Differential geometry: classical and modern

David Krejčiřík

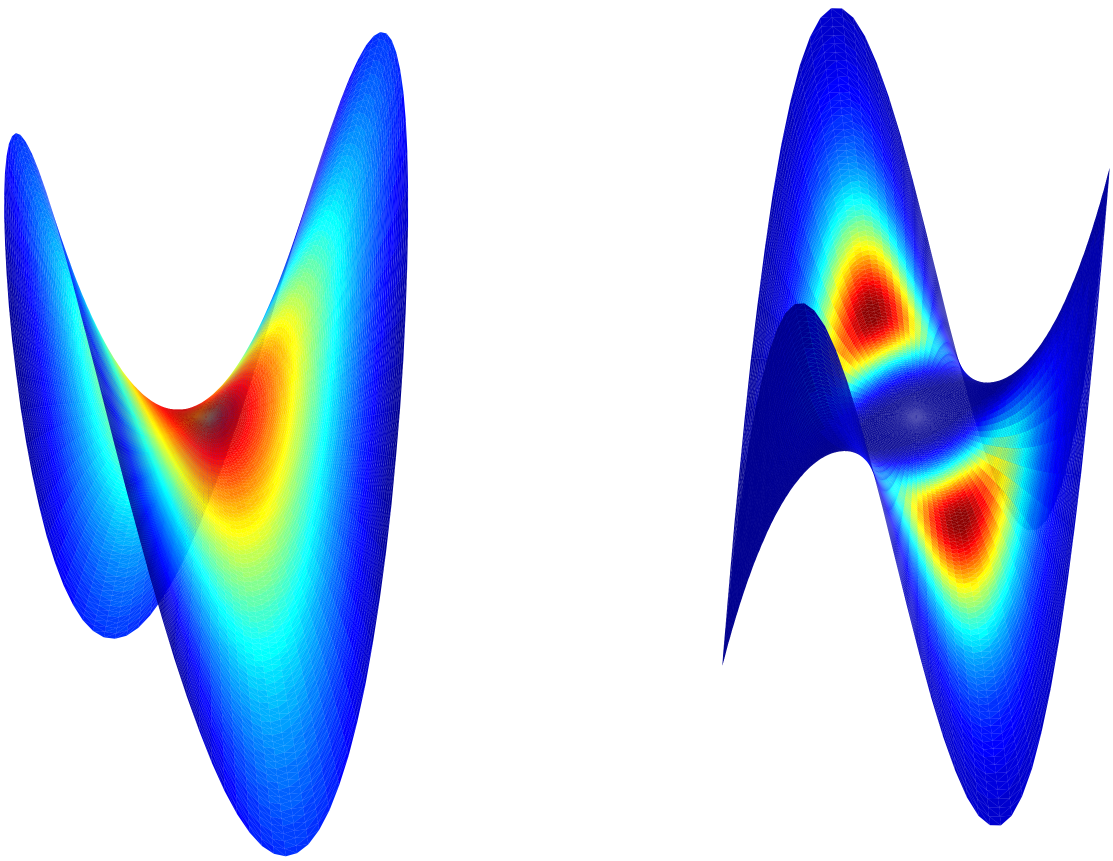
http://nsa.fjfi.cvut.cz/david

22 March 2025

ATHENS (one-week intensive) course delivered by the author at the *Faculty of Nuclear Sciences and Physical Engineering* of the *Czech Technical University in Prague* in 2025.

The most recent version of these lecture notes can be downloaded from the following link:

http://nsa.fjfi.cvut.cz/david/other/geometry.pdf



Contents

0 Introduction

The goal of the course is to acquaint students with classical notions of curves and surfaces in Euclidean spaces as well as with some modern aspects of Riemannian manifolds. The ultimate objective is to understand the geometric meaning of curvature and its intimate relationship to topology. The student will learn the basic apparatus of differential geometry suitable for further study of modern parts of mathematics and mathematical physics.

The course is mostly based on Klingenberg's textbook [4] as well as do Carmo's [2]. For a deeper understanding of differential geometry, I recommend Spivak's volumes [12, 10, 8, 9, 11].

Here we point out some special convention that we use.

- Natural numbers contain zero, *i.e.*, $\mathbb{N} := \{0, 1, 2, \dots\}$. $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$.
- Given $x, y \in \mathbb{R}^d$, $|x|$ and $x \cdot y$ denote the Euclidean norm and inner product, respectively.
- $I \subset \mathbb{R}$ is an interval (in general, bounded or unbounded; open, closed or semi-open).
- A° denotes the interior of a set A.

1 Curves

1.1 Inertia or straight lines

Let $\mathbb{R}^d$ be the Euclidean space of dimension $d \geq 1$. The simplest curve in $\mathbb{R}^d$ is the *straight line*. Parametrically, it can be realised as the image of the mapping

$$\gamma_1 : \mathbb{R} \to \mathbb{R}^d : \{t \mapsto (t, 0, \dots, 0)\} . \tag{1.1}$$

See Figure 1.1 for a planer realisation.

Figure 1.1: A straight line in the plane realised as the image of $\gamma_1, \gamma_2, \gamma_3$ with $d = 2$.

It is customary to think of the curve as the trajectory of a traveller in $\mathbb{R}^d$. Then $\gamma_1(t)$, $\gamma_1'(t)$ and $\gamma_1''(t)$ are respectively his position, velocity and acceleration at time $t \in \mathbb{R}$. Another mapping

$$\gamma_2 : \mathbb{R} \to \mathbb{R}^d : \{t \mapsto (2t, 0, \dots, 0)\}$$

gives the same image, but the speed of the second traveller is twice the velocity of the first traveller. Therefore it is useful to keep the information of parameterisation. Indeed, yet another mapping

$$\gamma_3 : \mathbb{R} \to \mathbb{R}^d : \{t \mapsto (t^3, 0, \dots, 0)\}$$

still copies the trajectory of the first and second travellers, but his velocity is highly non-linear. What is more, while the accelerations of the first and second travellers are identically zero ($|\gamma_1''(t)| = 0 = |\gamma_2''(t)|$ for all times t), the acceleration of the third traveller actually diverges ($|\gamma_3''(t)| = 6|t| \to \infty$ as $t \to \pm\infty$).

If $d \geq 2$, it is also possible to consider

$$\gamma_4 : \mathbb{R} \to \mathbb{R}^d : \{t \mapsto \tfrac{1}{\sqrt{2}}(t, t, 0, \dots, 0)\} , \tag{1.2}$$

which realises a straight-line traveller of the same (unit) speed as the first one, but with the straight-line trajectory rotated by 45 degrees. See Figure 1.2 for a planer realisation. The acceleration of the fourth traveller is also zero ($|\gamma_4''(t)| = 0$ for all times t).

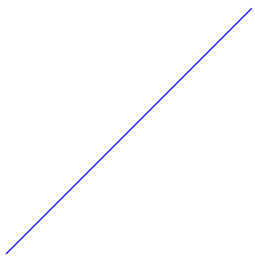


Figure 1.2: A straight line in the plane realised as the image of γ_4 with $d = 2$.

In general, inspired by Newton's first law of classical mechanics (the principle of inertia), we see that the acceleration is the right measure of *curvature* for constant-speed travellers on a straight line. More specifically, with the presence of no forces, the traveller remains at rest or in motion at a constant speed on a straight line.

For later purposes, we formalise a *straight line* in $\mathbb{R}^d$ to be the mapping $\gamma : I \to \mathbb{R}^d$, where $I \subset \mathbb{R}$ is an interval and $\gamma(t) := at + b$ with $a, b \in \mathbb{R}^d$ being constant vectors.

1.2 Forces or general curves

A *curve* is a deformation of the straight line. Continuing with the mechanical interpretation, the motion of the traveller is deformed by the force impressed.

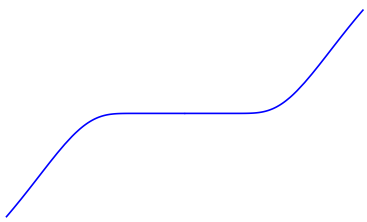
In general, given an open interval $I \subset \mathbb{R}$, any continuous mapping $\gamma : I \to \mathbb{R}^d$ is called a curve. However, the continuity is a very weak requirement in this geometric context. Indeed, fractal curves can be of a higher dimension than one (*e.g.*, the Brownian motion in $\mathbb{R}^d$ with $d \geq 2$ has a fractal dimension of 2). What is more, there exist space-filling curves (*i.e.*, their image reaches every point in $\mathbb{R}^d$). In order to avoid these pathological situations, the notion of differentiability comes into the play.

Definition 1.1. A *curve* is a smooth map $\gamma : I \to \mathbb{R}^d$.

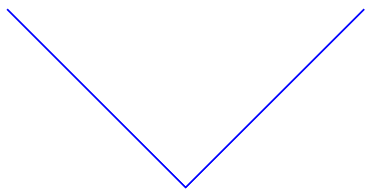
By a *smooth* (or *differentiable*) map we always mean *infinitely smooth*, *i.e.* $\gamma \in C^\infty(I)$.

If the interval I is not open, we still say that $\gamma : I \to \mathbb{R}^d$ is a *curve*, provided that the mapping γ is a restriction of a curve $\tilde{\gamma}$. More specifically, there exists an open interval $\tilde{I} \supset I$ and a curve $\tilde{\gamma} : \tilde{I} \to \mathbb{R}^d$ such that $\tilde{\gamma}|_I = \gamma$.

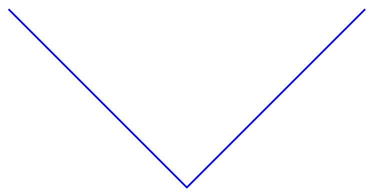
Of course, the parameterisation $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ of straight lines from the previous subsection are curves in the sense of Definition 1.1. Curved scenarios $\gamma_5, \gamma_6, \gamma_7$ are given in Figure 1.3. In particular, the smoothness of γ_7 is gained by a suitable reparameterisation of the non-smooth mapping γ_6 (the images of γ_6 and γ_7 coincide). In other words, slowing down the speed of the traveller to zero velocity at the corner of the image will make him move smoothly despite the sharp turning. The function $g : \mathbb{R} \to \mathbb{R}$ by itself is a curve in $\mathbb{R}$.



$\gamma_5(t) := (t, g(t))$
γ_5 is smooth
γ_5 is a curve
γ_5 is regular



$\gamma_6(t) := (t, |t|)$
γ_6 is not smooth
γ_6 is not a curve



$\gamma_7(t) := (g(t), |g(t)|)$
γ_7 is smooth
γ_7 is a curve
γ_7 is not regular

Figure 1.3: Mappings $\gamma_k : \mathbb{R} \to \mathbb{R}^2 : \{t \mapsto \gamma_k(t)\}$ with $g(t) := \begin{cases} -e^{-t^{-2}} & \Leftrightarrow t < 0\,, \\ 0 & \Leftrightarrow t = 0\,, \\ e^{-t^{-2}} & \Leftrightarrow t > 0\,. \end{cases}$

The image of the (smooth) curve γ_7 from Figure 1.3 is apparently not a "smooth line" according to our daily experience. To avoid these situations, we say that the curve γ of Definition 1.1 is *regular* if

$$\forall t \in I, \qquad \gamma'(t) \neq 0.$$

That is, γ is regular if, and only if, the traveller is never at rest. The curve γ_5 from Figure 1.3 is regular, while γ_7 is not (the notion of regularity for the non-smooth mapping γ_6 is not defined). The curve $g : \mathbb{R} \to \mathbb{R}$ is not regular.

The images of regular curves are "locally smooth lines" in the sense that locally they can be represented by unit-speed straight lines (1.1). In other words, the image of every regular curve is locally diffeomorphic to the image of a straight line. (A diffeomorphism is a bijective map, which is smooth and whose inverse is also smooth.)

Proposition 1.2 (Local linearisation of regular curves)**.** *Let $\gamma : I \to \mathbb{R}^d$ be a regular curve. For every $t_0 \in I^\circ$, there exists an open subinterval $I' \subset I$ containing t_0 and a diffeomorphism ϕ of a neighbourhood W of $\gamma(t_0) \in \mathbb{R}^d$ onto a neighbourhood $\phi(W)$ of $0 \in \mathbb{R}^d$ such that $\phi \circ \gamma : I' \to \mathbb{R}^2$ is an injective linear map. In fact, $(\phi \circ \gamma)(t) = (t, 0, \ldots, 0)$ for every $t \in I'$.*

Proof. Define

$$\tilde{\phi} : I \times \mathbb{R}^{d-1} \to \mathbb{R}^d \cong \mathbb{R} \times \mathbb{R}^{d-1} : \{(t, y) \mapsto \gamma(t) + (0, y)\} .$$

The Jacobi matrix of $\tilde{\phi}$ at $(t_0, 0)$ has the block-diagonal structure

$$\tilde{\phi}'(t_0, 0) = \begin{pmatrix} \gamma'(t_0) & 0 \\ 0 & 1 \end{pmatrix} ,$$

where 1 denotes the identity matrix of order $d - 1$. Since γ is regular, the Jacobi matrix is invertible. By the inverse function theorem, $\tilde{\phi}$ is a local diffeomorphism at $(t_0, 0)$. More specifically, there is an open neighbourhood U of $I \times \mathbb{R}^{d-1}$ containing $(t_0, 0)$ such that $\tilde{\phi} : U \to \tilde{\phi}(U)$ is a diffeomorphism. Clearly, we can choose U of the form $I' \times W$, where I' is an open subinterval of I containing t_0 and W is an open neighbourhood of $\mathbb{R}^{d-1}$ containing 0. Let us set $\phi := \tilde{\phi}^{-1}$. Since $\phi \circ \tilde{\phi}$ is an identity, in particular $(\phi \circ \tilde{\phi})(t, 0) = (t, 0)$ for every $t \in I'$. Consequently, $(\phi \circ \gamma)(t) = (t, 0)$. □

The proposition particularly claims that for any regular curve $\gamma : I \to \mathbb{R}^d$ and $t_0 \in I^\circ$ there exists an open subinterval $I' \subset I$ containing t_0 such that $\gamma|_{I'}$ is injective. Globally, however, regular curves can overlap, see Figure 1.4.

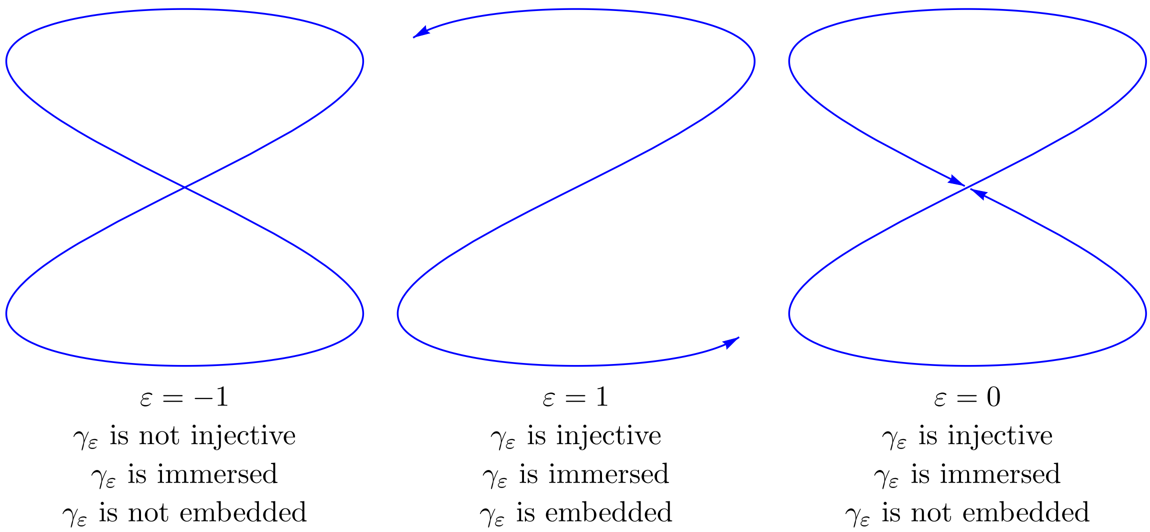


Figure 1.4: Curve $\gamma_\varepsilon : I_\varepsilon \to \mathbb{R}^2 : \{t \mapsto (\sin(2t), \cos(t))\}$ with $I_\varepsilon := (-\frac{\pi}{2} + \varepsilon, 3\frac{\pi}{2} - \varepsilon)$.

If the interval I is open, regular curves $\gamma : I \to \mathbb{R}^d$ are alternatively called *immersed curves*. A stronger requirement is an *embedded curve*, which is an injective immersed curve

$\gamma : I \to \mathbb{R}^d$ such that $\gamma : I \to \gamma(I)$ is a homeomorphism. (A homeomorphism is a bijective map, which is continuous and whose inverse is also continuous.) The curves γ_{-1} and γ_0 are not embedded, because their images around $0 \in \mathbb{R}^2$ "do not locally look like a smoothly deformed interval". Other examples of immersed but not embedded curves are depicted in Figure 1.5.

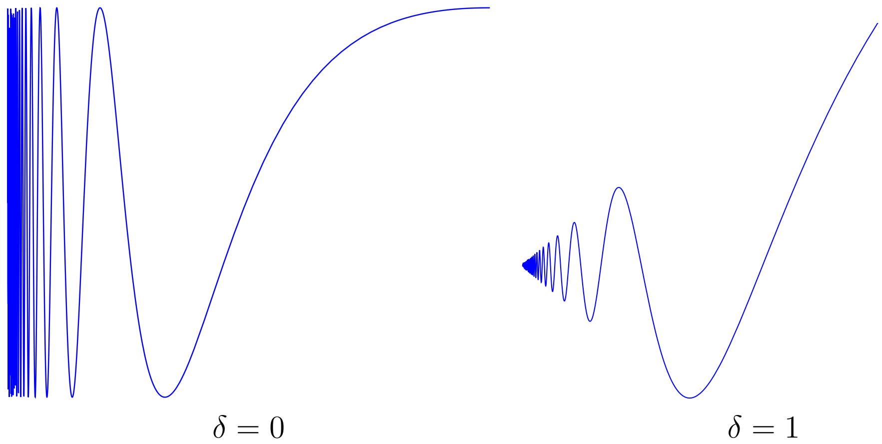


Figure 1.5: Immersed non-embedded curves $\gamma_\delta : (0,1) \to \mathbb{R}^2 : \{t \mapsto (t, t^\delta \sin(t^{-1}))\}$.

1.3 Acceleration or curvature

Inspired by the discussion of straight lines in Section 1.1, we intend to define the curvature of a general curve as the acceleration of a constant-speed traveller whose trajectory is the image of the curve. This requires that the traveller is never at rest (unless we allow for stationary curves whose image is just a point). In other words, we restrict to regular curves.

For regular curves, it is enough to restrict to unit-speed travellers, without loss of generality. (*Reparameterisation* of a curve $\gamma : I \to \mathbb{R}^d$ is a curve $\tilde\gamma := \gamma \circ \phi : \tilde I \to \mathbb{R}^d$, where $\tilde I \subset \mathbb{R}$ is another interval and $\phi : \tilde I \to I$ is a diffeomorphism. Obviously, γ and $\tilde\gamma$ have the same image.)

Proposition 1.3 (Regular travellers can be unit-speed)**.** *For every regular curve, there exists a unit-speed reparameterisation.*

Proof. Let $\gamma : I \to \mathbb{R}^d$ be the regular curve. Given any $t_0 \in I$, define the *arc-length* function $s : I \to [0, \infty)$ by

$$s(t) := \int_{t_0}^{t} |\gamma'(\tau)| \, \mathrm{d}\tau \,. \tag{1.3}$$

Since γ is regular, the derivative s' is positive. Consequently, by the inverse function theorem, the inverse $\phi := s^{-1}$ exists. Then $\tilde\gamma := \gamma \circ \phi$ is the desired unit-speed curve. Indeed, $|\tilde\gamma'(s(t))| = |\gamma'(t)|/s'(t) = 1$. □

A unit-speed curve is also said to be *parameterised by arc-length*. For unit-speed curves γ, it is customary to denote its parameter by s instead of t, *i.e.*, $\gamma : I \to \mathbb{R}^d : \{s \mapsto \gamma(s)\}$. Then the length of the interval $|I|$ coincides with the length of the curve.

Definition 1.4. The *curvature* of a unit-speed curve γ is the function $\kappa := |\gamma''|$.

The following observation shows that the curvature is the right measure of deviation of the curve from the straight line.

Proposition 1.5. *Let $\gamma : I \to \mathbb{R}^d$ be unit-speed. Then*

$$\kappa = 0 \qquad \Longleftrightarrow \qquad \gamma \text{ is a straight line.}$$

Proof. If γ is a straight line (*i.e.* $\gamma(s) = as + b$, where $a, b \in \mathbb{R}^d$ with $|a| = 1$), then $\kappa = 0$ identically. Conversely, if $\kappa = 0$, then by integration $\gamma(s) = as + b$, where $a, b \in \mathbb{R}^d$ with $|a| = 1$, so γ is a straight line. □

1.4 Reference or moving frames

To describe the traveller on the rotated straight line (1.2) (see also Figure 1.2), it is convenient to rotate the ambient system of Cartesian coordinates first, to end up with the simple realisation (1.1). Moreover, it might be useful to work in an inertial frame of reference, following the constant speed of the traveller. For the traveller on a (general) curve, it is similarly useful to work in a (generally non-inertial) frame of reference, moving along the curve and suitably being adapted to the curved trajectory.

To be more specific, we introduce the following general terminology. Given any regular curve $\gamma : I \to \mathbb{R}^d$, any non-zero vector which is parallel (respectively, perpendicular) to the velocity vector γ' is called *tangent* or *parallel* (respectively, *normal* or *perpendicular*) to γ.

Definition 1.6. Let $\gamma : I \to \mathbb{R}^d$ be a unit-speed curve.

- A *moving frame* along γ is a collection of d smooth mappings $E_1, \dots, E_d : I \to \mathbb{R}^d$ which is orthonormal (*i.e.*, $E_i(s) \cdot E_j(s) = \delta_{ij}$ for every $s \in I$ and $i, j = 1, \dots, d$).
- The moving frame is called *adapted* to the curve if its members are either tangent or perpendicular to the curve.
- The *Frenet frame* is a moving frame satisfying $\gamma^{(i)} \in \operatorname{span}(E_1, \dots, E_i)$ for $i = 1, \dots, d$.
- The *relatively parallel adapted frame* is an adapted moving frame whose normal members have tangential derivative.

Since γ is assumed to be unit-speed, the *tangent* vector field $T := \gamma'$ has norm one. Consequently, any Frenet frame $(E_1, \dots, E_d)$ satisfies $E_1 = \pm T$, so it is adapted.

Let $(E_1, \dots, E_d)$ be any moving frame of γ. For every $i \in \{1, \dots, d\}$, let us write

$$E_i' = A_i^{\ j} E_j \,,$$

where $A := (A_i^{\ j})_{i,j=1,\dots,d} \in \mathbb{R}^{d\times d}$ is a matrix and the *Einstein summation convention* is adopted on the right-hand side (the summation is understood over the same index that appears twice, once as an upper index and once as a lower index, in the present case over $j = 1, \dots, d$). Since (using that $E_i \cdot E_j = \delta_{ij}$)

$$A_i^{\ j} = E_i' \cdot E_j = (E_i \cdot E_j)' - E_i \cdot E_j' = -A_j^{\ i},$$

it follows that A is skew-symmetric. Consequently,

$$\begin{pmatrix} E_1 \\ \vdots \\ E_d \end{pmatrix}' = \begin{pmatrix} 0 & & * \\ & \ddots & \\ -* & & 0 \end{pmatrix} \begin{pmatrix} E_1 \\ \vdots \\ E_d \end{pmatrix}. \tag{1.4}$$

Here the asterisk denotes entries of the matrix we do not care about. The skew-symmetric matrix A admits an additional structure for the Frenet and relatively parallel adapted frames.

1.4.1 The relatively parallel adapted frame

Let $(E_1, \dots, E_d)$ be a relatively parallel adapted frame of the unit-speed curve γ. By changing the order of the members of the frame, we may assume that $E_1 := \pm T$. Then $E_i' \in \operatorname{span}(T)$ for every $i = 2, \dots, d$, because the normal members have tangential derivative. By the general skew-symmetry (1.4), it follows that there exist smooth functions $k_1, \dots, k_{d-1} : I \to \mathbb{R}$ such that

$$\begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_d \end{pmatrix}' = \begin{pmatrix} 0 & k_1 & \dots & k_{d-1} \\ -k_1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -k_{d-1} & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_d \end{pmatrix}. \tag{1.5}$$

Recalling Definition 1.4, we have

$$k_1^2 + \dots + k_{d-1}^2 = \kappa^2,$$

where κ is the curvature of γ. The feature of the relatively parallel adapted frame is that (due to the fact that the derivative of the normal members are tangential) the normal vectors rotate along the curve only whatever amount is necessary to remain normal.

It turns out that the relatively parallel adapted frame always exists. (Here $T_p\mathbb{R}^d$ denotes the the *tangent space* of $\mathbb{R}^d$ at point $p \in \mathbb{R}^d$, *i.e.*, the d-dimensional vector space consisting of pairs $(p, V) \in \{p\} \times \mathbb{R}^d$ and equipped with the structure $(p, V_1) + (p, V_2) := (p, V_1 + V_2)$ and $\alpha(p, V) := (p, \alpha V)$, where $\alpha \in \mathbb{R}$ and $V_1, V_2 \in \mathbb{R}^d$.)

Proposition 1.7 (Existence and uniqueness of the relatively parallel adapted frame)**.** *Let $\gamma : I \to \mathbb{R}^d$ be a unit-speed curve. Let $(E_1^0, \dots, E_d^0)$ be an orthonormal basis of $T_{\gamma(s_0)}\mathbb{R}^d$ for some $s_0 \in I$ such that $E_1^0 = \pm T(s_0)$. Then there exists a unique relatively parallel adapted frame $(E_1, \dots, E_d)$ of γ such that*

$$E_i(s_0) = E_i^0 \tag{1.6}$$

for every $i \in \{1, \dots, d\}$.

Proof. The proof is based on [6, App. A].

$\boxed{\text{Uniqueness.}}$ Let $(\tilde{E}_1, \dots, \tilde{E}_d)$ be another relatively parallel adapted frame of γ satisfying $\tilde{E}_i(s_0) = E_i^0$ for every $i \in \{1, \dots, d\}$. Then necessarily $\tilde{E}_1 = E_1$, because of the smoothness. At the same time, $M_i := \tilde{E}_i - E_i$ with $i \in \{2, \dots, d\}$ are also normal vector fields with tangential derivative. Consequently, $|M_i|^{2\prime} = 2M_i \cdot M_i' = 0$, which shows that the length of each M_i is constant. But $M_i(s_0) = 0$ due to (1.6), which implies $\tilde{E}_i = E_i$ identically on I.

$\boxed{\text{Local existence.}}$ To show the existence locally, one takes an arbitrary adapted moving frame $(F_1, \dots, F_d)$ of γ such that $F_1 = \pm T$. Such a frame always exists at least locally, say on the interval $I_0 := (s_0 - \varepsilon, s_0 + \varepsilon)$ with some positive ε. On this interval, it satisfies (1.4), or more specifically

$$\begin{pmatrix} F_1 \\ \vdots \\ F_d \end{pmatrix}' = \begin{pmatrix} 1 & b^\top \\ -b & B \end{pmatrix} \begin{pmatrix} F_1 \\ \vdots \\ F_d \end{pmatrix},$$

where $B \in \mathbb{R}^{(d-1)\times(d-1)}$ is skew-symmetric and $b \in \mathbb{R}^d$. Given a rotation (special orthogonal) matrix-valued function $R : I_0 \to \mathsf{SO}(d-1)\mathbb{R}^{(d-1)\times(d-1)}$ to be specified later, we rotate the normal components, that is, we consider the transformed moving frame

$$\begin{pmatrix} E_1 \\ \vdots \\ E_d \end{pmatrix} := \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \begin{pmatrix} F_1 \\ \vdots \\ F_d \end{pmatrix}. \tag{1.7}$$

Then

$$\begin{pmatrix} E_1 \\ \vdots \\ E_d \end{pmatrix}' = \begin{pmatrix} 1 & (Rb)^\top \\ -Rb & (R' + RB)R^\top \end{pmatrix} \begin{pmatrix} E_1 \\ \vdots \\ E_d \end{pmatrix}.$$

Comparing this equation with (1.5), we see that $(E_1, \dots, E_d)$ will be a relatively parallel adapted frame of $\gamma|_{I_0}$ satisfying the initial condition (1.6), provided that we choose R in such a way that

$$\begin{cases} R' + RB = 0 & \text{in} \quad I_0\,, \\ R = R_0 & \text{at} \quad s_0\,, \end{cases} \tag{1.8}$$

where $R_0 \in \mathbb{R}^{(d-1)\times(d-1)}$ is the (unique) orthogonal matrix satisfying

$$\begin{pmatrix} E_2^0 \\ \vdots \\ E_d^0 \end{pmatrix} = R_0 \begin{pmatrix} F_2(s_0) \\ \vdots \\ F_d(s_0) \end{pmatrix}.$$

It is well known that the system (1.8) admits a unique smooth solution R.

$\boxed{\text{Global existence.}}$ Let J be any open precompact subinterval of I containing the point s_0. The interval can be covered by a finite number of open intervals of equal length, for each of which there exists a family of relatively parallel adapted frames by the local construction above. To get the global relatively parallel adapted frame on J satisfying the desired initial condition at s_0, we can patch together the local ones by employing the local frame already constructed on I_0 and the freedom of choosing the initial condition in the problem analogous to (1.6) for the covering subintervals. Smoothness at the point where they link together is a consequence of the uniqueness part. Since there is the desired relatively parallel adapted frame on *any* open precompact subinterval J of I, the result follows. $\square$

In summary, every curve possesses a relatively parallel adapted frame. In fact, there are infinitely many such frames whenever $d \geq 3$ (because there is an infinite number of vectors perpendicular to γ at the initial point s_0).

1.4.2 The Frenet frame

Not every regular curve possesses a Frenet frame, see Figure 1.6. Indeed, since γ'' lies in the x^1x^3-plane for negative times and in the x^1x^2-plane for positive times, there exists no smooth version of E_2. Notice that $\kappa(0) = 0$ for this curve.

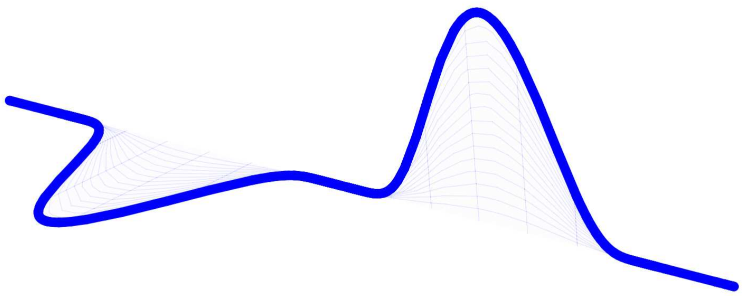


Figure 1.6: A curve in $\mathbb{R}^3$ without the Frenet frame, $\gamma(t) := \begin{cases} (t, 0, e^{-t^{-2}}) & \Leftrightarrow t < 0\,, \\ (0,0,0) & \Leftrightarrow t = 0\,, \\ (t, e^{-t^{-2}}, 0) & \Leftrightarrow t > 0\,. \end{cases}$

To construct a Frenet frame, let us restrict to unit-speed curves $\gamma : I \to \mathbb{R}^d$ satisfying the non-degeneracy condition

$$\forall s \in I\,, \qquad \gamma'(s), \gamma''(s), \dots, \gamma^{(d-1)}(s) \quad \text{are linearly independent.} \tag{1.9}$$

A necessary condition is that the curvature κ vanishes nowhere in I.

Proposition 1.8 (Existence and uniqueness of the distinguished Frenet frame)**.** *Let $\gamma : I \to \mathbb{R}^d$ be a unit-speed curve and assume* (1.9)*. Then there exists a unique Frenet frame $(E_1, \dots, E_d)$ of γ satisfying, for all $s \in I$ and $k \in \{1, \dots, d-1\}$,*

(i) $\gamma'(s), \gamma''(s), \dots, \gamma^{(k)}(s)$ *and* $E_1(s), \dots, E_k(s)$ *have the same orientation and*

(ii) $E_1(s), \dots, E_d(s)$ *has the positive orientation.*

Proof. The claim follows by applying the Gram–Schmidt orthogonalisation process to the linearly independent vector fields $\gamma', \gamma'', \dots, \gamma^{(d-1)}$. In detail, we set $E_1 := \gamma'$ and

$$E_k := \frac{\tilde{E}_k}{|\tilde{E}_k|} \qquad \text{with} \qquad \tilde{E}_k := \gamma^{(k)} - E_1 \cdot \gamma^{(k)} E_1 - \cdots - E_{k-1} \cdot \gamma^{(k)} E_{k-1} \tag{1.10}$$

for every $k \in \{2, \dots, d-1\}$. Finally, we define $E_d := \pm(E_1 \times \cdots \times E_{d-1})$ with the sign $\pm$ chosen in such a way that $E_1, \dots, E_d$ is positively oriented (*i.e.* it has the same orientation as the canonical basis of $\mathbb{R}^d$). Recall that the cross-product $E := E_1 \times \cdots \times E_{d-1}$ is the unique vector satisfying

$$\forall F \in \mathbb{R}^d\,, \qquad E \cdot F = \det(E_1, \dots, E_{d-1}, F)\,.$$

Obviously, the moving frame constructed in this way is a Frenet frame. Notice that (i) is also satisfied because

$$\begin{pmatrix} E_1 \\ \vdots \\ E_k \end{pmatrix} = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ * & & c_k \end{pmatrix} \begin{pmatrix} \gamma' \\ \vdots \\ \gamma^{(k)} \end{pmatrix}, \tag{1.11}$$

where $c_1 := 1$ and $c_k := 1/|\tilde{E}_k|$ for $k \geq 2$. Indeed, since the transfer matrix is lower triangular and the diagonal elements are positive, its determinant is positive. □

The Frenet frame of Proposition 1.8 is called the *distinguished Frenet frame*.

Let us see how the equation (1.4) looks for the distinguished Frenet frame. For $k = d-1$, equation (1.11) reads

$$\begin{pmatrix} E_1 \\ \vdots \\ E_{d-1} \end{pmatrix} = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ * & & c_{d-1} \end{pmatrix} \begin{pmatrix} \gamma' \\ \vdots \\ \gamma^{(d-1)} \end{pmatrix}. \tag{1.12}$$

Differentiating and using an inversion of the relationship (1.12), we obtain

$$\begin{aligned}
\begin{pmatrix} E_1 \\ \vdots \\ E_{d-1} \end{pmatrix}' &= \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ * & & c_{d-1} \end{pmatrix} \begin{pmatrix} \gamma'' \\ \vdots \\ \gamma^{(d)} \end{pmatrix} + \begin{pmatrix} c_1' & & 0 \\ & \ddots & \\ *' & & c_{d-1}' \end{pmatrix} \begin{pmatrix} \gamma' \\ \vdots \\ \gamma^{(d-1)} \end{pmatrix} \\
&= \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ * & & c_{d-1} \end{pmatrix} \begin{pmatrix} * & c_2^{-1} & & & & 0 \\ & * & c_3^{-1} & & & \\ & & \ddots & \ddots & & \\ & & & * & c_{d-1}^{-1} & 0 \\ * & & & & * & * \end{pmatrix} \begin{pmatrix} E_1 \\ \vdots \\ E_d \end{pmatrix} \\
&\quad + \begin{pmatrix} c_1' & & 0 \\ & \ddots & \\ *' & & c_{d-1}' \end{pmatrix} \begin{pmatrix} * & & 0 \\ & \ddots & \\ * & & * \end{pmatrix} \begin{pmatrix} E_1 \\ \vdots \\ E_{d-1} \end{pmatrix} \\
&= \begin{pmatrix} * & c_1 c_2^{-1} & & & & 0 \\ & * & c_2 c_3^{-1} & & & \\ & & \ddots & \ddots & & \\ & & & * & c_{d-2} c_{d-1}^{-1} & 0 \\ * & & & & * & c_{d-1} * \end{pmatrix} \begin{pmatrix} E_1 \\ \vdots \\ E_d \end{pmatrix} \\
&\quad + \begin{pmatrix} * & & 0 & 0 \\ & \ddots & & \vdots \\ * & & * & 0 \end{pmatrix} \begin{pmatrix} E_1 \\ \vdots \\ E_d \end{pmatrix}.
\end{aligned}$$

Denoting $\kappa_1 := c_1 c_2^{-1}, \dots, \kappa_{d-2} := c_{d-2} c_{d-1}^{-1}$ and $\kappa_d := c_{d-1} *$, the necessary skew-symmetry of A eventually implies

$$\begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_d \end{pmatrix}' = \begin{pmatrix} 0 & \kappa_1 & & 0 \\ -\kappa_1 & \ddots & \ddots & \\ & \ddots & \ddots & \kappa_{d-1} \\ 0 & & -\kappa_{d-1} & 0 \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_d \end{pmatrix}.$$

These are the celebrated *Frenet–Serret equations*. The functions $\kappa_1 \dots \kappa_{d-2}$ are called the *curvature functions* of γ. Note that $\kappa_1, \dots, \kappa_{d-2}$ are positive, while κ_1 may change sign. One has $\kappa_1 = \kappa > 0$ if $d \geq 3$, while κ_1 may change sign if $d = 2$.

1.4.3 Low dimensions

Let $\gamma : I \to \mathbb{R}^d$ be unit-speed. Recall the notation $T := \gamma'$ for the unit tangent vector field.

$\boxed{d = 2}$ In the planar case, $T = ({\gamma^1}', {\gamma^2}')$ and there is a unique choice of the unit normal vector field $N := (-{\gamma^2}', {\gamma^1}')$, which makes the moving frame (T, N) positively oriented. It is a relatively parallel adapted frame of γ satisfying

$$\begin{pmatrix} T \\ N \end{pmatrix}' = \begin{pmatrix} 0 & k \\ -k & 0 \end{pmatrix} \begin{pmatrix} T \\ N \end{pmatrix}, \tag{1.13}$$

where the curvature function $k := k_1 = \kappa_1$ is actually defined by the equations, namely $k := T' \cdot N = \det(\gamma', \gamma'')$. What is more, (T, N) is also the distinguished Frenet frame of γ (the hypothesis (1.9) is automatically satisfied if $d = 2$). Of course, $|k| = \kappa$. In this planar context, k is occasionally called the *signed curvature* of γ. In practice, k is positive if the curve is "turning toward" N, see Figure 1.7.

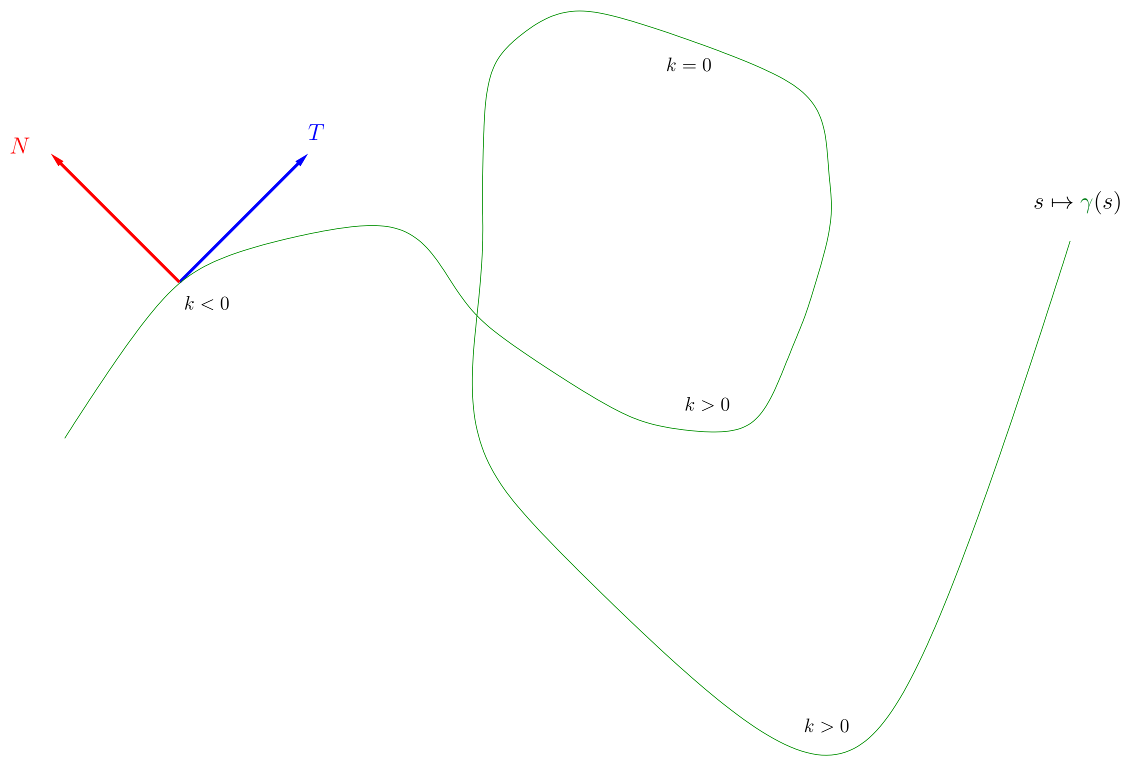


Figure 1.7: The signed curvature of a plane curve.

$\boxed{d = 3}$ By Proposition 1.7, the relatively parallel adapted frame (T, M_1, M_2) exists for any space unit-speed curve $\gamma : I \to \mathbb{R}^3$. It satisfies

$$\begin{pmatrix} T \\ N_1 \\ N_2 \end{pmatrix}' = \begin{pmatrix} 0 & k_1 & k_2 \\ -k_1 & 0 & 0 \\ -k_2 & 0 & 0 \end{pmatrix} \begin{pmatrix} T \\ N_1 \\ N_2 \end{pmatrix},$$

where $k_1, k_2 : I \to \mathbb{R}^3$ are smooth functions.

As for the Frenet frame, the planar simplicity does not extend to higher dimensions. In particular, a condition is inevitable to guarantee the existence of a Frenet frame (recall Figure 1.6). In the spatial realisation, the hypothesis (1.9) reduces to the requirement that the acceleration γ'' vanishes nowhere in I (which is equivalent to the fact that the curvature κ is positive). Under this condition, the distinguished Frenet frame is given by (T, N, B), where $N := \gamma''/\kappa$ is the *principal normal* and $B := T \times N$ is the *binormal*. The Frenet–Serret equations read

$$\begin{pmatrix} T \\ N \\ B \end{pmatrix}' = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix} \begin{pmatrix} T \\ N \\ B \end{pmatrix}, \tag{1.14}$$

where $\tau := \det(\gamma', \gamma'', \gamma''')/\kappa^2$ is the *torsion* of γ.

Proposition 1.9. *Let $\gamma : I \to \mathbb{R}^3$ be a unit-speed curve with $\kappa > 0$. Then*

$$\tau = 0 \qquad \Longleftrightarrow \qquad \gamma \textit{ is a plane curve.}$$

Proof. If γ is a plane curve, *i.e.*, the image $\gamma(I)$ is contained in a plane, then the plane is actually the *osculating plane* determined by the vectors T and N. Therefore B is constant, which implies $\tau = 0$ due to the equation $B' = -\tau N$. Conversely, if $\tau = 0$, then B is constant. Therefore, $(\gamma \cdot B)' = \gamma' \cdot B = 0$, which implies that $\gamma(I)$ is contained in a plane normal to B. □

In summary, if $\kappa > 0$, the distinguished Frenet frame (T, N, B) is the relatively parallel adapted frame (T, N_1, N_2) if, and only, if $\tau = 0$ (*i.e.*, γ is a plane curve). In general (but still under the hypothesis $\kappa > 0$), the two frames are related by a suitable rotation of the normal vectors:

$$\begin{pmatrix} N_1 \\ N_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} N \\ B \end{pmatrix}, \tag{1.15}$$

where $\theta : I \to \mathbb{R}$ is a smooth function. This represents a parameterisation of the orthogonal matrix R from (1.7). Then the differential equation of (1.8) is equivalent $\theta' = \tau$. Thus, the normal vectors of any relatively parallel adapted frame of γ are rotated with respect to the distinguished Frenet frame with the angle being a primitive of the torsion.

1.4.4 Closed curves

Now, the question arises why to ever consider the Frenet frame, which requires the extra hypothesis (1.9) whenever $d \geq 3$ (in particular, $\kappa > 0$ if $d = 3$), while the relatively parallel adapted frame exists for any regular curve. Furthermore, the relatively parallel adapted frame exists for any curve $\gamma : I \to \mathbb{R}^d$ under the modest regularity $\gamma \in C^2(I)$ (in fact, even $\gamma \in C^{1,1}(I)$ is considered in [6]), while the Frenet frame requires $\gamma \in C^d(I)$.

An immediate answer is the simplicity of the construction of the Frenet frame (just by computing derivatives of a given curve, no need to solve any differential equation). More importantly (but related to the previous observation), the construction of the Frenet frame is *local*. This becomes fundamental when considering *closed curves*, *i.e.* curves $\gamma : [a, b] \to \mathbb{R}^d$ with $a < b$ such that

$$\forall k \in \mathbb{N}, \qquad \gamma^{(k)}(a) = \gamma^{(k)}(b).$$

The most famous example of a closed curve is a circle, see Figure 1.8 and Example 1.10.

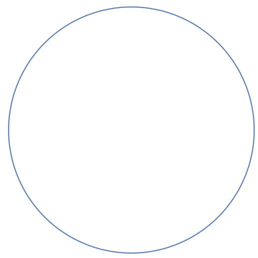


Figure 1.8: A circle as a canonical example of a closed curve.

For closed curves in $\mathbb{R}^3$ with $\kappa > 0$, the distinguished Frenet frame is smooth at $\gamma(a) = \gamma(b)$. In particular, $N(a) = N(b)$ and $B(a) = B(b)$. This is due to the local nature of its construction.

On the other hand, the relatively parallel adapted frame (T, N_1, N_2) is a non-local object (in the sense that it is constructed through solving a differential equation on γ), so it is generally not smooth at $\gamma(a) = \gamma(b)$. To be more specific, consider a space curve γ with $\kappa > 0$. Let $N_1(a) = N(a)$ and $N_2(a) = B(a)$. Then (1.15) yields

$$\begin{aligned} N_1(b) &= \cos(\theta(b))N(b) - \sin(\theta(b))B(b) = \cos(\theta(b))N_1(a) - \sin(\theta(b))N_2(a)\,, \\ N_2(b) &= \sin(\theta(b))N(b) + \cos(\theta(b))B(b) = \sin(\theta(b))N_1(a) + \cos(\theta(b))N_2(a)\,, \end{aligned}$$

where $\theta(s) = \int_a^s \tau$. However, it is rare (and indeed generally not true, see Figure 1.9) that $\int_a^b \tau \in 2\pi\mathbb{Z}$. Nevertheless, it is true for plane curves, in particular for the circle of Figure 1.8.

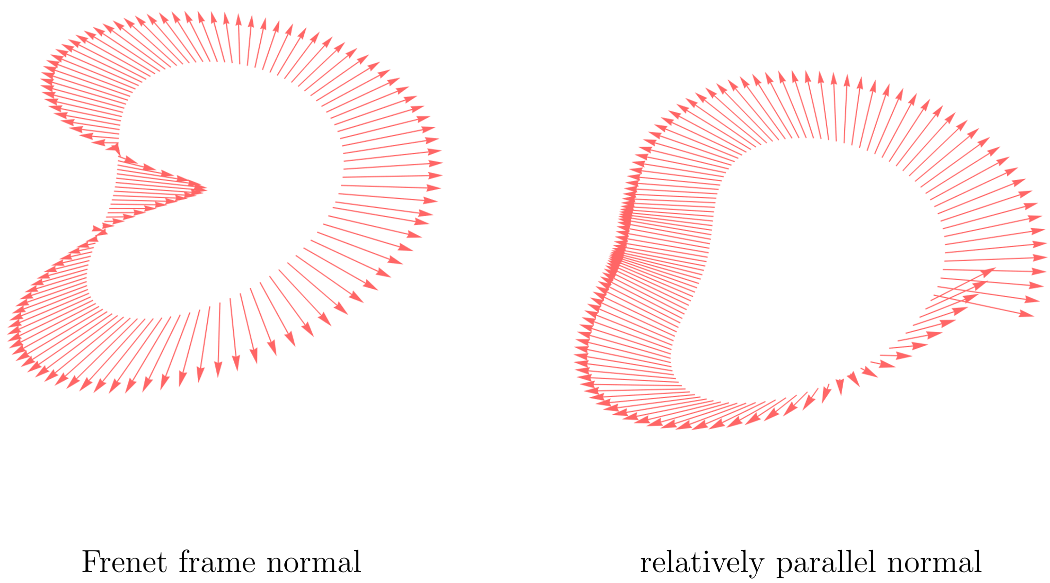


Figure 1.9: The parallel curve γ_r of Example 1.13 with $R = 1$, $\theta(s) = s$ and $r = -0.7$ (the red curve in Figure 1.11) and its principal normal vector field N (left) versus relatively parallel normal vector field N_1 with $N_1(0) = N(0)$. The latter is obviously non-smooth at $\gamma_r(0) = \gamma_r(2\pi R)$.

A closed curve $\gamma : [a, b] \to \mathbb{R}^d$ is called *simple* if $\gamma|_{[a,b)}$ is injective.

Example 1.10 (Circle in space)**.** As an example of a simple closed curve, let us consider the circle of radius $R > 0$ in $\mathbb{R}^3$:

$$\gamma : [0, 2\pi R] \to \mathbb{R}^3 : \left\{ s \mapsto \left(R\cos\left(\frac{s}{R}\right), R\sin\left(\frac{s}{R}\right), 0 \right) \right\} .$$

The division by R is introduced in order to make the parameterisation unit-speed. The distingusihed Frenet frame reads

$$\begin{aligned} T &= \gamma' = \left(-\sin\left(\frac{s}{R}\right), \cos\left(\frac{s}{R}\right), 0 \right), \\ N &= \frac{\gamma''}{|\gamma''|} = \left(-\cos\left(\frac{s}{R}\right), -\sin\left(\frac{s}{R}\right), 0 \right), \\ B &= T \times N = (0, 0, 1) . \end{aligned}$$

As expected, the curvature of γ is the inverse of the radius,

$$\kappa = |\gamma''| = \frac{1}{R} .$$

The torsion is identically equal to zero, because γ is contained in the x^1x^2-plane. Explicitly, by the third of the Frenet–Serret equations (1.14), $\tau = -N \cdot B' = 0$. $\diamondsuit$

1.5 Parallel curves and their length

To apply the concepts of moving frames, let us recall this classical mathematical puzzle:

> *Imagine that a cable is wrapped around the equator of Earth. If the cable should be raised 1 metre above the ground level, how much extra cable would one need?*

The answer is counter-intuitive to non-mathematicians. Indeed, if R is the radius of Earth, the solution reads $2\pi(R + 1) - 2\pi R = 2\pi$, which is merely around 6 metres. In particular, the solution does not depend on R.

Let us consider the puzzle in the context of general unit-speed curves $\gamma : I \to \mathbb{R}^d$ with I bounded. (The image $\gamma(I)$ is not necessarily a circle, moreover γ does not need to be closed.) Let $(T, N_1, \dots, N_{d-1})$ be a relatively parallel adapted frame of γ. Given a rotation matrix-valued function $R : I \to \mathsf{SO}(d-1) \subset \mathbb{R}^{(d-1)\times(d-1)}$, consider the transformed moving frame (recall (1.7), where however the role of rotation is inverted)

$$\begin{pmatrix} T \\ M_1 \\ \vdots \\ M_{d-1} \end{pmatrix} := \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \begin{pmatrix} T \\ N_1 \\ \vdots \\ N_{d-1} \end{pmatrix} .$$

Let M denote one of its normal members $M_1, \dots, M_{d-1}$.

Without loss of generality, let us take

$$M := M_1 = R_1{}^j N_j ,$$

where $R = (R_i^{\ j})_{i,j=1,\dots,d-1}$. To simplify the notation, set $\Theta^j := R_1^{\ j}$ and call

$$\Theta := (\Theta^1, \dots, \Theta^{d-1})$$

the *twisting vector*. In summary, one has $M = \Theta^j N_j = \Theta^1 N_1 + \dots + \Theta^{d-1} N_{d-1}$, where $|\Theta|^2 = (\Theta^1)^2 + \dots + (\Theta^{d-1})^2 = 1$ (because of the orthogonality of R).

Given $r \in \mathbb{R}$, consider the *parallel* curve

$$\gamma_r : I \to \mathbb{R}^d : \{s \mapsto \gamma(s) + rM(s)\}. \tag{1.16}$$

Obviously, $\gamma_0 = \gamma$ and γ_r is a curve parallel to γ at distance $|r|$. Let us make a restriction on r to guarantee that γ_r is regular. To this purpose, recalling (1.5), we compute

$$\gamma_r' = (1 - r\, k_j \Theta^j)\, T + r\, \Theta^j N_j\,.$$

Consequently,

$$|\gamma_r'|^2 = (1 - r\, k_j \Theta^j)^2 + r^2\, (\Theta^j N_j)^2\,. \tag{1.17}$$

By the Schwarz inequality in $\mathbb{R}^{d-1}$, $|k_j \Theta^j| \le |k||\Theta| = |k|$, where $k := (k_1, \dots, k_{d-1})$ is called the *curvature vector*. Hence, a sufficient condition to guarantee that γ_r is regular reads $k \in L^\infty(I)$ and

$$|r|\, |k| < 1. \tag{1.18}$$

Here the restriction on the boundedness of the curvature vector is fundamental if the interval I is allowed to be open, see Figure 1.10.

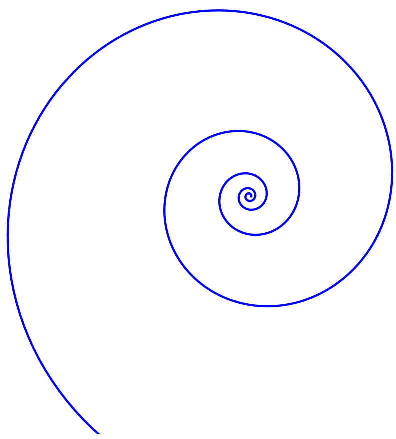


Figure 1.10: A spiral $\gamma : (0,1) \to \mathbb{R}^2 : \{t \mapsto (e^{t^{-1}} \cos(t^{-1}), e^{t^{-1}} \sin(t^{-1}))\}$ as an example of a curve with unbounded curvature.

We are interested in the length

$$L(\gamma_r) := \int_I |\gamma_r'(s)|\, \mathrm{d}s\,.$$

Since γ is unit-speed, one obviously has $L(\gamma) = |I|$. How is this length changed for non-zero r? From (1.17), we deduce the general formula

$$L(\gamma_r) - L(\gamma) = \int_I \left(\sqrt{(1 - r\, k_j(s) \Theta^j(s))^2 + r^2\, (\Theta^j(s) N_j(s))^2} - 1 \right) \mathrm{d}s.$$

$\boxed{d=2}$ In the planar situation $\gamma : I \to \mathbb{R}^2$ corresponding to the puzzle above, there is no twisting. More specifically, one can choose $\Theta = \Theta^1 = 1$ and $N_1 = N$, where N is the normal vector of (1.13). Consequently,

$$L(\gamma_r) - L(\gamma) = -r \int_I k(s)\,\mathrm{d}s\,, \tag{1.19}$$

where k is the signed curvature of γ. The integral of the curvature function k is called the *total curvature* of γ. The following theorem provides a beautiful relationship between a local geometry of the curve (curvature) and its global properties. (In German, *Umlauf* and *Satz* mean "rotation" and "theorem", respectively.)

Theorem 1.11 (Umlaufsatz [Hopf 1935])**.** *Let $\gamma : I \to \mathbb{R}^2$ be a simple closed unit-speed curve. Then*

$$\int_I k(s)\,\mathrm{d}s = \pm 2\pi\,,$$

where the sign depends solely on the orientation of the curve.

Proof. We only give an idea of the proof. Let $\theta : I =: [a,b] \to \mathbb{R}$ be a continuous, piecewise differentiable function satisfying

$$T(s) = \gamma'(s) = (\cos\theta(s), \sin\theta(s))\,.$$

Such a function is called the *tangent angle* of γ. Since $k = \theta'$, one has

$$\int_I k(s)\,\mathrm{d}s = \theta(b) - \theta(a)\,.$$

Since $\theta(s)$ is the angle that $T(s)$ makes with the first axis of $\mathbb{R}^2$ and γ is closed, it is not difficult to see that $\theta(b) - \theta(a) \in 2\pi\mathbb{Z}$. The most difficult part of the proof is to show that for *simple* closed curves, the only possibility is $\theta(b) - \theta(a) = \pm 2\pi$. The plus sign corresponds to *positively* (or *counterclockwise*) oriented curves, meaning that the traveller has the curve interior of γ to the left. □

As a consequence of this theorem and (1.19), we get the ultimate result generalising the puzzle to arbitrary plane curves.

Corollary 1.12. *Let γ be a simple closed unit-speed plane curve. Given any real number r satisfying* (1.18)*, let γ_r be a parallel curve at distance $|r|$. Then*

$$L(\gamma_r) - L(\gamma) = \mp 2\pi r\,.$$

We see that the difference is indeed independent of the curvature of γ. Moreover, recall that the parallel curve γ_r was constructed using the normal N of the distinguished (positively oriented) Frenet frame (T, N). If γ is positively oriented, then N points to the interior of γ. Hence, we get the expected property that the plus sign corresponds to parallel curves located in the exterior of γ.

Example 1.13 (Parallel curves to the circle in space)**.** Recall Example 1.10. Since γ is a plane curve, the Frenet frame (T, N, B) is also the relatively parallel adapted frame of γ. We define

$$\begin{pmatrix} M_1 \\ M_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} N \\ B \end{pmatrix},$$

where $\theta : I \to \mathbb{R}$ is a smooth function, and set $M := M_1$. Finally, we consider the parallel curve γ_r defined by (1.16). The regularity hypothesis (1.18) is satisfied provided that $|r| < R$. Several scenarios are depicted in Figure 1.11.

As an example of the general non-smoothness of the relatively parallel adapted frame for closed curves, see Figure 1.9. ♢

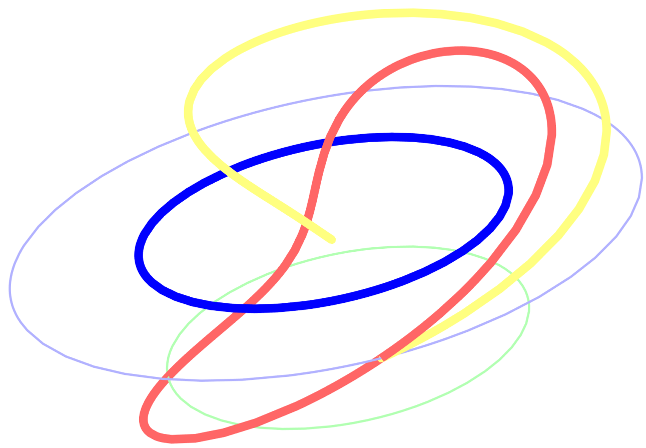


Figure 1.11: Parallel curves of Example 1.13 with $R = 1$: $r = 0$ (dark blue); $\theta = 0$ and $r = -0.7$ (light blue); $\theta = \frac{\pi}{2}$ and $r = 0.7$ (light green); $\theta(s) = s$ and $r = -0.7$ (red); $\theta(s) = \frac{s}{2}$ and $r = -0.7$ (yellow)

1.6 The isoperimetric inequality

In this subsection, we look at another classical problem in global theory of plane curves. Indeed, this is perhaps the oldest global theorem in differential geometry.

> *Among all closed simple closed plane curves with a fixed length, which one encloses the largest area?*

The well-known answer is the *circle* (of the same length).

Since the interior of the closed curve is a domain (*i.e.*, open connected set) the fact can alternatively be stated that among all planar domains of a given perimeter, the disk has the

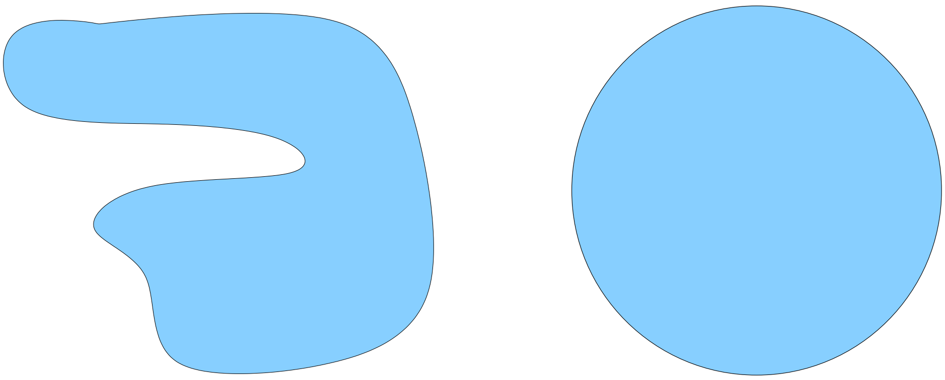


Figure 1.12: An arbitrary domain (left) to be compared with the disk (right).

largest area (see Figure 1.12). In fact, this is a more general statement, because the domain needs to be neither smooth nor simply connected. That is,

$$\max_{|\partial\Omega|=\mathrm{const}} |\Omega| = |B| \,, \tag{1.20}$$

where the maximum is taken over all bounded domains $\Omega \subset \mathbb{R}^2$ of the fixed perimeter $|\partial\Omega| = \mathrm{const}$, B denotes the disk of the same perimeter as Ω (*i.e.* $|\partial B| = |\partial\Omega| = \mathrm{const}$) and $|\Omega|$ denotes the area of Ω. It is indeed an inequality because (1.20) is equivalent to the statement

$$\forall\Omega,\ |\partial\Omega| = \mathrm{const}, \qquad |\Omega| \leq |B| \,. \qquad\qquad (|\partial B| = |\partial\Omega| = \mathrm{const}) \tag{1.21}$$

Moreover, the inequality becomes equality if, and only if, $\Omega = B$.

By scaling, (1.20) is equivalent to the *isochoric inequality* stating that among all planar domains of a given area, the disk has the smallest perimeter. That is,

$$\min_{|\Omega|=\mathrm{const}} |\partial\Omega| = |\partial B| \,, \tag{1.22}$$

where the minimum is taken over all bounded domains $\Omega \subset \mathbb{R}^2$ of a fixed area $|\Omega| = \mathrm{const}$ and now B denotes the disk of the same area as Ω (*i.e.* $|B| = |\Omega| = \mathrm{const}$). Again, one is concerned with an inequality because (1.22) is equivalent to the statement

$$\forall\Omega,\ |\Omega| = \mathrm{const}, \qquad |\partial\Omega| \geq |\partial B| \,. \qquad\qquad (|B| = |\Omega| = \mathrm{const}) \tag{1.23}$$

Moreover, the inequality becomes equality if, and only if, $\Omega = B$.

Since the perimeter and area of a disk is known explicitly, the two inequalities (1.21) and (1.23) can be stated as a unique inequality (without any further constraints on the domain Ω).

Theorem 1.14 (Isoperimetric inequality [Dido 900 BC])**.** *For any bounded domain $\Omega \subset \mathbb{R}^2$, one has*

$$|\partial\Omega|^2 - 4\pi|\Omega| \geq 0 \,. \tag{1.24}$$

Moreover, the inequality becomes equality if, and only if, Ω is a disk.

Indeed, if R denotes the radius of B, then the isoperimetric constraint requires $2\pi R = |\partial\Omega|$, while (1.21) states that $|\Omega| \leq \pi R^2$; eliminating R, we arrive at (1.24).

These two classical geometric optimisation problems were known to ancient Greeks. Indeed, they are usually attributed to the legendary queen of Carthage Dido (roughly 900 BC). However, a first "proof" was given by Steiner in 1838 AD. Actually, he provided five different "proofs", but all of them were wrong, in the sense that he *a priori* assumed that a solution should exist. This was eventually settled by Weierstrass in 1879. See [1] for an overview of this mathematical adventure. The analogous statements hold in higher dimensions as well.

Proof of Theorem 1.14. Following [2, Thm. 1.7.1], we give an elementary proof in the initial setting of $\partial\Omega$ being a (smooth) simple closed curve.

Define

$$x^1_{\min} := \inf_{x\in\Omega} x^1\,, \quad x^1_{\max} := \sup_{x\in\Omega} x^1\,, \quad x^2_{\min} := \inf_{x\in\Omega} x^2\,.$$

Set $r := (x^1_{\max} - x^1_{\min})/2$. Translate Ω in $\mathbb{R}^2$ in such a way that (the translated domain, denoted by the same symbol) $\Omega \subset \{x \in \mathbb{R}^2 : |x^1| < r \ \wedge \ x^2 > r\}$, see Figure 1.14. Let $\gamma : [0, l] \to \mathbb{R}^2$ with $l := |\partial\Omega|$ be a positively oriented simple closed unit-speed curve tracing the boundary $\partial\Omega$. That is, the interior of γ equals the domain Ω and the image $\gamma([0, l])$ coincides with the boundary $\partial\Omega$.

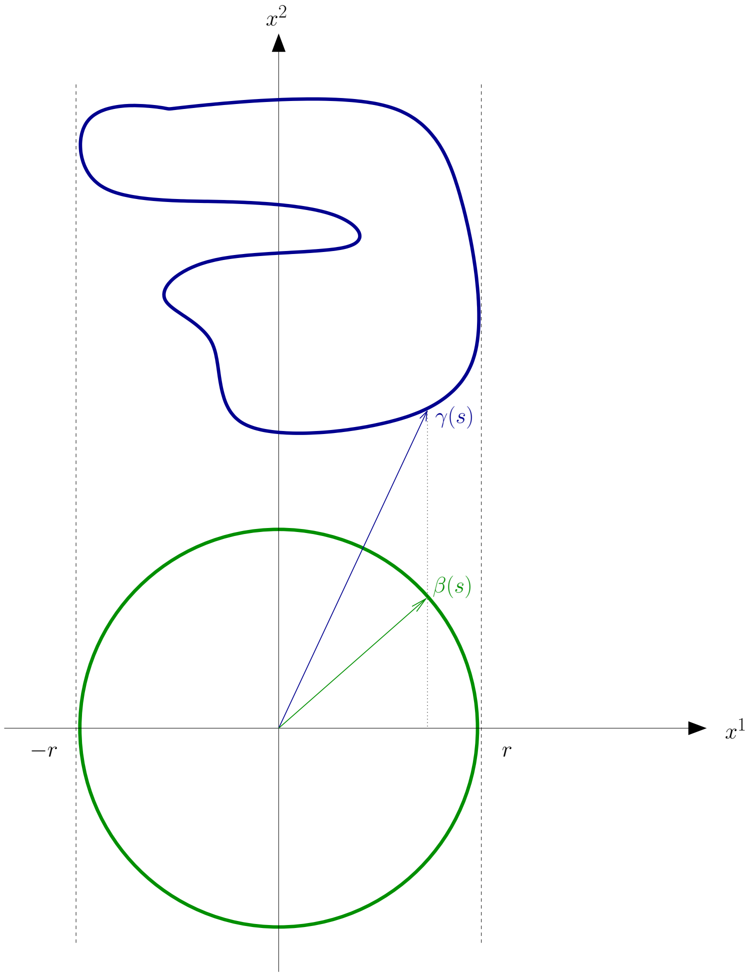


Figure 1.13: The geometry of the proof of Theorem 1.14.

Let $B_r(0)$ be the open disk centred as the origin of $\mathbb{R}^2$. There exists a smooth function $\beta^2 : [0,l] \to \mathbb{R}^2$ such that $\beta := (\beta^1, \beta^2)$ with $\beta^1 := \gamma^1$ is a curve tracing the boundary $\partial B_r(0)$. Note that β is not necessarily unit-speed.

Let us recall the divergence theorem

$$\int_\Omega \operatorname{div} F = \int_{\partial\Omega} F \cdot n \,,$$

where $F : \Omega \to \mathbb{R}^d$ is any smooth vector field and n is the outward unit normal of Ω. If (T, N) is the distinguished Frenet frame of γ, then $N = (-\gamma^{2\prime}, \gamma^{1\prime})$ is inward pointing. Choosing $F(x) := (x^1, 0)$, $F(x) := (0, x^2)$ or $F(x) := \frac{1}{2}(x^1, x^2)$, we respectively obtain

$$|\Omega| = \int_0^l \gamma^1(s)\, \gamma^{2\prime}(s)\, \mathrm{d}s = -\int_0^l \gamma^2(s)\, \gamma^{1\prime}(s)\, \mathrm{d}s = \frac{1}{2}\int_0^l \left[\gamma^1(s)\, \gamma^{2\prime}(s) - \gamma^2(s)\, \gamma^{1\prime}(s)\right] \mathrm{d}s \,.$$

Exactly the same formula holds for γ tracing the boundary $\partial\Omega$ even if it is not unit-speed.

Applying the area formula above, one has

$$|\Omega| = \int_0^l \gamma^1(s)\, \gamma^{2\prime}(s)\, \mathrm{d}s \qquad \text{and} \qquad |B_r(0)| = -\int_0^l \beta^2(s)\, \beta^{1\prime}(s)\, \mathrm{d}s \,.$$

Using $|B_r(0)| = \pi r^2$, we write

$$\begin{aligned} |\Omega| + \pi r^2 &= \int_0^l \left[\gamma^1(s)\, \gamma^{2\prime}(s) - \beta^2(s)\, \beta^{1\prime}(s)\right] \mathrm{d}s \\ &\leq \int_0^l \left|\gamma^1(s)\, \gamma^{2\prime}(s) - \beta^2(s)\, \beta^{1\prime}(s)\right| \mathrm{d}s \\ &\leq \int_0^l \sqrt{\gamma^1(s)^2 + \beta^2(s)^2}\, \sqrt{\gamma^{2\prime}(s)^2 + \beta^{1\prime}(s)^2}\, \mathrm{d}s \qquad (1.25) \\ &= \int_0^l \sqrt{\beta^1(s)^2 + \beta^2(s)^2}\, \mathrm{d}s \\ &= lr \,. \end{aligned}$$

Here the second estimate is due to the Schwarz inequality and the last but one equality employs that we have chosen $\beta^1 = \gamma^1$ (together with the fact that γ is unit speed). The last equality employs that β is a circle of radius r. Using the inequality between the geometric and arithmetic means, we obtain

$$\sqrt{|\Omega|}\sqrt{\pi r^2} \leq \frac{|\Omega| + \pi r^2}{2} \leq \frac{lr}{2} \,. \qquad (1.26)$$

Squaring, we establish (1.24).

Now assume that equality holds in (1.24). Then equality must hold everywhere in (1.25) and (1.26). From the latter, we deduce $|\Omega| = \pi r^2$ and $l = 2\pi r$. In particular, the "horizontal width" r of Ω depends only on the perimeter $|\partial\Omega|$. From (1.25) (since the equality in the Schwarz inequality holds if, and only if, one of the vectors is a multiple of the other), we obtain

$$(\gamma^1, \beta^2) = \lambda\, (\gamma^{2\prime}, -\beta^{1\prime}) \,, \qquad (1.27)$$

where λ is a real number. Consequently, recalling that $\beta^1 = \gamma^1$,

$$r^2 = (\gamma^1)^2 + (\beta^2)^2 = \lambda^2 \left[({\gamma^2}')^2 + ({\beta^1}')^2\right] = \lambda^2 .$$

Hence $\lambda = \pm r$. From (1.27), we deduce $\gamma^1 = \pm r\, {\gamma^2}'$. Since r is independent of the rotation of Ω, we can interchange γ^1 and γ^2 in the last relation and obtain $\gamma^2 = \pm r\, {\gamma^1}'$. Consequently,

$$(\gamma^1)^2 + (\gamma^2)^2 = r^2 ,$$

so γ is a circle, as we wished. $\square$

1.7 Exercises

1. Let $\gamma : I \to \mathbb{R}^2$ be a regular curve, not necessarily unit-speed. Show that its signed curvature is given by
$$k = \frac{\det(\gamma', \gamma'')}{|\gamma'|^3}\,.$$

2. Let $\gamma : I \to \mathbb{R}^3$ be a regular curve, not necessarily unit-speed. Show that its curvature and torsion are respectively given by
$$\kappa = \frac{|\gamma' \times \gamma''|}{|\gamma'|^3} \qquad \text{and} \qquad \tau = \frac{\det(\gamma', \gamma'', \gamma''')}{|\gamma' \times \gamma''|^2}\,.$$

3. Compute the curvature of the curve γ_5 from Figure 1.3.

4. Compute the curvature of the curve γ_ε from Figure 1.4.

5. Compute the curvature of the curve γ_δ from Figure 1.5.

6. Compute the curvature and torsion of the curve γ from Figure 1.6.

7. Compute the curvature and torsion of the curve γ from Figure 1.10.

8. Compute the curvature and torsion of the curves γ_r from Figure 1.11.

2 Surfaces

To simplify the presentation we restrict to (two-dimensional) surfaces in $\mathbb{R}^3$.

2.1 Parameterised surfaces

The generic notation $U \subset \mathbb{R}^2$ will always mean an open subset in this section. By a surface, we mean a smooth deformation of U. In the same way as a curve was a smooth deformation of the one-dimensional interval I. In contrast to curves, however, we include the condition on regularity in the very definition.

Definition 2.1. A *surface* is a smooth map $p : U \to \mathbb{R}^3$ satisfying $\operatorname{rank} p' = 2$.

Here $p'(u)$ denotes the Jacobi matrix of p at point $u \in U$ with respect to the standard bases in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Recall (*cf* [7]) that the *derivative* of any smooth function $f : \mathbb{R}^n \to \mathbb{R}^m$ at point $x \in \mathbb{R}^n$ is the linear transformation $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Df(x)h}{|h|} = 0 .$$

The *Jacobi matrix* $f'(x)$ of f at point $x \in \mathbb{R}^n$ is the matrix of $Df(x)$ with respect to the standard bases in $\mathbb{R}^n$ and $\mathbb{R}^m$. For practical computations, note that $Df = (\nabla f)^\top$.

Since $2 = \dim \mathbb{R}^2 = \dim \ker Dp(u) + \operatorname{rank} Dp(u)$, we see that the condition $\operatorname{rank} p' = 2$ means that $Dp(u) : \mathbb{R}^2 \to \mathbb{R}^3$ is injective for every $u \in U$. By the inverse function theorem, it follows that p is a local diffeomorphism. In our case, the Jacobi matrix reads

$$p'(u) = \begin{pmatrix} \partial_1 p^1(u) & \partial_2 p^1(u) \\ \partial_1 p^2(u) & \partial_2 p^2(u) \\ \partial_1 p^3(u) & \partial_2 p^3(u) \end{pmatrix} .$$

The geometric meaning of the regularity condition $\operatorname{rank} p' = 2$ is that the *tangent* vector fields $\partial_1 p, \partial_2 p$ are linearly independent at every point of U, see Figure 2.1.

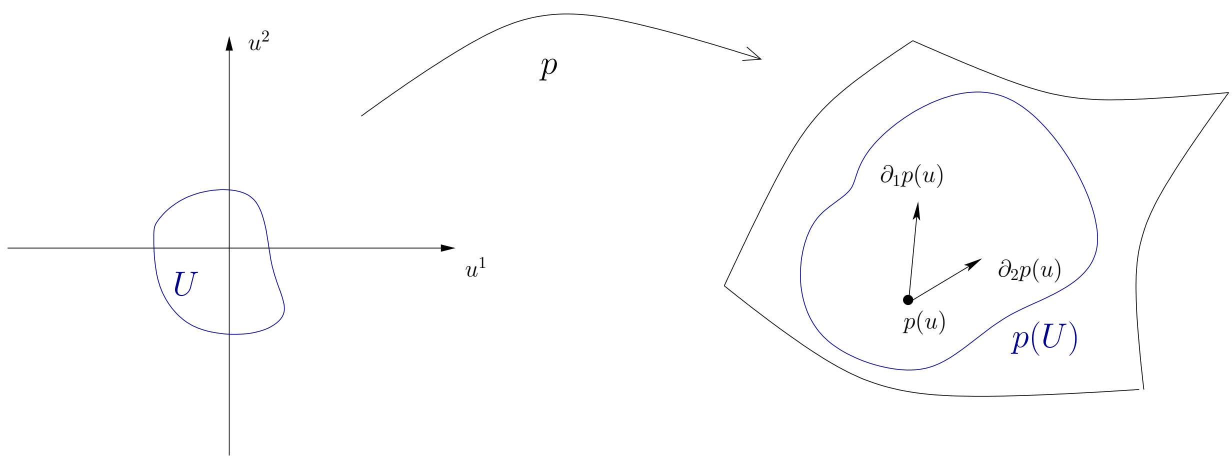


Figure 2.1: A parameterised surface.

Reparameterisation of a surface $p : U \to \mathbb{R}^3$ is another surface $\tilde{p} := p \circ \phi : \tilde{U} \to \mathbb{R}^3$, where $\tilde{U} \subset \mathbb{R}^2$ is another open set and $\phi : \tilde{U} \to U$ is a diffeomorphism. Obviously, p and $\tilde{p}$ have the same image. If $\det \phi' > 0$ (respectively, $\det \phi' < 0$), the reparameterisation is said to be *orientation-preserving* (respectively, *orientation-reversing*).

2.2 Examples

Example 2.2 (Sphere)**.** The most famous "surface" is certainly the *sphere* (see Figure 2.2)

$$\mathbb{S}_R := \{(x^1, x^2, x^3) \in \mathbb{R}^3 : \ (x^1)^2 + (x^2)^2 + (x^3)^2 = R^2\} .$$

In quotation marks because, globally, it does not fall into the category of our (parameterised) surfaces from Definition 2.1. The northern and southern hemispheres without the equator can be respectively parameterised via

$$p_\pm : U_R \to \mathbb{R}^3 : \left\{(u^1, u^2) \mapsto \left(u^1, u^2, \sqrt{R^2 - (u^1)^2 - (u^2)^2}\right)\right\} ,$$

where $U_R := \{u \in \mathbb{R}^2 : \ |u| < R\}$ is the planar open disk of radius R. Alternatively, using spherical coordinates, we define

$$p : (0, \pi) \times (0, 2\pi) \to \mathbb{R}^3 : \{(\theta, \varphi) \mapsto (R \sin\theta \cos\varphi, R \sin\theta \sin\varphi, R \cos\theta)\} ,$$

which parameterises the whole sphere, except for the meridian (also visualised in Figure 2.2)

$$\{(R \sin\theta, 0, R \cos\theta) : \ \theta \in [0, \pi]\} .$$

If $R = 1$, we abbreviate $\mathbb{S}_1 =: \mathbb{S}$. $\diamondsuit$

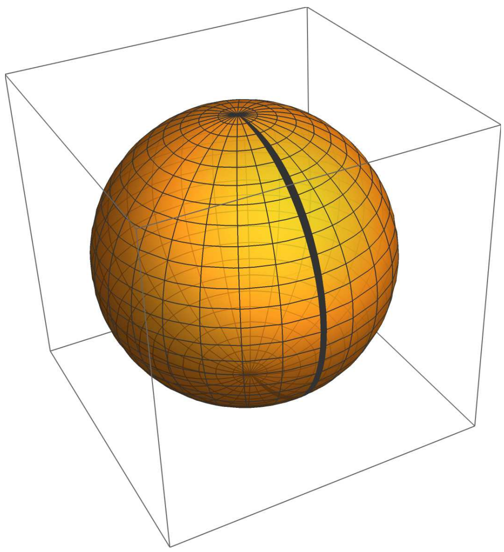


Figure 2.2: Sphere as the boundary of a ball.

Example 2.3 (Torus)**.** Another famous "surface" is the *torus*, which is obtained by rotating a circle of radius $r > 0$ about a straight line belonging to the plane of the circle and at a distance $R > r$ away from the centre:

$$\mathbb{T}_{R,r} := \left\{ (x^1, x^2, x^3) \in \mathbb{R}^3 : \ \left[\sqrt{(x^1)^2 + (x^2)^2} - R \right]^2 + (x^3)^2 = r^2 \right\} .$$

Except for a set formed by the union of a small circle and a great circle, a suitable parameterisation is given by

$$p : (0, 2\pi) \times (0, 2\pi) \to \mathbb{R}^3 : \{ (\phi, \varphi) \mapsto ((R + r\cos\phi)\cos\varphi, (R + r\cos\phi)\sin\varphi, r\sin\phi) \} ,$$

see Figure 2.3.

If we are not interested in specific values of the radii, we simply write $\mathbb{T}$ for a torus. $\diamondsuit$

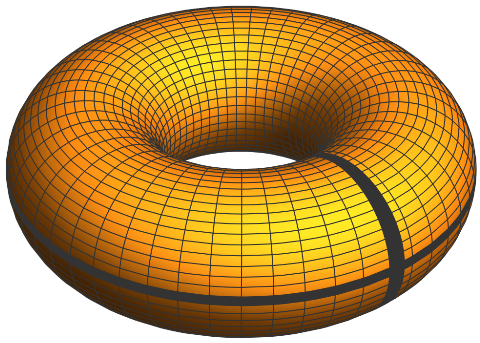


Figure 2.3: Torus as the boundary of a donut.

Example 2.4 (Ruled surfaces)**.** Recall Example 1.13 and Figure 1.11 about parallel curves γ_r to the circle in $\mathbb{R}^3$. Letting the distance function r in the definition (1.16) variable, we obtain a special class of *ruled surfaces*

$$p : (0, 2\pi R) \times (-r_{\max}, r_{\max}) \to \mathbb{R}^3 : \{ (s, r) \mapsto \gamma(s) + r M(s) \} .$$

(What is the restriction on the positive number $r_{\max}$ to make p the (regular) surface? See Exercise 2.9.4.) The surface variants of the parallel curves realisations from Figure 1.11 are depicted in Figure 2.4 (following the same colour scheme). $\diamondsuit$

Example 2.5 (Surfaces of revolution)**.** Consider the surface in $\mathbb{R}^3$ generated by revolving the image of a plane curve (contained in the x^1x^3-plane) $\gamma : [0, a) \to \mathbb{R}^2 : \{ s \mapsto (r(s), z(s)) \}$ with $a \in (0, \infty) \cup \{\infty\}$ about the x^3-axis (see Figure 2.5). Assume that γ is injective and unit-speed (*i.e.*, $r'^2 + z'^2 = 1$). Moreover, assume $r(s) > 0$ for $s \in [0, a)$ and $\lim_{s \to a} \gamma(s) \neq (0, 0)$. Finally, in order to ensure the smoothness at the *pole* $o := (0, 0, 0)$, assume $r(0) = 0$ and

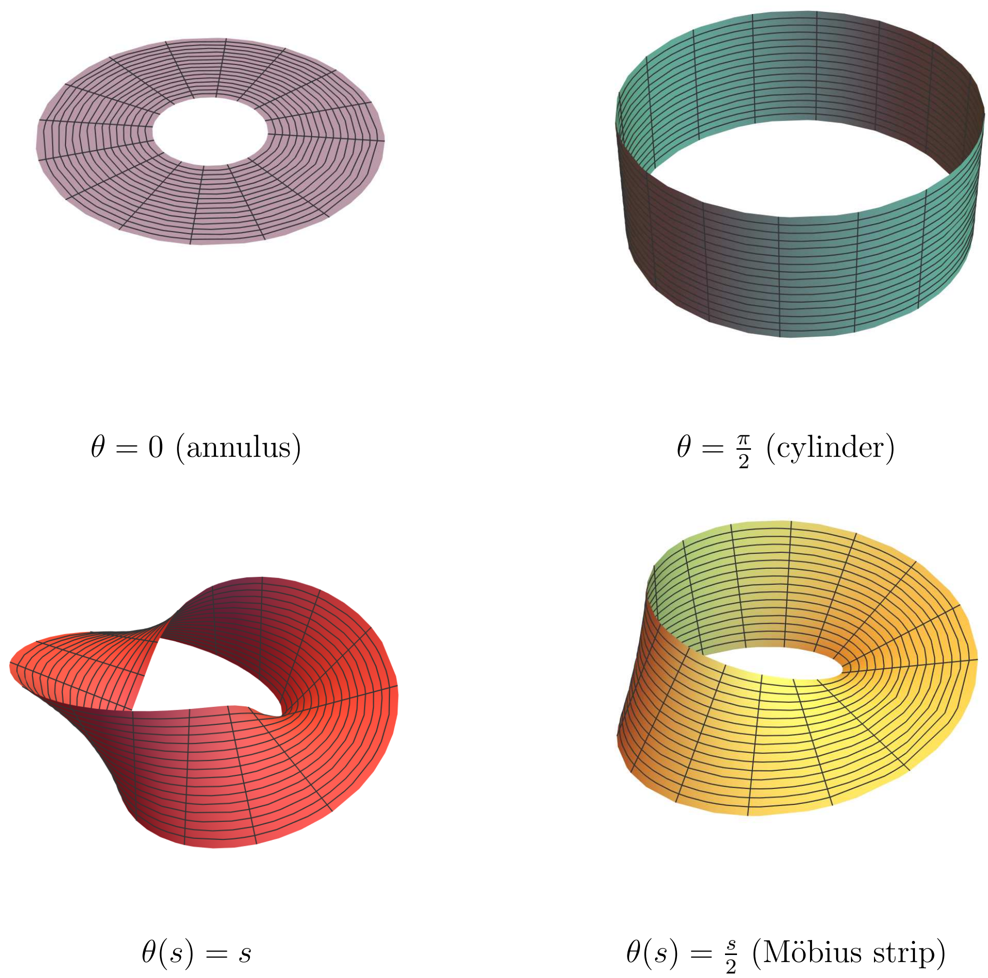


Figure 2.4: Ruled surfaces of Example 2.4 .

$r'(0) = 1$ (which implies $z'(0) = 0$). The natural parameterisation of such a surface (except for the pole and the *meridian* $s \mapsto (r(s), 0, z(s))$) is given by

$$p : (0, a) \times (0, 2\pi) \to \mathbb{R}^3 : \{(s, \varphi) \mapsto (r(s)\cos\varphi, r(s)\sin\varphi, z(s))\} \ . \qquad (2.1)$$

Since $|\partial_1 p \times \partial_2 p| = r > 0$, it is a (regular) surface.

Particular situations of the surfaces of revolution are the plane ($r(s) := s$ and $z(s) := 0$ with $a := \infty$) and the sphere ($r(s) := R\sin(\frac{s}{R})$ and $z(s) := R - R\cos(\frac{s}{R})$ with $a := \pi R$). $\diamondsuit$

Other examples of surfaces can be found in Figure 2.9.

2.3 The first fundamental form or metric

Let $p : U \to \mathbb{R}^3$ be a surface. Then the tangent vectors $\partial_1 p(u), \partial_2 p(u)$ are linearly independent at every point $u \in U$. We introduce the *tangent space* of p at $u \in U$ by

$$T_u p := \{p(u)\} \times [(Dp)(u)](\mathbb{R}^2) = \{p(u)\} \times \operatorname{span}\big(\partial_1 p(u), \partial_1 p(u)\big) \ .$$

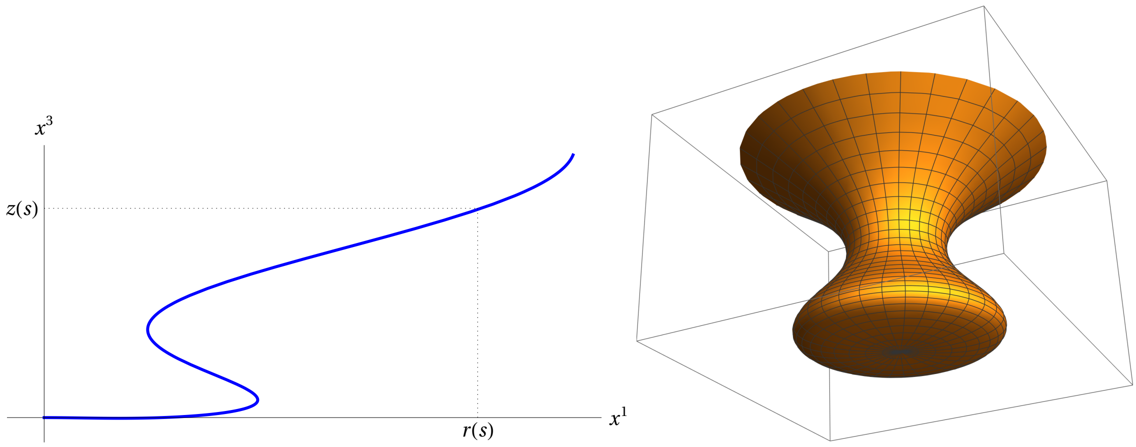


Figure 2.5: The generating plane curve γ and the generated surface of revolution.

Clearly, $T_u p \subset T_{p(u)}\mathbb{R}^3$. In fact, $\dim(T_u p) = 2$ and the geometrical interpretation of $T_u p$ is the tangent plane to $p(U)$ at $p(u)$. It is customary to identify $T_u p$ with the two-dimensional subspace $[(Dp)(u)](\mathbb{R}^2) \subset \mathbb{R}^3$.

The disjoint union of all the tangent spaces of p is called the *tangent bundle*:

$$Tp := \dot{\bigcup_{u\in U}} T_u p\,.$$

Definition 2.6. The *metric* (or the *first fundamental form*) of p at $u \in U$ is the bilinear form

$$g_u : T_u p \times T_u p \to \mathbb{R} : \Big\{ \big((p(u), X), (p(u), Y)\big) \mapsto X \cdot Y \Big\}\,.$$

In other words, the metric associates to two tangent vectors their inner product inherited by the ambient space $\mathbb{R}^3$. We also use the symbol g to denote the map $g : Tp \times Tp \to \mathbb{R}$, which is the union of all g_u.

The matrix representation of the metric with respect to the frame $\big(\partial_1 p, \partial_1 p\big)$ is the matrix $\big(g_{ij}\big)_{i,j=1,2} := \big(g(\partial_i p, \partial_j p)\big)_{i,j=1,2}$. One has

$$g_{ij} = \partial_i p \cdot \partial_j p. \tag{2.2}$$

In matrix formalism, this is equivalent to $\big(g_{ij}\big)_{i,j=1,2} = (\nabla p)(\nabla p)^\top = (Dp)^\top(Dp)$.

Clearly, the matrix is positive definite and $u \mapsto g_{ij}(u)$ is smooth. What is more, if $X, Y : U \to \mathbb{R}^3$ are (smooth) tangential vector fields (*i.e.*, $X(u), Y(u) \in T_u p$ for every $u \in U$), then $u \mapsto g_{p(u)}(X(u), Y(u))$ is smooth.

Occasionally, we write $\langle X, Y\rangle$ instead of $g(X, Y)$, which is justified by the fact that g satisfies the properties of an inner product.

It is customary to denote by g^{ij} the entries of the matrix $\big(g^{ij}\big)_{i,j=1,2}$ inverse to $\big(g_{ij}\big)_{i,j=1,2}$.

The metric g plays the role of the squared norm $|\gamma'|^2 = \gamma' \cdot \gamma'$ for a curve γ. Without loss of generality, one can always choose $|\gamma'| = 1$ (recall Proposition 1.3), so the notion of the metric for curves would be trivial. For surfaces, however, this is a fundamental object. In particular, it enables one to measure lengths of tangential vectors and angles between them.

Example 2.7 (The metric of surfaces of revolution)**.** Recalling Example 2.5, in particular (2.1), we have

$$\begin{aligned} \partial_1 p(s,\varphi) &= \big(r'(s)\cos\varphi, r'(s)\sin\varphi, z'(s)\big)\,, \\ \partial_2 p(s,\varphi) &= \big(-r(s)\sin\varphi, r(s)\cos\varphi, 0\big)\,. \end{aligned}$$

Consequently, $\partial_1 p \cdot \partial_1 p = r'^2 + z'^2 = 1$, $\partial_2 p \cdot \partial_2 p = r^2$, $\partial_1 p \cdot \partial_2 p = 0$, and therefore

$$(g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}.$$

Note that the metric is diagonal and independent of the second (angular) variable. ♢

2.4 The second fundamental form

Let $p : U \to \mathbb{R}^3$ be a surface. The tangent vector fields $\partial_1 p, \partial_2 p$ gives rise to the unit normal vector field

$$n := \frac{\partial_1 p \times \partial_2 p}{|\partial_1 p \times \partial_2 p|}. \tag{2.3}$$

The function $n : U \to \mathbb{R}^3$ (with image identified with the sphere $\mathbb{S}$) is called the *Gauss map*. The moving frame $(\partial_1 p, \partial_2 p, n)$ is called the *Gauss frame* of p.

Differentiating the identity $|n|^2 = 1$, it follows that $n \perp \partial_i n$ for $i = 1, 2$. In other words, $[(Dn)(u)](\mathbb{R}^2) \subset [(Dp)(u)](\mathbb{R}^2)$. This justifies the following definition.

Definition 2.8. The *Weingarten map* (or the *shape operator*) of p at $u \in U$ is the endomorphism

$$L_u : T_u p \to T_u p : \big\{\big(p(u), X\big) \mapsto \big(p(u), -[(Dn)(u)] \circ [(Dp^{-1})(u)]X\big)\big\}\,.$$

The *second fundamental form* is the associated bilinear form

$$h_u : T_u p \times T_u p \to \mathbb{R} : \big\{\big((p(u), X), (p(u), Y)\big) \mapsto X \cdot L_u Y\big\}\,.$$

As for the metric, we also use the symbol h to denote the map $h : Tp \times Tp \to \mathbb{R}$, which is the union of all h_u. Similarly, we introduce L.

The matrix representation of the second fundamental with respect to the frame $\big(\partial_1 p, \partial_1 p\big)$ is the matrix $\big(h_{ij}\big)_{i,j=1,2} := \big(h(\partial_i p, \partial_j p)\big)_{i,j=1,2}$. Using that $(Dp^{-1})(Dp) = 1$ (identity in $\mathbb{R}^2$), one has

$$h_{ij} = -\partial_i p \cdot \partial_j n.$$

From the identity $h_{ij} = -\partial_j(\partial_i p \cdot n) + \partial_j \partial_i p \cdot n = \partial_j \partial_i p \cdot n$, it is clear that h is symmetric.

The matrix representation of the Weingarten map with respect to the frame $\big(\partial_1 p, \partial_1 p\big)$ is the matrix $\left(L^i{}_j\right)_{i,j=1,2}$ satisfying $L\partial_j p = L^i{}_j \partial_i p$. One has

$$L^i{}_j = g^{ik} h_{kj}.$$

Because both g and h are symmetric, one also has $L^i{}_j = h_{jk} g^{ki}$. This matrix is not necessarily symmetric, but L is self-adjoint with respect to the inner product $\langle\cdot,\cdot\rangle$. This follows immediately by definition and the symmetry of h.

Example 2.9 (The second fundamental form of surfaces of revolution)**.** Recalling Examples 2.5 and 2.7, we have

$$\partial_1 p(s,\varphi) \times \partial_2 p(s,\varphi) = \big(-r(s)z'(s)\cos\varphi, -r(s)z'(s)\sin\varphi, r(s)r'(s)\big)\,.$$

Consequently,

$$\begin{aligned} n(s,\varphi) &= \big(-z'(s)\cos\varphi, -z'(s)\sin\varphi, r'(s)\big)\,, \\ \partial_1 n(s,\varphi) &= \big(-z''(s)\cos\varphi, -z''(s)\sin\varphi, r''(s)\big)\,, \\ \partial_2 n(s,\varphi) &= \big(z'(s)\sin\varphi, -z'(s)\cos\varphi, r'(s)\big)\,. \end{aligned}$$

Since $\partial_1 p \cdot \partial_1 n = -r'z'' + r''z'$, $\partial_2 p \cdot \partial_2 n = -rz'$, $\partial_1 p \cdot \partial_2 n = 0$, we get

$$(h_{ij}) = \begin{pmatrix} r'z''-r''z' & 0 \\ 0 & rz' \end{pmatrix} \qquad \text{and} \qquad (L^i{}_j) = \begin{pmatrix} r'z''-r''z' & 0 \\ 0 & \dfrac{z'}{r} \end{pmatrix}.$$

Again, both the second fundamental form and the Weingarten map are diagonal and independent of the second (angular) variable. ♢

2.5 Curvatures

Since L is self-adjoint, its spectrum is real. The eigenvalues of L are denoted by κ_1, κ_2 and called the *principal curvatures* of p, *i.e.*,

$$\sigma(L) = \{\kappa_1, \kappa_2\}.$$

The corresponding eigenvectors $E_1, E_2 \in Tp$ are called the *principal directions* of p, *i.e.*,

$$LE_i = \kappa_i E_i$$

for $i = 1, 2$. By the spectral theorem, the eigenvectors can be chosen to be orthonormal.

The principal curvatures determine two important curvature functions of p:

$$\begin{aligned} &\text{the } \textit{Gauss curvature} && K := \kappa_1\kappa_2 && = \det(L)\,, \\ &\text{the } \textit{mean curvature} && H := \frac{\kappa_1+\kappa_2}{2} && = \frac{\operatorname{tr}(L)}{2}\,. \end{aligned}$$

Note that $K = 0$ and $H = 0$ if, and only if, $\kappa_1 = 0$ and $\kappa_2 = 0$.

We say that the surface $p : U \to \mathbb{R}^3$ is *planar* if the unit normal vector field n given in (2.3) is constant ($i.e.$, $Dn = 0$). Obviously, p is planar if, and only if, the image $p(U)$ is contained in a plane in $\mathbb{R}^3$. The following observation shows that the principal curvatures are the right measure of deviation of the surface from being planar.

Proposition 2.10. *Let $p : U \to \mathbb{R}^d$ be a surface. Then*

$$\kappa_1 = \kappa_2 = 0 \qquad \Longleftrightarrow \qquad p \text{ is planar.}$$

Proof. The equivalence follows immediately by observing that p is planar if, and only if, the Weingarten tensor $L = 0$ identically. □

Example 2.11 (The second fundamental form of surfaces of revolution)**.** Recalling Examples 2.5, 2.7 and 2.9, we immediately obtain the principal curvatures

$$\kappa_1 = r'z'' - r''z' \,, \qquad \kappa_2 = \frac{z'}{r} \,.$$

In particular, for the plane we get $\kappa_1 = \kappa_2 = 0$, in agreement with Proposition 2.10. For the sphere, we get $\kappa_1 = \kappa_2 = \frac{1}{R}$. ♢

2.6 Curves on surfaces

By a *curve on the surface* $p : U \to \mathbb{R}^3$ we mean a curve $\gamma : I \to \mathbb{R}^3$ which can be written in the form $\gamma = p \circ \beta$, where $\beta : I \to U \subset \mathbb{R}^2$ is a curve in U.

One has the following very intuitive interpretation of the principal curvatures.

Proposition 2.12. *Let $p : U \to \mathbb{R}^d$ be a surface and E_1, E_2 its principal directions. Given $u \in U$, consider the vector*

$$X_\theta := \cos\theta \, E_1(u) + \sin\theta \, E_2(u) \,, \qquad \theta \in [0, 2\pi) \,.$$

Let γ_θ be the unit-speed plane curve on p obtained by the intersection of the plane spanned by the vectors $n(u)$ and X_θ with the surface image $p(U)$ such that $\gamma_\theta(0) = p(u)$. Consider its curvature $k_\theta(0) := -\gamma_\theta'(0) \cdot N'(0)$, where N is the normal member of the relatively parallel adapted frame of γ_θ lying in the plane. Then

$$\{\kappa_1(u), \kappa_2(u)\} = \left\{ \max_\theta k_\theta(0), \min_\theta k_\theta(0) \right\} \,,$$

where κ_1, κ_2 are the principal curvatures of p.

Proof. See Figure 2.6 for the geometric setting. By definition, $X_\theta \in T_u p$, but we also identify it with a vector in $\mathbb{R}^3$ below. There exists $\varepsilon > 0$ such that the unit-speed curve $\gamma_\theta : (-\varepsilon, \varepsilon) \to \mathbb{R}^3$ satisfies $\gamma(0) = p(u)$ and $\gamma'(0) = X_\theta$. One has $N(0) := \pm n(0)$. Writing $\gamma_\theta = p \circ \beta$, one has $\gamma_\theta' = \beta^{i\prime} \partial_i p$. (To simplify, we write $\partial_i p$ instead of the correct $\partial_i p \circ \beta$, and similarly below.) Consequently,

$$\begin{aligned} k_\theta(0) &= -\beta^{i'}(0)\partial_i p(u) \cdot (\pm)\beta^{j'}(0)\partial_j n(u) \\ &= \pm\beta^{i'}(0)\, h_{ij}(u)\, \beta^{j'}(0) \\ &= \pm h_u[X_\theta] \\ &= \pm(\kappa_1(u)\cos^2\theta + \kappa_2(u)\sin^2\theta)\,, \end{aligned}$$

where $h[\cdot]$ denotes the quadratic form associated with the bilinear form $h(\cdot,\cdot)$. Since the derivative of $k_\theta(0)$ with respect to θ is equal to zero if, and only if, $\theta \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$, we see that the principal curvatures $\kappa_1(u) = \mp k_0(0) = \mp k_\pi(0)$ and $\kappa_2(u) = \mp k_{\frac{\pi}{2}}(0) = \mp k_{\frac{3\pi}{2}}(0)$ are the extremal values (minima and maxima) of the function $\theta \mapsto k_\theta(0)$. $\square$

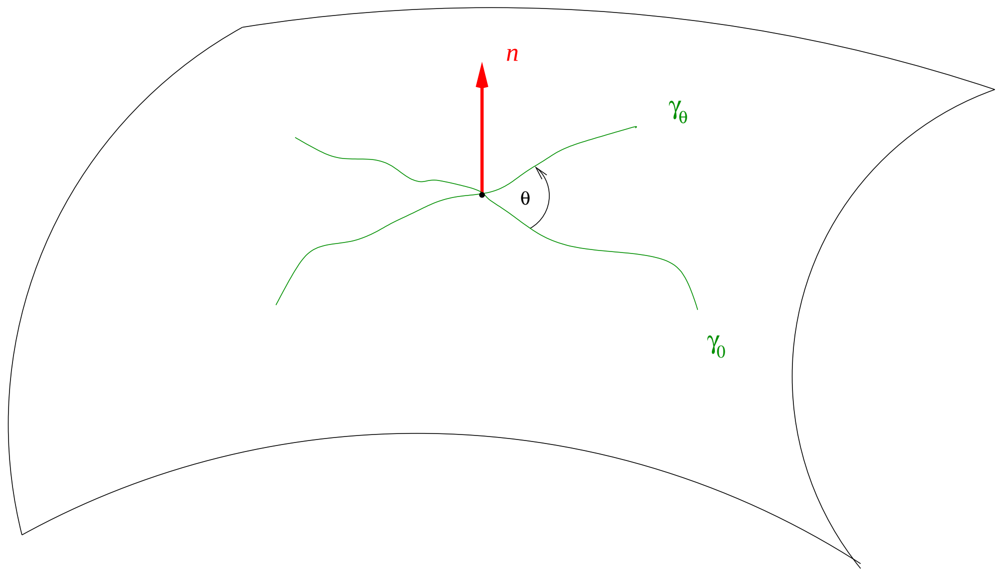


Figure 2.6: Normal sections define a family of plane curves γ_θ depending on an angle parameter θ. The principal curvatures are the maximal and minimal values of the (signed) curvatures k_θ of this family:

$$\{\kappa_1, \kappa_2\} = \left\{\max_\theta k_\theta(0), \min_\theta k_\theta(0)\right\}\,.$$

Given any vector $X \in T_u p$, the value $\kappa(X) := h[X] = h[-X]$ is called the *normal curvature* of p in the direction $\pm X$. As we have seen, $|\kappa(X)|$ equals the curvature of the plane curve obtained by intersecting the surface with plane spanned by the vectors $n(u)$ and X.

Now, let $\gamma : I \to \mathbb{R}^3$ be any unit-speed curve on the surface $p : U \to \mathbb{R}^3$. The acceleration $\gamma''(s)$ with $s \in I$ has a component parallel to $p(U)$ at $\gamma(s)$ and a component perpendicular to $p(U)$ at $\gamma(s)$. The former is perpendicular to the velocity $\gamma'(s)$, because the curve is assumed to be unit-speed. Consequently, if n is the unit normal vector field given by (2.3), we have the decomposition

$$\gamma'' = n \cdot \gamma''\, n + (n \times \gamma') \cdot \gamma''\, (n \times \gamma')\,, \tag{2.4}$$

where the first (respectively, second) term on the right-hand side is the normal (respectively, parallel) component. Inspired by this decomposition, we define two (signed) curvature functions of γ:

$$\begin{aligned} &\text{the } \textit{normal curvature} && \kappa_n := n \cdot \gamma''\,, \\ &\text{the } \textit{geodesic curvature} && \kappa_g := (n \times \gamma') \cdot \gamma''\,. \end{aligned} \tag{2.5}$$

Of course, we have the identity

$$\kappa^2 = \kappa_n^2 + \kappa_g^2 \,,$$

where $\kappa = |\gamma''|$ is the curvature of γ introduced previously. Furthermore, the normal curvature of γ is the normal curvature of p in the direction γ'.

Proposition 2.13. *Let $\gamma : I \to \mathbb{R}^3$ be a unit-speed curve on the surface $p : U \to \mathbb{R}^3$. Then*

$$\kappa_n = h[\gamma'] \,,$$

where κ_n is the normal curvature of γ and h is the second fundamental form of p.

Proof. Let $\beta : I \to U$ be a curve such that $\gamma = p \circ \beta$. Then $\gamma' = \beta^{i\prime} \partial_i p$. Moreover, if n is the unit normal vector field given by (2.3), then $\hat{n}' = \beta^{i\prime} \partial_i n$ with $\hat{n} := n \circ \beta$. Consequently,

$$h[\gamma'] = \beta^{i\prime} h_{ij} \beta^{j\prime} = -\beta^{i\prime}\, \partial_i p \cdot \partial_j n \, \beta^{j\prime} = -\gamma' \cdot \hat{n}' = \gamma'' \cdot \hat{n} = \kappa_n \,.$$

$\square$

Definition 2.14. A unit-speed curve $\gamma : I \to \mathbb{R}^3$ on the surface $p : U \to \mathbb{R}^3$ is called a *geodesic* if its geodesic curvature vanishes:

$$\gamma \text{ is a geodesic} \qquad \Longleftrightarrow \qquad \kappa_g = 0 \,.$$

Obviously, γ is a geodesic if, and only if, its acceleration γ'' is perpendicular to the surface.

2.7 Extrinsic versus intrinsic

Let us start by recalling the "distance-preserving" transformations in the Euclidean space. More specifically, a map $\phi : \mathbb{R}^d \to \mathbb{R}^d$ is called an *isometry* (or *congruence*) if it preserves the metric structure (distances) of the Euclidean space, *i.e.* $|\phi(x) - \phi(y)| = |x - y|$ for every $x, y \in \mathbb{R}^d$. Any isometry is a composition of an orthogonal motion and a translation. More specifically,

$$\phi(x) = Rx + x_0 \,,$$

where $x_0 \in \mathbb{R}^d$ and $R : \mathbb{R}^d \to \mathbb{R}^d$ is a linear transformation satisfying $|Rx| = |x|$ for every $x \in \mathbb{R}^d$. If $\det R = 1$ (respectively, $\det R = -1$), then ϕ is said to be *orientation-preserving* (respectively, *orientation-reversing*). An example of an orientation-preserving (respectively, orientation-reversing) isometry in $\mathbb{R}^2$ is the rotation $(x^1, x^2) \mapsto (-x^2, x^1)$ (respectively, the reflection $(x^1, x^2) \mapsto (x^1, -x^2)$).

Let $p : U \to \mathbb{R}^3$ be a surface and denote by κ_1, κ_2, K and H its principal, Gauss and mean curvatures, respectively. Given an isometry $\phi : \mathbb{R}^3 \to \mathbb{R}^3$, the composition $\tilde{p} := \phi \circ p$ is also a surface and we denote by $\tilde{\kappa}_1, \tilde{\kappa}_2$, $\tilde{K}$ and $\tilde{H}$ its principal, Gauss and mean curvatures, respectively. The relationship between the curvature functions of p and $\tilde{p}$ is given in the following proposition.

Proposition 2.15. *Let $p : U \to \mathbb{R}^3$ be a surface, $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ and isometry and $\tilde{p} := \phi \circ p$. Then*

$$\tilde{\kappa}_1 = \pm\kappa_1 \,, \qquad \tilde{\kappa}_2 = \pm\kappa_2 \,, \qquad \tilde{K} = K \,, \qquad \tilde{H} = \pm H \,,$$

where the signs are positive (respectively, negative) if ϕ is orientation-preserving (respectively, orientation-reversing).

Proof. By the chain rule, $D\tilde{p}(u) = D\phi(p(u))Dp(u) = RDp(u)$. (In components, writing $\phi^\alpha(x) = R^\alpha{}_\beta\, x^\beta + x_0^\beta$, this is equivalent to $\partial_i\tilde{p}^\alpha(u) = R^\alpha{}_\beta\, \partial_i p^\beta(u)$, where $i \in \{1,2\}$ and $\alpha, \beta \in \{1,2,3\}$.) In particular, $\partial_i\tilde{p} = R\partial_i p$. It follows that

$$\tilde{g}_{ij} := \partial_i\tilde{p} \cdot \partial_j\tilde{p} = (R\partial_i p) \cdot (R\partial_j p) = \partial_i p \cdot \partial_j p = g_{ij} \,,$$

where g is the metric of p. At the same time,

$$\partial_1\tilde{p} \times \partial_2\tilde{p} = (R\partial_1 p) \times (R\partial_2 p) = \det(R)\, (R^{-1})^\top (\partial_1 p \times \partial_2 p) \,, \tag{2.6}$$

where the second identity is valid for any invertible matrix (see Exercise 2.9.5). In our case, R is orthogonal (*i.e.*, $(R^{-1})^\top = R$), therefore

$$\tilde{n} := \frac{\partial_1\tilde{p} \times \partial_2\tilde{p}}{|\partial_1\tilde{p} \times \partial_2\tilde{p}|} = \det(R)\, Rn \,,$$

where n is given by (2.3). Consequently,

$$\tilde{h}_{ij} := -\partial_i\tilde{p} \cdot \partial_j\tilde{n} = -\det(R)\, (R\partial_i p) \cdot (R\partial_j n) = -\det(R)\, \partial_i p \cdot \partial_j n = \det(R)\, h_{ij} \,,$$

where h is the second fundamental form of p. Finally, we get

$$\tilde{L}^i{}_j := \tilde{g}^{ik}\tilde{h}_{kj} = \det(R)\, g^{ik}h_{kj} = \det(R)\, L^i{}_j \,,$$

where L is the Weingarten map of p. From this relationship between the Weingarten maps, we immediately deduce the relationships between the curvatures. □

We see that the Gauss curvature K is the only one of the curvature functions which does not change sign under the orientation-reversing isometries. For this reason, we say that K is an *isometric invariant* of p.

We now investigate a much deeper invariance of K. To simplify the notation, we denote the partial derivative of a function f by a comma (*i.e.* $f_{,i} := \partial_i f$, $f_{,ij} := \partial_j\partial_i f$, *etc.*). The range of indices being $i, j, \dots \in \{1,2\}$.

Definition 2.16. The *Christoffel symbols* of a surface $p : U \to \mathbb{R}^3$ are the functions

$$\Gamma^l_{ij} := \frac{1}{2}\, g^{kl} \left(-g_{ij,k} + g_{jk,i} + g_{ki,j}\right) \,, \tag{2.7}$$

where g is the metric of p.

The formula is easy to remember by writing $\Gamma_{ij}^{l} = g^{kl}\,\Gamma_{ijk}$ with $\Gamma_{ijk} := \frac{1}{2}(-g_{ij,k}+g_{jk,i}+g_{ki,j})$. Indeed, we start with minus $g_{ij,k}$ and then cyclically permute the indices with pluses. Notice the symmetry $\Gamma_{ij}^{l} = \Gamma_{ji}^{l}$.

Lemma 2.17. *Let $p : U \to \mathbb{R}^3$ be a surface. Then*

$$n_{,i} = -h_{ij}\, g^{jk}\, p_{,k} \qquad \text{and} \qquad p_{,ij} = \Gamma_{ij}^{k}\, p_{,k} + h_{ij}\, n\,,$$

where g and h are the metric and the second fundamental form of p, respectively, and n is the unit normal vector field given by (2.3).

Proof. Since the norm of n is constant, it follows that $n \perp n_{,i}$. Consequently, $n_{,i} = A_i{}^k p_{,k}$ with some coefficients $A_i{}^k : \mathbb{R} \to \mathbb{R}$ to be determined. Taking the inner product with $p_{,j}$, we get

$$-h_{ij} = n_{,i} \cdot p_{,j} = A_i{}^k p_{,k} \cdot p_{,j} = A_i{}^k g_{kj}\,.$$

Multiplying by g^{jl}, we get $-h_{ij} g^{jl} = A_i{}^k g_{kj} g^{jl} = A_i{}^l$. This proves the first identity.

To prove the second identity, we note that $(p_{,1}(u), p_{,2}(u), n(u))$ span $T_{p(u)}\mathbb{R}^3$. Consequently, $p_{,ij} = A_{ij}^{k}\, p_{,k} + B_{ij}\, n$ with some coefficients $A_{ij}^{k}, B_{ij} : \mathbb{R} \to \mathbb{R}$ to be determined. By taking the inner product with n, we see that $B_{ij} = h_{ij}$. By taking the inner product with $p_{,l}$, we get

$$p_{,ij} \cdot p_{,l} = A_{ij}^{k}\, g_{kl}\,.$$

Multiplying by g^{lm}, we get the symmetry $A_{ij}^{m} = p_{,ij} \cdot p_{,l}\, g^{lm} = A_{ji}^{m}$. Furthermore,

$$g_{ij,k} = (p_{,i} \cdot p_{,j})_{,k} = p_{,ik} \cdot p_{,j} + p_{,i} \cdot p_{,jk} = A_{ik}^{l}\, g_{lj} + A_{jk}^{l}\, g_{li}\,.$$

A cyclical permutation of the indices yields

$$\begin{aligned} g_{jk,i} &= A_{ji}^{l}\, g_{lk} + A_{ki}^{l}\, g_{lj}\,, \\ g_{ki,j} &= A_{kj}^{l}\, g_{li} + A_{ij}^{l}\, g_{lk}\,. \end{aligned}$$

Consequently (the respective colours cancel),

$$-g_{ij,k} + g_{jk,i} + g_{ki,j} = 2A_{ij}^{l}\, g_{lk}\,.$$

From this identity, it follows that $A_{ij}^{l} = \Gamma_{ij}^{l}$. ∎

Definition 2.18. The *curvature tensor* of a surface $p : U \to \mathbb{R}^3$ is the collection of functions

$$R_{ijkl} := R_{ijk}{}^{\rho}\, g_{\rho l} \qquad \text{with} \qquad R_{ijk}{}^{l} := \Gamma_{jk,i}^{l} - \Gamma_{ik,j}^{l} + \Gamma_{i\rho}^{l}\, \Gamma_{jk}^{\rho} - \Gamma_{j\rho}^{l}\, \Gamma_{ik}^{\rho}\,, \tag{2.8}$$

where g and Γ_{ij}^{k} are the metric and Christoffel symbols of p, respectively.

As a consequence of this definition and Lemma 2.17, we get the following relationship.

Lemma 2.19 (Gauss equation)**.** *Let $p : U \to \mathbb{R}^3$ be a surface. Then*

$$R_{ijkl} = h_{il} h_{jk} - h_{ik} h_{jl} \,,$$

where R_{ijkl} and h are the curvature tensor and the second fundamental form of p, respectively.

Proof. Using Lemma 2.17 (the second identity twice and then the first identity), we have

$$\begin{aligned} p_{,ijl} &= \Gamma^k_{ij,l}\, p_{,k} + \Gamma^k_{ij}\, p_{,kl} + h_{ij,l}\, n + h_{ij}\, n_{,l} \\ &= \Gamma^k_{ij,l}\, p_{,k} + \Gamma^k_{ij}\, (\Gamma^\rho_{kl}\, p_{,\rho} + h_{kl}\, n) + h_{ij,l}\, n + h_{ij}\, n_{,l} \\ &= \Gamma^k_{ij,l}\, p_{,k} + \Gamma^k_{ij}\, (\Gamma^\rho_{kl}\, p_{,\rho} + h_{kl}\, n) + h_{ij,l}\, n + h_{ij}\, (-h_{l\rho} g^{\rho k} p_{,k}) \,. \end{aligned}$$

By taking the inner product with $p_{,m}$, we get

$$p_{,ijl} \cdot p_{,m} = \left(\Gamma^k_{ij,l} + \Gamma^\rho_{ij}\, \Gamma^k_{\rho l} - h_{ij} h_{l\rho} g^{\rho k}\right) g_{km} \,.$$

Interchanging the indices i and l, we also have

$$p_{,lji} \cdot p_{,m} = \left(\Gamma^k_{lj,i} + \Gamma^\rho_{lj}\, \Gamma^k_{\rho i} - h_{lj} h_{i\rho} g^{\rho k}\right) g_{km} \,.$$

Subtracting these two identities, we deduce

$$R_{lij}{}^k = \Gamma^k_{ij,l} - \Gamma^k_{lj,i} + \Gamma^\rho_{ij}\, \Gamma^k_{\rho l} - \Gamma^\rho_{lj}\, \Gamma^k_{\rho i} = (h_{ij} h_{l\rho} - h_{lj} h_{i\rho}) g^{\rho k} \,,$$

which is equivalent to the desired claim. □

Now we are in a position to establish the main result of this section. (The translation from Latin would mean "totally remarkable theorem", which shows how much surprised Gauss was.)

Theorem 2.20 (Theorema Egregium [Gauss 1828])**.** *Let $p : U \to \mathbb{R}^3$ be a surface. Then*

$$K = \frac{R_{1221}}{\det(g_{ij})} = \frac{1}{2}\, R_{ijkl}\, g^{il} g^{jk} \,, \tag{2.9}$$

where K, R_{ijkl} and g are the Gauss curvature, the curvature tensor and the metric of p, respectively.

Proof. By Lemma 2.19 and the formula $K = \det(L) = \det(h_{ij}) \det(g^{kl})$, we have

$$R_{1221} = h_{11} h_{22} - h_{12} h_{21} = \det(h_{ij}) = K \det(g_{kl}) \,,$$

which yields the first equality of (2.9). The second equality follows by symmetries of the curvature tensor (see Exercise 2.9.6). Indeed,

$$\begin{aligned} R_{ijkl}\, g^{il} g^{jk} &= R_{1221}\, g^{11} g^{22} + R_{2112}\, g^{22} g^{11} + R_{1212}\, g^{12} g^{21} + R_{2121}\, g^{21} g^{12} \\ &= R_{1221}\, (g^{11} g^{22} + g^{22} g^{11} - g^{12} g^{21} - g^{21} g^{12}) \\ &= 2\, R_{1221}\, \det(g^{ij}) \,, \end{aligned}$$

which establishes the second equality of (2.9). □

The theorem says that the Gauss curvature can be computed merely from the knowledge of the metric and its derivatives. The metric is determined by tangent vectors only, so the Gauss curvature can be computed by a "resident on the surface", without knowledge of (or reference to) the shape of the surface in $\mathbb{R}^3$. That is why the Gauss curvature is an *intrinsic* quantity of the surface.

This is indeed remarkable, because the Gauss curvature was defined in terms of the second fundamental form, which is dependent on how the surface "sits in space". Indeed, by an orientation-reversing isometry, the second fundamental form and the other curvatures differ by a sign (recall Proposition 2.15 and its proof). That is why the mean and principal curvatures are *extrinsic* quantities of the surface.

2.8 Geodesics

Recall (Definition 2.14) that a unit-speed curve $\gamma = p : I \to \mathbb{R}^3$ on the surface $p : U \to \mathbb{R}^3$ is a geodesic if, and only if, its geodesic curvature κ_g vanishes. The geodesic curvature was defined via the unit normal vector field n, see (2.5). However, the following claim reveals that geodesics are actually intrinsic quantities of the surface.

Proposition 2.21 (Geodesic equations)**.** *Let $\gamma = p \circ \beta : I \to U \to \mathbb{R}^3$ be a unit-speed curve on the surface $p : U \to \mathbb{R}^3$. Then*

$$\gamma \text{ is a geodesic} \qquad \Longleftrightarrow \qquad \beta^{k''} + \Gamma^k_{ij}\beta^{i'}\beta^{j'} = 0\,.$$

Proof. Recalling the abuse of notation of writing $p_{,i}$ instead of $p_{,i} \circ \beta$, we have

$$\kappa_g = n \times \gamma' \cdot \gamma'' = n \times \beta^{i'} p_{,i} \cdot (\beta^{j'} p_{,j})' = n \times \beta^{i'} p_{,i} \cdot (\beta^{j''} p_{,j} + \beta^{j'}\beta^{k'} p_{,jk})\,.$$

Using Lemma 2.17, we get

$$\begin{aligned}
\kappa_g &= p_{,i} \times p_{,j} \cdot n\, \beta^{i'}\beta^{j''} + p_{,i} \times (\Gamma^l_{jk}\, p_{,l} + h_{jk} n) \cdot n\, \beta^{i'}\beta^{j'}\beta^{k'} \\
&= p_{,1} \times p_{,2} \cdot n\, (\beta^{1'}\beta^{2''} - \beta^{2'}\beta^{1''}) + p_{,1} \times p_{,2} \cdot n\, (\Gamma^2_{jk}\, \beta^{1'} - \Gamma^1_{jk}\, \beta^{2'})\beta^{j'}\beta^{k'} \\
&= |p_{,1} \times p_{,2}| \left[\beta^{1'}(\beta^{2''} + \Gamma^2_{jk}\, \beta^{j'}\beta^{k'}) - \beta^{2'}(\beta^{1''} + \Gamma^1_{jk}\, \beta^{j'}\beta^{k'})\right].
\end{aligned}$$

From this identity, it is evident that the geodesic equations imply $\kappa_g = 0$. Conversely, if $\kappa_g = 0$, then it follows that

$$(\beta^{1''} + \Gamma^1_{jk}\, \beta^{j'}\beta^{k'}, \beta^{2''} + \Gamma^2_{jk}\, \beta^{j'}\beta^{k'}) = c\, (\beta^{1'}, \beta^{2'})\,, \tag{2.10}$$

where $c : I \to \mathbb{R}$ is a function to be determined. The fact that γ is unit-speed is equivalent to $\beta^{i'} g_{ij} \beta^{j'} = 1$. Differentiating this identity, we obtain

$$2\beta^{i''} g_{ij}\beta^{j'} + g_{ij,k}\beta^{i'}\beta^{j'}\beta^{k'} = 0\,.$$

Taking the inner product (with respect to the metric g) of both sides of (2.10) with the vector (β^1, β^2), using the last identity and recalling Definition 2.16, we get

$$\begin{aligned}
c &= (\beta^{l''} + \Gamma^l_{jk}\, \beta^{j'}\beta^{k'})\, g_{lm}\, \beta^{m'} \\
&= -\frac{1}{2}\, g_{lm,k}\, \beta^{l'}\beta^{m'}\beta^{k'} + \frac{1}{2}\, (-g_{jk,m} + g_{km,j} + g_{mj,k})\, \beta^{j'}\beta^{k'}\beta^{m'} \\
&= 0\,.
\end{aligned}$$

Since c vanishes, (2.10) yields the geodesic equations. $\square$

The following observation demonstrates that geodesics represents a generalisation of straight lines in the Euclidean plane. (Recall that our definition of a straight line requires that it is a constant-speed curve.)

Proposition 2.22. *If the surface $p : U \to \mathbb{R}^3$ is planar, then geodesics are straight lines.*

Proof. Without loss of generality, we can take the parameterisation $p(u^1, u^2) := (u^1, u^2)$. Then $g_{ij} = \delta_{ij}$ and $\Gamma^k_{ij} = 0$ identically. Therefore, solutions of the geodescis equations are straight lines. $\square$

Example 2.23 (Geodesics on surfaces of revolution)**.** Recalling Examples 2.5 and 2.7, we find that all the Christoffel symbols vanish except for

$$\Gamma^1_{22} = -rr' \qquad \text{a} \qquad \Gamma^2_{12} = \Gamma^2_{21} = \frac{r'}{r} \,.$$

Writing $(\beta^1(t), \beta^2(t)) =: (s(t), \varphi(t))$, the geodesic equations read

$$s''(t) - r(s(t))\, r'(s(t))\, \varphi'(t)^2 = 0 \qquad \text{a} \qquad \varphi''(t) + 2\,\frac{r'(s(t))}{r(s(t))}\, s'(t)\, \varphi'(t) = 0 \,. \tag{2.11}$$

For the second differential equation, we immediately get the general solution

$$\varphi'(t) = \frac{c_1}{r(s(t))^2}$$

where $c_1 \in \mathbb{R}$. Let us focus on geodesics starting at a point $(s(0), \varphi(0)) = (s_0, \varphi_0) \in (0, \infty) \times (0, 2\pi)$ with velocity $(s'(0), \varphi'(0)) = (\pm 1, 0)$. The latter requires that the initial direction is within the meridian $\varphi(t) = \varphi_0$, while $s'(0) = \pm 1$ is necessary to ensure that the geodesic is unit-speed at this point. Then $c_1 = 0$, so $\varphi' = 0$ identically. Plugging this result to the first equation of (2.11), we get $s''(t) = 0$, which gives the general solution

$$s(t) = c_2 t + c_3 \,,$$

where $c_2, c_3 \in \mathbb{R}$. The initial conditions yield $c_3 = s_0$ and $c_2 = \pm 1$. In summary, the searched geodesic is the meridian curve $\varphi(t) = \varphi_0$ parameterised by arc-length (see the blue curves emanating from the pole in Figure 2.7).

For the particular examples of the plane and the sphere, any point can be considered as the pole with respect to which it is the surface of revolution. For the plane, we therefore reproduce the result of Proposition 2.22 that the geodesics are straight lines. For the sphere, the geodesics are the great circles (for instance, the equator), see Figure 2.8. $\diamondsuit$

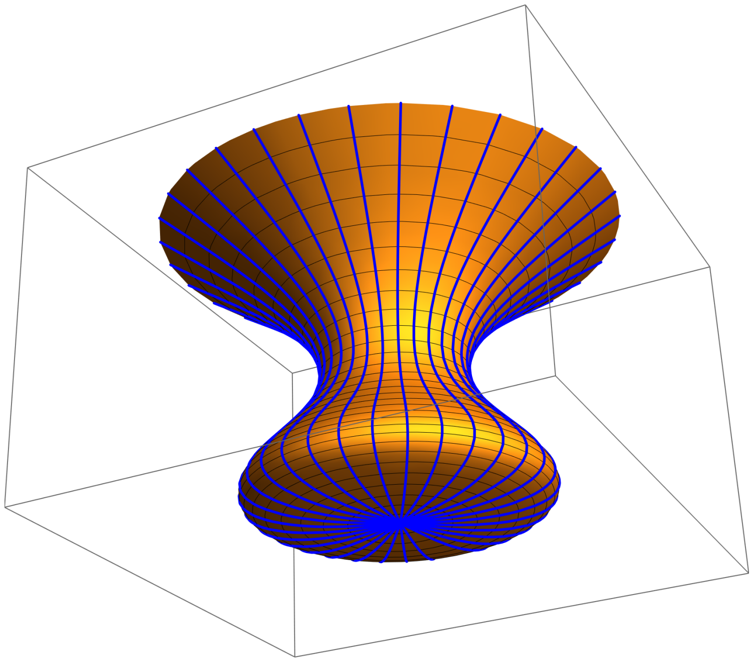


Figure 2.7: Geodesics emanating from the pole of a surface of revolution.

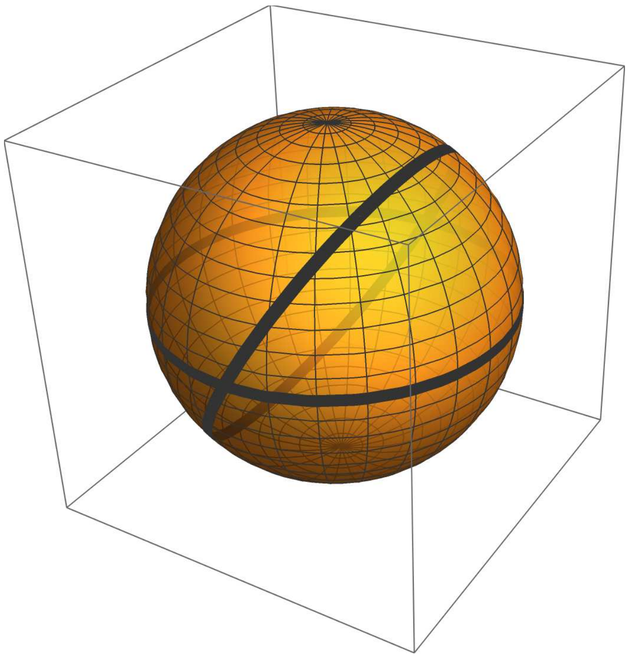


Figure 2.8: Thick blue curves as examples of great circles on the sphere.

2.9 Exercises

1. Compute the curvature and torsion of the meridian from Figure 2.2.
2. Show that the parameterisations $p_\pm$ and p of Example 2.2 are (regular) surfaces. Compute their fundamental forms and curvatures.
3. Show that the parameterisation p of Example 2.3 is a (regular) surface. Compute its fundamental forms and curvatures.
4. Find a sufficient condition (involving $r_{\max}$) which guarantees that p of Example 2.4 is

a (regular) surface. Compute its fundamental forms and curvatures.

5. Prove the second equality of (2.6). More generally, given any $u, v \in \mathbb{R}^3$ and $A \in \mathbb{R}^{3\times 3}$ show that
$$(Au) \times (Av) = \operatorname{cof}(A)\,(u \times v)\,,$$
where $\operatorname{cof}(A)$ denotes the *cofactor* matrix of A.

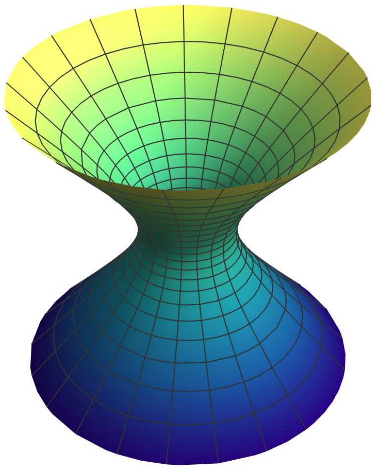
   Hint: Establish first the following facts.

   (a) $(u \times v)_i = \varepsilon_{ijk}\, u^j v^k$.
   (b) $(\operatorname{cof} A)_i^{\ j} = \frac{1}{2}\,\varepsilon_{iab}\,\varepsilon^{jcd}\, A^a{}_c\, A^b{}_d\,.$
   (c) $\varepsilon_{iab}\,\varepsilon^{icd} = \delta_a^c \delta_b^d - \delta_a^d \delta_b^c$.

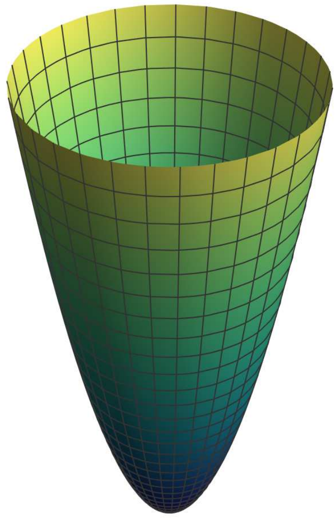
   Here
$$\varepsilon_{ijk} := \begin{cases} +1 & \text{if} \quad (ijk) \text{ is an even permutation of } (1,2,3)\,, \\ -1 & \text{if} \quad (ijk) \text{ is an even permutation of } (1,2,3)\,, \\ 0 & \text{if} \quad \text{any two indices are equal.} \end{cases}$$
is the *Levi–Civita symbol*

6. Use Lemma 2.19 to deduce the following symmetries of the curvature tensor:
$$\begin{aligned} R_{ijkl} &= -R_{jikl} && \text{(interchange of the first two indices)}\,, \\ R_{ijkl} &= -R_{ijlk} && \text{(interchange of the last two indices)}\,, \\ R_{ijkl} &= R_{klij} && \text{(interchange of the first and second pair of indices)}\,, \\ R_{ijkl} + R_{jkil} + R_{kijl} &= 0 && \text{(cyclic permutation of the first three indices)}\,. \end{aligned}$$
The first three symmetries imply that the curvature tensor is totally determined by $R_{1221} = R_{2112} = -R_{1212} = -R_{2121}$; all the other coefficients are equal to zero. The last symmetry is called the (*first* or *algebraic*) *Bianchi identity.*

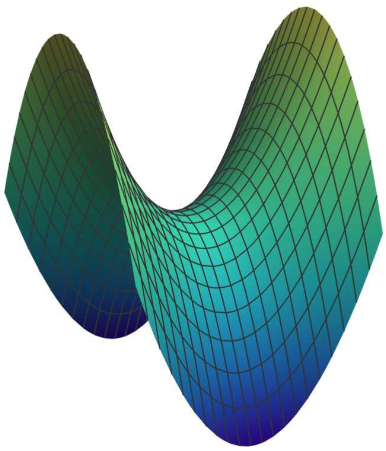
7. Compute fundamental forms and curvatures of the surfaces from Figure 2.9.



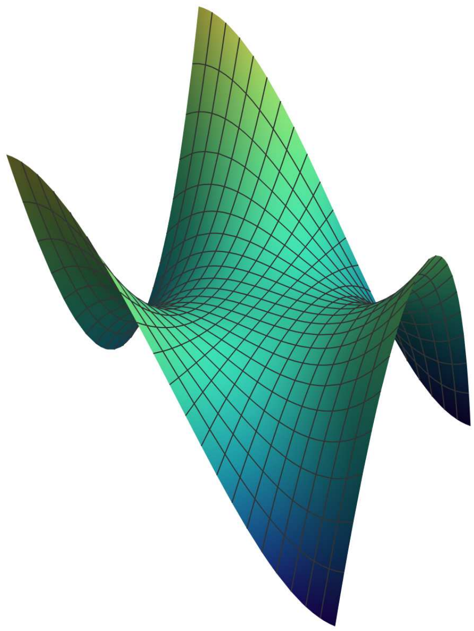
hyperboloid of one sheet
$x^2 + y^2 - z^2 = 1$



paraboloid
$z = x^2 + y^2$



hyperbolic paraboloid
$z = y^2 - x^2$



monkey saddle
$z = x^3 - 3xy^2$

Figure 2.9: Examples of the most famous surfaces.

3 Manifolds

Our Definition 2.1 of (parameterised) surfaces has several defects.

- First, the restriction to two-dimensional surfaces is not really important. It is easy to extend Definition 2.1 to "surfaces" of **arbitrary dimensions and codimensions** $p : U \subset \mathbb{R}^n \to \mathbb{R}^d$, with any $2 \leq n \leq d$, just by imposing the requirement $\operatorname{rank} p' = n$.
- Second, more importantly, it is possible to extend the definition to a surface being a collection of such parameterisations (covering the sphere for example). In this way, it is possible to go **from local to global** properties of a surface.
- The major defect of Definition 2.1, however, is the dependence on the ambient space $\mathbb{R}^3$ (or $\mathbb{R}^d$ in the more general setting above). The main idea of manifolds is to **stay intrinsic**.

Quoting [3, § 0.1], although the necessity of an abstract idea of surface is clear since Gauss (1777–1855), it was nearly a century before such an idea attained the definitive form that we present here.

3.1 Topological manifolds

Intuitively, a *manifold* is a metric space M which "locally looks" like the Euclidean space. This is formalised by the following requirement:

$$\forall p \in M, \quad \exists \text{ neighbourhood } U \subset M \quad \& \quad n \in \mathbb{N}, \qquad U \text{ is homeomorphic } \mathbb{R}^n. \tag{3.1}$$

Homeomorphic means that there exists a bijective map $\phi : U \to \mathbb{R}^n$ which is continuous and whose inverse ϕ^{-1} is also continuous, see Figure 3.1.

By definition, a *neighbourhood* of a point $p \in M$ is a subset $U \subset M$ which contains an open set containing p. From properties of bijective continuous maps, it follows that the neighbourhood in the definition above must necessarily be open. Instead of considering homeomorphisms from open subsets of M onto $\mathbb{R}^n$, it suffices to consider homeomorphisms onto open subsets of $\mathbb{R}^n$.

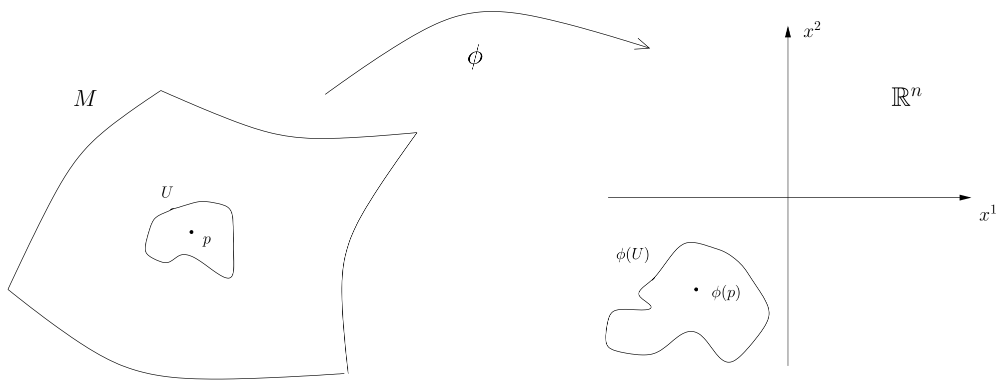


Figure 3.1: A manifold is an object which locally looks like the Euclidean space.

It is not difficult to see that the natural number n in our definition is unique. We denote $\dim M =: n$ and say that M is n-dimensional (metric) manifold. If we want to indicate that M has dimension n, we also write M^n.

In Definition 2.1, the role of the homeomorphism ϕ is played by the inverse of the parameterisation, *i.e.* $\phi = p^{-1}$. This leads to some reverse of notation, namely p now stands for a point on the manifold (*i.e.* on the image of the surface $p : U \to \mathbb{R}^3$ before). Moreover, now U and $\phi(U)$ are the ancient $p(U)$ and U, respectively. A point in the Euclidean space $\mathbb{R}^n$ is typically denoted by $x = (x^1, \dots, x^n)$ (instead of $u = (u^1, u^2)$ before).

The simplest example of a manifold is the Euclidean space $\mathbb{R}^n$ (or any open subset of it). The only connected 1-dimensional manifolds are the real line $\mathbb{R}$ (or any open interval of it) and the circle $\mathbb{S}^1 := \{x \in \mathbb{R}^2 : |x| = 1\}$. Another canonical example is the n-dimensional version of the sphere of Example 2.2.

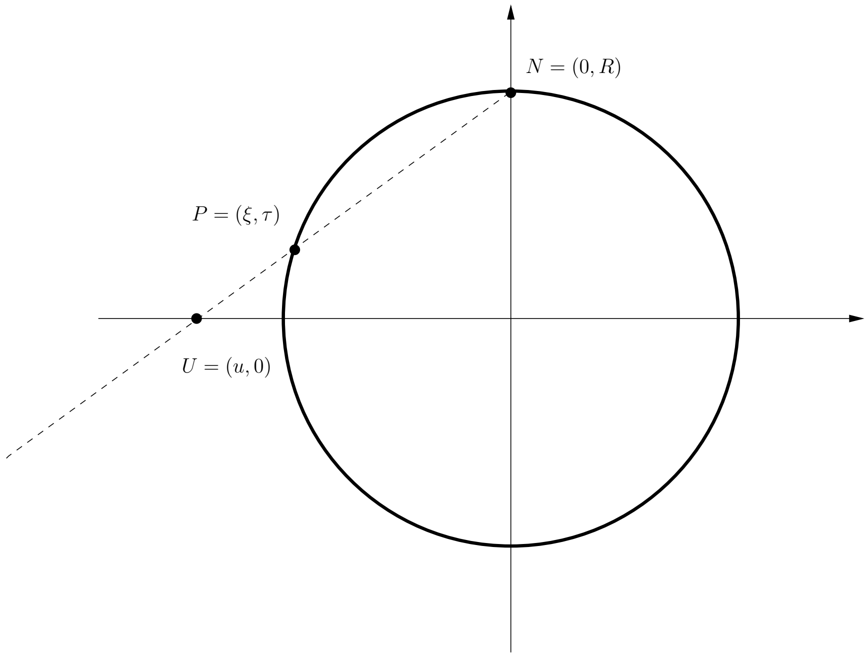


Figure 3.2: Stereographic projection.

Example 3.1 (n-sphere)**.** By the *n-dimensional sphere* (or simply *n-sphere*) of radius $R > 0$ we mean the following subset of the Euclidean space $\mathbb{R}^{n+1}$:

$$\mathbb{S}^n_R := \left\{(x^1, \dots, x^{n+1}) \in \mathbb{R}^{n+1} : \ (x^1)^2 + \dots + (x^{n+1})^2 = R^2\right\} .$$

Using n-dimensional generalisations of the parameterisations of Example 2.2 (*i.e.*, either the Cartesian or hyperspherical coordinates), it is possible to cover the n-sphere by several charts. A more elegant way, which leads to merely two charts, is the *stereographic projection*

$$\sigma : \mathbb{S}^n_R \setminus \{N\} \to \mathbb{R}^n : \{P \mapsto u\} ,$$

where $N := (0, \dots, 0, R)$ is the north pole (the second chart would be obtained by using the south pole instead) and $u := (u^1, \dots, u^n)$ is the point in the hyperplane $\{x^{n+1} = 0\}$ that we obtain as the intersection of this hyperplane with the line connecting the points N and P, see Figure 3.2. Elementary geometric considerations lead to the formula

$$\sigma(P) = \frac{R\xi}{R - \tau} , \qquad \text{where} \qquad P = (\xi, \tau) \in \mathbb{R}^n \times \mathbb{R} ,$$

and consequently

$$\sigma^{-1}(u) = \left(\frac{2R^2 u}{|u|^2 + R^2}, R\,\frac{|u|^2 - R^2}{|u|^2 + R^2} \right) .$$

The latter represents a parameterisation of $\mathbb{S}^n_R$ except for the north pole.

If $R = 1$, we abbreviate $\mathbb{S}^n_1 =: \mathbb{S}^n$. To be consistent with Example 2.2, we also write $\mathbb{S}^2_R =: \mathbb{S}_R$ if $n = 2$. $\diamondsuit$

More generally, in our definition of manifold above, it is possible to replace "metric" by "topological". In this more general setting, it is necessary to exclude some really pathological creatures and to this purpose to assume two extra properties:

1. M is *Hausdorff* (every two distinct points have disjoint neighbourhoods);
2. M is *completely separable* (its topology has a countable basis).

The two properties are not "local properties". More specifically, although they are both preserved by homeomorphisms (and the Euclidean space $\mathbb{R}^n$ satisfies both), the local condition (3.1) does not imply that M is Hausdorff and completely separable.

By definition of a topological manifold, there exists a family $\{\phi_i\}_{i\in J}$ of homeomorphisms $\phi_i : U_i \subset M \to \mathbb{R}^n$ covering M, *i.e.*,

$$\bigcup_{i\in J} U_i \supset M\,,$$

where J is an index set (possibly uncountable). The topological hypotheses above guarantee that the index set can actually be chosen to be countable. Moreover, we can assume that the covering is *locally finite*, *i.e.*, for every $p \in M$ there exists a neighbourhood W such that $W \cap U_i \neq \varnothing$ for only a finite number of indices i from the index set J.

3.2 Differentiable manifolds

On a general (metric or topological) manifold, the notion of a continuous function $f : M \to \mathbb{R}$ makes sense, but the differentiability does not. (Recall that "smooth" or "differentiable" always means C^∞-smooth.) Indeed, consider two homeomorphisms $\phi : U \to \mathbb{R}^n$ and $\psi : V \to \mathbb{R}^n$ such that $U \cap V \neq \varnothing$. Then the function $f \circ \phi^{-1} : \mathbb{R}^n \to \mathbb{R}$ can be differentiable, while it is not necessarily true that $f \circ \psi^{-1}$ is also differentiable. Indeed, this is clear from writing

$$f \circ \psi^{-1} = f \circ \phi^{-1} \circ (\phi \circ \psi^{-1})$$

and choosing $\phi := h \circ \psi$, where $h : \mathbb{R}^n \to \mathbb{R}^n$ is a homeomorphism that is not differentiable. To develop differential calculus on manifolds, it is necessary to adorn our manifolds with an additional structure.

An *atlas* (or *differentiable structure*) $\mathcal{A}$ for a manifold M is a family of homeomorphisms whose domains cover M and any pair of elements $\phi, \psi \in \mathcal{A}$ are *smoothly compatible* in the sense that the *transition* maps

$$\left.\begin{aligned} \psi \circ \phi^{-1} &: \ \phi(U \cap V) \to \psi(U \cap V) \\ \phi \circ \psi^{-1} &: \ \psi(U \cap V) \to \phi(U \cap V) \end{aligned}\right\} \quad \text{are } C^\infty. \tag{3.2}$$

A particular member $\phi \in \mathcal{A}$ with domain U ($\ni p$) is called a *chart* (or *coordinate system* on U) for M (at point p). If we want to indicate the domain $U \ni p$ (called the *coordinate neighbourhood* of p), we also write (ϕ, U) for the chart. A customary (but occasionally confusing) notation for the individual chart is the letter x, in order to identify the point $p \in M$ with the point $x(p) \in \mathbb{R}^n$, which has "coordinates" $(x^1(p), \dots, x^n(p))$.

It remains to think a little bit about the (non-)uniqueness of the atlas choice: In general, we have many choices of different atlases, which lead to the same differentiable structure for M. However, every atlas for the manifold M is contained in a uniquely determined *maximal* atlas (which is not contained in any strictly larger smooth atlas).

A *smooth* (or *differentiable*) manifold is a topological manifold endowed with a maximal atlas. Schematically,

$$\text{differentiable manifold} \;:=\; \text{manifold} + \text{maximal atlas.}$$

(The adjectives "smooth" or "differentiable" will typically be omitted.)

3.3 Functions

Let M, N be two manifolds, $\dim M = n$ and $\dim N = m$. A function $F : M \to N$ is called *smooth* (or *differentiable*) at point $p \in M$ if for every coordinate systems (x, U) for M at $p \in M$ and (y, V) for N at $F(p) \in N$ the function

$$y \circ F \circ x^{-1} : \; \mathbb{R}^n \to \mathbb{R}^m$$

is differentiable at $x(p) \in \mathbb{R}^n$, see Figure 3.3. The fact that this definition is good, *i.e.* independent of the choice of charts, follows from the condition of compatibility (3.2). Note that the function F is necessarily continuous.

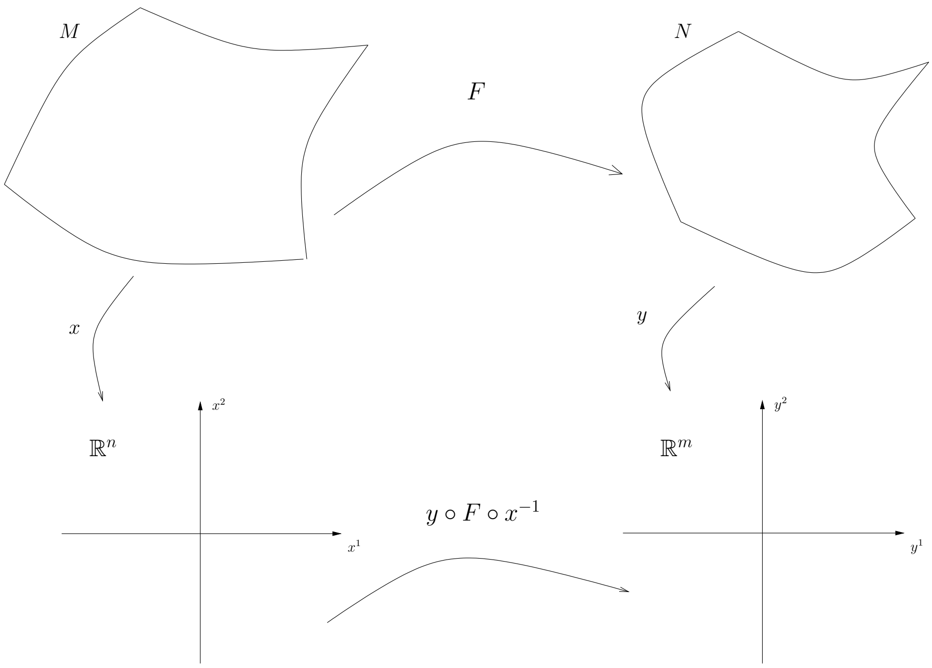


Figure 3.3: A function between manifolds and their charts.

In the case that there is a smooth function $F : M \to N$ which is a diffeomorphism, we say that the manifolds M, N are *diffeomorphic*.

Given a smooth function $f : M \to \mathbb{R}$ and a coordinate system (x, U), we define

$$\frac{\partial f}{\partial x^i} := \partial_i (f \circ x^{-1}) \circ x, \tag{3.3}$$

where $\partial_i g$ denotes the usual partial derivative of $g : \mathbb{R}^n \to \mathbb{R}$ with respect to the i-th variable. The expression (3.3) defines a function from $U \subset M$ to $\mathbb{R}$. In an arbitrary point $p \in M$, we have

$$\left.\frac{\partial f}{\partial x^i}\right|_p := \frac{\partial f}{\partial x^i}(p) = \partial_i (f \circ x^{-1})(x(p)) \,. \tag{3.4}$$

The function $\hat{f} := f \circ x^{-1} : x(U) \subset \mathbb{R}^n \to \mathbb{R}$ is called the *coordinate representation* of the function f. (A confusion begins when the two functions f and $\hat{f}$ are identified, which is a very common practice of differential geometry.)

Now, consider two coordinate systems (x, U) and (y, V) for the same manifold M. On the intersection $U \cap V$, we have an analogy of the classical *chain rule*

$$\frac{\partial f}{\partial y^i} = \frac{\partial x^j}{\partial y^i} \frac{\partial f}{\partial x^j} \,. \tag{3.5}$$

Here we continue to use the Einstein summation convention, the range of indices being $i, j = 1, \dots, n$.

We have

$$(x \circ y^{-1})'(y(p)) = \left(\frac{\partial x^j}{\partial y^i}(p)\right)_{j,i=1,\dots,n} = \begin{pmatrix} \frac{\partial x^1}{\partial y^1}(p) & \dots & \frac{\partial x^1}{\partial y^n}(p) \\ \vdots & & \vdots \\ \frac{\partial x^n}{\partial y^1}(p) & \dots & \frac{\partial x^n}{\partial y^n}(p) \end{pmatrix},$$

where on the left-hand side there is the Jacobi matrix of the mapping $x \circ y^{-1}$ at point $y(p)$ and on the right-hand side there is the $n \times n$ matrix of formula (3.5). It is a regular matrix and one has

$$[(x \circ y^{-1})'(y(p))]^{-1} = \left(\frac{\partial x^j}{\partial y^i}(p)\right)^{-1} = \left(\frac{\partial y^i}{\partial x^j}(p)\right) = (y \circ x^{-1})'(x(p)).$$

3.4 Immersions and embeddings

More generally, consider a smooth function $F : M^n \to N^m$, a coordinate system (x, U) for M^n at $p \in M^n$ and a coordinate system (y, V) for N^m at $F(p) \in N^m$. We define the *rank* of the function F at the point $p \in U$ to be the rank of the $m \times n$ Jacobi matrix of the mapping $y \circ F \circ x^{-1} : \mathbb{R}^n \to \mathbb{R}^m$ at the point $x(p)$:

$$\operatorname{rank} F(p) := \operatorname{rank} (y \circ F \circ x^{-1})'(x(p)) \,. \tag{3.6}$$

The definition does not depend on the choice of the coordinate systems. Recalling definition (3.3), we get the alternative expression

$$\begin{aligned} \operatorname{rank} F(p) &= \operatorname{rank} \left(\frac{\partial (y^j \circ F)}{\partial x^i}(p)\right)_{(j,i)\in\{1,\dots,m\}\times\{1,\dots,n\}} \\ &= \operatorname{rank} \begin{pmatrix} \frac{\partial (y^1 \circ F)}{\partial x^1}(p) & \dots & \frac{\partial (y^1 \circ F)}{\partial x^n}(p) \\ \vdots & & \vdots \\ \frac{\partial (y^m \circ F)}{\partial x^1}(p) & \dots & \frac{\partial (y^m \circ F)}{\partial x^n}(p) \end{pmatrix}. \end{aligned}$$

A point $p \in M$ is called a *critical point* of the function F (and $F(p)$ is a *critical value* of F) if $\operatorname{rank} F(p) < m$; otherwise p is called a *regular* point of F (and $F(p)$ is a *regular value* of F). If M has at most countably many components then (according to Sard's theorem) the critical values of F form a set o measure zero in N.

The most interesting situation is the setting $n \leq m$, when the following definition makes sense:

$$F : M^n \to N^m \quad \text{is an } \textit{immersion} \qquad :\Longleftrightarrow \qquad \operatorname{rank} F = n\,.$$

Since $n = \dim M^n = \dim\ker F(p) + \operatorname{rank} F(p)$, where $\dim\ker F(p)$ is defined in analogy of the rank in (3.6) (more specifically, as the dimension of the kernel of the derivative $D(y \circ F \circ x^{-1})$ at $x(p)$), we see that $F : M^n \to N^m$ is an immersion if, and only if, the derivative $D(y \circ F \circ x^{-1})(x(p)) : \mathbb{R}^n \to \mathbb{R}^m$ is injective (or, equivalently, the Jacobi matrix $(y \circ F \circ x^{-1})'(x(p)$ induced an injective mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$). By the inverse function theorem, we get that the function F is a local diffeomorphism.

Various scenarios for curve immersions can be seen in Figures 1.3, 1.4 and 1.5. (Recall that $\gamma : I \to \mathbb{R}^d$ with I open is an immersion if, and only if, the curve is regular.)

A stronger property is the following definition:

$$F : M^n \to N^m \text{ is an } \textit{embedding} \qquad :\Longleftrightarrow \qquad \begin{cases} F \text{ is an injective immersion}\,, \\ F : M \to F(N) \text{ is a homeomorphism}\,. \end{cases}$$

Here the second property means that F is a topological embedding (the homeomorphism is understood with respect to the subspace topology). (Occasionally [5], by an embedding it is merely understood an injective immersion and our property is called a *proper embedding.* Later, we shall see that our terminology is more consistent.)

The curve γ_1 from Figure 1.3 is an embedding, while the curves γ_{-1} and γ_0 from Figure 1.4 and the curves from Figure 1.5 are not embedded.

A subset $M \subset N$ is called an *immersed submanifold* if the *inclusion map* $\iota : M \to N : \{p \mapsto p\}$ is an immersion. Here it is important that M can be equipped with a topology and a differential structure which differ from the topology and differential structure induced by the superset N.

If the inclusion map ι is in addition an embedding, then $M \subset N$ is called an *embedded submanifold* or simply *submanifold.*

Immersed submanifolds are counter-intuitive from several respects, but they are important for certain application (*e.g.*, for considering self-intersecting curves). They usually arise in the following way. Let I, N be smooth manifolds and let us consider the injective immersion $f : I \to N$ (*e.g.*, the curve $\gamma_0 : I_0 \to \mathbb{R}^2$ from Figure 1.4). On the image $M := f(I) \subset N$ we define the topology by declaring that $U \subset M$ is open if, and only if, $f^{-1}(U) \subset I$ is open. With this topology, M is clearly a topological submanifold homeomorphic to I. Moreover, there is a unique smooth structure on it such that $f : I \to M$ is diffeomorphism (the smooth charts are just the maps of the form $\varphi \circ f^{-1}$, where φ is a smooth coordinate chart of I). With this topology and smooth structure, $\iota : M \to N$ is clearly a smooth immersion, because it can be realised as a composition of a diffeomorphism and an immersion:

$$\iota : M \xrightarrow{f^{-1}} I \xrightarrow{f} N\,.$$

In summary, M is an immersed submanifold of N.

If M is an immersed submanifold, then it follows from the procedure above that the inclusion map $\iota : M \to N$ is an injective immersion. Immersed submanifolds can therefore be characterised as images of injective immersions. It is evident that every embedded submanifold is automatically an immersed submanifold. The converse is of course false: An immersed submanifold M is embedded if, and only if, its topology is induced by the superset N.

In a certain sense, it is always possible to restrict ourselves to embedded submanifolds in a Euclidean space, because (by the Whitney embedding theorem) any smooth *connected* manifold M can be embedded to $\mathbb{R}^m$ with a sufficiently large m.

3.5 Vectors and the tangent space

Next we would like to define the notion of tangent vectors on manifolds. For this purpose, we are inspired by the surfaces of Section 2. There, given a surface $p : U \to \mathbb{R}^3$, we can understand the tangent vector at a point $p \cong p(u) \in p(U) =: M$ as the velocity γ' of a curve $\gamma := p \circ \beta : I \to U \to \mathbb{R}^3$ passing by the point, see Figure 3.4. Since we do not have at our disposal the support of the ambient space (to define the velocity of the curve $\gamma : I \to M$ as a derivative), we have to find a characteristic property of the tangent vector which will substitute for the idea of velocity.

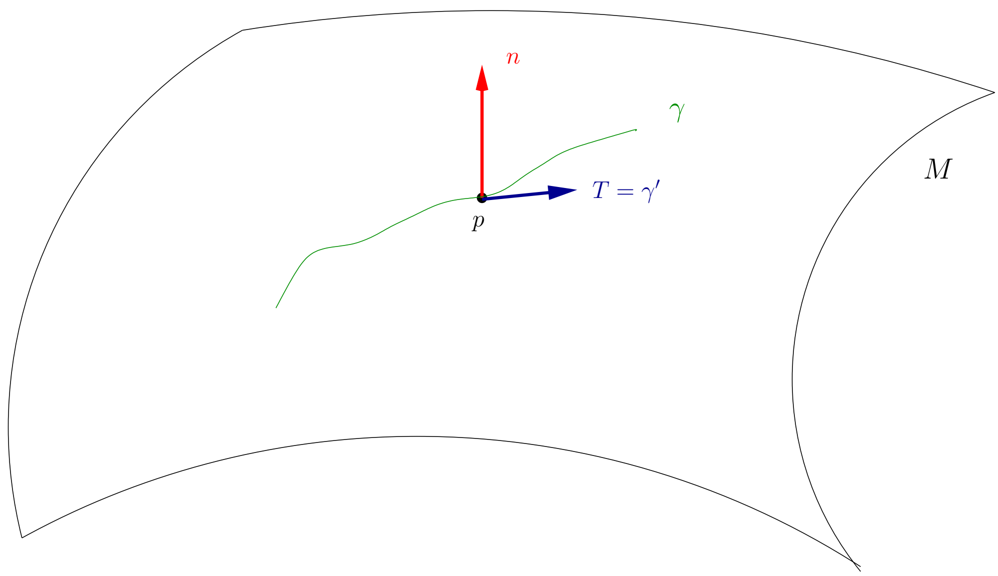


Figure 3.4: The tangent vector as the velocity of a curve on a surface embedded in $\mathbb{R}^3$.

Let $\gamma : (-\varepsilon, \varepsilon) \to M \subset \mathbb{R}^3$ with $\varepsilon > 0$ be a curve such that $\gamma(0) = p$. Let us write $T := \gamma'(0) = ({\gamma^1}'(0), {\gamma^2}'(0), {\gamma^3}'(0))$. Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a smooth function defined in a neighbourhood of the point p. Considering the restriction of the function on the curve, the directional derivative with respect to the vector T reads

$$(f \circ \gamma)'(0) = (\partial_k f)(p)\, {\gamma^k}'(0) = (T^k\, \partial_k)\big|_p f = T \cdot \nabla\big|_p f\,.$$

Therefore, the directional derivative with respect to T is an operator on functions that depends uniquely on T. This is the characteristic property that we are going to use to define tangent vectors on manifolds.

Definition 3.2. Let $\gamma : (-\varepsilon, \varepsilon) \to M$ with $\varepsilon > 0$ be a smooth curve such that $\gamma(0) = p$. The *tangent vector to the curve* γ at 0 is the function

$$\gamma'(0) : C^\infty(M) \to \mathbb{R} : \{f \mapsto (f \circ \gamma)'(0)\} .$$

A *tangent vector* at $p \in M$ is the tangent vector to some smooth curve $\gamma : (-\varepsilon, \varepsilon) \to M$ with $\varepsilon > 0$ at 0 satisfying $\gamma(0) = p$. The set of all tangent vectors to M at p is called the *tangent space* of M at p and denoted by T_pM.

Let us choose a coordinate system (x, U) at p and write $(x \circ \gamma)(t) =: (\gamma^1(t), \dots, \gamma^n(t))$. Using definition (3.3), one has

$$\gamma'(0)f = (f \circ x^{-1} \circ x \circ \gamma)'(0) = \frac{\partial f}{\partial x^k}(p)\, \gamma^{k\prime}(0) = \gamma^{k\prime}(0) \left.\frac{\partial}{\partial x^k}\right|_p f .$$

Omitting the function f, we get a coordinate expression of the tangent vector $\gamma'(0)$:

$$\gamma'(0) = \gamma^{k\prime}(0) \left.\frac{\partial}{\partial x^k}\right|_p . \tag{3.7}$$

From this, we see that $\left.\frac{\partial}{\partial x^k}\right|_p$ is the tangent vector at 0 to the “coordinate curve”

$$t \mapsto x^{-1}(0, \dots, t, \dots, 0) ,$$

where t stands on the k-th place.

From the formula (3.7), we see that the tangent vector to the curve γ at 0 depends only on the derivative of γ in a coordinate system. Using the usual operations on functions, it follows that T_pM is a vector space. Since the vectors

$$\left.\frac{\partial}{\partial x^1}\right|_p , \dots , \left.\frac{\partial}{\partial x^n}\right|_p$$

form a (*coordinate*) basis of the tangent space T_pM, we see that $\dim T_pM = \dim M = n$. This linear structure of T_pM does not depend on the choice of the coordinate system.

From the formula (3.7), we also see that T_pM can be identified with the set of n-tuples $(\xi^1, \dots, \xi^n)$ with respect to a coordinate system $(x^1, \dots, x^n)$, which transform according to the formula

$$\tilde{\xi}^i = \left.\frac{\partial \tilde{x}^i}{\partial x^j}\right|_p \xi^j$$

when passing to another coordinate system $(\tilde{x}^1, \dots, \tilde{x}^n)$, *cf* (3.5).

A linear map $X : C^\infty(M) \to \mathbb{R}$ is called a *derivation* at $p \in M$ if (in addition to the linearity) it satisfies the *Leibniz rule*

$$\forall f, g \in C^\infty(M) , \qquad X(fg) = f(p)X(g) + g(p)X(f) . \tag{3.8}$$

The tangent space T_pM can be identified with the set of all derivations at point $p \in M$.

3.6 Push-forwards

Let M, N be two smooth manifolds and consider a smooth function $F : M \to N$. For an arbitrary point $p \in M$, let us define the *push-forward* $F_{*p} : T_pM \to T_{F(p)}N$ by

$$\forall X \in T_pM,\ f \in C^\infty(N)\,, \qquad (F_{*p}\,X)(f) := X(f \circ F)\,. \tag{3.9}$$

The transformation F_{*p} is clearly linear and it is a derivation at $F(p)$ in the sense of definition (3.8). If F is diffeomorphism, then F_{*p} is an isomorphism. If P is another smooth manifold and $G : N \to P$ is another smooth function, then

$$(G \circ F)_{*p} = G_{*F(p)} \circ F_{*p} : T_pM \to T_{(G\circ F)(p)}P\,.$$

Let us explore how push-forwards look in coordinates. Let (x, U) be a coordinate system at $p \in M^n$ and (y, V) be a coordinate system at $F(p) \in N^m$. Denote

$$\begin{aligned}\hat{f} &:= f \circ y^{-1} : y(V) \to \mathbb{R}\,,\\ \hat{F} &:= y \circ F \circ x^{-1} : x(U \cap F^{-1}(V)) \to y(V)\,,\end{aligned}$$

coordinate representations of the functions f and F, respectively. Then

$$\begin{aligned}\left(F_{*p}\,\frac{\partial}{\partial x^i}\bigg|_p\right)(f) &= \frac{\partial}{\partial x^i}\bigg|_p (f \circ F)\\ &= \partial_i(f \circ F \circ x^{-1})(x(p))\\ &= \partial_i(\hat{f} \circ \hat{F})(x(p))\\ &= (\partial_j \hat{f})(\hat{F}(p))\,(\partial_i \hat{F}^j)(x(p))\\ &= (\partial_i \hat{F}^j)(x(p))\,\frac{\partial}{\partial y^j}\bigg|_{F(p)} (f)\,.\end{aligned}$$

Thus, omitting the auxiliary function f,

$$F_{*p}\,\frac{\partial}{\partial x^i}\bigg|_p = (\partial_i \hat{F}^j)(x(p))\,\frac{\partial}{\partial y^j}\bigg|_{F(p)}\,,$$

from which we see that the matrix of the map F_{*p} with respect to the coordinate bases $\left(\frac{\partial}{\partial x^1}\big|_p, \dots, \frac{\partial}{\partial x^n}\big|_p\right)$ and $\left(\frac{\partial}{\partial y^1}\big|_{F(p)}, \dots, \frac{\partial}{\partial y^m}\big|_{F(p)}\right)$ of tangent spaces T_pM and $T_{F(p)}N$, respectively, is the Jacobi matrix

$$\mathsf{M}(F_*) = \begin{pmatrix} \partial_1 \hat{F}^1 & \dots & \partial_n \hat{F}^1\\ \vdots & & \vdots\\ \partial_1 \hat{F}^m & \dots & \partial_n \hat{F}^m \end{pmatrix}(x(p)) = \hat{F}'(x(p)) = \left(\frac{\partial(y^j \circ F)}{\partial x^i}(p)\right)_{(j,i)\in\{1,\dots,m\}\times\{1,\dots,n\}}.$$

(This is a reason why the push-forward F_{*p} is occasionally called the differential of the function F at p and denoted by $\mathrm{d}F_p$.)

3.7 Curves on manifolds

Let us look more closely on curves on manifolds and establish a relationship between their tangent vectors (Section 3.5) and push-forwards (Section 3.6).

Let γ be a smooth *curve* on a manifold M, *i.e.* smooth mapping $\gamma : I \to M$, where $I \subset \mathbb{R}$ is an open interval.

As for the curves in the Euclidean space, γ can be self-intersecting and have "corners", see Figures 1.3 and 1.4. The corners can be prohibited by the additional hypothesis that γ is an immersion, *i.e.* the image $\gamma(I)$ is an immersed submanifold of M. The smooth curve $\gamma : I \to M$ is an immersion (and its image $\gamma(I)$ is an immersed submanifold) if, and only if, γ is *regular* (*i.e.* $\gamma'(t) \neq 0$ for all $t \in I$). The self-intersections can be prohibited by the stronger requirement that the image $\gamma(I)$ is an embedded submanifold of M.

Let $t_0 \in I$. Let us denote by the canonical letter t the identity map $t : I \to \mathbb{R} : \{t \mapsto t\}$, which represents a coordinate system of the manifold I, and write

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0} := \left.\frac{\partial}{\partial t}\right|_{t_0}$$

for the standard coordinate basis of $T_{t_0}I$. On smooth functions $f : M \to \mathbb{R}$, the tangent vector to the curve γ at t_0 acts as follows (recall Definition 3.2):

$$\gamma'(t_0)f = (f \circ \gamma)'(t_0) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0} (f \circ \gamma) = \gamma_{*t_0} \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0} f\,.$$

Omitting the function f, we get

$$\boxed{\gamma'(t_0) = \gamma_{*t_0} \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0}} \in T_{\gamma(t_0)}M\,.$$

An alternative notation is $\dot\gamma(t_0)$ or $\left.\frac{\mathrm{d}\gamma}{\mathrm{d}t}\right|_{t_0}$. Moreover, the abbreviation $\frac{\mathrm{d}\gamma}{\mathrm{d}t} := \left.\frac{\mathrm{d}\gamma}{\mathrm{d}t}\right|_{t}$ is commonly used, where the letter t on the right-hand side denotes both the identity chart as well as one special point $t \in I$.

If the interval I is not open, we still say that $\gamma : I \to M$ is a *curve*, provided that the mapping γ is a restriction of a curve $\tilde\gamma$. More specifically, there exists an open interval $\tilde I \supset I$ and a curve $\tilde\gamma : \tilde I \to M$ such that $\tilde\gamma|_I = \gamma$.

3.8 The tangent bundle and vector fields

The *tangent bundle* of a manifold M is a disjoint union of all tangent spaces:

$$\boxed{TM := \dot{\bigcup_{p\in M}} T_pM\,.}$$

We consider an element of this disjoint union to be an ordered pair (p, X), where $p \in M$ and $X \in T_pM$. Define the *projection* map

$$\pi : TM \to M : \{(p, X) \mapsto p\}\,.$$

The tangent bundle can be endowed with a natural topology and differential structure to make it a smooth manifold. One has $\dim TM = 2 \dim M$ and π is smooth. Indeed, for the smooth chart $\phi : TM \to \mathbb{R}^{2n}$ one can choose

$$\phi(p, X) := \big(x^1(p), \dots, x^2(p), X^1, \dots, X^n\big)\,,$$

where (x, U) is a coordinate system of M at p and

$$X = X^i \left.\frac{\partial}{\partial x^n}\right|_p .$$

A physical interpretation of the tangent bundle is the phase space in classical mechanics.

A *vector field* X on a manifold M is a *section* of the tangent bundle TM, which means a continuous map $X : M \to TM : \{p \mapsto X_p\}$ such that $X_p \in T_pM$. (Here we write the value of X at p as X_p or $X|_p$ instead of $X(p)$ to avoid conflict with the action of a vector on a function p.) Given a coordinate system (x, U), we can write

$$X_p =: X^i(p) \left.\frac{\partial}{\partial x^i}\right|_p ,$$

where $X^i : U \to \mathbb{R}$ are called *component functions* of X with respect to the given chart x.

A vector field X is smooth if, and only if, the component functions are smooth for any chart (x, U). At the same time, a vector field X is smooth if, and only if, the function Xf is smooth for any smooth function $f : M \to \mathbb{R}$. For instance, the *coordinate vector field*

$$\frac{\partial}{\partial x^i} : U \to TU : \left\{ p \mapsto \left.\frac{\partial}{\partial x^i}\right|_p \right\}$$

is a smooth vector field.

Any vector can be extended to a vector field. More specifically, for any point $p \in M$ and a vector $X \in T_pM$, there exists a smooth vector field $\tilde{X}$ on M satisfying $\tilde{X}_p = X$.

The set of all smooth vector fields on M is denoted caligraphically as follows:

$$\mathcal{T}(M) := \{X \in C^\infty(M, TM) : \; X_p \in T_pM\} .$$

Let $f : M \to N$ be a smooth function and $f_{*p} : T_pM \to T_{f(p)}N$ the corresponding push-forward. Then $f_* : TM \to TN$ denotes the union of all f_{*p}. For this reason, we shall often omit the point p in the notation of push-forward.

3.9 Covectors and cotangent space

A *covector* at $p \in M$ is an element of the *cotangent space*

$$T_p^*M := (T_pM)^* ,$$

which is by definition the dual space of the tangent space T_pM. A covector is thus a linear functional $\omega : T_pM \to \mathbb{R}$. Given an arbitrary basis $(E_1, \dots, E_n)$ of T_pM, the functionals $(\varphi^1, \dots, \varphi^n)$ defined by the relations $(\forall j, k = 1, \dots, n)$

$$\varphi^j(E_k) = \delta^j_k$$

form a (so-called *dual*) basis of T_p^*M. Thus $\dim T_p^*M = \dim T_pM$.

The *cotangent bundle* of a manifold M is a disjoint union of all cotangent spaces:

$$T^*M := \dot{\bigcup_{p \in M}} T_p^*M .$$

A *covector field* ω on a manifold M is a section of the cotangent bundle T^*M, which means a continuous map $\omega : M \to T^*M : \{p \mapsto \omega_p\}$ such that $\omega_p \in T_p^*M$. (Here we again write the value of ω at p as ω_p or $\omega|_p$ instead of $\omega(p)$ to avoid conflict with the action of a covector on a vector p.) The section ω is smooth if, and only if, the function $\omega(X) : M \to \mathbb{R} : \{p \mapsto \omega_p(X_p)\}$ is smooth for all vector fields on M. The set of all smooth covector fields on M is denoted caligraphically:

$$\mathcal{T}^*(M) := \left\{\omega \in C^\infty(M, T^*M) : \ \omega_p \in T_p^*M\right\}.$$

Let $f : M \to \mathbb{R}$ be a smooth function. A *differential* $\mathrm{d}f$ of the function f is the smooth covector field satisfying

$$\forall X \in T_pM\,, \qquad \mathrm{d}f_p(X) := X(f)\,. \tag{3.10}$$

In particular, if (x, U) is a coordinate system at $p \in M$, then $\mathrm{d}x^i$ for $i = 1, \dots, n$ are smooth covector fields on U, called *coordinate covector fields*. From the definition, we get

$$\mathrm{d}x^i\big|_p \left(\frac{\partial}{\partial x^j}\bigg|_p \right) = \delta^i_j\,.$$

It follows that $(\mathrm{d}x^1|_p, \dots, \mathrm{d}x^n|_p)$ is a basis of the cotangent space T_p^*M, which is dual to the coordinate basis $(\frac{\partial}{\partial x^1}\big|_p, \dots, \frac{\partial}{\partial x^n}\big|_p)$. For an arbitrary section of T^*U, we therefore have the unique decomposition

$$\omega_p = \omega_i(p)\ \mathrm{d}x^i\big|_p\,,$$

where $\omega_i(p) : U \to \mathbb{R}$ are called the *component functions* of the covector field ω with respect to the chart x. A covector field ω is smooth if, and only if, the component functions are smooth for any chart. We have the following formula with a classical flavour:

$$\mathrm{d}f = \frac{\partial f}{\partial x^i}\,\mathrm{d}x^i\,.$$

3.10 Pull-backs

Let M, N be two smooth manifolds and consider a smooth function $F : M \to N$. Recall that, for an arbitrary point $p \in M$, we have defined the push-forward $F_{*p} : T_pM \to T_{F(p)}N$ by formula (3.9). Since F_{*p} is a linear map between vector spaces, there exists a dual map

$$F_p^* := (F_{*p})^* : \ T^*_{F(p)}N \to T_p^*M$$

called *pull-back* and satisfying

$$\forall \omega \in T^*_{F(p)}\,, \quad X \in T_pM\,, \qquad (F_p^*\omega)(X) := \omega(F_{*p}X)\,.$$

In general (except for the special situation when F is a diffeomorphism), it is not true that the push-forward image $F_{*p}X$ of a smooth vector field X on M is a smooth vector field on N. The surprising fact about pull-backs is that they always pull smooth covector fields back to smooth covector fields. Indeed, let ω be a smooth section of T^*N. Then we can define the smooth section $F^*\omega$ of the cotangent bundle T^*M by

$$(F^*\omega)_p := F_p^*\,\omega_{F(p)} = \omega_{F(p)} \circ F_{*p}$$

or, more precisely, $(F^*\omega)_p(X) := \omega_{F(p)}(F_{*p}X_p)$ for any vector field X on M.

The symbol $f^* : T^*M \to T^*N$ denotes the union of all f_p^*. For this reason, we shall often omit the point p in the notation of pull-back.

3.11 Orientation

We say that a manifold M is *orientable* if it admits an atlas $\{\phi_i\}_{i\in J}$ such that

$$\det\left((\phi_j \circ \phi_i^{-1})'\right) > 0 \tag{3.11}$$

for every pair of indices i, j from an index set J satisfying that the intersection of the domains of ϕ_i and ϕ_j is not empty. In the opposite case, we say that M is *non-orientable.*

If M is orientable, a choice of an atlas satisfying (3.11) is called an *orientation* of M. With this choice, M is said to be *oriented.* Two atlases satisfying (3.11) *determine the same orientation* if their union again satisfies (3.11). It is not difficult to verify that if M is orientable and connected there exist exactly two distinct orientations on M.

Now, let M, N be two manifolds and $\varphi : M \to N$ a diffeomorphism. It is easy to verify that M is orientable if, and only if, N is orientable. Let us assume in addition that M, N are oriented and connected. Then, given the orientation on M, the diffeomorphism φ induces an orientation on N, which may or may not coincide with the initial orientation of N. In the first (respectively, second) case we say that φ *preserves* (respectively, *reverses*) the orientation.

Proposition 3.3. *If M can be covered by two charts in such a way that the intersection of their domains is connected, then M is orientable.*

Proof. Let us assume that M is covered by two coordinate neighbourhoods U_1 and U_2 such that $U_1 \cap U_2$ is connected. Since the determinant of the Jacobi matrix corresponding to the change of coordinates is non-zero, it cannot change sign on the intersection $U_1 \cap U_2$. If it is negative at a single point, it suffices to change the sign of one of the coordinates to make it positive at that point, hence on $U_1 \cap U_2$. □

This proposition can be used to show that $\mathbb{R}^n$ (covered by the identity map) is orientable. At the same time, the n-dimensional sphere $\mathbb{S}^n_R$ of Example 3.1 (covered by two stereographic projections) is orientable. A notoriously known example of a non-orientable manifold is the Möbius strip from Figure 2.4.

3.12 Partition of unity

An immensely useful tool in differential geometry is the so-called *partition of unity.* It enables one a passage from local to global properties of manifolds. Its existence holds in the full generality of topological manifolds satisfying the two axioms of Section 3.1 (*i.e.*, the Hausdorff axiom about the disjoint neighbourhoods and the axiom about the complete separability).

Let $\{V_i\}_{i\in J} \subset M$ be a family of open sets covering M. Here J is an index set that can be finite or infinite (but countable). Thanks to the topological axioms from Section 3.1, we can also assume that the covering is locally finite. Moreover, by a potential refinement of the covering, we can also assume that each of the open sets V_i is contained in a coordinate neighbourhood of M.

Definition 3.4. A family $\{f_i\}_{i\in J} \subset C^\infty(M)$ is a *partition of unity subordinate to* the cover $\{V_i\}_{i\in J} \subset M$ if:

(1) $\forall i \in J, \quad 0 \leq f_i \leq 1;$

(2) $\forall i \in J, \quad \operatorname{supp} f_i \subset V_i;$

(3) $\forall p \in M, \quad \sum_{i\in J} f_i(p) = 1.$

As we have already mentioned, such a family of functions always exists. (Of course, it is not unique, see Figure 3.5.)

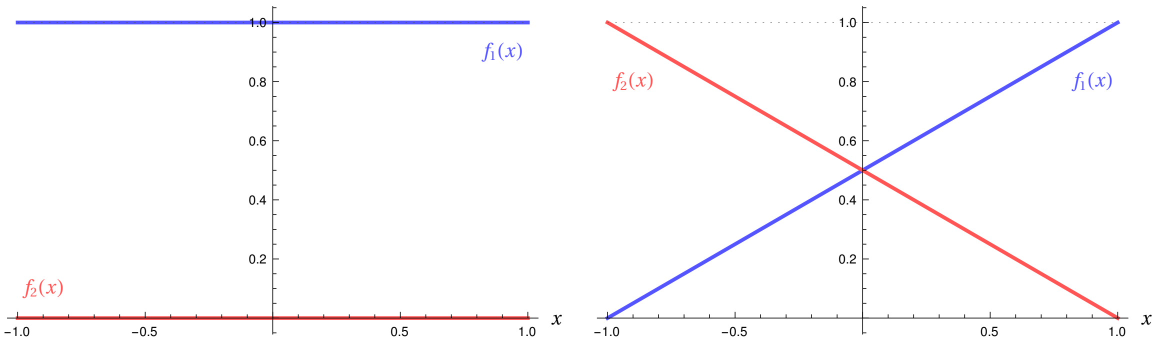


Figure 3.5: Two different partitions of unity $\{f_1, f_2\}$ subordinate to the double cover $\{I, I\}$ of the interval $I := (-1, 1)$.

3.13 Tensors

An element of $\mathcal{T}(M)$, *i.e.* a vector field, is occasionally called a *contravariant* vector field or also a *contravariant tensor of order* 1. An element of $\mathcal{T}^*(M)$, *i.e.* a covector field, is then called a *covariant* vector field or also a *covariant tensor of order* 1. This a warning that worse things are to come.

A smooth covector field $\omega \in \mathcal{T}^*(M)$ maps a smooth vector field $X \in \mathcal{T}(M)$ to a smooth function:

$$\omega : \mathcal{T}(M) \to C^\infty(M) : \{X \mapsto \omega(X)\}\,.$$

Moreover, this identification is *linear over* $C^\infty(M)$:

$$\forall X, Y \in \mathcal{T}(M),\ f, g \in C^\infty(M), \qquad \omega(fX + gY) = f\,\omega(X) + g\,\omega(Y)\,.$$

The notion of a covariant tensor is a generalisation of these observations.

Definition 3.5. A (covariant) *tensor* T of order $k \in \mathbb{N}^*$ on a manifold M is a mapping

$$T : \underbrace{\mathcal{T}(M) \times \cdots \times \mathcal{T}(M)}_{k\text{-times}} \to C^\infty(M)\,,$$

which is multilinear over $C^\infty(M)$. The set of all tensors of order k on M is denoted by $\mathcal{T}^k(M)$.

The multilinearity over $C^\infty(M)$ means the linearity in $C^\infty(M)$ in every argument, *i.e.*,

$$T(X_1,\dots,fX_i+gY,\dots,X_k) = fT(X_1,\dots,X_i,\dots,X_k) + gT(X_1,\dots,Y,\dots,X_k)$$

for arbitrary functions $f,g \in C^\infty(M)$ and vector fields $X_1,\dots,X_k,Y \in \mathcal{T}(M)$.

Definition 3.5 is consistent with the aforementioned motivation. Indeed, $\mathcal{T}^1(M) = \mathcal{T}^*(M)$. For consistency, we also define $\mathcal{T}^0(M) := C^\infty(M)$, *i.e.* smooth functions are understood as covariant tensors of order 0.

- Every tensor T has a *superlocal* character in the sense that $T(X_1,\dots,X_k)$ at $p \in M$ depends only on the values of the vector fields $X_1,\dots,X_k$ at the point p. Indeed, if (x,U) is a coordinate system at p, then the multilinearity implies

  $$T(X_1,\dots,X_k) = X_1^{j_1}\dots X_k^{j_k}\, T_{j_1,\dots,j_k},$$

  where $X_l^{j_l} \in C^\infty(U)$ are the component functions of the vector field $X_l \in \mathcal{T}(M)$, *i.e.* $X_l = X_l^{j_l}\frac{\partial}{\partial x^{j_l}}$, and $T_{j_1,\dots,j_k} := T\left(\frac{\partial}{\partial x^{j_1}},\dots,\frac{\partial}{\partial x^{j_k}}\right) \in C^\infty(U)$ are the component functions of the tensor T. From this formula, we see that the value $T(X_1,\dots,X_k)$ at p depends only on the values of the component functions at p.

- Vice versa, associating n^k smooth functions $T^{(i)}_{j_1,\dots,j_k} : U_i \to \mathbb{R}$ to every chart (x_i,U_i) of an atlas of M, one can introduce a global function $T : \mathcal{T}(M)^k \to C^\infty(M)$ by setting

  $$T(X_1,\dots,X_k)|_p := \sum_{i\in J} f_i(p)\, X_1^{j_1(i)}(p)\dots X_k^{j_k(i)}(p)\; T^{(i)}_{j_1,\dots,j_k}(p)\,,$$

  where $\{f_i\}_{i\in J}$ is a partition of unity subordinate to the cover $\{U_i\}_{i\in J}$ and $X_l^{j_l(i)}(p)$ are the component functions of the vector field $X_l \in \mathcal{T}(M)$ in the chart (x_i,U_i). Such a function represents a tensor if (and only if) the functions $T^{(i)}_{j_1,\dots,j_k}$ "transform correctly" on intersections of charts. More specifically, if $T_{j_1,\dots,j_k}$ and $\tilde{T}_{j_1,\dots,j_k}$ correspond to charts (x,U) and $(\tilde{x},\tilde{U})$, then on $U\cap\tilde{U}$ one has to have (*cf* (3.5))

  $$\tilde{T}_{j_1,\dots,j_k} = \frac{\partial x^{l_1}}{\partial \tilde{x}^{j_1}}\dots\frac{\partial x^{l_l}}{\partial \tilde{x}^{j_k}}\, T_{l_1,\dots,l_k}\,. \tag{3.12}$$

  Quoting [12, Chap. 4], classical differential geometry books are filled with monstrosities like this equation.

Given $T \in \mathcal{T}^k(M)$ and $S \in \mathcal{T}^l(M)$, we define the *tensor product* $T\otimes S \in \mathcal{T}^{k+l}(M)$ by

$$(T\otimes S)(X_1,\dots,X_k,X_{k+1},\dots,X_{k+l}) := T(X_1,\dots,X_k)\,S(X_{k+1},\dots,X_{k+l}), \tag{3.13}$$

where $X_1,\dots,X_{k+l} \in \mathcal{T}(M)$ are arbitrary vector fields.

In a similar manner, it is possible to introduce *contravariant tensors* of order k (as multilinear mappings over $C^\infty(M)$ from $\mathcal{T}^*(M)^k$ to $C^\infty(M)$) and even mixed tensors of arbitrary orders. For reasons of economy, we restrict ourselves to covariant tensors. Moreover, on Riemannian manifold (to which we restrict below), it is always possible to identify a contravariant tensor with a covariant one (and vice versa).

3.14 Riemannian manifolds

A *Riemannian manifold* is a smooth manifold additionally endowed with an inner product, which enables us to measure lengths and angles. Therefore we have the following analogy:

linear algebra:	vector space	$\supset$	inner product space
differential geometry:	smooth manifold	$\supset$	Riemannian manifold

Let us start by specifying what we mean by the inner product.

Definition 3.6. A (Riemannian) *metric* on M is a tensor g of order 2 satisfying, $\forall p \in M$,

(i) $\forall X, Y \in T_pM, \quad g_p(X,Y) = g_p(Y,X);$ (symmetry)

(ii) $\forall \underset{X \neq 0}{X \in T_pM}, \quad g_p(X,X) > 0.$ (positive definiteness)

Without the knowledge of tensors from Section 3.13, the definition can be understood pragmatically as follows. In every point $p \in M$, let us consider a function $g_p : T_pM \times T_pM \to \mathbb{R}$, which is linear (over $\mathbb{R}$) in each of the arguments, satisfies the symmetry (i) and the positive definiteness (ii). In addition, we require a smooth dependence on the point $p \in M$ in the following sense:

$$\forall \text{ chart } (x,U) \text{ at } p,\ i,j \in \{1,\dots,n\}, \qquad p \mapsto g_p\left(\frac{\partial}{\partial x^i}\Big|_p, \frac{\partial}{\partial x^j}\Big|_p\right) \in C^\infty(U)\,.$$

The function $g : \mathcal{T}(M) \times \mathcal{T}(M) \to C^\infty(M)$ is consequently obtained by setting $g(X,Y)|_p := g_p(X_p, Y_p)$.

Since the function $g_p : T_pM \times T_pM \to \mathbb{R}$ associates to every point $p \in M$ an inner product on the tangent space T_pM, it is customary to write

$$g(X,Y) =: \langle X, Y \rangle_g\,.$$

If the choice of metric is clear from context, the index g is omitted.

Definition 3.7. A *Riemannian manifold* is a manifold endowed with a metric.

We write (M,g), where M is the manifold and g is the metric on M.

The following claim shows that every manifold can be upgraded to a Riemannian manifold.

Proposition 3.8. *Every manifold admits a metric.*

Proof. The claim is an analogy of the classical fact that every finite-dimensional vector space admits an inner product (for every finite-dimensional vector space is isomorphic to some Euclidean space, in which we have the standard inner product). The proof for manifolds is analogous by using the local identification of the manifold with the Euclidean space and the partition of unity (recall Section 3.12) to define the metric globally. Indeed, on every set V_i from the global cover it is clear (using that M is locally diffeomorphic to $\mathbb{R}^n$) how to introduce the metric $g_p^{(i)}$ at any point $p \in V_i$. For every point $p \in M$ and $X, Y \in T_pM$, we then define

$$g_p(X,Y) := \sum_{i \in J} f_i(p)\, g_p^{(i)}(X,Y)\,.$$

It is easy to verify that this construction defines a metric on M. $\square$

As in the Euclidean geometry, we introduce:

- *length (norm)* of a vector X: $\quad |X| := \sqrt{\langle X, X\rangle}$;
- *angle* between vectors $X, Y \neq 0$: $\quad \theta \in [0,\pi]$ satisfying $\cos\theta = \dfrac{\langle X,Y\rangle}{|X||Y|}$;
- $X \perp Y$ (vectors X, Y are *orthogonal*) $\quad :\Longleftrightarrow \quad \langle X,Y\rangle = 0$ (the angle is $\frac{\pi}{2}$).

Let $\gamma : I \to M$ be a curve on a Riemannian manifold M. Recall (see Section 3.7) that the tangent vector γ' to the curve γ at $t_0 \in I$ is defined via the push-forward $\gamma'(t_0) := \gamma_{*t_0} \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t_0}$, where t denotes the identity map $t : I \to \mathbb{R} : \{t \mapsto t\}$. The *length of the curve* γ is defined by

$$L(\gamma) := \int_I |\gamma'(t)|\, \mathrm{d}t\,. \tag{3.14}$$

The length of the curve restricted to a subinterval $I' \subset I$ is defined analogously.

Let $(E_1, \dots, E_n)$ be an arbitrary *local frame* on TM, *i.e.* n smooth vector fields defined on an open set $U \subset M$ such that $(E_1|_p, \dots, E_n|_p)$ is a basis on T_pM for every $p \in U$. Let $(\varphi^1, \dots, \varphi^n)$ be a *dual frame* on T^*M, *i.e.* n smooth covector fields satisfying $\varphi^i(E_j) = \delta^i_j$. Then we can locally write

$$g = g_{ij}\ \varphi^i \otimes \varphi^j = g_{ij}\ \varphi^i\varphi^j\,, \qquad g_{ij} := \langle E_i, E_j\rangle\,,$$

where the second equality employs the symmetry of g together with the definition of the *symmetric product*

$$\varphi^i\varphi^j := \frac{1}{2}(\varphi^i \otimes \varphi^j + \varphi^j \otimes \varphi^i)\,.$$

In particular, if $(x^1, \dots, x^n)$ are local coordinates on M, then $(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n})$ is a coordinate frame on TM, $(\mathrm{d}x^1, \dots, \mathrm{d}x^n)$ is the dual frame on T^*M and we have

$$g = g_{ij}\ \mathrm{d}x^i\,\mathrm{d}x^j. \qquad g_{ij} := \langle \tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^j}\rangle. \tag{3.15}$$

3.15 Riemannian submanifolds

Recall that the simplest example of a differentiable manifold is the Euclidean space $\mathbb{R}^n$ endowed with the identity map $x : \mathbb{R}^n \to \mathbb{R}^n : \{p \mapsto p\}$ forming an atlas. In these (global) coordinates, the usual inner product on $\mathbb{R}^n$ corresponds to the *Euclidean metric*

$$\bar{g} = \delta_{ij}\ \mathrm{d}x^i\mathrm{d}x^j\,.$$

In the special situation of surfaces $p : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ from Section 2, the image $p(U)$ is an immersed submanifold of $\mathbb{R}^3$. Now we would like to endow it with a natural Riemannian metric of Definition 3.6 and relate it with the previously introduced metric of p from Definition 2.6. The meaning of "natural" is that the Riemannian metric of $p(U)$ is inherited from the ambient space $\mathbb{R}^3$.

More generally, given a Riemannian manifold $(\tilde{M}, \tilde{g})$, let $M \subset \tilde{M}$ be an immersed submanifold defined through the immersion $\iota : M \to \tilde{M}$. The metric $\tilde{g}$ of the ambient space $\tilde{M}$ naturally defines an *induced metric* g on M by

$$g := \iota^* \tilde{g}. \tag{3.16}$$

This formula precisely means

$$\forall p \in M \,, \; \forall X, Y \in T_p M \,, \qquad (\iota_p^* \, \tilde{g})_p(X, Y) := \tilde{g}_{\iota(p)}(\iota_{*p} X, \iota_{*p} Y) \,.$$

The induced metric is thus a restriction of the ambient metric $\tilde{g}$ on vectors tangent to the submanifold M.

This definition leads us to the extension of some notions familiar from linear algebra to manifolds. Let (M, g) and $(\tilde{M}, \tilde{g})$ be two Riemannian manifolds. An *isometry* from (M, g) to $(\tilde{M}, \tilde{g})$ is a diffeomorphism $\varphi : M \to \tilde{M}$ such that $\varphi^* \tilde{g} = g$. (In the case of the immersion ι above, where g is the induced metric, we thus speak of an *isometric immersion*, because $\iota : M \to \iota(M)$ is an isometry.) We say that (M, g) and $(\tilde{M}, \tilde{g})$ are *isometric* (or *isometrically equivalent*) if there exists an isometry between (M, g) and $(\tilde{M}, \tilde{g})$. A mapping $\varphi : M \to \tilde{M}$ is a *local isometry* if for every point $p \in M$ there exists a neighbourhood U such that $\varphi|_U$ is an isometry to an open set in $\tilde{M}$. If the manifolds (M, g) a $(\tilde{M}, \tilde{g})$ have the same dimension, then $\varphi : M \to \tilde{M}$ is a local isometry if, and only if, $\varphi^* \tilde{g} = g$.

Let us now look at the special case $\tilde{M} = \mathbb{R}^d$ with $d \geq n$, when M is immersed to the Euclidean space $\mathbb{R}^d$ endowed with the Euclidean metric $\bar{g}$. In local coordinates x of M (and using the identity map for the ambient manifold $\mathbb{R}^d$), one has (*cf* Section 3.6)

$$g(X, Y) = \delta_{\alpha\beta} \, (\iota_* X)^\alpha \, (\iota_* Y)^\beta = \underbrace{\delta_{\alpha\beta} \, \frac{\partial \iota^\alpha}{\partial x^i} \, \frac{\partial \iota^\beta}{\partial x^j}}_{g_{ij}} \, X^i \, Y^j \,,$$

where $X, Y \in \mathfrak{T}(M)$, $\alpha, \beta = 1, \dots, d$ and $i, j = 1, \dots, n$. In practice, computations are most conveniently carried out in terms of a local parameterisation (see Figure 2.1)

$$p : U \to \mathbb{R}^d \qquad \text{with} \qquad p(U) \subset M \,,$$

where $U \subset \mathbb{R}^n$ and $p(U) \subset M$ are open set. In the coordinates $x := p^{-1} \circ \iota$ on $p(U)$, one has

$$\frac{\partial \iota^\alpha}{\partial x^i} = \partial_i (\iota^\alpha \circ x^{-1}) \circ x = (\partial_i p^\alpha) \circ x \,,$$

and therefore

$$g_{ij} = \delta_{\alpha\beta} \; (\partial_i p^\alpha) \circ x \; (\partial_j p^\beta) \circ x \,.$$

Identifying the maps $(\partial_i p^\alpha) \circ x$ and $(\partial_i p^\alpha)$, we can simply write

$$g_{ij} = \partial_i p \cdot \partial_j p \,, \tag{3.17}$$

where the dot $\cdot$ denotes the inner product in $\mathbb{R}^d$. This is just the classical formula (2.2). In matrix formalism, $(g_{ij}) = (\nabla p)(\nabla p)^\top$.

3.16 Integration

The notion of Riemannian metric enables one to integrate functions on Riemannian manifolds. In particular, a notion of volume of a compact manifold can be introduced.

Let M be an oriented Riemannian manifold, $p \in M$ and consider a chart x at p from the given orientation. Let $(E_1, \dots, E_n)$ be a positively oriented orthonormal basis of the tangent space T_pM (*i.e.*, $(E_1, \dots, E_n)$ has the same orientation as $(\frac{\partial}{\partial x^1}\big|_p, \dots, \frac{\partial}{\partial x^n}\big|_p)$). Every vector of the coordinate basis can be decomposed into the other basis as follows:

$$X_i := \frac{\partial}{\partial x^i}\bigg|_p = A_i^{\ j} E_j \,,$$

where $A_i^{\ j} \in \mathbb{R}$. Then

$$g_{ik}|_p := \langle X_i, X_k \rangle|_p = A_i^{\ j} A_k^{\ l} \, \langle E_j, E_l \rangle|_p = \sum_{j=1}^{n} A_i^{\ j} A_k^{\ j} \,.$$

Or, in a matrix formalism, $(g_{ik}|_p) = AA^T$, where $A := (A_i^{\ j})$. Let $\mathrm{vol}(X_1, \dots, X_n)$ denote the volume of the parallelepiped formed by the vectors $X_1, \dots, X_n$ v T_pM. Then

$$\mathrm{vol}(X_1, \dots, X_n) = \mathrm{vol}(E_1, \dots, E_n)\, \det(A_i^{\ j}) = \det(A_i^{\ j}) = \sqrt{\det(g_{ij})}\bigg|_p \,,$$

because $\mathrm{vol}(E_1, \dots, E_n) = 1$. Let us now consider another chart y at p from the given orientation of M and let us write

$$Y_i := \frac{\partial}{\partial y^i}\bigg|_p \,, \qquad \tilde{g}_{ik}|_p := \langle Y_i, Y_k \rangle|_p \,.$$

Then

$$\sqrt{\det(g_{ij})}\bigg|_p = \mathrm{vol}(X_1, \dots, X_n) = J\,\mathrm{vol}(Y_1, \dots, Y_n) = J\,\sqrt{\det(\tilde{g}_{ij})}\bigg|_p \,,$$

where

$$J := \det\left(\frac{\partial y^i}{\partial x^j}\right) = \det\left((y \circ x^{-1})'\right) > 0$$

is the determinant of the Jacobi matrix of the change of coordinates.

Let $\Omega \subset M$ be a domain (*i.e.* open connected set), which is precompact (*i.e.* its closure is compact). Let us assume that $\Omega \subset U$, where U is the domain of the chart x from the given orientation of M. Let us assume in addition that $x(\Omega) \subset \mathbb{R}^n$ is Jordan measurable, *i.e.* the boundary $\partial x(\Omega)$ has measure zero in $\mathbb{R}^n$ (the notion of measure zero in $\mathbb{R}^n$ is invariant by diffeomorphism). Let us define the *volume* of Ω by the integral

$$\mathrm{vol}(\Omega) := \int_{x(\Omega)} \left(\sqrt{\det(g_{ij})}\right) \circ x^{-1}(\hat{x})\; \mathrm{d}\hat{x}^i \dots \mathrm{d}\hat{x}^n \,. \tag{3.18}$$

It is customary to denote the variable $\hat{x} \in \mathbb{R}^n$ by the same letter x of the chart $x : U \subset M \to \mathbb{R}^n$. What is worse, it is customary to omit the composition with x^{-1} at $\hat{x}$.

The formula (3.18) is well defined. Indeed, if Ω is contained in a domain of another chart y from the given orientation, then the change of variables for multidimensional integrals yields (keeping the same notation as above)

$$\int_{x(\Omega)} \left(\sqrt{\det(g_{ij})} \circ x^{-1}\right)(x)\, \mathrm{d}x^i \dots \mathrm{d}x^n = \int_{y(\Omega)} \left(\sqrt{\det(\tilde{g}_{ij})} \circ y^{-1}\right)(y)\, \mathrm{d}y^i \dots \mathrm{d}y^n \,.$$

Hence $\mathrm{vol}(\Omega)$ does not depend on the choice of coordinates. Here the hypothesis of the orientability of M enters by guaranteeing that $\mathrm{vol}(\Omega)$ does not change sign by the change of coordinates.

The integral of a smooth function $f : \Omega \to \mathbb{R}$ can be introduced analogously:

$$\int_\Omega f := \int_{x(\Omega)} \left(f \sqrt{\det(g_{ij})}\right) \circ x^{-1}(x)\, \mathrm{d}x^i \dots \mathrm{d}x^n \,.$$

Then, of course, $\mathrm{vol}(\Omega) = \int_\Omega 1$.

For a domain Ω which is not necessarily contained in the domain of one chart, the integral can be introduced with help of the partition of unity as follows:

$$\int_\Omega f := \sum_{i\in I} \int_{V_i\cap\Omega} f_i\, f \,.$$

Here $\{f_i\}_{i\in I}$ is the partition of unity subordinated to the cover $\{V_i\}_{i\in I}$ from Section 3.12, but additionally from the given orientation of M.

If M is compact, the definitions above enable one to introduce the integral $\int_M f$ for any smooth function $f : M \to \mathbb{R}$, and in particular $\mathrm{vol}(M) := \int_M 1$.

3.17 The Gauss–Bonnet theorem

Finally, we present (without proof) the most fundamental and important theorem in differential geometry, at least as far as two-dimensional manifolds are concerned. It provides an astonishing link between a local object of the manifold (its Gauss curvature) and its global topological properties.

Let (M, g) be a two-dimensional Riemannian manifold. Given a local chart x, let g_{ij} denote the $2^2 = 4$ component functions of the metric tensor g, see (3.15). Using Definition 2.16, we introduce the $2^3 = 8$ Christoffel symbols Γ^k_{ij}. Using Definition 2.18, we introduce the $2^4 = 16$ component functions of the curvature tensor R_{ijkl}. Finally, using *Theorem Egregium* (Theorem 2.20), we introduce the Gauss curvature K of M via (2.9):

$$K = \frac{1}{2}\, R_{ijkl}\, g^{il} g^{jk} \,.$$

Then everything is defined intrinsically (*i.e.* no embedding of M in $\mathbb{R}^3$ needed) and invariantly (*i.e.* independent of the choice of coordinates) provided that we verify that R_{ijkl} are indeed component functions of a tensor $R : \mathcal{T}(M)^4 \to \mathbb{R}$. We leave this routine (but monstrous) verification as Exercise 3.18.1.

Now, let us assume in addition that M is compact. A (smooth) *triangulation* of M is a collection $\{\Omega_i\}_{i=1}^{N_f}$ such that

$$\bigcup_{i=1}^{N_f} \overline{\Omega}_i = M \,,$$

where $N_f < \infty$ and Ω_i are curved triangles (*i.e.* a curved polygon with exactly three edges and three vertices) with disjoint interiors such that the intersection of any pair of triangles is either empty or a single vertex of each or a single edge of each. Such a triangulation always exists. The *Euler characteristic* of M (with respect to the given triangulation) is the number

$$\chi(M) := N_v - N_e + N_f \,,$$

where

$$\begin{aligned} N_v &:= \text{number of vertices}\,, \\ N_e &:= \text{number of edges}\,, \\ N_f &:= \text{number of faces (= number of triangles)}\,. \end{aligned}$$

The Euler characteristic is a purely topological invariant, which is independent of the choice of triangulation. One has $\chi(\mathbb{S}) = 2$ for the sphere, $\chi(\mathbb{T}) = 0$ for the torus and in general $\chi(M) = 2 - 2\mathtt{g}$ for any oriented manifold of genus $\mathtt{g}$ (*i.e.* homeomorphic to the sphere $\mathbb{S}$ with $\mathtt{g} \geq 0$ handles attached).

Theorem 3.9 (Gauss–Bonnet)**.** *Let M be a compact oriented two-dimensional Riemannian manifold. Then*

$$\int_M K = 2\pi\,\chi(M)\,.$$

The significance of the Gauss–Bonnet theorem cannot be overstated. Let us mention just a couple of direct consequences. (For simplicity, we denote by the symbol $\sim$ the relation of homeomorphism of two manifolds.)

Corollary 3.10.

$$\begin{array}{llll} \text{(a)} & M \sim \mathbb{S} \quad (\mathtt{g}=0) & \Longrightarrow & K > 0 \textit{ somewhere}\,; \\ \text{(b)} & M \sim \mathbb{T} \quad (\mathtt{g}=1) & \Longrightarrow & K = 0 \textit{ everywhere} \\ & & & \vee\ (K > 0 \textit{ somewhere} \ \wedge\ K < 0 \textit{ somewhere})\,; \\ \text{(c)} & \textit{another topology} \quad (\mathtt{g} \geq 2) & \Longrightarrow & K < 0 \textit{ somewhere}\,. \end{array}$$

Corollary 3.11.

$$\begin{array}{llll} \text{(a)} & K > 0 & \Longrightarrow & M \sim \mathbb{S} \quad (\mathtt{g}=0)\,; \\ \text{(b)} & K \leq 0 & \Longrightarrow & \mathtt{g} \geq 1\,. \end{array}$$

Much of the effort in contemporary Riemannian geometry is aimed at generalising the Gauss–Bonnet theorem and its topological consequences to higher dimensions.

3.18 Exercises

1. Verify that the curvature tensor from Definition 2.18 is invariantly defined, meaning that it satisfies (*cf* (3.12))

$$\tilde{R}_{ijkl} = \frac{\partial x^\alpha}{\partial \tilde{x}^i}\frac{\partial x^\beta}{\partial \tilde{x}^j}\frac{\partial x^\gamma}{\partial \tilde{x}^k}\frac{\partial x^\delta}{\partial \tilde{x}^l}\, R_{\alpha\beta\gamma\delta}\,,$$

   where $R_{\alpha\beta\gamma\delta}$ (respectively, $\tilde{R}_{ijkl}$) are the functions computed from the metric via formulae (2.7) and (2.8) in a chart x (respectively, $\tilde{x}$). As the first step, verify that the Christoffel symbols obey the transformation rule

$$\tilde{\Gamma}^k_{ij} = \frac{\partial x^\alpha}{\partial \tilde{x}^i}\frac{\partial x^\beta}{\partial \tilde{x}^j}\frac{\partial \tilde{x}^k}{\partial x^\gamma}\,\Gamma^\gamma_{\alpha\beta} + \frac{\partial \tilde{x}^k}{\partial x^\rho}\frac{\partial^2 x^\rho}{\partial \tilde{x}^i \partial \tilde{x}^j} \tag{3.19}$$

   (*i.e.*, they do not transform as a tensor), where $\Gamma^\gamma_{\alpha\beta}$ (respectively, $\tilde{\Gamma}^k_{ij}$) are the functions computed from the metric via formula (2.7) in the chart x (respectively, $\tilde{x}$).

2. Verify that the geodesic equation from Proposition 2.21 is invariantly defined. More specifically, let $\gamma : I \to M$ be a curve and, using the notation of the previous exercise, denote its component functions in charts x and $\tilde{x}$ by $\gamma^\alpha := (x \circ \gamma)^\alpha$ and $\tilde{\gamma}^i := (\tilde{x} \circ \gamma)^i$, respectively. Show that

$$\tilde{\gamma}^{k\prime} + \tilde{\Gamma}^k_{ij}\,\tilde{\gamma}^{i\prime}\tilde{\gamma}^{j\prime} = \frac{\partial \tilde{x}^k}{\partial x^\alpha}\left(\gamma^{\alpha\prime\prime} + \Gamma^\alpha_{\rho\sigma}\,\gamma^{\rho\prime}\gamma^{\sigma\prime}\right)\,.$$

   Hint: Use (3.19) and the identity

$$\frac{\partial^2 \tilde{x}^k}{\partial x^\rho \partial x^\sigma} + \frac{\partial^2 x^\alpha}{\partial \tilde{x}^i \partial \tilde{x}^j}\frac{\partial \tilde{x}^k}{\partial x^\alpha}\frac{\partial \tilde{x}^i}{\partial x^\rho}\frac{\partial \tilde{x}^j}{\partial x^\sigma} = 0\,,$$

   which can be verified by differentiating

$$\frac{\partial \tilde{x}^i}{\partial x^\rho}\frac{\partial x^\rho}{\partial \tilde{x}^j} = \delta^i_j\,.$$

References

[1] V. Blåsjö, *The isoperimetric problem*, Am. Math. Mon. **112** (2005), 526–566.

[2] M. P. do Carmo, *Differential geometry of curves and surfaces*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, 1976.

[3] ______, *Riemannian geometry*, Birkhaüser, Boston, 1992.

[4] W. Klingenberg, *A course in differential geometry*, Springer-Verlag, New York, 1978.

[5] O. Kowalski, *Úvod do Riemannovy geometrie*, Univerzita Karlova v Praze, Praha, 1995.

[6] D. Krejčiřík and K. Zahradová, *Quantum strips in higher dimensions*, Oper. Matrices **14** (2020), 635–665.

[7] M. Spivak, *Calculus on manifolds*, Addison-Wesley Publishing Company, United States of America, 1965.

[8] ______, *A comprehensive introduction to differential geometry*, vol. III, Publish or Perish, Berkeley, Calif., 1975.

[9] ______, *A comprehensive introduction to differential geometry*, vol. IV, Publish or Perish, Boston, Mass., 1975.

[10] ______, *A comprehensive introduction to differential geometry*, vol. II, Publish or Perish, Houston, Texas, 1999.

[11] ______, *A comprehensive introduction to differential geometry*, vol. V, Publish or Perish, Boston, Mass., 1999.

[12] ______, *A comprehensive introduction to differential geometry*, vol. I, Publish or Perish, Houston, Texas, 2005.