continuous maps then the pair (f, g) is called a *homotopy equivalence* if $f \circ g \sim \mathrm{id}_X$ and $g \circ f \sim \mathrm{id}_Y$. The spaces X and Y are then called *homotopy equivalent*, or to be of the same *homotopy type*, and denoted by $X \simeq Y$. A space X will be called *contractible*, if $X \simeq \{*\}$ or equivalently $\mathrm{id}_X \sim \mathrm{const}_*$. For a space X and a subspace A, a *homotopy relative* A is a homotopy H such that $H_t|_A$ is constant for all t. We say a subspace $A \subseteq X$ is a *(strict) deformation retract* if there exists a homotopy relative A from id_X to a retraction $r : X \to A$.

Simplicial Complexes

A finite *simplicial complex* $\mathbf{K}$ is a pair $(\mathbf{V}, \mathbf{S})$ where $\mathbf{V}$ is a finite set, called the vertices and $\mathbf{S}$, the *simplices*, is a set of finite non-empty subsets of $\mathbf{V}$, with the following properties:

(i) $\{x\} \in \mathbf{S}$ for all $x \in \mathbf{V}$

(ii) If $\sigma \in \mathbf{S}$ and $\tau \subseteq \sigma$ then $\tau \in \mathbf{S}$, such a $\tau \subset \sigma$ is called a face of σ.

A $\sigma \in \mathbf{S}$ with $(q+1)$-many elements is called a *q-simplex* or q-dimensional simplex of $\mathbf{K}$. We then say a simplicial complex has dimension q, if the maximal dimension of its simplices is q. The 0-simplices and 1-simplices will be called *vertices* and *edges*. Note that a simplex is determined by its vertices.

For a given simplicial complex $\mathbf{K}$ we have an associated topological space $|\mathbf{K}|$ the *geometric realization*. We start with $\mathbb{R}^{\mathbf{V}}$ and build $|\mathbf{K}|$ as a subspace of this product. We define $|\mathbf{K}|$ to be the set of functions $\varphi := \mathbf{V} \to [0, 1]$ with the two properties:

(i) $\{x \in \mathbf{V} | \varphi(x) > 0\}$ is a simplex of $\mathbf{K}$.

(ii) $\sum_{x \in V} \varphi(x) = 1$

and we equip $|\mathbf{K}|$ with the subset topology of the product space $[0, 1]^{\mathbf{V}}$.

For a topological space X and a covering $\mathcal{U} := \{U_i\}_{i \in I}$. The *nerve* of $\mathcal{U}$ is the simplicial complex $\mathcal{N}(\mathcal{U}) = (\mathbf{V}, \mathbf{S})$, with vertices given by the index set $\mathbf{S} = I$, and the set $\{i_0, \dots, i_k\} \in \mathbf{S} :\Leftrightarrow \bigcap_{j=0}^{k} U_{i_j} \neq \emptyset$. If this covering is the union of ε-balls we will call this the *ε-Čech complex*.

At last I want to show the nerve lemma, as the proof requires a bit more preparation. For a full proof and a bit more information I strongly recommend reading [Hat02, Section 4.G] or [Koz08] for a more category theoretic approach.

III.8 Theorem: Nerve Lemma
If $\mathcal{U}$ is an open cover of a paracompact Hausdorff space X, such that every non-empty intersection of finitely many sets in $\mathcal{U}$ is contractible, then X is homotopy equivalent to the nerve $\mathcal{N}(\mathcal{U})$

Now we can combine Theorem III.2 and the Nerve Lemma III.8 to see.

III.9 Corollary: *For a submanifold S and a covering $\mathcal{U}$ as defined in Theorem III.2 we have the nerve of $\mathcal{U}$ is homotopy equivalent to S.*

Proof. The Theorem III.2 shows that the covering is open and star-shaped in each fibre, in particular we have every intersection of finitely many balls around sampling points is convex, by the requirement of inequality (3.2), $\varepsilon < r_{cvx}$, and thus contractible. So we can apply the Nerve Lemma III.8 to see $S \simeq \mathcal{N}(\mathcal{U})$. □

3.3. Corollaries and special cases

We will now spend some time to examine the conditions we demanded in Theorem III.2 more detailed and compare them to the results, given in [NSW08], where the manifold M is always the Euclidean space $\mathbb{R}^n$.

First note that the condition (3.5), which asserts that inequality (3.13) still holds if we substitute a with ε, is equivalent to inequality (3.1), if the manifold M has constant (sectional) curvature. Though in general it is hard to check whether the inequality (3.6)

$$\frac{\mathrm{tn}_{\mathcal{K}}(\varepsilon)}{\mathrm{sn}_{\kappa}(\varepsilon)} \cdot \left(\frac{\mathrm{md}_{\kappa}(\varepsilon) - \mathrm{md}_{\kappa}(r)}{\mathrm{tn}_{\kappa}(\tau)} + \mathrm{sn}_{\kappa}(r) \right) \leq \mathrm{tn}_{\mathcal{K}} \left(\frac{\varepsilon - s}{2} \right) \tag{3.6}$$

is satisfied, one can see that in various special cases it can be simplified to a quadratic inequality.

The first special case we look at is the situation, where the points are taken from the submanifold without sampling errors.

III.10 Corollary: *Let M, S as in Theorem III.2, but let $\bar{x} := (x_1, \ldots, x_N) \subseteq S$ be chosen such that $\bigcup_{i=1}^{N} \mathcal{B}_s(x_i)$ is a covering of S. If the following conditions hold:*

$$\varepsilon \leq \min\{r_{cvx}, \frac{r_{inj}}{2}, \tau\} \tag{3.22}$$

$$2\varepsilon \leq \frac{\pi}{\sqrt{\mathcal{K}}} \text{ in the case of } \mathcal{K} > 0 \tag{3.23}$$

$$\mathrm{tn}_{\mathcal{K}}(\varepsilon) \leq \mathrm{tn}_\kappa(\tau) \text{ and} \tag{3.24}$$
$$\frac{\mathrm{tn}_{\mathcal{K}}(\varepsilon)\,\mathrm{tn}_\kappa(\frac{\varepsilon}{2})}{\mathrm{tn}_{\mathcal{K}}(\frac{\varepsilon-s}{2})} \leq \mathrm{tn}_\kappa(\tau) \tag{3.25}$$

Then $\mathcal{U} := \bigcup_{i=1}^N \mathcal{B}_\varepsilon(x_i)$ is diffeomorphic to $\mathcal{B}_\tau(S)$, and $\exp^{-1}(\mathcal{U}) \cap \mathcal{B}_\tau(S)$ is open and fibrewise star-shaped with the zero section as its centre.

Proof. The only thing to show is that (3.6) is equivalent to (3.25). To see this we apply the equation $\mathrm{md}_\kappa(x) = \mathrm{sn}_\kappa(x)\,\mathrm{tn}_\kappa(x/2)$ in Lemma II.7 to (3.6) with $r = 0$

$$\frac{\mathrm{tn}_{\mathcal{K}}(\varepsilon)}{\mathrm{sn}_\kappa(\varepsilon)} \cdot \left(\frac{\mathrm{md}_\kappa(\varepsilon)}{\mathrm{tn}_\kappa(\tau)}\right) \leq \mathrm{tn}_{\mathcal{K}}\left(\frac{\varepsilon - s}{2}\right)$$

and obtain

$$\frac{\mathrm{tn}_{\mathcal{K}}(\varepsilon)}{\mathrm{tn}_\kappa(\tau)}\,\mathrm{tn}_\kappa(\frac{\varepsilon}{2}) \leq \mathrm{tn}_{\mathcal{K}}\left(\frac{\varepsilon - s}{2}\right)$$

□

Another interesting special case, which we want to examine is the case of complete manifolds with constant sectional curvature.

III.11 Corollary: *Let M be a simply connected complete Riemannian manifold with constant sectional curvature κ, and S a closed submanifold, that has condition number τ. Furthermore for $r \geq 0$ we have points $\bar{x} := (x_1, \ldots, x_N) \in \mathcal{B}_r(S)$ such that $S \subseteq \mathcal{B}_s(\bar{x})$, where $0 < s < \varepsilon$. Now suppose the following conditions hold:*

$$\varepsilon + r \leq \tau \tag{3.26}$$
$$\frac{1}{\mathrm{cs}_\kappa(\varepsilon)} \cdot \left(\frac{\mathrm{md}_\kappa(\varepsilon) - \mathrm{md}_\kappa(r)}{\mathrm{tn}_\kappa(\tau)} + \mathrm{sn}_\kappa(r)\right) \leq \mathrm{tn}_\kappa\left(\frac{\varepsilon - s}{2}\right) \tag{3.27}$$

Then $\mathcal{U} := \bigcup_{i=1}^N \mathcal{B}_\varepsilon(x_i)$ is diffeomorphic to $\mathcal{B}_\tau(S)$, and $\exp^{-1}(\mathcal{U}) \cap \mathcal{B}_\tau(S)$ is open and fibrewise star-shaped with the zero section as its centre.

Proof. First note that the manifold M is isometric to $\mathbb{M}_\kappa^n$ and Myers' Diameter Estimate II.15 asserts that $\tau \leq \frac{\pi}{\sqrt{\kappa}}$ in the case of positive κ. In addition we know $\tau \leq \mathrm{diam}(\mathbb{M}_\kappa^n) = r_{inj}^{(\mathbb{M}_\kappa^n)}$, combined with condition (3.26) this implies $r \leq r_{inj}$. Next note that in the case of non-positive curvature the convexity radius is infinity, and otherwise equal to the injectivity radius. At last we note that (3.6) is equivalent to (3.27). □

Now take a closer look at the setting examined in [NSW08], where the surrounding manifold M is just the n-dimensional Euclidean space and S being a submanifold of it.

Specializing the previous result to $\kappa = 0$ we get the following corollary.

III.12 Corollary: *Let S be a closed submanifold of $\mathbb{R}^n$ with condition number τ. For $r > 0$ let $\bar{x} := (x_1, \ldots, x_N) \subseteq \mathcal{B}_r(S)$ such that $S \subseteq \bigcup_{i=1}^N \mathcal{B}_r(x_i)$. If the following conditions are met,*

$$\varepsilon + r \leq \tau \tag{3.28}$$

$$\varepsilon \in \left(\frac{\tau}{2} - \sqrt{\frac{\tau^2}{4} - 3r\tau + r^2}, \frac{\tau}{2} + \sqrt{\frac{\tau^2}{4} - 3r\tau + r^2} \right) \text{ which requires} \tag{3.29}$$

$$0 \leq r \leq \frac{\sqrt{9} - \sqrt{8}}{2} \tau \tag{3.30}$$

Then $\mathcal{U} := \bigcup_{i=1}^N \mathcal{B}_\varepsilon(x_i)$ is diffeomorphic to $\mathcal{B}_\tau(S)$, and $\exp^{-1}(\mathcal{U}) \cap \mathcal{B}_\tau(S)$ is open and fibrewise star-shaped with the zero section as its centre.

Proof. With the definitions of sn_κ, cs_κ, tn_κ and md_κ we have an equivalent condition

$$\frac{\mathrm{tn}_0(\varepsilon)}{\mathrm{sn}_0(\varepsilon)} \cdot \left(\frac{\mathrm{md}_0(\varepsilon) - \mathrm{md}_0(r)}{\mathrm{tn}_0(\tau)} + \mathrm{sn}_0(r) \right) \leq \mathrm{tn}_0 \left(\frac{\varepsilon - s}{2} \right),$$

and with $r = s$ this results in

$$\frac{\varepsilon^2 - r^2}{2\tau} + r \leq \frac{\varepsilon - r}{2} \tag{3.31}$$

$$\varepsilon^2 - r^2 + 2\tau r \leq \varepsilon\tau - r\tau \tag{3.32}$$

$$\varepsilon^2 - \tau\varepsilon + 3r\tau - r^2 \leq 0 \tag{3.33}$$

If we regard the left hand side as a function with parameter ε, we see this is a parabola with a minimum at $\tau/2$ and the inequality holds as long as

$$\varepsilon \in \left(\frac{\tau}{2} - \sqrt{\frac{\tau^2}{4} - 3r\tau + r^2}, \frac{\tau}{2} + \sqrt{\frac{\tau^2}{4} - 3r\tau + r^2} \right).$$

This requires the discriminant to be non-negative, thus

$$r^2 - 3\tau r + \frac{\tau^2}{4} \geq 0. \tag{3.34}$$

This is positive if r is smaller than the lower root of this polynomial:

$$r \leq \frac{3 - 2\sqrt{2}}{2}\tau \tag{3.35}$$

or equivalently

$$r \leq \frac{\sqrt{9} - \sqrt{8}}{2}\tau. \tag{3.36}$$

This is a slightly less optimal result compared to the version shown in [NSW08, Proposition 7.2], this is due to the fact that Niyogi, Smale and Weinberger use better estimates that allows them to extend the upper bound for t further than $\varepsilon - s$, which was just a consequence of the triangle inequality. The authors in comparison use the Theorem of Pythagoras and some more elaborate worst case scenarios, which I could, unfortunately, not generalize to the setting of Riemannian geometry. $\square$

And at last the most special case, where the sample points are drawn from the submanifold itself, this corresponds to a real world situation, where the sampling error would be negligible.

III.13 Corollary: *Let S be a closed submanifold of $\mathbb{R}^n$ with condition number τ. And let $\bar{x} := (x_1, \ldots, x_N) \subseteq S$ such that $S \subseteq \bigcup_{i=1}^N \mathcal{B}_s(x_i)$. If the following conditions are met,*

$$\varepsilon \leq \tau \tag{3.37}$$

$$\varepsilon \in \left(\frac{\tau - \sqrt{\tau^2 - 4\tau s}}{2}, \frac{\tau + \sqrt{\tau^2 - 4\tau s}}{2} \right) \text{ which requires} \tag{3.38}$$

$$s \leq \frac{\tau}{4}, \tag{3.39}$$

we have $\mathcal{U} := \bigcup_{i=1}^N \mathcal{B}_\varepsilon(x_i)$ being diffeomorphic to $\mathcal{B}_\tau(S)$ and $\exp^{-1}(\mathcal{U}) \cap \mathcal{B}_\tau(S)$ is open and star-shaped with the zero section as its centre.

Proof. The only thing to show is that (3.38) and (3.39) imply (3.6). The condition (3.6)

$$\frac{\mathrm{tn}_0(\varepsilon)}{\mathrm{sn}_0(\varepsilon)} \cdot \left(\frac{\mathrm{md}_0(\varepsilon) - \mathrm{md}_0(r)}{\mathrm{tn}_0(\tau)} + \mathrm{sn}_0(r) \right) \leq \mathrm{tn}_0\left(\frac{\varepsilon - s}{2} \right)$$

with $r = 0$, $\mathrm{sn}_0(t) = \mathrm{tn}_0(t) = t$, $\mathrm{md}_0(t) = \frac{t^2}{2}$ results in

$$\frac{\varepsilon^2}{\tau} \leq \varepsilon - s$$

or equivalently

$$\varepsilon^2 - \tau\varepsilon + \tau s \leq 0.$$

again this is a parabola with minimum at $\tau/2$ and holds for

$$\varepsilon \in \left(\frac{\tau - \sqrt{\tau^2 - 4\tau s}}{2}, \frac{\tau + \sqrt{\tau^2 - 4\tau s}}{2} \right).$$

This requires the discriminant to be non-negative, so we see

$$\tau^2 - 4\tau s \geq 0$$

and with $\tau > 0$ we get

$$s \leq \frac{\tau}{4}.$$

□

On a final note: One sees that the in Corollaries III.12 and III.13 the conditions (3.29) and (3.38) introduce lower bounds for the parameter ε in each case. Which should be no surprise as the parameter r and s prevent ε-balls from intersecting with S, if ε is small.

IV. Manifold Learning from Random Samples

We will start with some observations on the packing and covering numbers, and that one can be used to estimate the other. We will define an extended packing and extended covering number to analyse the situation of a submanifold S of a manifold M, which can be packed/covered by balls in M. And we will relate all concepts of packing and covering numbers.

The next section is used to derive analogue high-confidence estimates as in the paper [NSW08]. Where we start with a probability measure μ on a metric space X that has the property for all s-balls with centres in the subspace Y there exists a $k_s > 0$ such that $k_s \leq \inf\{\mu(\mathcal{B}_s(y))|y \in Y\}$. And then we will relate the number of sample points $x_1, \ldots, x_N$, this constant k_s, the extended covering number $\mathrm{ecov}_s(Y \subseteq X, \mathrm{d})$ to the probability of the sample points being s-dense, i.e. $Y \subseteq \bigcup_{i=1}^{N} \mathcal{B}_s(x_i)$. Then we will provide two estimates of the extended packing number one by the probability measure constant k_s, which we can apply to the situation of Theorem III.2. The other estimate we can prove uses the volume form and the second fundamental form of S and again we apply this to Theorem III.2 to get a high confidence result.

4.1. Packing and Covering Numbers in Metric Spaces

We begin with some observations of metric spaces and establish the terms of packing and covering number. For a metric space (X, d) we have the *s-packing number* defined as the maximum number of s-balls that do not overlap. And the *s-covering number* is the minimal number of s-balls one needs to cover X. We will denote them by $\mathrm{pck}_s(X, \mathrm{d})$ and $\mathrm{cov}_s(X, \mathrm{d})$ respectively. For a subset $Y \subseteq X$ we define the *extended s-packing number*, denoted $\mathrm{epck}_s(Y \subseteq X, \mathrm{d})$, to be the maximal number of s-balls in X with centres in Y that do not overlap. And analogously we define the *extended s-covering number* as the minimal number of s-balls in X

with centres in Y that cover Y. We will denote it by $\mathrm{ecov}_s(Y \subseteq X, \mathrm{d})$.

IV.1 Remark: on packing numbers

Let M be a Riemannian manifold and S a submanifold. Then we have to different ways to endow S with a metric. We have the intrinsic distance in S denoted by d^S and we can restrict the distance functional of M to S, denoted by $\mathrm{d}\,|_S$, to get another metric. Note that $\mathrm{d}^M\,|_S \le \mathrm{d}^S$ and therefore $\mathcal{B}^S_r(x) \subseteq \mathcal{B}_r(x) \cap S$, which

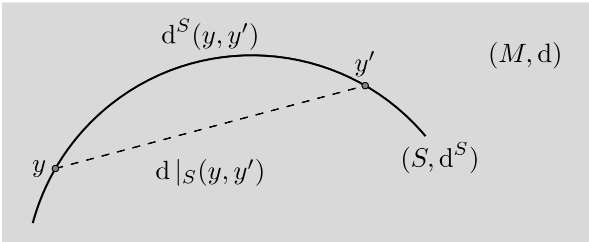


Figure IV.1.: The distance measured in S compared to the distance in M

implies the following on packing numbers and covering numbers.

$$\mathrm{pck}_s(S, \mathrm{d}^M\,|_S) \le \mathrm{pck}_s(S, \mathrm{d}^S) \qquad \mathrm{cov}_s(S, \mathrm{d}^M\,|_S) \le \mathrm{cov}_s(S, \mathrm{d}^S) \tag{4.1}$$

Or more informally there are more S-balls that fit in S, than restricted M-balls of the same radius, because they are smaller; and one needs less M-balls to cover S, as they are bigger.

IV.2 Proposition: relations of packing and covering numbers

Let M be a Riemannian manifold and S a submanifold, with respective Riemannian metrics g and g^S, and the respective distance functions d *and* d^S*. Then the (extended) packing numbers of S and the (extended) covering numbers of S meet the following relations:*

$$\begin{array}{ccccc}
\mathrm{pck}_s(S, \mathrm{d}^S) & \leqslant & \mathrm{cov}_s(S, \mathrm{d}^S) & \leqslant & \mathrm{pck}_{\frac{s}{2}}(S, \mathrm{d}^S) \\
\rotatebox{90}{$\leqslant$} & & \rotatebox{90}{$\leqslant$} & & \rotatebox{90}{$\leqslant$} \\
\mathrm{pck}_s(S, \mathrm{d}\,|_S) & \leqslant & \mathrm{cov}_s(S, \mathrm{d}\,|_S) & \leqslant & \mathrm{pck}_{\frac{s}{2}}(S, \mathrm{d}\,|_S) \\
\boxed{\rotatebox{90}{$\leqslant$}} & & \parallel & & \boxed{\rotatebox{90}{$\leqslant$}} \\
\mathrm{epck}_s(S \subseteq M, \mathrm{d}) & \leqslant & \mathrm{ecov}_s(S \subseteq M, \mathrm{d}) & \leqslant & \mathrm{epck}_{\frac{s}{2}}(S \subseteq M, \mathrm{d})
\end{array}$$

Figure IV.2.: A diagram of the relations between the different notions of packing and covering numbers of a submanifold S of M.

Proof. First note that we will denote by $\mathcal{B}_r(y)$ the balls in M, by $\mathcal{B}_r^S(y)$ the balls in S and by $\mathcal{B}_s|_S(y) := \mathcal{B}_s(y) \cap S$, the balls in M restricted to S. And we have the following inclusions

$$\mathcal{B}_s(y) \subseteq \mathcal{B}_s|_S(y) \subseteq \mathcal{B}_s(y).$$

A proof for the horizontal inequalities in the first and second line of Figure IV.2 can be found in [NSW08, Lemma 5.2]. The vertical inequalities between first and second line are consequences of $\mathrm{d}\,|_S \leq \mathrm{d}^S$.

The boxed inequalities are a consequence of $\mathcal{B}_s(y) \cap S \subseteq \mathcal{B}_s(y)$ and the fact that the $\mathcal{B}_s(y)$ may overlap outside of S. And the equality in the middle follows from the definition of $\mathcal{B}_s|_S(y)$. So if one has a minimal covering $\mathcal{B}_s|_S(y)_{i=1,\dots,n}$ then $(\mathcal{B}_s(y_i))_{i=1,\dots,n}$ is a minimal covering too, and vice versa.

Now only the last horizontal line of inequalities remains to be shown. The first one

$$\mathrm{pck}_s(S, \mathrm{d}\,|_S) \leq \mathrm{cov}_s(S, \mathrm{d}\,|_S)$$

is a consequence of the more general fact that any packing of S has less (or equal) balls than a covering of balls with the same radius. Which can be seen as follows. For a given packing every centre of an s-ball has to be contained in one of the covering balls. But no covering ball can contain more than one packing ball centre, because then such balls would overlap by triangle inequality.

At last the inequality at the bottom right is obtained as follows: Let $y_1, \dots, y_n$ be the centres of a maximal packing of S with $\mathcal{B}_{\frac{s}{2}}(y_i)$. Then $\mathcal{U} := \bigcup_{i=1}^n \mathcal{B}_s(y_i)$ covers all of S, because assume indirectly that there exists a $y \in S \setminus \mathcal{U}$. Then as a result of the triangle inequality we have

$$\mathcal{B}_{\frac{s}{2}}(y) \cap \bigcup_{i=1}^{n} \mathcal{B}_{\frac{s}{2}}(y_i) = \emptyset$$

which is a contradiction to $\mathrm{pck}_{\frac{s}{2}}(S, \mathrm{d}\,|_S)$ being maximal. □

4.2. Probability Theory applied to Manifold Learning

Following the proof from [NSW08, Lemma 5.1] we obtain:

IV.3 Lemma: estimate on the number of points
Let (X, d) be a metric space and $Y \subseteq X$ a subspace. And let μ be a probability

measure on X, for all $s > 0$ we define k_s by

$$k_s := \inf_{y \in Y} \{\mu(\mathcal{B}_s(y))\}. \tag{4.2}$$

And let $\bar{x} := x_1, \ldots, x_N \in X$ be identically and independently distributed drawn with respect to μ, or shorthand μ-i.i.d. drawn. Then

$$\mathbb{P}(Y \not\subseteq \mathcal{B}_s(\bar{x})) \leq \mathrm{ecov}_{\frac{s}{2}}(Y \subseteq X, \mathrm{d})(1 - k_{\frac{s}{2}})^N.$$

Note that if $k_s = 0$ then the estimate above is tautological. And if $Y = \emptyset$ we have $\mathrm{ecov}_s = 0$ *for all s, so the inequality above is also true.*

Proof. At first I want to clarify that y will always denote a point in Y, furthermore we will write $\mathcal{B}_r(x) := \{x' \in X | \, \mathrm{d}(x, x') < r\}$ and $\mathcal{B}_r|_Y(y) := \{y' \in Y | \, \mathrm{d}(y, y') < r\}$. The latter can also be realized as $\mathcal{B}_r(y) \cap Y$. Let $c := \mathrm{ecov}_{\frac{s}{2}}(Y \subseteq X, \mathrm{d})$ and $\bar{y} := y_1, \ldots, y_c \in Y$ such that $Y \subseteq \mathcal{B}_{\frac{s}{2}}(\bar{y})$ then the probability, that one of the x_i does *not* lie in one of the balls $\mathcal{B}_{\frac{s}{2}}(y_j)$ is given by:

$$\mathbb{P}\big(x_i \notin \mathcal{B}_{\frac{s}{2}}(y_j)\big) = 1 - \mu\big(\mathcal{B}_{\frac{s}{2}}(y_j)\big) \leq 1 - k_{\frac{s}{2}}.$$

This implies the probability that none of the $x_1, \ldots, x_N$ is in $\mathcal{B}_{\frac{s}{2}}(y_j)$ is

$$\mathbb{P}\big(\{\bar{x}\} \cap \mathcal{B}_{\frac{s}{2}}(y_j) = \emptyset\big) = [1 - \mu\big(\mathcal{B}_{\frac{s}{2}}(y_j)\big)]^N \leq [1 - k_{\frac{s}{2}}]^N,$$

and then

$$\begin{aligned}
\mathbb{P}\big(Y \not\subseteq \mathcal{B}_s(\bar{x})\big) &\leq \mathbb{P}\big(\exists j : \{\bar{x}\} \cap \mathcal{B}_{\frac{s}{2}}(y_j) = \emptyset\big) \leq \mathbb{P}\big(\bigcup_{j=1}^{c} [\{\bar{x}\} \cap \mathcal{B}_{\frac{s}{2}}(y_j) = \emptyset]\big) \\
&\leq \sum_{j=1}^{c} \mathbb{P}\big(\{\bar{x}\} \cap \mathcal{B}_{\frac{s}{2}}(y_j) = \emptyset\big) \leq \sum_{j=1}^{c} [1 - k_{\frac{s}{2}}]^N \\
&= c \cdot [1 - k_{\frac{s}{2}}]^N
\end{aligned}$$

□

Note that Niyogi, Smale and Weinberger use this to derive a lower bound for the number of sampling points. So if we have a $\delta > 0$ such that

$$\mathrm{ecov}_{\frac{s}{2}}(S \subseteq M, \mathrm{d}) \cdot [1 - k_{\frac{s}{2}}]^N \leq \delta.$$

Then we have $\mathbb{P}(S \subseteq \mathcal{B}_s(\bar{x})) \geq 1 - \delta$ if the number of sample points

$$N \geq \frac{\log(\mathrm{ecov}_{\frac{s}{2}}(S \subseteq M, \mathrm{d})) - \log(\delta)}{\log(\frac{1}{1-k_{\frac{s}{2}}})}. \tag{4.3}$$

IV.4 Lemma: *Let* (X, d) *be a manifold and* $Y \subseteq X$ *a subspace, let* μ *be a probability measure on* X*, let* k_s *be as previously defined. Then the extended packing number* $\mathrm{epck}_s(Y \subseteq X, \mathrm{d})$ *can be estimated by:*

$$\mathrm{epck}_s(Y \subseteq X, \mathrm{d}) \leq \frac{1}{k_s}, \tag{4.4}$$

where $\frac{1}{k_s} = \infty$ *in the case of* $k_s = 0$.

Proof. Let $c := \mathrm{epck}_s(Y \subseteq X, \mathrm{d})$ and $y_1, \ldots, y_c \in Y$ be the centres of a maximal extended packing of S with M-balls of radius s. Then we get

$$1 \geq \mu\big(\bigcup_{i=1}^{c} \mathcal{B}_s(y_i)\big) = \sum_{i=1}^{c} \mu\big(\mathcal{B}_s(y_i)\big) \geq c \cdot k_s = \mathrm{epck}_s(Y \subseteq X, \mathrm{d}) \cdot k_s.$$

□

IV.5 Corollary: *Let* M *be a complete Riemannian manifold with bounded sectional curvature* $\kappa \leq \sec \leq \mathcal{K}$ *and* S *a closed submanifold with condition number* τ*. Let* $r_{inj}, r_{cvx}, r, \varepsilon > 0$ *such that the inequalities* (3.1)-(3.6) *in Theorem III.2 hold. And let* μ *be a probability measure on* M *with* $k_s := \inf\{\mu(\mathcal{B}_s(y)) | y \in S\}$*. Let* $\bar{x} := (x_1, \ldots, x_N) \subseteq \mathcal{B}_r(S)$ *points i.i.d. drawn according to* μ*, if*

$$\frac{[1 - k_{\frac{s}{2}}]^N}{k_{\frac{s}{4}}} \leq \delta. \tag{4.5}$$

then for $\mathcal{U} := \mathcal{B}_\varepsilon(\bar{x})$ *we have* $\exp^{-1}(\mathcal{U}) \cap \mathcal{B}_\tau(S)$ *is an open fibrewise star-shaped neighbourhood with centre* 0*, with high confidence, i.e. the complementary event is true with probability less than* δ.

Proof. With Proposition IV.2 we see that $\mathrm{ecov}_{\frac{s}{2}}(S \subseteq M, \mathrm{d}) \leq \mathrm{epck}_{\frac{s}{4}}(S \subseteq M, \mathrm{d})$ and then we apply Lemma IV.4 to see that $\mathrm{epck}_{\frac{s}{4}}(S \subseteq M, \mathrm{d}) \leq \frac{1}{k_{\frac{s}{4}}}$. So

$$\mathrm{epck}_{\frac{s}{4}}(S \subseteq M, \mathrm{d}) \cdot [1 - k_{\frac{s}{2}}]^N \leq \frac{1}{k_{\frac{s}{4}}} \cdot [1 - k_{\frac{s}{2}}]^N \leq \delta$$

asserts $S \not\subseteq \bigcup_{i=1}^{N} \mathcal{B}_s(x_i)$ with probability less than δ. So Theorem III.2 applies with probability higher than $1-\delta$. □

IV.6 Lemma: estimate on the packing number with the volume form
Let M be a Riemannian manifold with injectivity radius $r_{inj} > 0$, and S be a closed k-submanifold then by equation (2.16) we have an upper bound on the sectional curvature of S given by

$$\omega := \sup_{y \in S} \left(\sec_y^M + 2\| \mathrm{II}_y \|^2 \right) < \infty.$$

then the packing number $\mathrm{pck}_s(S, \mathrm{d}^S)$ can be estimated by:

$$\mathrm{pck}_s(S, \mathrm{d}^S) \leq \frac{\mathrm{vol}(S)}{\frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} \int_0^s \mathrm{sn}_\omega^{n-1}(t)dt}$$

In the case of $\omega > 0$ we have to require that $s < \frac{\pi}{\sqrt{\omega}}$.

Proof. Let $c := \mathrm{pck}_s(S, \mathrm{d}^S)$ and $y_1, \ldots, y_c \in S$ be the centres of a maximal packing of S with S-balls of radius s. Then we get

$$\mathrm{vol}(S) \geq \mathrm{vol}\left(\bigcup_{i=1}^{c} \mathcal{B}_s^S(y_i) \right) = \sum_{i=1}^{c} \mathrm{vol}\left(\mathcal{B}_s^S(y_i) \right) = c \cdot \mathrm{vol}\left(\mathcal{B}_s^S(y_i) \right).$$

Now we apply Theorem II.17 and the formula for the volume of a ball in the model space with constant sectional curvature (2.71) to get

$$\mathrm{vol}(S) \geq c \cdot \frac{2\sqrt{\pi^n}}{\Gamma(\frac{n}{2})} \cdot \int_0^s \mathrm{sn}_\omega^{n-1}(t)dt.$$

□

Again with Proposition IV.2 we are able to deduce a high confidence version of our central theorem.

IV.7 Corollary: *Let M be a complete Riemannian manifold with bounded sectional curvature $\kappa \leq \sec \leq \mathcal{K}$. Let S a closed submanifold with condition number τ. Let $r_{inj}, r_{cvx}, r, \varepsilon > 0$ such that the inequalities (3.1)-(3.6) in Theorem III.2 hold. And let μ be a probability measure on M, such that for all $s > 0$ there exists a $k_s > 0$ such that*

$$k_s < \inf\{\mu(\mathcal{B}_s(y)) | y \in S\}.$$

Let $\bar{x} := (x_1, \ldots, x_N) \subseteq \mathcal{B}_r(S)$ points i.i.d. drawn according to μ, if

$$\frac{\operatorname{vol}(S)}{\frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\int_0^{\frac{s}{4}} \operatorname{sn}_\omega^{n-1}(t)dt} \cdot [1 - k_{\frac{s}{2}}]^N \leq \delta,$$

where $\omega := \mathcal{K} + 2\| \operatorname{II} \|^2$ and for $\omega > 0$ we have to require that $s < \frac{\pi}{\sqrt{\omega}}$. Then the probability that for $\mathcal{U} := \mathcal{B}_\varepsilon(\bar{x})$ we have $\exp^{-1}(\mathcal{U}) \cap \mathcal{B}_\tau(S)$ is an open fibre-wise star-shaped neighbourhood of the zero section with high confidence, i.e. with probability higher $1 - \delta$.

Proof. With Proposition IV.2 we see that $\operatorname{ecov}_{\frac{s}{2}}(S \subseteq M, \mathrm{d}) \leq \operatorname{pck}_{\frac{s}{4}}(S, \mathrm{d}^S)$ and then we apply Lemma IV.6 to see that $\operatorname{pck}_{\frac{s}{4}}(S, \mathrm{d}^S) \leq \frac{\operatorname{vol}(S)}{\frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\int_0^{\frac{s}{4}} \operatorname{sn}_\omega^{n-1}(t)dt}$. These two results combined show

$$\operatorname{ecov}_{\frac{s}{4}}(S \subseteq M, \mathrm{d}) \cdot [1 - k_{\frac{s}{2}}]^N \leq \frac{\operatorname{vol}(S) \cdot [1 - k_{\frac{s}{2}}]^N}{\frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\int_0^{\frac{s}{4}} \operatorname{sn}_\omega^{n-1}(t)dt} \leq \delta,$$

which asserts $S \not\subseteq \bigcup_{i=1}^N \mathcal{B}_s(x_i)$ with probability less than δ. So Theorem III.2 applies with probability higher than $1 - \delta$. $\square$

Bibliography

[AE06] H. Amann and J. Escher. *Analysis II.* Number v. 2 in Grundstudium Mathematik. Birkhäuser Basel, 2006.

[Car09] Gunnar Carlsson. Topology and data. *Bull. Amer. Math. Soc. (N.S.)*, 46(2):255–308, 2009. `doi:10.1090/S0273-0979-09-01249-X`.

[CC89] S.S. Chern and L. Cesari. *Global differential geometry.* Studies in mathematics. Mathematical Association of America, 1989.

[CES75] J. Cheeger, D.G. Ebin, and American Mathematical Society. *Comparison Theorems in Riemannian Geometry.* AMS Chelsea Publishing Series. AMS Chelsea Pub., 1975.

[Cha06] I. Chavel. *Riemannian Geometry: A Modern Introduction.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2006.

[dC92] M.P. do Carmo. *Riemannian geometry.* Theory and Applications Series. Birkhäuser, 1992.

[Die08] T. Dieck. *Algebraic topology.* EMS textbooks in mathematics. European Mathematical Society, 2008.

[Gri12] Worldwide LHC Computing Grid. Welcome to the worldwide lhc computing grid. http://wlcg.web.cern.ch/, 2012. [Online; accessed 20-December-2012].

[Gün60] Paul Günther. Einige Sätze über das Volumenelement eines Riemannschen Raumes. *Publ. Math. Debrecen*, 7:78–93, 1960.

[Hal11] Stefan Haller. Differentialgeometrie, 2011. Skriptum zur Vorlesung. URL: `http://www.mat.univie.ac.at/~stefan/files/DGIII2011/DG.pdf`.

[Hat02] A. Hatcher. *Algebraic Topology.* Cambridge University Press, 2002. URL: `http://www.math.cornell.edu/~hatcher/AT/ATpage.html`.

[Jos08] Jürgen Jost. *Riemannian geometry and geometric analysis.* Springer, 5th edition, 2008.

[Kar87] H. Karcher. *Riemannian comparison constructions.* Preprint. Sonderforschungsbereich 256. Nichtlineare Partielle Differentialgleichungen. Rheinische Friedrich-Wilhelms- Universität Bonn. Sonderforschunbgsber., Univ., 1987. URL: `http://people.math.uni-bonn.de/karcher/RiemCompConst.pdf`.

[Kli82] W. Klingenberg. *Riemannian geometry.* De Gruyter studies in mathematics. W. de Gruyter, 1982.

[Koz08] D.N. Kozlov. *Combinatorial Algebraic Topology.* Algorithms and computation in mathematics. Springer, 2008.

[Lan96] S. Lang. *Differential and Riemannian Manifolds.* Graduate Texts in Mathematics. Springer, 1996.

[Lee03] John M. Lee. *Introduction to Smooth Manifolds.* Graduate Texts in Mathematics. Springer, 2003.

[Lee09] Jeffrey M. Lee. *Manifolds and Differential Geometry.* Graduate Studies in Mathematics Series. American Mathematical Society, 2009.

[LJ11] Stan Z. Li and Anil K. Jain, editors. *Handbook of Face Recognition, 2nd Edition.* Springer, 2011.

[LPM03] Ann B. Lee, Kim Steenstrup Pedersen, and David Mumford. The nonlinear statistics of high-contrast patches in natural images. *International Journal of Computer Vision*, 54(1-3):83–103, 2003.

[LZ08] Tong Lin and Hongbin Zha. Riemannian manifold learning. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(5):796 –809, may 2008. `doi:10.1109/TPAMI.2007.70735`.

[LZL06] Tony Lin, Hongbin Zha, and Sang Uk Lee. Riemannian manifold learning for nonlinear dimensionality reduction. In *IN PROCEEDINGS OF ECCV*, 2006.

[Mey89] W. T. Meyer. Toponogov's theorem and applications. *Lecture Notes, College on Differential Geometry, Trieste*, pages 1–51, 1989. URL: `http://www.math.uni-muenster.de/u/meyer/publications/topo.ps`.

[NSW08] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry*, 39:419–441, 2008. `doi:10.1007/s00454-008-9053-2`.

[NSW11] P. Niyogi, S. Smale, and S. Weinberger. A topological view of unsupervised learning from noisy data. *SIAM J. Comput.*, 40(3):646–663, June 2011. `doi:10.1137/090762932`.

[Pet06] P. Petersen. *Riemannian geometry*. Number Bd. 10 in Graduate texts in mathematics. Springer, 2006.

[Spa94] E.H. Spanier. *Algebraic Topology*. McGraw-Hill series in higher mathematics. Springer-Verlag, 1994.

[vHvdS98] J. H. van Haeteren and A. van der Schaaf. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society B: Biological Sciences*, 265(1394):359366, 1998. PMC1688904. URL: `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1688904&rendertype=abstract`.

[Wik12] Wikipedia. Nonlinear dimension reduction — Wikipedia, the free encyclopedia, 2012. [Online; accessed 20-December-2012]. URL: `http://en.wikipedia.org/wiki/Nonlinear_dimensionality_reduction`.

[ZCPR00] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey, 2000.

Curriculum Vitae

Personal Information

Name	Martin Heuschober
Date of Birth	12th of August, 1984
Place of Birth	Linz, Austria

Education

since 10/2003	Diploma studies of Mathematics at the University of Vienna
06/2003	Matura (school leaving examination)
1995-2003	Bundesrealgymnasium Fadingerstraße, Linz
1990-1995	Volksschule VS 25, Linz

Employment

10/2006-07/2007	Tutor for undergraduate students
10/2008-07/2009	at University of Vienna
10/2009-07/2010	