2.3 Arc length and Riemannian distance

In this section we define the *arc length* of (piecewise smooth) curves. This notion in turn on *Riemannian* manifolds allows to define a notion of distance between two points p and q as the infimum of the arc length of all curves connecting p and q. We will show that this *Riemanninan distance function* encodes the topology of the manifold.

To begin with we introduce some notions on curves. If $I = [a, b]$ is a closed interval we call a curve $\alpha : [a, b] \to M$ piecewise smooth if there is a partition $a = t_0 < t_1 < \cdots < t_k = b$ of the interval $[a, b]$ such that all the restrictions $\alpha|_{[t_i,t_{i+1}]}$ are smooth. Thus at each of the so-called *break points* t_i the curve may well have two distinct velocity vectors $\alpha'(t_i^-)$ and $\alpha'(t_k^+)$, representing the left- resp. right-sided derivative at t_i. If I is an arbitrary interval we call $\alpha : I \to M$ *piecewise smooth* if for all $a < b$ in I the restriction $\alpha|_{[a,b]}$ is piecewise smooth (in the previous sense). In particular, the break points have no cluster point in I.

2.3.1 Definition (Arc length). *Let $\alpha : [a, b] \to M$ be a piecewise smooth curve into a SRMF M. We define the* arc length *(or* length*, for short) of α by*

$$L(\alpha) := \int_a^b \|\alpha'(s)\|\, ds. \tag{2.3.1}$$

We recall that $\|\alpha'(s)\| = |\langle\alpha'(t), \alpha'(t)\rangle|^{1/2}$ and in coordinates $\|\alpha'(s)\| = |g_{ij}(\alpha(s))\, \frac{d(x^i\circ\alpha)}{ds}(s)$ $\frac{d(x^j\circ\alpha)}{ds}(s)|^{1/2}$.
On Riemannian manifolds the arc length behaves much as in the Euclidean setting, cf. [9, Ch. 1]. However in the semi-Riemannian case with an indefinite metric there are new effects. For example null curves always have vanishing arc length.
A reparametrisation of α is a piecewise smooth function $h : [c, d] \to [a, b]$ such that $h(c) = a$ and $h(d) = b$ (orientation preserving) or $h(c) = b$ and $h(d) = a$ (orientation reversing). If $h'(t)$ does not change sign then h is monotone and precisely as in [9, 1.1.2, 1.1.3] we have.

2.3.2 Lemma (Parametrisations).

(i) The arc length 2.3.1 of a curve is invariant under monotonous reparametrisations.

(ii) If $\|\alpha'(t)\| > 0$ for all t then α possesses a strictly monotonous reparametrisation h such that for $\beta := \alpha \circ h$ we have $\|\beta'(s)\| = 1$ for all s. Such a parametrisation is called parametrisation by arc length.

Let now be $p \in M$ and let U be a normal neighbourhood of p. The function

$$r : U \to \mathbb{R}^+, \quad r(q) := \|\exp_p^{-1}(q)\| \tag{2.3.2}$$

is called the *radius function* of M at p. In Riemannian normal coordinates we have

$$r(q) = |-\sum_{i=1}^{r}(x^i)^2(q) + \sum_{j=r+1}^{n}(x^j)^2(q)|^{1/2}. \tag{2.3.3}$$

Hence r is smooth except at its zeroes, hence off p as well as off the local null cone at p. In a normal neighbourhood the radius is given exactly by the length of radial geodesics as we prove next.

2.3.3 Lemma (Radius in normal neighbourhoods). *Let r be the radius in a normal neighbourhood U of a point p in a SRMF M. If σ is the radial geodesic form p to some $q \in U$ the we have*

$$L(\sigma) = r(q). \tag{2.3.4}$$

Proof. Let $v = \sigma(0)$ then by 2.1.15 $v = \exp_p^{-1}(q)$. Now since σ is a geodesic $\|\sigma'\|$ is constant and we have

$$L(\sigma) = \int_0^1 \|\sigma'(s)\|\, ds = \int_0^1 \|v\|\, ds = \|v\| = \|\exp_p^{-1}(q)\| = r(q). \tag{2.3.5}$$

□

From now on we will until the end of this chapter *exclusively deal with Riemanninan manifolds*. Indeed it is only in the Riemannian case that the topology of the manifold is completely encoded in the metric.

2.3.4 Proposition (Radial geodesics are locally minimal). *Let M be a* Riemannian *manifold and let U be a normal neighbourhood of a point $p \in M$. If $q \in U$, then the radial geodesic $\sigma : [0, 1] \to M$ from p to q is the unique shortest piecewise smooth curve in U from p to q, where uniqueness holds up to monotone reparametrisations.*

Proof. Let $c : [0, 1] \to M$ a piecewise smooth curve form p to q and set $s(t) = r(c(t))$, with r the radius function of (2.3.2). Since $\exp_p$ is a diffeomorphism we may for $t \neq 0$ uniquely write c in the form

$$c(t) = \exp_p(s(t)v(t)) =: f(s(t), t), \tag{2.3.6}$$

where v is a curve in T_pM with $\|v(t)\| = 1$ for all t. (This amounts to using polar coordinates in T_pM.) Here $f(s, t) = \exp_p(sv(t))$ is a two-parameter map on a suitable domain and the function $s : (0, 1] \to \mathbb{R}^+$ is piecewise smooth. (Indeed we may suppose w.l.o.g. that $s(t) \neq 0$, i.e., $c(t) \neq p$ for all $t \in (0, 1]$ since otherwise we may define t_0 to be the last parameter value when $c(t_1) = p$ and replace c by $c|_{[t_0,1]}$.)
Now except for possibly finitely many values of t we have (cf. (2.1.19) and below)

$$\frac{dc(t)}{dt} = \frac{\partial f}{\partial s}(s(t), t)\, s'(t) + \frac{\partial f}{\partial t}(s(t), t). \tag{2.3.7}$$

From $f(s, t) = \exp_p(sv(t))$ we have

$$\frac{\partial f}{\partial s} = (T_{sv(t)} \exp_p)(v(t)) \quad \text{and} \quad \frac{\partial f}{\partial t} = (T_{sv(t)} \exp_p)(sv'(t)) \tag{2.3.8}$$

and by the Gauss lemma 2.1.21 we find

$$\langle \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle(s,t) = s\langle v(t), v'(t)\rangle = s\frac{1}{2}\frac{\partial}{\partial t}\underbrace{\|v(t)\|^2}_{=1} = 0 \tag{2.3.9}$$

and so $\frac{\partial f}{\partial s} \perp \frac{\partial f}{\partial t}$. Similarly we obtain via the Gauss lemma that $\|\frac{\partial f}{\partial s}\|^2 = \langle v(t), v(t)\rangle = 1$ which using (2.3.7) implies $\|\frac{dc}{dt}\|^2 = |s'(t)|^2 + \|\frac{\partial f}{\partial t}\|^2 \geq s'(t)^2$. This now gives for all $\varepsilon > 0$ that

$$\int_\varepsilon^1 \|c'(t)\|\, dt \geq \int_\varepsilon^1 |s'(t)|\, dt \geq \int_\varepsilon^1 s'(t)\, dt = s(1) - s(\varepsilon) \tag{2.3.10}$$

and hence in the limit $\varepsilon \to 0$ we find $L(c) \geq s(1) = r(q) = L(\sigma)$, where the final equality is due to 2.3.3.

Let conversely $L(c) = L(\sigma)$ then in the estimate (2.3.10) we have to have equality everywhere which enforces $\frac{\partial f}{\partial t}(s(t), t) = 0$ for all t. But since $T_{sv(t)} \exp_p$ is bijective this implies $v' \equiv 0$ and hence v has to be constant.
Moreover we need to have $|s'(t)| = s'(t) > 0$ and hence $c(t) = \exp_p(s(t)v)$ is a monotonous reparametrisation of $\sigma(t)$: Indeed by 2.1.15 we have $\sigma(t) = \exp_p(t \exp_p^{-1}(q))$ and moreover $\exp_p^{-1}(q) = \exp_p^{-1}(c(1)) = s(1)v$ and so $\sigma(t) = \exp_p(ts(1)v)$. Now with $h(t) := \frac{s(t)}{s(1)}$ we have $\sigma(h(t)) = \exp_p(\frac{s(t)}{s(1)}\, s(1)v) = \exp_p(s(t)v) = c(t)$. □

In $\mathbb{R}^n$ the distance between two points $d(p,q) = \|p - q\|$ is at the same time the length of the shortest curve between these two points, i.e., the straight line segment from p to q. But this ceases to be true even in the simple case of $\mathbb{R}^2 \setminus \{(0,0)\}$. Here there is no shortest path between the two points $p = (-1, 0)$ and $q = (1, 0)$. However, the infimum of the arc length of all paths connecting theses two points clearly equals the Euclidean distance 2 between p and q. ♣ `insert Figure` ♣ This idea woks also on general Riemannian manifolds and we start with the following definition.

2.3.5 Definition (Riemannian distance). *Let M be a* Riemannian *manifold and let p, $q \in M$. We define the set of 'permissible' paths connecting p and q by*

$$\Omega(p,q) := \{\alpha : \alpha \text{ is a piecewise smooth curve from } p \text{ to } q\}. \tag{2.3.11}$$

The Riemannian distance *$d(p,q)$ between p and q is then defined to be*

$$d(p,q) := \inf_{\alpha \in \Omega(p,q)} L(\alpha). \tag{2.3.12}$$

Furthermore we define just as in $\mathbb{R}^n$ the *ε-neighbourhood* of a point p in a RMF M for all $\varepsilon > 0$ by

$$U_\varepsilon(p) := \{q \in M : \ d(p,q) < \varepsilon\}. \tag{2.3.13}$$

Now in ε-neighbourhoods the usual Euclidean behaviour of geodesics holds true, more precisely that is:

2.3.6 Proposition (ε-neighbourhoods). *Let M be a RMF and let $p \in M$. Then for ε sufficiently small we have*

(i) $U_\varepsilon(p)$ is a normal neighbourhood.

(ii) The radial geodesic σ from p to any $q \in U_\varepsilon(p)$ is the unique shortest piecewise smooth curve in *M connecting p with q. In particular, we have $L(\sigma) = r(q) = d(p,q)$.*

Note the fact that the radial geodesic in 2.3.6(ii) is *globally* the shortest curve from p to q and in moreover that it is smooth.

Proof. Let U be a normal neighbourhood of p and denote by $\tilde{U}$ the corresponding neighbourhood of $0 \in T_pM$. Then for ε sufficiently small $\tilde{U}$ contains the star shaped open set $\tilde{N} = \tilde{N}_\varepsilon(0) := \{v \in T_pM : \|v\| < \varepsilon\}$. Therefore $N := \exp_p(\tilde{N})$ is also a normal neighbourhood.

By 2.3.4 for any $q \in N$ the radial geodesic σ from p to q is the unique shortest piecewise smooth curve *in* N from p to q. Moreover by 2.3.3 we have $L(\sigma) = r(q)$. Since $\sigma'(0) = \exp_p^{-1}(q) \in \tilde{N}$ we have $r(q) = \|\exp_p^{-1}(q)\| < \varepsilon$.

To finish the proof is suffices to show the following assertion:

$$\text{Any piecewise } \mathcal{C}^\infty\text{-curve } \alpha \text{ in } M \text{ starting in } p \text{ and leaving } N \text{ satisfies } L(\alpha) \geq \varepsilon. \quad (2.3.14)$$

Indeed then σ is the unique shortest piecewise smooth curve *in* M connecting p and q. But this implies that $r(q) = L(\sigma) = d(p,q)$ and we have shown (ii) but also $d(p,q) < \varepsilon$. Moreover for $q \notin N$ by (2.3.14) we have $d(p,q) \geq \varepsilon$. So in total $N = U_\varepsilon(p)$ implying (i). So it only remains to prove (2.3.14).

To begin with let $0 < a < \varepsilon$. We then have $\tilde{N}_a(0) \subseteq \tilde{N}$ and hence $\exp_p(\tilde{N}_a(0)) \subseteq N$. Now since α leaves N is also has to leave $\exp_p(\tilde{N}_a(0))$. Let $t_0 = \sup\{t : \alpha(t) \in \exp_p(\tilde{N}_a(0))\}$. Then $\alpha|_{[0,t_0]}$ is a curve connecting p with a point $q \in \partial \exp_p(\tilde{N}_a(0))$ which stays entirely in the slightly bigger normal neighbourhood $\exp_p(\tilde{N}_{a+\delta}(0))$ of p where δ is chosen such that $a + \delta < \varepsilon$. Now by 2.3.3 and 2.3.6 we have $L(\alpha) \geq L(\alpha_{[0,t_0]} \geq r(q) = a$. Finally for $a \to \varepsilon$ we obtain $L(\alpha) \geq \varepsilon$. □

2.3.7 Remark (Normal ε-neighbourhoods). Observe that the above proof also shows that any normal neighbourhood contains an ε-neighbourhood which itself is normal.

2.3.8 Example (Cylinder). Let M be the cylinder of 2.1.4 and denote by L any vertical line in M. Hence if $p \in M \setminus L$ then $M \setminus L$ is a normal neighbourhood of p and so $M \setminus L$ is convex. By 2.1.15 the radial geodesic σ (cf. Figure 2.6) is the unique shortest piecewise smooth curve from p to q in $M \setminus L$. However, the curve τ is obviously a shorter curve from p to q in M. *But τ leaves $M \setminus L$.*

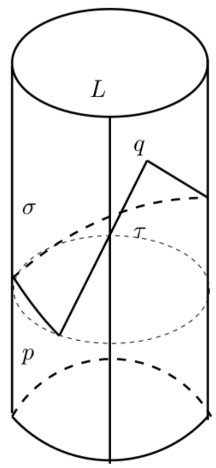

Figure 2.6: Convex neighbourhood on the cylinder

If p in M is arbitrary then the largest normal ε-neighbourhood of p is $U_\pi(p)$. This fact is most vividly seen from the (isometric hence equivalent) picture of the unwinded cylinder, see Figure 2.7. If $q \in U_\pi(p)$ then the radial geodesic σ form p to q is the unique shortest curve in all of M. If r is a point in $M \setminus U_\pi(p)$ there is still a shortest curve from p to r in M. If, however, r lies on the line L through the antipodal point p' of p then this curve is non-unique.

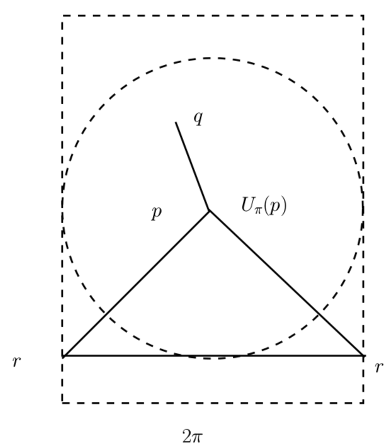

Figure 2.7: The largest normal ε-neighbourhood on the cylinder

2.3.9 Theorem (Riemannian distance). *Let M be a conected RMF. Then the Riemannian distance function $d : M \times M \to \mathbb{R}$ is a metric (in the topological sense) on M. Furthermore the topology induced by d on M coincides with the manifold topology of M.*

Proof. First of all note that d is finite: let $p, q \in M$ then by connectedness there exists a piecewise C^1-curve α from connecting p and q. and so $d(p,q) \leq L(\alpha) < \infty$.
Now we show that d actually is a metric on M.

Positive definiteness: Clearly $d(p,q) > 0$ for all p, q. Now assume $d(p,q) = 0$. We have to show that $p = q$. Suppose to the contrary that $p \neq q$. Then by the Hausdorff property of M there exists a normal neighbourhood U of p not containing q. But then by 2.3.7 U contains a normal ε-neighbourhood $U_\varepsilon(p)$ and so by (2.3.14) $d(p,q) \geq \varepsilon > 0$.

Symmetry: $d(p,q) = d(q,p)$ since $L(\alpha) = L(t \mapsto \alpha(-t))$.

Triangle inequality: Let $p, q, r \in M$. For $\varepsilon > 0$ let $\alpha \in \Omega(p,q)$, $\beta \in \Omega(q,r)$ such that $L(\alpha) < d(p,q) + \varepsilon$ and $L(\beta) < d(q,r) + \varepsilon$. Now define $\gamma = \alpha \cup \beta$. Then γ connects p with r and we have

$$d(p,r) \leq L(\gamma) = L(\alpha) + L(\beta) < d(p,q) + d(q,r) + 2\varepsilon \tag{2.3.15}$$

and since ε was arbitrary we conclude $d(p,r) \leq d(p,q) + d(q,r)$.

Finally by 2.2.7 the normal neighbourhoods provide a basis of neighbourhoods of p and by 2.3.7 every normal neighbourhood contains some $U_\varepsilon(p)$. Conversely every sufficiently small ε-neighbourhood of p is open and so d generates the manifold topology of M. □

We remark that there is also a proof which does not suppose the Riemannian metric to be smooth and the statement of the theorem remains true also for Riemannian metrics that are merely continuous.

2.3.10 Remark (Minimising curves). By the definition of the Riemannian distance function a piecewise smooth curve σ from p to q is a curve of minimal distance between these points if $L(\sigma) = d(p,q)$. In this case we call σ a *minimising curve*. In general there can be several such curves between a given pair of points; just consider the meridians running from the north pole to the south pole of the sphere.
Observe that every segment of a minimising curve form p to q is itself minimising (between its respective end points). Otherwise, by the triangle inequality there would be a shorter curve between p and q. ♣ `insert figure` ♣.

We finally complete our picture concerning the relationship between geodesics and minimising curves. By 2.1.15 radial geodesics are the unique minimising curves which connect the center of a normal neighborhood U to any point $q \in U$ and *stay within* U. Moreover by making U smaller, more precisely by considering a normal ε-neighborhood, radial geodesics become globally minimising curves, cf. 2.3.6. On the other hand geodesics that become to long, e.g. after leaving a normal neighborhood of its starting point need no longer be minimising; just again think of the sphere. However, if a curve is minimising it has to be a geodesic and hence it is even smooth and not merely piecewise smooth, as the following statement says.

2.3.11 Corollary (Minimising curves are geodesics). *Let p and q be points in a RMF M. Let α be a minimising curve from p to q then α is (up to monotone reparametrisations) a geodesic from p to q.*

Proof. Let $\alpha : [0,1] \to M$ a minimising piecewise smooth curve between $\alpha(0) = p$ and $\alpha(1) = q$. We may now find a finite partition $I = \cup_{i=1}^{k} I_i$ such that each segment $\alpha_i := \alpha|_{I_i}$ is contained in a convex set. We may also suppose w.o.l.g. that every α_i is non constant since otherwise we can leave out the interval I_i. By 2.3.10 every α_i is minimising and hence by 2.1.15 a monotone reparametrisation of a geodesics. Hence by patching togehter these reparametrisations we obtain a possibly broken geodesic σ with break points at the end points of the I_i. We have $L(\sigma) = L(\alpha) = d(p,q)$. So the corollary follows from the following statement which shows that there are actually no break points.

If a geodesic segment c_1 ending at r and a geodesic segment c_2 starting at r combine to give a minimising curve segment c, then c is an (unbroken) geodesic. (2.3.16)

Observe that the intuitive idea behind this statement is that rounding off a corner of γ near r would make γ shorter ♣ `insert figure` ♣.
To formally prove (2.3.16) we once again choose a convex neighborhood U around r. Then the end part of c_1 and the starting part of c_2 combine to a minimising curve $\bar{c}$ in U. Since U is normal for each of its points and particular for some $r' \neq r$ on $\bar{c}$ it follows by 2.1.15 that $\bar{c}$ is a radial geodesic. So $\bar{c}$ and hence c has no break point at r. □