RIEMANNIAN GEOMETRY–AN INTRODUCTORY COURSE

NOTES TO A COURSE BY E.P. VAN DEN BAN AND E. LOOIJENGA

Some conventions. In this course *differentiable* shall mean 'differentiable up to any order' (often written as C^∞). So if $U \subset \mathbb{R}^m$ is open, then a differentiable function $f : U \to \mathbb{R}$ is one for which all (iterated) partial derivatives exist. But as a rule all maps encountered here are tacitly assumed to have that property whenever it makes sense, unless the contrary is explicitly stated. So when you are told here that a certain map is differentiable, then this is just for emphasis.

We also recall that given any subset $D \subset \mathbb{R}^m$, a function $f : D \to \mathbb{R}$ is said to be *differentiable* if it is the restriction of a differentiable function defined on an open subset of $\mathbb{R}^m$ containing D. For instance, $f : [a, b] \to \mathbb{R}$ is differentiable if there exists an $\varepsilon > 0$ such that $f : (a - \varepsilon, b + \varepsilon) \to \mathbb{R}$ is so.

CONTENTS

0. Curves and surfaces in the plane and in space

Curves in $\mathbb{R}^m$. An *arc in* $\mathbb{R}^m$ will be simply a (differentiable) map from an interval in $\mathbb{R}$ to $\mathbb{R}^m$: $\gamma : I \to \mathbb{R}^m$, $t \in I \mapsto (\gamma^1(t), \dots, \gamma^m(t))$. Its *velocity vector* resp. *acceleration vector* at $t \in I$ is the first resp. second order derivative

$$\dot\gamma(t) := \begin{pmatrix} \dot\gamma^1(t) \\ \dot\gamma^2(t) \\ \vdots \\ \dot\gamma^n(t) \end{pmatrix} \in \mathbb{R}^m, \quad \ddot\gamma(t) := \begin{pmatrix} \ddot\gamma^1(t) \\ \ddot\gamma^2(t) \\ \vdots \\ \ddot\gamma^n(t) \end{pmatrix} \in \mathbb{R}^m.$$

Often we are more concerned with the geometric properties of the image of γ rather than with γ itself. These properties will certainly be invariant under *reparameterization*. To explain what this means, let us recall from real analysis that if $h : I \to \mathbb{R}$ is a differentiable function on an open interval whose derivative is positive everywhere, then h maps I bijectively onto an interval I' and has a differentiable inverse $h^{-1} : I' \to I$. Our notion of differentiability shows that this is still true if we suppress 'open': I can be any interval. A reparameterization of $\gamma : I \to \mathbb{R}^m$ is now simply an arc $\gamma' : I' \to \mathbb{R}^m$ of the form $\gamma \circ h^{-1}$, with $h : I \cong I'$ as above. Notice that $\gamma = \gamma' \circ h$ and so

$$\dot\gamma(t) = \dot\gamma'(h(t))\frac{dh}{dt}(t).$$

Arc length. In what follows we use the *standard inner product* on $\mathbb{R}^m$ defined by $\langle x, y\rangle = x^1y^1 + \cdots x^my^m$. The vector space $\mathbb{R}^m$, when endowed with this inner product, is also called *m-dimensional Euclidean space* and will be denoted $\mathbb{E}^m$. The *norm* on $\mathbb{E}^m$ is defined as usual by $\|x\| := \sqrt{\langle x, x\rangle}$. This makes $\mathbb{E}^m$ a metric space by letting the distance between $p, q \in \mathbb{E}^m$ be $\|p - q\|$.

Exercise 1. On $\mathbb{E}^m$ there is also a notion of *angle*: two rays r, r' emanating from $p \in \mathbb{E}^m$ have an angle in $[0, \pi]$. Define this in terms of the inner product.

Exercise 2. A *Euclidean transformation* of $\mathbb{E}^m$ is a map of the form $p \in \mathbb{E}^m \mapsto p_0 + \sigma(p) \in \mathbb{E}^m$ with $p_0 \in \mathbb{E}^m$ and $\sigma \in O(m)$. It is called a *translation* if σ is the identity.
(a) Prove that a Euclidean transformation is an isometry.
(b) Prove that the Euclidean transformations make up a group with respect to composition. Show that the translations form a normal subgroup with factor group $O(m)$.

Exercise 3. (a) Prove that collinearity in $\mathbb{E}^m$ is preserved under isometries: if p, q, r lie on a line, then so do their images under an isometry.
(b) Prove that angles are preserved under isometries. Conclude that an isometry of $\mathbb{E}^m$ maps a triangle onto a congruent triangle.
(c) Now show that any isometry of $\mathbb{E}^m$ which fixes the origin O and the unit vectors $e_1, \dots, e_m$ is in fact the identity.
(d) Prove that any isometry of $\mathbb{E}^m$ is in fact a Euclidean transformation.

For an arc $\gamma : [a, b] \to \mathbb{E}^m$ there is a notion of length. A subdivision $\mathcal{V} = (a = t_0 < t_1 < \cdots < t_N = b)$ of $[a, b]$ determines a piecewise-linear approximation of γ that takes the same values on the division points and is affine-linear on each subinterval $[t_{i-1}, t_i]$. To be precise, it takes $t \in [t_{i-1}, t_i]$ to $(t_i - t)/(t_i - t_{i-1})\gamma(t_{i-1}) + (t - t_{i-1})/(t_i - t_{i-1})\gamma(t_i)$, but this formula is not important to us; what matters here is only that the total length $l(\mathcal{V})$ of this approximation equals $\Sigma_{i=1}^N \|\gamma(t_i) - \gamma(t_{i-1})\|$. This length has a limit as the *mesh* $m(\mathcal{V}) := \max\{t_i - t_{i-1} | i = 1, \dots N\}$ tends to zero:

Proposition 0.1. *We have*

$$\lim_{m(\mathcal{V})\to 0} l(\mathcal{V}) = \int_a^b \|\dot\gamma(t)\| dt.$$

This limit, called the arc length *of* γ*, and denoted* $L(\gamma)$*, does not change after reparameterization.*

Proof. For every $\varepsilon > 0$ there exists a $\delta > 0$, such that for all $t, t' \in [a, b]$ with $|t - t'| < \delta$ we have

$$\|\gamma(t') - \gamma(t) - \dot\gamma(t).(t' - t)\| < \varepsilon|t' - t|.$$

and hence

$$\Big(\|\gamma(t') - \gamma(t)\| - \|\dot\gamma(t)\|.|t' - t|\Big) < \varepsilon|t' - t|.$$

So if we assume that $m(\mathcal{V}) < \delta$, then this holds for $(t, t') = (t_{i-1}, t_i)$, $i = 1, \dots, N$ and hence

$$\begin{aligned} &\Big|l(\mathcal{V}) - \Sigma_{i=1}^N \|\dot\gamma(t_{i-1})\|.(t_i - t_{i-1})\Big| = \\ &\qquad \Big|\Sigma_{i=1}^N \Big(\|\gamma(t_i) - \gamma(t_{i-1})\| - \|\dot\gamma(t_i)\|.(t_i - t_{i-1})\Big)\Big| \\ &\qquad\qquad < \Sigma_{i=1}^N \varepsilon(t_i - t_{i-1}) = \varepsilon(b - a). \end{aligned}$$

The sum in the lefthand side tends to $\int_a^b \|\dot\gamma(t)\| dt$ as $m(\mathcal{V}) \to 0$, and so the first part of the proposition follows. The second part is a consequence of the chain rule:

$$L(\gamma) = \int_a^b \|\dot\gamma(t)\| dt = \int_a^b \|\dot\gamma'(h(t))\| . |\frac{dh}{dt}(t)| dt =$$
$$= \int_a^b \|\dot\gamma'(h(t))\| . \frac{dh}{dt} dt = \int_{a'}^{b'} \|\dot\gamma'(t')\| dt' = L(\gamma').$$

□

If $\gamma : I \to \mathbb{E}^m$ has the property that $\|\dot\gamma\|$ is constant 1, then the preceding implies that for closed interval $[a, b] \subset I$ the arc length of $\gamma | [a, b]$ equals $b - a$. We therefore then say that γ is *parameterized by arc length*. Such a parametrization often exists:

Lemma 0.2. *Assume that the arc $\gamma : I \to \mathbb{E}^m$ has the property that $\dot\gamma$ is never zero. Then γ can be reparameterized by arc length and such a reparameterization is unique up to a time translation.*

Proof. Choose $t_0 \in I$, and consider the arc length function

$$h(t) := \int_{t_0}^t \|\dot\gamma(\tau)\| d\tau.$$

Its derivative in t is $\|\dot\gamma(t)\| > 0$ and so h defines a diffeomorphism from I onto an interval I'. Let $\gamma' := \gamma \circ h^{-1}$, so that $\gamma = \gamma' \circ h$. By the chain rule,

$$\dot\gamma(t) = \dot\gamma'(h(t)) \frac{dh}{dt}(t).$$

Taking the norm at both sides yields $\|\dot\gamma'\| \equiv 1$ and so γ' is parameterized by arc length.

This argument also shows that, in case γ already happened to be parameterized by arc length, h has the property that its derivative is constant equal to 1 and so must be of the form $h(t) = t +$ constant. This proves the uniqueness clause. □

Curvature. We here consider only arcs parameterized by arc length. Given such an arc γ, then we often write $e_1(t)$ for the unit vector $\dot\gamma(t)$. We begin with a simple proposition.

Proposition 0.3. *If $\gamma : I \to \mathbb{E}^m$ is parameterized by arc length, then $\ddot\gamma(t) \perp \dot\gamma(t)$ for all $t \in I$.*

Proof. By assumption we have that $\langle \dot\gamma(t), \dot\gamma(t) \rangle \equiv 1$. Differentiating this identity with respect to t yields $\langle \ddot\gamma(t), \dot\gamma(t) \rangle \equiv 0$. □

We shall give a geometric interpretation of $\|\ddot\gamma(t)\|$ in this case.

Definition 0.4. The *curvature* of γ at t_0 is $\|\ddot{\gamma}(t_0)\|$. If this vanishes in t_0, we say that γ has a *flex point* at t_0. If not, then we put

$$r(t_0) := \|\ddot{\gamma}(t_0)\|^{-1}, \quad e_2(t_0) := \frac{\ddot{\gamma}(t_0)}{\|\ddot{\gamma}(t_0)\|} \text{ (a unit vector).}$$

The plane spanned by $e_1(t_0), e_2(t_0)$, but displaced over $\gamma(t_0)$ (so $\gamma(t_0) + \mathbb{R}e_1(t_0) + \mathbb{R}e_2(t_0)$), is called the *osculating plane* of γ at t_0.

The latter is the plane in which γ stays up to second order at t_0, because

$$\gamma(t) = \gamma(t_0) + (t - t_0)e_1(t_0) + \frac{(t - t_0)^2}{2r(t_0)}e_2(t_0) + O((t - t_0)^3).$$

In order to explain the curvature terminology, let us assume that $m = 2$. Since we want to understand the metric properties of γ at a given point $t_0 \in I$, we make a time translation and assume $t_0 = 0$. Next we compose γ with a Euclidean transformation if necessary in order to arrange that $\gamma(0) = 0$, $\dot{\gamma}(t_0) = e_1$. We also assume that $\ddot{\gamma}(0) = r(0)^{-1}e_2$. A circle in $\mathbb{E}^2$ through the origin that has there the same tangent line as γ will have its center on the line spanned by e_2, say at re_2. The parameterization of that circle by arc length is

$$\delta(t) = r\sin(t/r)e_1 + r(1 - \cos(t/r))e_2.$$

Notice that $\delta(0) = 0$, $\dot{\delta}(0) = e_1$ and $\ddot{\delta}(0) = r^{-1}e_2$. So in order that $\ddot{\delta}(0) = \ddot{\gamma}(0)$ we want that $r = r(0)$. We then get a circle which approximates γ at 0 up to order three: $\delta(t) = \gamma(t) + O(t^3)$. We call it the *osculating circle* of γ at 0. Its radius is $r(0)$ and its center is $r(0)e_2$, or more intrinsically,

$$\gamma(0) + r(0)e_2(0) = \gamma(0) + \|\ddot{\gamma}(0)\|^{-2}\ddot{\gamma}(0).$$

If we agree that the curvature of a circle of radius r is r^{-1} (which is reasonable from a geometric perspective), then the geometric content of $\|\ddot{\gamma}(0)\|$ is that of curvature.

Remark 0.5. Observe that these notions are indeed Euclidean invariants in the following sense: if $T : \mathbb{E}^m \to \mathbb{E}^m$ is Euclidean transformation, then γ and $T\gamma$ have the same curvature and T sends the osculating plane of γ at t_0 to the one of $T\gamma$ at t_0.

Surfaces in space. We now focus on surfaces in $\mathbb{E}^3$. Let us first give a definition that is easy to grasp (when the inner product is irrelevant we revert to the notation $\mathbb{R}^3$).

Definition 0.6 (Surface in $\mathbb{R}^3$)**.** A subset $S \subset \mathbb{R}^3$ is called a *differentiable surface* if it is locally the graph of a differentiable function. Precisely, if $p \in S$, then after a permutation of the coordinates, there exist an open subsets $U \subset \mathbb{R}^2$ and $I \subset \mathbb{R}$ and a (differentiable) function $f : U \to I$ such that $S \cap (U \times I)$ contains p and is the graph of f.

This definition has also some disadvantages. For instance, it is not a priori clear that a linear transformation takes a differentiable surface to a another such. From this point of view the following definition is better:

Definition 0.7 (Surface in $\mathbb{R}^3$, second version)**.** A subset $S \subset \mathbb{R}^3$ is called a *surface* if it is locally the zero set of a differentiable function without singular points. Precisely, if $p \in S$, then there exist an open neighborhood $N \subset \mathbb{R}^3$ of p and a differentiable function $\phi : N \to \mathbb{R}$ with the property that its three partial derivatives have no common zero and $S \cap N$ is the zero set of ϕ.

If S satisfies the first definition, then it clearly satisfies the second: take $N = U \times I$ and $\phi(x, y, z) = z - f(x, y)$. To see the converse, apply the inverse function theorem.

Tangent plane. The surface S has at $p \in S$ a tangent plane T_pS. We usually think of this plane as having its origin at p, but occasionally we displace this plane to the origin so that it becomes a linear subspace of $\mathbb{R}^2$. In terms of the first definition this plane is given as the graph of the best first order approximation of f:

$$(x, y) \mapsto (x_0 + (x - x_0)\frac{\partial f}{\partial x}(x_0, y_0), y_0 + (y - y_0)\frac{\partial f}{\partial y}(x_0, y_0)),$$

where $p = (x_0, y_0, z_0)$. In terms of the second definition it is the kernel of the derivative of ϕ at p displaced to p, i.e., the set of $(x, y, z) \in \mathbb{R}^3$ satisfying

$$(x - x_0)\frac{\partial \phi}{\partial x}(p) + (y - y_0)\frac{\partial \phi}{\partial y}(p) + (z - z_0)\frac{\partial \phi}{\partial z}(p) = 0.$$

Exercise 4. Prove that the tangent plane has the property that if $\gamma : I \to S$ is an arc with $\gamma(t_0) = p$, then the line $t \mapsto p + (t - t_0)\dot{\gamma}(t_0)$ stays in T_pS. In particular, $\dot{\gamma}(t_0)$, when displaced over p, may be thought of as a vector of T_pS. Prove that every vector of T_pS is so obtained. (This leads to an even more intrinsic definition of T_pS as the space of first order displacements of p inside S.)

Tangent planes of a surface appear as the domains of derivatives of a function on that surface:

Lemma-definition 0.8. *Let $S \subset \mathbb{R}^3$. We say that a function $\phi : S \to \mathbb{R}$ is* differentiable *if is satisfies the following two equivalent properties:*

(i) *every $p \in S$ admits an open neighborhood N in $\mathbb{R}^3$ such that $\phi|S \cap N$ is the restriction of a differentiable function $\tilde{\phi}$ on N,*

(ii) *if S is given at $p \in S$ as the graph of an a differentiable function (say by $z = z(x, y)$), then $f(x, y) := \phi(x, y, z(x, y))$ is differentiable in (x, y).*

In that case we have for every $v = v_1e_1 + v_2e_2 + v_3e_3 \in T_pS$ (where we regard the latter as a subspace of $\mathbb{R}^3$),

$$v_1\partial_x\tilde{\phi}(p) + v_2\partial_y\tilde{\phi}(p) + v_3\partial_z\tilde{\phi}(p) = v_1\partial_x f(p) + v_2\partial_y f(p)$$

and the linear function $T_pS \to \mathbb{R}$ defined by this expression only depends on ϕ; we call it the derivative *of ϕ at p and denote it by $D_p\phi : T_pS \to \mathbb{R}$ (or by $d\phi(p)$, in which case we usually call it the* differential *of ϕ at p).*

Proof. That (i) implies (ii) is clear. If we are in the situation of (ii), then $f(x, y) = \phi(x, y, z(x, y))$, but now regarded as a function in x, y, z (which happens to be constant in z) is an extension $\tilde{\phi}$ of ϕ as desired in (i). The rest is an exercise. □

This notion extends in an obvious manner to a map $\Phi : S \to \mathbb{R}^N$: it is said to be differentiable simply if each component is and the derivative $D_p\Phi$ is now a linear map from T_pS to $\mathbb{R}^N$. If $N = 3$ and $\Phi(S)$ lies in a surface $S' \subset \mathbb{R}^3$, then this derivative takes its values in $T_{\Phi(p)}S'$. So when Φ is viewed as map $S \to S'$, then its derivative at p may be viewed as a linear map $D_p\Phi : T_pS \to T_{\Phi(p)}S'$.

The Gauß map. Our goal is to investigate properties of surfaces in $\mathbb{E}^3$ that are Euclidean invariants in the sense of Remark 0.5.

Let $S \subset \mathbb{E}^3$ be a surface. We suppose S endowed with an *orientation*, which we here define as being given by a unit normal field N along S which is differentiable, when viewed as a map from S to the unit sphere $S^2 \subset \mathbb{E}^3$. (This need not exist, witness the Möbius band, and if such an N exists, then $-N$ is another choice, called the *opposite orientation*.) We call

$$N : S \to S^2 \subset \mathbb{E}^3$$

the *Gauß map* of the oriented surface S.

Let us observe that when $q \in S^2$, the tangent plane T_qS^2, when viewed as a subspace of $\mathbb{E}^3$, is the orthogonal complement of q. If $p \in S$ and $q = N(p)$, then this orthogonal complement is just T_pS. Hence the derivative of N at p may be regarded as an endomorphism of T_pS. We will find that its derivative tells us a lot about the way S is curved.

Since we wish to study S near a given point $p \in S$, we may (after perhaps performing a Euclidean transformation) assume that $p = 0$ and that $N(p) = e_3$ (so that T_pS is given by $z = 0$). Hence S is near p given as the graph of a function $f : U \to \mathbb{R}$ defined on a neighborhood U of $(0,0)$ in $\mathbb{E}^2$ with $f, \partial_x f, \partial_y f$ all vanishing in $(0,0)$. So the Taylor development of f at p begins with a quadratic form in x and y:

$$f(x,y) := \tfrac{1}{2}\partial_x^2 f(0,0)x^2 + \partial_x\partial_y f(0,0)xy + \tfrac{1}{2}\partial_y^2 f(0,0)y^2 + O(\|(x,y)\|^3).$$

In terms of our parameterization of S, the Gauß map N is given by

$$N(x,y) := \frac{-\partial_x f(x,y)e_1 - \partial_y f(x,y)e_2 + e_3}{\sqrt{\partial_x f(x,y)^2 + \partial_y f(x,y)^2 + 1}}.$$

The derivative of N at p is a linear transformation of the plane $z = 0$ to itself whose matrix turns out to be minus the Hessian matrix of f at $(0,0)$:

Lemma 0.9. *The derivative matrix of* $N = (N_1, N_2, N_3)$ *at* $p = (0,0,0)$ *is*

$$\begin{pmatrix} \partial_x N_1(0,0) & \partial_y N_1(0,0)) \\ \partial_x N_2(0,0) & \partial_y N_2(0,0)) \\ \partial_x N_3(0,0) & \partial_y N_3(0,0)) \end{pmatrix} = -\begin{pmatrix} \partial_x^2 f(0,0) & \partial_y\partial_x f(0,0) \\ \partial_x\partial_y f(0,0) & \partial_y^2 f(0,0) \\ 0 & 0 \end{pmatrix}.$$

Proof. An easy computation. □

It is worthwhile to state this in coordinate independent form as follows: the derivative of N at p is a linear map $D_pN : T_pS \to T_pS$ and we have

$$\langle D_pN(v), v'\rangle = -\partial_v\partial_{v'} f(0,0)$$

(if we let v and v' run over e_1, e_2 we get the above matrix identity). This shows in particular that the left hand side is symmetric in v and v' (in other words, D_pN is self-adjoint). This quadratic form was introduced by Gauß.

Definition 0.10 (Gauß)**.** The *second fundamental form* of S at p is the quadratic form on T_pS defined by $(v, v') \in T_pS \times T_pS \mapsto \langle D_pN(v), v'\rangle \in \mathbb{R}$.

We investigate its geometric properties. Given a unit vector $v \in T_pS$, let $\gamma : I \to S$ be an arc parameterized by arc length with $\gamma(0) = p$ and $\dot\gamma(0) = v$. Recall that the curvature of γ at 0 is by definition the length of the vector $\ddot\gamma(0)$. From the fact that $\gamma_3(t) = f(\gamma_1(t), \gamma_2(t))$, we easily derive that

$$\ddot\gamma_3(0) = \partial_v\partial_v f(0,0).$$

The lefthand side is the vertical component of $\ddot\gamma(0)$ and the right hand side equals $-\langle D_pN(v), v\rangle$. So we may write this as

$$\langle\ddot\gamma(0), N(p)\rangle = -\langle D_pN(v), v\rangle.$$

The absolute value of the left hand side is maximal if and only if $\ddot\gamma(0)$ is proportional to $N(p)$. So this amounts to the osculating plane of γ at 0 being $\gamma(0) + \mathbb{R}\dot\gamma(0) + \mathbb{R}N(p)$. It is also equivalent to $\ddot\gamma_1(0) = \ddot\gamma_2(0) = 0$. We sum up our findings in coordinate free terms and use the occasion to make a definition.

Lemma-definition 0.11. *Let $\gamma : I \to S$ be an arc parameterized by arc length. We define the* normal curvature *of γ relative to S at $t \in I$ as $\langle \ddot{\gamma}(t), N \circ \gamma(t) \rangle$. Its absolute value equals the curvature of γ at t if and only if the osculating plane of γ at t contains $N \circ \gamma(t)$; if that is the case for all $t \in I$, then we say that γ is a* geodesic.

For every unit tangent vector $v \in T_pS$ there is an arc $\gamma : I \to S$ parameterized by arc length with $0 \in I$, $\gamma(0) = p$, $\dot{\gamma}(0) = v$ and $\ddot{\gamma}(0)$ normal to S at p. This γ is unique up to second order and in that case its curvature at 0 is given by $|\langle \partial_v N(p), N(p) \rangle|$.

Recall that a symmetric matrix admits an orthonormal basis of eigen vectors (in particular, its eigen values are all real). So if we denote by $k_1 \geq k_2$ the eigen values of D_pN, then after a rotation around the z axis we can arrange that e_i is an eigen vector with eigen value k_i $(i = 1, 2)$. So then

$$\partial_x^2 f(0,0) = -k_1,\ \partial_y \partial_x f(0,0) = 0,\ \partial_y \partial_y f(0,0) = -k_2.$$

In particular, the surface S is near p given up to order two as the graph of the quadratic function $z = -\frac{1}{2}k_1 x^2 - \frac{1}{2}k_2 y^2$.

Definition 0.12. The *principal curvatures* of S at p are the two eigen values of D_pN. We say that p is an *umbilical point* of S if the two principal curvatures are equal. If they differ, then the tangent plane T_pS gets decomposed into two mutually perpendicular lines, called the *lines of curvature*.

The *Gauß curvature of S at p* is the product of the principal curvatures (which is in fact the determinant of D_pN) and is usually denoted $K(p)$. We say that p is *elliptic, parabolic, hyperbolic* according to whether the Gauß curvature is there > 0, $= 0$, < 0.

Notice that if we reverse orientation (i.e., N is replaced by $-N$), then the principal curvatures change sign, but the other notions introduced here (Gauß curvature, umbilical point, lines of curvature) are unaffected.

Since the Gauß curvature of S at p is the determinant of the Jacobian matrix $D_pN : T_pN \to T_pN = T_{N(p)}S^2$, its absolute value $|K(p)|$ may be understood as the scalar by which the area of a small area element on S at p is multiplied, when N is applied to it.

The average of the principal curvatures $\frac{1}{2}(k_1 + k_2)$ is called the *mean curvature of S at p*.

Exercise 5. (a) Suppose that there is an arc through $p \in S$ parameterized by arc length which has zero curvature. Prove that the Gauß curvature of S at p is ≤ 0.
(b) Let S be surface in $\mathbb{E}^3$ invariant under the translations in a fixed direction. Prove that such a surface has zero Gauß curvature.

Exercise 6. Compute the Gauß curvature of a sphere in $\mathbb{E}^3$ of radius R. Give an a priori reason for the fact that it is constant.

Exercise 7. The standard torus T in $\mathbb{E}^3$ with radii $R > r > 0$ is obtained as follows. Take the unit circle in the plane spanned by e_1 and e_3 centered at

Re_1 and of radius r. Then T is the surface of revolution of this circle around the e_3-axis. Describe the locus where T has positive, zero and negative Gauß curvature. Give an a priori reason for the fact that these subsets are invariant under rotation around the e_3-axis.

Exercise 8. Let $0 < a < b$ and let $f : (a, b) \to \mathbb{R}$ be a differentiable function with $\dot{f} > 0$. The *surface of revolution* S_f defined by f is the surface defined by $z = f(\sqrt{x^2 + y^2})$. Let $p : (a, b) \to S_f$ be given by $p(t) := te_1 + f(t)e_3 \in S_f$.
(a) Show that e_2 and $\dot{p}(t)$ span lines of curvature at $p(t)$ (when defined).
(b) Choose N to be the unit normal field over S that has positive inner product with e_3. Compute the normal curvature of the arc through $p(t)$ that traverses the circle through $p(t)$ in the plane $z = f(t)$ centered at $f(t)e_3$.
(c) Prove that the Gauß curvature at $\gamma(t)$ has the same sign as $\ddot{f}(t)$.

Intrinsic surface geometry. By this we refer to notions and properties that only depend on the notion of length on S as measured along the surface. Since for a differentiable arc $\gamma : [a, b] \to S$ its length is given by $L(\gamma) = \int_a^b \|\dot{\gamma}(t)\| dt$, this means that we only need to know the norm of any tangent vector of S. But norm and inner product determine each other and so this is equivalent to knowing for every $p \in S$ the inner product on the tangent space T_pS. Gauß proved that the curvature named after him is intrinsic (we will prove that later), this in contrast with the mean curvature.

Let us clarify the preceding in terms of our model, where S is given as $z = f(x, y)$. The tangent space of S at $(x, y, f(x, y))$ is, when displaced to the origin, the span of the vectors $\tilde{e}_1(x, y) := e_1 + \partial_x f(x, y)e_3$ and $\tilde{e}_2(x, y) = e_2 + \partial_y f(x, y)e_3$. These are the images of the vectors e_1 and e_2 under the derivative of the map $(x, y) \mapsto (x, y, f(x, y))$. We have $\langle \tilde{e}_1(x, y), \tilde{e}_1(x, y)\rangle = 1 + (\partial_x f(x, y))^2$, $\langle \tilde{e}_2(x, y), \tilde{e}_2(x, y)\rangle = 1 + (\partial_y f(x, y))^2$ and $\langle \tilde{e}_1(x, y), \tilde{e}_2(x, y)\rangle = \partial_x f(x, y)\partial_y f(x, y)$. This is better expressed by writing the inner product in the tangent spaces as a quadratic form:

$$(dx)^2 + (dy)^2 + (df)^2 = E(x, y)(dx)^2 + 2F(x, y)dxdy + G(x, y)(dy)^2,$$

where

$$\begin{aligned} E(x, y) &:= 1 + (\partial_x f(x, y))^2, \\ F(x, y) &:= \partial_x f(x, y)\partial_y f(x, y), \\ G(x, y) &:= 1 + (\partial_y f(x, y))^2. \end{aligned}$$

Here the expressions dx and dy are to be understood as *differentials*. This is a rather subtle concept, usually passed over in a basic calculus course, and so we here take some time to explain (for a bona fide definition—which is indispensable later anyway—we refer to the Notes of a Differentiable Manifolds course or any text book on that subject).

A vector field in an open subset of (x, y)-plane can be given $v(x, y) = v_1(x, y)e_1 + v_2(x, y)e_2$, but since we prefer to write ∂_x for the constant unit vector field e_1 and likewise ∂_y for e_2, we write v as $v(x,y) = v_1(x,y)\partial_x + v_2(x,y)\partial_y$. We define dx as a function on such vector fields: it assigns to ∂_x the value 1 and to ∂_y the value 0 and is 'linear with respect to coefficients' in the sense that it assigns to the vector field v above the function $v_1(x,y)$. The differential dy is similarly defined. In a sense the pair of differentials (dx, dy) is a basis dual to the basis of vector fields (∂_x, ∂_y).

Now f defines a function of vector fields by the rule $v \mapsto \partial_v f$. Since it assigns to ∂_x the function $\partial_x f$ and to ∂_y the function $\partial_y f$, it can be represented by $\partial_x f dx + \partial_y f dy$. This is what we call the *total differential* of f. It is a function on vector fields which we denote by df.

The derivative of $(x,y) \mapsto (x, y, f(x,y))$ maps the vector field v on S to the vector field $\tilde{v}$ on S given by $\tilde{v}(x,y) = v(x,y) + df(v(x,y))\partial_z$. If $v' = v'(x,y)$ is another vector field, then

$$\langle \tilde{v}, \tilde{v}' \rangle = \langle v, v' \rangle + df(v).df(v') = dx(v).dx(v') + dy(v).dy(v') + df(v).df(v')$$

as functions in (x,y). In particular,

$$\langle \tilde{v}, \tilde{v} \rangle = dx(v)^2 + dy(v)^2 + df(v)^2,$$

which is indeed the value of the quadratic form $E(dx)^2 + 2F dx dy + G(dy)^2$ on v. So the latter expression defines an inner product g on vector fields by

$$g(v, v') := E dx(v) dx(v') + F dx(v) dy(v') + F dy(v) dx(v') + G dy(v) dy(v')$$
$$(= E v_1 v_1' + F v_1 v_2' + F v_2 v_1' + G v_2 v_2').$$

A differentiable arc $\gamma = (\gamma_1, \gamma_2) : [a,b] \to \mathbb{R}^2$ mapping to the domain of f is the projection of the arc $\tilde{\gamma} : [a,b] \to S$ with $\tilde{\gamma}_3(t) = f(\gamma_1(t), f_2(t))$ and we have that the norm of $\dot{\tilde{\gamma}}(t)$ is the norm of $\dot{\gamma}(t)$ measured with g: $\sqrt{g(\dot{\gamma}(t), \dot{\gamma}(t))}$ so that

$$L(\tilde{\gamma}) = \int_a^b \sqrt{g(\dot{\gamma}(t), \dot{\gamma}(t))}\,dt.$$

The upshot of the preceding the discussion is that for the purposes of intrinsic geometry, we may as well work with an open subset U of $\mathbb{R}^2$ on which we are given the inner product defined by the quadratic form $E(x,y)(dx)^2 + 2F(x,y)dxdy + G(x,y)(dy)^2$. There is no need anymore to assume that it has the special form above: it is enough to know that it defines an inner product at all points. This means that E, F and $EG - F^2$ must be positive everywhere. We then say that g is a *Riemannian metric* on U. This observation leads to the more general (hence more abstract) notion of a *Riemannian manifold* that we will introduce in the next section. It was conceived by Riemann himself in his inaugural lecture in 1854. Its proper formulation requires the a priori simpler differential-topological notion of a *manifold*, which itself took more than half a century to evolve. For us this means that as of this point we need to assume that you have some basic acquaintance with differential-topology.

1. Riemannian manifolds

We recall that from now on we assume the reader is familiar with the basic notions of differential topology. The *Smooth Manifolds* notes on
`<http://www.math.uu.nl/people/looijeng/coursenotes.html>`
more than suffice. On that webpage you will also find a synopsis (in development) tailored to what we will need here.

If M is a submanifold of $\mathbb{E}^N$, then the tangent space of M at $p \in M$ can be understood as a linear subspace of $\mathbb{E}^N$ and that makes it an inner product space. Thus the tangent bundle $TM \to M$ acquires an inner product. If $\gamma : I \to M$ is a differentiable arc M, then its arc length is entirely determined by the this inner product, as we can see from its definition that this only requires knowing the norm of the tangent vectors $\dot\gamma(t)$. This suggests:

Definition 1.1. Een *Riemannian metric* on a manifold M is an inner product g_p on every tangent space T_pM which depends differentiably on p (in the sense that it assumes on the domain of a chart (U,κ) the form $\sum_{i,j} g_{ij} d\kappa^i d\kappa^j$, where for every $p \in U$, $(g_{ij}(p))$ is a positive definite symmetric matrix and $g_{ij} : U \to \mathbb{R}$ is differentiable.) A manifold endowed with a Riemann metric is called a *Riemannian manifold*.

This notion is clearly so devised that given such a Riemannian manifold M, we can define the *arc length* $L(\gamma)$ of an arc $\gamma : [a,b] \to M$ as the integral van $\|\dot\gamma\|$ over $[a,b]$. Proposition 0.1 (or rather its proof) shows that arc length is insensitive under reparametrization. The range of this definition is extended in a straightforward manner to *piecewise differentiable arcs* (continuous arcs that fail to be differentiable in only finitely many points). We use this to show that the Riemann metric of a connected Riemannian manifold metrizes the underlying topological space.

Exercise 9. Prove that a Riemannian manifold which is diffeomorphic to an open interval, is in fact isometric to an open interval in $\mathbb{E}^1$. (Hint: parameterise by arc length.)

Proposition 1.2. *Let* M *be a connected Riemann metric. Then a metric* $d : M \times M \to \mathbb{R}$ *on* M *is defined (whose underlying topology is the given topology of* M*) by letting* $d(p,q)$ *be the infimum of the arc lengths of piecewise differentiable arcs in* M *that connect* p *with* q.

Proof. Let us first verify that d defines a metric. The only nonobvious property to check here is that $p \neq q$ implies $d(p,q) > 0$. For this we choose a chart $\kappa : U \to \mathbb{R}^m$ at p with $q \notin U$, $\kappa(p) = 0$ and such that $\kappa(U)$ contains the closed unit ball $B \subset \mathbb{R}^m$. The Riemann metric g is then written $\sum_{i,j} g_{ij} d\kappa^i d\kappa^j$. We compare this with the Riemann metric g' on U that is provided by $\mathbb{E}^m = \mathbb{R}^m$: $g' = \sum_i (d\kappa^i)^2$. We claim that there exist positive constants c_{min}, c_{max} such that for every tangent vector $v \in TM|\kappa^{-1}B$, we have

$$c_{\min}^2 g'(v,v)\rangle \leq g(v,v) \leq c_{\max}^2 g'(v,v).$$

To see the existence of c: think of $B \times S^{m-1}$ as the set of unit tangent vectors in $T\mathbb{E}^m|B$. This is evidently compact and hence so is its image $K \subset TU$ under the derivative $D(\kappa^{-1}) : \kappa(U) \times \mathbb{R}^m \to TU$. The function $v \in K \mapsto g(v,v)$ is continuous and so has a minimum $c_{\min}$ and a maximum $c_{\max}$. Both are positive and a simple homogeneity argument shows that they have the desired property.

Any piecewise differentiable arc γ in M from p to q has an initial piece γ' that connects p with some point $q' \in \partial B$. Thus $L(\gamma) \geq L(\gamma')$, and the latter is estimated from below by the integral of $c_{\min}^{-1}\|(\kappa\dot{\gamma}')(t)\|$, i.e., by $c_{\min}^{-1}$ times the arc length of $\kappa\gamma'$ in $\mathbb{E}^m$. Since $\kappa\gamma'$ is an arc from the origin to a point on the unit sphere, we find that $L(\gamma') \geq c_{\min}^{-1}$. We conclude that $d(p,q) \geq c_{\min}^{-1} > 0$.

That same argument shows that in fact for *any* pair $q, q' \in \kappa^{-1}(B)$ we have $d(q,q') \geq c_{\min}^{-1}\|\kappa(q) - \kappa(q')\|$. On the other hand, a parameterization of the straight line connecting $\kappa(q)$ with $\kappa(q')$ (which lies in B) composed with κ^{-1} yields an arc from q to q' whose arc length is $\leq c_{\max}$ times the distance between $\kappa(q)$ and $\kappa(q')$ so that we also have $d(q,q') \leq c_{\max}\|\kappa(q)-\kappa(q')\|$. It follows that the d-topology on $\kappa^{-1}B$ equals the topology of B induced by κ . So the d-topology is the given topology. □

A map f from a Riemannian manifold (M,g) to a Riemannian manifold (N,h) is called an *isometric immersion* if its derivative respects the inner product of tangent vectors: $h(D_pf(v), D_pf(w)) = g(v,w)$ for all $v,w \in T_pM$, $p \in M$. Such a map is an immersion indeed, for if $v \in T_pM$ is such that $D_pf(v) = 0$, then $g(v,v) = 0$, and hence $v = 0$. We may also reverse this reasoning: given an immersion $f : M \to N$ and a Riemann metric h on N, then we have an *induced* Riemann metric $g = f^*h$ on M defined by $g(v,w) = h(D_pf(v), D_pf(w))$. A special case is a submanifold of a Euclidean space (that gave rise to the notion of Riemann metric in the first place). We thus obtain essentially all possible Riemann metrics, for according to a deep result of John Nash every Riemannian manifold admits a local-isometric embedding in some $\mathbb{E}^N$.

An *isometric embedding* resp. *isometry* is an isometric immersion, that is also an embedding resp. a diffeomorphism. The isometries of a Riemannian manifold (M,g) onto itself make up a group (under composition) $\mathrm{Aut}(M,g)$, its *isometry group*. In general this group is small, but below we list some examples for which this group is in the first three cases quite large.

Examples 1.3. (i) *Flat space.* Euclidean space $\mathbb{E}^m$ has as its isometries the transformations that are an orthogonal transformation followed by a translation (i.e., $x \in \mathbb{E}^m \mapsto x + a \in \mathbb{E}^m$ for some $a \in \mathbb{E}^m$). If a Riemannian m-manifold is isometric to $\mathbb{E}^m$, we sometimes also call it *flat m-space*.

(ii) *Round spheres.* The unit sphere $\mathbb{S}^m$ of $\mathbb{E}^{m+1}$, endowed with the induced Riemann metric. Any orthogonal transformation of $\mathbb{E}^{m+1}$ preserves $\mathbb{S}^m$ and yields an isometry of it. The same applies to $\mathbb{S}^m(r) \subset \mathbb{E}^{m+1}$, the sphere of radius r. If a Riemannian m-manifold is isometric to $\mathbb{S}^m(r)$, we sometimes also call it a *round* m*-sphere of radius* r.

(iii) *Hyperbolic space.* Consider the quadratic form $Q(x, y, z) = x^2+y^2-z^2$ on $\mathbb{R}^3$. The zero set of $Q + 1$ is a two-sheeted hyperboloid. Let $\mathbb{H}^2$ be the sheet that lies in the half space $z > 0$, hence is given by $z = \sqrt{x^2 + y^2 + 1}$. The form Q determines a quadratic form $q := (dx)^2 + (dy)^2 - (dz)^2$ on the tangent space $\mathbb{R}^3$. This fails to be a Riemann metric (as it is not positive definite), but its restriction to $\mathbb{H}^2$ is: for $p \in \mathbb{H}^2$, $T_p\mathbb{H}^2$ is the plane in $T_p\mathbb{R}^3$ defined by putting $dQ_p = 2xdx + 2ydy - 2zdz$ equal to zero. So $dz = (xdx + ydy)/z$ on $T\mathbb{H}^2$ and substituting this in q yields

$$q_{|T\mathbb{H}^2} = (1 - \frac{x^2}{z^2})(dx)^2 + 2\frac{xy}{z^2}dxdy + (1 - \frac{y^2}{z^2})(dy)^2.$$

The right hand side is also equal to $z^{-2}\{(dx)^2 + (dy)^2 + (ydx + xdy)^2\}$, and hence positive definite. Every linear transformation of $\mathbb{R}^3$ which preserves $\mathbb{H}^2$, preserves Q, and hence also the Riemann metric on $\mathbb{H}^2$. Consequently such a transformation yields an isometry of $\mathbb{H}^2$. We call $\mathbb{H}^2$ the *hyperbolic plane*. The preceding is generalized in a straightforward manner to the case of arbitrary dimension m: we can define *hyperbolic* m*-space* as

$$\mathbb{H}^m = \{x \in \mathbb{R}^{m+1} : x^{m+1} = \sqrt{(x^1)^2 + \cdots + (x^m)^2 + 1}\}$$

with the obvious Riemann metric. Every linear transformation of $\mathbb{R}^{m+1}$ which preserves $\mathbb{H}^m$ induces in the latter an isometry.

(iv) *Flat* m*-tori.* The unit circle $\mathbb{S}^1$ in $\mathbb{E}^2$ yields in its m-th power an embedding $(\mathbb{S}^1)^m \subset \mathbb{E}^{2m}$. If we parameterise every factor by arc length: $(\theta^1, \dots, \theta^m)$, then the metric on $(\mathbb{S}^1)^m$ is simply $(d\theta^1)^2 + \cdots + (d\theta^m)^2$. So if we now think of $(\theta^1, \dots, \theta^m)$ as the coordinates of $\mathbb{E}^m$, then we see that $(\mathbb{S}^1)^m$ is also obtained from $\mathbb{E}^m$ as its orbit space under the group of translations generated by $2\pi e_i$, $i = 1, \dots, m$. We may in fact take any additive subgroup $L \subset \mathbb{E}^m$ generated by a basis of $\mathbb{E}^m$ (such a subgroup is often called a *lattice*)and find that the quotient $\mathbb{E}^m/L$ is a compact Riemannian manifold. This manifold is diffeomorphic but not necessarily isometric to $(\mathbb{S}^1)^m$. If a Riemannian m-manifold is isometric to $\mathbb{E}^m/L$ for some lattice L, we sometimes also call it a *flat* m*-torus*.

Exercise 10. Let M be one of the Riemannian manifold described in Examples 1.3-(i),(ii),(iii). Prove that for any pair $p, p' \in M$ and orthonormal bases $(e_1, \dots, e_m)$ resp. $(e'_1, \dots, e'_m)$ of T_pM resp. $T_{p'}M$ there exists exactly *one* isometry of the described type that maps p to p' and whose derivative at p sends e_i to e'_i $(i = 1, \dots m)$. (We shall later show that an isometry of a connected Riemann manifold is completely determined by its derivative at any pre-assigned point so that there are no other isometries than the ones specified. Notice that in all three cases the group in question has dimension $m + \dim O(m) = \frac{1}{2}m(m+1)$.)

Riemannian geometry is by definition concerned with properties that are invariant under isometries, in other words, properties that can be expressed in principle solely in terms of the topology, the differentiable structure, and the Riemann metric .

We remark that every compact manifold M admits a Riemann metric: if $\{\phi_i : M \to [0,1]\}_{i=1}^n$ is a partition of 1 for which the support of every ϕ is contained in the domain of a chart κ_i, then

$$\sum_i \phi_i(p)\langle d\kappa_i(v), d\kappa_i(w)\rangle,$$

($p \in M$ en $v, w \in T_pM$) is a Riemann metric on M. (This proof can be generalized to the case when M, instead of being compact, has only a countable basis for its topology.)

Remark 1.4. There are many areas outside mathematics where Riemannian spaces occur naturally. Here is an example in classical mechanics. Let be given N particles in $\mathbb{E}^3$ with masses $m_1, \dots, m_N$. If $q_i(t) \in \mathbb{E}^3$ is the position of the ith particle at time t, then $q(t) = (q_1(t), \dots, q_N(t)) \in (\mathbb{E}^3)^N$ describes the position of the whole system. The mechanical constraints of the system (for instance, some of these particles could be connected by rigid rods of negligible mass) force $q(t)$ to lie on a fixed subset $M \subset (\mathbb{E}^3)^N$, called its *configuration space*. Let us assume that M is a submanifold of $(\mathbb{E}^3)^N$. The velocities $\dot{q}_i(t)$ make up a vector $\dot{q}(t) = (\dot{q}_1(t), \dots, \dot{q}_N(t)) \in (\mathbb{E}^3)^N$ that is in fact a tangent vector of M at $q(t)$. The kinetic energy of the system at time t is $\sum_i \frac{1}{2}m_i\|\dot{q}_i(t)\|^2$. So this suggests to endow $(\mathbb{E}^3)^N$ with the Riemann metric $g(v, v') = \sum_i m_i\langle v_i, v_i'\rangle$ and then to restrict this metric to M. The kinetic energy is now given by $v \in TM \mapsto \frac{1}{2}g(v, v)$.

2. Connections

Definition 2.1. Let $\xi : E \to M$ be a vector bundle. A *connection* ∇ on ξ is a prescription for differentiating a section of ξ with respect to a vector field on M. Precisely, given a vector field X on an open subset U of M, then ∇_X assigns to any section s of $\xi|U$ another section $\nabla_X s$ of $\xi|U$ (so that $\nabla_X : \mathcal{E}(U,\xi) \to \mathcal{E}(U,\xi)$), in such a manner that the following four conditions are fulfilled:

local nature: for $U' \subset U$ open we have $\nabla_X s|U' = \nabla_{X|U'}(s|U')$,
linearity: ∇_X is $\mathbb{R}$-linear,
variance in X**:** for every function $f : U \to \mathbb{R}$, $\nabla_{fX} = f\nabla_X$,
Leibniz rule: for every function $f : U \to \mathbb{R}$ we have

$$\nabla_X(fs) = f\nabla_X s + X(f)s.$$

The first condition expresses the fact that the value of $\nabla_X s$ in a given p only depends on the restriction of X and s to an arbitrary (small) neighborhood of p. We shall shortly see that much more is true, namely that $(\nabla_X s)(p) \in E_p$ only depends on $D_p s(X_p) \in T_{s(p)}E$. We shall also give a more geometric characterization of a connection.

A very simple (but important) example is the following. Take the trivial bundle $\mathbb{R} \times M \to M$ (so that a section over $U \subset M$ is just a function $f : U \to \mathbb{R}$. Then $\nabla_X f := X(f)$ defines a connection on this bundle.

Let be given a trivialization of $\xi|U$, so that the we have a basis $(e_1, \dots, e_r)$ of sections of $\xi|U$ (r is the rank of ξ). Assume also that U is the domain of a chart of M and write ∂_i for the basis vectorfields $\partial/\partial\kappa^i$ (so that for a function f on U we have $df = \sum_i \partial_i(f) d\kappa^i$). Then we define functions $\Gamma^\sigma_{\rho,i}$ on U the equation

$$\nabla_{\partial_i} e_\rho = \sum_{\sigma=1}^{r} \Gamma^\sigma_{\rho,i} e_\sigma.$$

The functions $\Gamma^\sigma_{\rho,i}$ are called the *Christoffel symbols* of the connection relative to the given bases. They completely determine the connection on U because any vector field X on U is written as $\sum_i X^i \partial_i$ and any section s of $\xi|U$ is written as $\sum_\rho s^\rho e_\rho$ so that the defining properties of a connection yield

$$\nabla_X s = \sum_{\sigma=1}^{r} \Big(\sum_{i=1}^{m} \sum_{\rho=1}^{r} X^i \Gamma^\sigma_{\rho,i} s^\rho + X(s^\sigma) \Big) e_\sigma. \tag{1}$$

Notice that in the right hand side X and s only enter via their coefficients X^i, s^ρ and the derivatives $X(s^\rho)$. This is made more transparant by introducing the so-called *connection matrix* of ∇ relative to the basis of sections $(e_1, \dots, e_r)$. It is the matrix of differentials on U defined by

$$A = (A^\sigma_\rho) \quad \text{with} \quad A^\sigma_\rho := \sum_{i=1}^{m} \Gamma^\sigma_{\rho,i} d\kappa^i.$$

This matrix is no longer dependent on the chart κ, for we have $\nabla_X e_\rho = \sum_\sigma A^\sigma_\rho(X)$ (a differential being evaluated on vector field). Hence

$$\nabla_X s = \sum_{\sigma=1}^{r} \Big(\sum_{\rho=1}^{r} s^\rho A^\sigma_\rho(X) + X(s^\sigma) \Big) e_\sigma, \text{ or even shorter}$$

$$\nabla_X s = A(X)(s) + \sum_\sigma X(s^\sigma) e_\sigma. \tag{1'}$$

This leads to the following important observation: *the value of* $\nabla_X s$ *in* p *only depends on* $D_p s(X_p) \in T_{s(p)}E$. In fact, we obtain for every $e \in E_p$ a linear retraction $r_e : T_e E \to T_e E_p = E_p$ characterized by $(\nabla_X s)(p) = r_e D_p s(X_p)$. In terms of the above trivialization (which identifies E_U with $\mathbb{R}^r \times U$ and hence $T_e E$ with $\mathbb{R}^r \times T_p M$) this retraction is at e given by $(u, v) \mapsto A_p(v)(e) + u$.

It is easy to see that conversely, any $r \times r$ matrix of differentials on U defines a connection on the trivialized vector bundle $\mathbb{R}^r \times U \to U$.
We say that a local section s of ξ is *flat* if for every vector field X we have $\nabla_X s = 0$. It suffices to verify this for a basis of vector fields, for instance, in terms of the above chart, for $\partial_1, \dots, \partial_m$. We then see that the r component functions $s^1, \dots, s^r$ must satisfy mr first order differential equations and this indicates that for $m > 1$ we cannot expect nonzero flat sections to exist. In the (extraordinary) case that any point of M, ξ has a basis of flat local sections, we say that ∇ is *flat*.

If a trivialization $(e_1, \dots, e_r)$ of $\xi|U$ is given, then the connection on $\xi|U$ with zero matrix of differentials defines a connection having the constant sections (i.e., with the s^ρ's constant) as flat sections, and so this connection is flat.

A sum of two connections is never a connection (unless the dimension of M or the rank of E is zero). But a convex linear combination is. In fact, if $\nabla^1, \dots, \nabla^k$ are connections on ξ, and $\phi_1, \dots, \phi_k$ are functions on M with sum constant 1, then it is easy to see that $\phi_1\nabla^1 + \cdots + \phi_k\nabla^k$ is a connection as well. From this it follows that a connection on a bundle always exists, at least when M is compact: we just saw that a local trivialization determines a local (flat) connection and by means of a partition of 1 such local connections can be used to produce one on all of ξ.

If $\dim M = 1$, then the connection is always flat:

Proposition 2.2. *If I is an interval and $a \in I$, then for every $e \in E_a$ there is precisely one flat section $s_e : I \to E$ with $s_e(a) = e$. If we let e run over E_a, then these sections define a trivialization $I \times E_a \to E$ of ξ.*

Proof. We choose a basis $e_1, \dots, e_r$ of sections of ξ over a neighborhood of a. If $s = \sum_\rho s^\rho e_\rho$ is a local section at a, then according to formula (1) the condition $\nabla_{\partial/\partial t} s = 0$ boils down to s obeying the system of ordinary differential equations

$$0 = \sum_{\rho=1}^{r} \Gamma_\sigma^\rho s_\rho + \dot{s}_\sigma, \ \sigma = 1, \dots, r.$$

We write this in vector notation:

$$\frac{ds}{dt} = -\Gamma(s),$$

where $\Gamma = (\Gamma_\rho^\sigma)$ is a matrix valued function in a single variable. According to the theory of ordinary differential equations this system has a unique local solution if we fix the value in a: $s(a) = e$. If we denote this solution by s_e, then $v \mapsto s_e$ is linear in v (for if $c, c' \in \mathbb{R}$ and $e, e' \in E_a$, then $cs_e + c's_{e'}$ also solves the system and has initial value $ce + c'e'$). Since $s_e(a) = e$, we have that for t in a neighborhood U_a of a in I, $e \mapsto s_e(t)$ is a linear isomorphism of E_a onto E_t. So the flat sections define a trivialization of ξ over E_a. This implies the assertion on U_a. By letting a run over I we find a covering of I

by solution intervals. The local unicity of the corresponding solutions makes these agree on overlaps and thus the assertion follows for all of I. □

Pull-back of a connection. Let us return to the general case of a vector bundle $\xi : E \to M$. Suppose we have a manifold N and a map $f : N \to M$. We recall that the *pull-back of* ξ *along* f, $f^*\xi$, is a vector bundle on N whose total space f^*E is the fiber product of E and N over M: the set of pairs $(e, q) \in E \times N$ with $e \in E_{f(q)}$. This is a submanifold of $E \times N$ and the projection on the second factor $f^*E \to N$ makes it a vector bundle. So if s is a section of $\xi|U$, then $q \in f^{-1}U \mapsto (sf(q), q) \in f^*E$ is a section of $f^*\xi|f^{-1}U$ (which we shall denote by f^*s). It is easy to see that if $(e_1, \dots, e_r)$ is a basis of sections of ξ over an open subset $U \subset M$, then $(f^*e_1, \dots, f^*e_r)$ is a basis of sections of $f^*\xi$ over an open subset $f^{-1}U \subset N$.

We claim that a connection ∇ on ξ determines a connection (denoted $f^*\nabla$) on $f^*\xi$. The simplest (but perhaps also the least useful) description is in terms of the linear retractions $r_e : T_eE \to E_{\pi(e)}$: any point of $(f^*E)_q$ is given by some $e \in E_{f(q)}$ and we then take as our retraction the composite

$$T_{(e,q)}f^*E \subset T_eE \times T_qN \xrightarrow{\text{proj}} T_eE \xrightarrow{r_e} E_{f(q)} = (f^*E)_q.$$

This is more concretely described in terms of connection forms: if $A = (A_\rho^\sigma)$ is the connection matrix of ∇ relative to a basis $(e_1, \dots, e_r)$ of $\xi|U$, then the matrix of pulled back differentials $f^*A = (f^*A_\rho^\sigma)$ is the connection matrix of $f^*\nabla$ relative to the basis $(f^*e_1, \dots, f^*e_r)$ of $f^*\xi|f^{-1}U$. To spell this out: if Y is a vector field on some open $V \subset f^{-1}U$, then for any $t = \sum_\rho t^\rho f^*e_\rho \in \mathcal{E}(V, f^*\xi)$,

$$(f^*\nabla)_Y t = \sum_\sigma \Big(\sum_\rho t^\rho (f^*A_\rho^\sigma)(Y) + Y(t^\sigma) f^*e_\sigma \Big).$$

This also shows that thus is defined a connection on $f^*\xi$.

Let us apply this to the case of an arc $\gamma : [a, b] \to M$ from $p = \gamma(a)$ to $q = \gamma(b)$. According to the preceding $\gamma^*\xi$ has a connection $\gamma^*\nabla$. By Proposition 2.2 the flat sections of $\gamma^*\nabla$ trivialize the bundle $\gamma^*\xi$; in particular we get an isomorphism of the fiber E_p onto the fiber E_q. This is called *parallel transport* along γ. Here we may even take γ piecewise differentiable (but continuous), for then we simply take the composites of the parallel transports associated to the differentiable pieces.

In general the parallel transport from p to q along γ depends on γ and not just on p and q. For instance, if we have a loop: $p = q$, then the automorphism of E_p thus obtained (called the *holonomy* of the connection along this loop) need not be the identity, even if the image of γ is contained in a small neighborhood of p. The deviation from the identity is measured infinitesimally by the *curvature form* R which enters in the next proposition.

Proposition 2.3. *Let* X *and* Y *be vector fields at* $p \in M$ *and let* s *be a local section of* ξ *at* p. *Then the value of the expression*

$$R(X, Y)s := (\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]})s$$

at p *only depends on* X, Y *and* s *at* p *(and not on any of their derivatives).*

Proof. We show that the dependence of $(R(X,Y)s)(p)$ on s is linear and via $s(p)$. According to the definition

$$\nabla_X \nabla_Y fs = \nabla_X(f\nabla_Y s + Y(f)s) = f\nabla_X\nabla_Y s + X(f)\nabla_Y s + Y(f)\nabla_X s + XY(f)s.$$

Subtract from this the equality obtained by interchanging X and Y and subsequently also the identity

$$\nabla_{[X,Y]} fs = f\nabla_{[X,Y]} s + [X,Y](f)s.$$

We obtain

$$R(X,Y)fs = fR(X,Y)s + XY(f)s - YX(f)s - [X,Y](f)s$$

and the last three terms cancel so that $R(X,Y)fs = fR(X,,Y)s$. If s' is a local section at p which takes in p the same value as s, then $u := s - s'$ can be written $\sum_\rho u^\rho e_\rho$ with $u^\rho(p) = 0$. So $R(X,Y)(s - s') = \sum_\rho u^\rho R(X,Y)e_\rho$ and clearly the last expression is zero in p. So the dependence of $(R(X,Y)s)(p)$ on s is via $s(p)$ and is linear in $s(p)$.

In the same way we find that its dependence on X and Y is via X_p and Y_p and is linear in each. □

We now may regard R as a map which assigns to every pair of tangent vectors $v, v' \in T_pM$ the endomorphism $e \in E_p \mapsto R(v,v')e \in E_p$. This map is bilinear and, as the definition clearly shows, antisymmetric. Hence R can be viewed as a vector bundle homomorphism $\wedge^2\tau_M \to \mathrm{End}(\xi)$, or equivalently, as a 2-form 'with values in $\mathrm{End}(\xi)$', (i.e., as a section of $\mathrm{End}(\xi) \otimes \wedge^2\tau_M^*$). We call R the *curvature form* of the connection.

The interpretation of R as an infinitesimal holonomy to which we alluded before is now as follows: given $v, v' \in T_pM$, choose a chart for M at p, and let for small $t, t' \in \mathbb{R}$, $\gamma_{t,t'}$ be the arc which (relative to the chart) traverses the parallellogram $(p, p + tv, p + tv + t'v', p + t'v')$. If we denote the holonomy of this arc by $\Phi_{t,t'}$, then

$$\Phi_{t,t'} = \mathbf{1} + tt'R(v,v') + o(|tt'|).$$

You are invited to check this yourself.

In the following exercises $\xi : E \to M$ is a vector bundle endowed with a connection ∇.

Exercise 11. Let $p \in M$. Prove that the holonomies of piecewise differentiable arcs from p to p make up a subgroup of $GL(E_p)$ (which is called the *holonomy group* of ξ at p).

Exercise 12. (a) Prove that the dual bundle ξ^* has a unique connection ∇ with the property that for every local vector field X and every pair of local sections s and s^* of $\xi|U$ resp. $\xi^*|U$ we have $X(s^*(s)) = (\nabla_X s^*)(s) + s^*(\nabla_X s)$.
(b) Let $\xi' = (\pi' : E' \to M)$ be another vector bundle endowed with a

connection ∇'. Prove that $\xi \otimes \xi'$ and $\mathrm{Hom}(\xi, \xi')$ have connections ∇'' and ∇''' characterized by

$$\nabla''_X(s \otimes s') = (\nabla_X s) \otimes s' + s \otimes \nabla'_X s'$$
$$\nabla'''_X(\phi)(s) = \phi(\nabla_X s) - \nabla'_X(\phi(s)).$$

Exercise 13. We denote the local sections of the vector bundle $\xi \otimes \wedge^k \tau^*_M$ by $\mathcal{E}^k(\xi)$.
(a) Show that the connection can be regarded as a map $D : \mathcal{E}(\xi) \to \mathcal{E}^1(\xi)$ which (i) is locally defined, (ii) is $\mathbb{R}$-linear and (iii) obeys the Leibniz rule $D(f.s) = fD(s) + s \otimes df$.
(b) Show that D extends uniquely to a map $D : \mathcal{E}^\bullet(\xi) \to \mathcal{E}^\bullet(\xi)$ which (i) is locally defined, (ii) is $\mathbb{R}$-linear and (iii) obeys the generalized Leibniz rule $D(s \otimes \alpha) = D(s) \wedge \alpha + s \otimes d\alpha$.
(c) We regard the curvature form R of ∇ as an element of $\mathcal{E}^2(M, \mathrm{End}(\xi))$. Prove that for every $\omega \in \mathcal{E}^\bullet(\xi)$ we have $DD(\omega) = R \wedge \omega$.
(d) Prove that $D(R) = 0$. (This is called the *Bianchi identity*.)

Exercise 14. Suppose the connection is flat.
(a) Prove that there is a bundle atlas for ξ which has constant transition functions.
(b) Prove that ∇ has curvature constant zero.

Exercise 15. Consider the trivial vector bundle $\mathbb{R}^r \times M \to M$ with connection given by the $r \times r$-matrix of differentials $A = (A^\sigma_\rho)$. Prove that the curvature of this connection is given by the $r \times r$-matrix of 2-forms $dA + A \wedge A$, where $(A \wedge A)^\sigma_\rho = (\sum_\mu A^\sigma_\mu \wedge A^\mu_\rho)$. (You may use that if α, β are differentials, then $d\alpha(X, Y) = X(\alpha(Y)) - Y(\alpha(X)) - \alpha([X, Y])$ and $(\alpha \wedge \beta)(X, Y)) = \alpha(X)\beta(Y) - \alpha(Y)\beta(X)$.)

Exercise 16. Prove a local converse of Exercise 14-(b): if M is an open hypercube and ∇ has constant zero curvature, then the connection is flat. More precisely, there is a trivialization of ξ for which the connection matrix becomes identically zero. (Hint: Project M along the last coordinate onto a hypercube of one dimension less and apply 2.2 to the fibers of this projection.)

Remark 2.4. In Yang-Mills theory a potential is best understood as a connection on a vector bundle. The curvature form of that connection has then the interpretation as the associated force field.

The simplest (and guiding) example is the electromagnetic field. In that case, we have a trivial complex vector bundle of rank 1 over Minkowski space M. Following our discussion a connection is then given (in terms of a local trivialization) by a differential A on M with complex coefficients so that is its curvature $F := dA$ (a 2-form which is independent of the local trivialization). The Maxwell equations identify this 2-form as the electromagnetic field. In that case A purely imaginary, so that $(2\pi\sqrt{-1})^{-1}A$ is an ordinary (real) differential.

3. The Levi-Civita connection

In this chapter M is a Riemannian manifold met Riemann metric g. We begin with what is often refered to as the main theorem of Riemannian Geometry.

Theorem 3.1 (Levi-Civita connection). *There is precisely one connection* ∇ *on* TM *obeying the following two identities for any triple* X, Y, Z *of vector fields on any given open subset* $U \subset M$*:*

symmetry: $\nabla_X Y - \nabla_Y X - [X, Y] = 0$ *and*
flatness metric: $X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$,

This connection is called the Riemannian *or* Levi-Civita *connection.*

Proof. We first check that these two conditions fix ∇. For X, Y, Z as in the theorem we find:

$$\begin{aligned} & X(g(Y, Z)) + Y(g(Z, X)) - Z(g(X, Y)) = \\ &= g(\nabla_X Y + \nabla_Y X, Z) + g(Y, \nabla_X Z - \nabla_Z X) + g(X, \nabla_Y Z - \nabla_Z Y) \text{ (flatness metric)} \\ &= 2g(\nabla_X Y, Z) - g([X, Y], Z) + g(Y, [X, Z]) + g(X, [Y, Z]) \text{ (by symmetry).} \end{aligned}$$

This implies

$$(\dagger) \quad g(\nabla_X Y, Z) = \frac{1}{2}\Big(X(g(Y, Z)) + Y(g(X, Z)) - Z(g(X, Y)) + g([X, Y], Z) - g(Y, [X, Z]) - g(X, [Y, Z])\Big).$$

This indeed fixes $\nabla_X Y$ (take for U the domain of a chart and let Z run over the coordinate vector fields $\partial_1, \dots, \partial_m$). If conversely, we define $\nabla_X Y$ by this formula, then it is easily verified that thus is defined a connection which satisfies the symmetry property and metric flatness. □

Formula (†) allows us to compute the Christoffel symbols of the Levi-Civita connection in terms of a chart κ. If the metric is written as $g = \sum_{i,j} g_{i,j} d\kappa^i d\kappa^j$, then the formula implies

$$g(\nabla_{\partial_i} \partial_j, \partial_k) = \frac{1}{2}\Big(\partial_i g_{j,k} + \partial_j g_{i,k} - \partial_k g_{i,j}\Big).$$

As the Christoffel symbols of the Levi-Civita connection relative to $\partial_1, \dots, \partial_m$ are defined by the identity $\nabla_{\partial_i} \partial_j = \sum_{k=1}^m \Gamma^l_{j,i} \partial_l$, we have that $g(\nabla_{\partial_i} \partial_j, \partial_k) = \sum_{l=1}^m \Gamma^l_{j,i} g_{l,k}$. Substituting this in the above formula yields

$$\sum_{l=1}^m \Gamma^l_{j,i} g_{l,k} = \frac{1}{2}\Big(\partial_i g_{j,k} + \partial_j g_{i,k} - \partial_k g_{i,j}\Big).$$

So if the matrix inverse to $(g_{k,l})$ is denoted $(g^{l,k})$, then we find what is called the *second Christoffel identity*:

$$\text{(2}^{\text{nd}}\text{ Chr. id)} \qquad \Gamma^k_{i,j} = \frac{1}{2}\sum_{l=1}^{m}(\partial_i g_{j,l} + \partial_j g_{i,l} - \partial_l g_{i,j})g^{l,k}.$$

This formula shows that the Christoffel symbol $\Gamma^k_{i,j}$ is symmetric in i and j. It is a reflection of the symmetry of the Levi-Civita connection.

Let us apply this to the simple example of Euclidean space $\mathbb{E}^m$: then $(g_{k,l} = \delta_{k,l})$ is constant and hence all Christoffel symbols vanish. It follows that

$$\text{(Cov. der. in } \mathbb{E}^m\text{)} \qquad \nabla_X Y = \sum_{i,j=1}^{m} X^i.(\partial_i Y^j)\partial_j = \sum_{j=1}^{m} X(Y^j)\partial_j,$$

where $X = \sum X^i\partial_i$ and $Y = \sum Y^i\partial_i$.

Remark 3.2. According to Exercise 12, the bundle $\tau^*M \otimes \tau^*M$ inherits from the Levi-Civita connection a connection (which we shall also denote ∇). Let us regard the metric g as a section of this bundle. If X is a vector field on an open $U \subset M$, then the covariant derivative $\nabla_X g$ is a section of $\tau^*M \otimes \tau^*M$ characterized by $(\nabla_X g)(Y, Z) = X(g(Y, Z)) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z)$. The flatness of g just says that this expression vanishes and so now we understand why that property is thus named: it amounts to saying that g is flat as a section of the bundle $\tau^*M \otimes \tau^*M$.

The following exercise shows that the symmetry property of the Levi-Civita connection can be understood somewhat similarly:

Exercise 17. Via Exercise 12 τ^*_M inherits a connection from the Levi-Civita connection. Hence by Exercise 13 we have an operator $D : \mathcal{E}^1_M = \mathcal{E}^0(\tau^*_M) \to \mathcal{E}^1(\tau^*_M) = \mathcal{E}^0(\tau^*_M \otimes \tau^*_M)$. Its composition with 'taking the exterior product', $\wedge : \mathcal{E}^0(\tau^*_M \otimes \tau^*_M) \to \mathcal{E}^2_M$, yields a map $\wedge D : \mathcal{E}^1_M \to \mathcal{E}^2_M$. Prove that this equaling the exterior derivative is equivalent to the symmetry property of the Levi-Civita connection.

Curvature tensor. The most important invariant of the Riemannian manifold M is the curvature form $R(X, Y)Z$ of its Levi-Civita connection (here X, Y, Z are all vector fields on M with a common domain). We know that R is alternating in its first two arguments, but it turns out to have some other (anti)symmetry properties as well.

Proposition 3.3. *We have*

Bianchi-identity: $R(X, Y)Z + R(Y, Z)X + R(Z, X)Y = 0$,
antisymmetry: $g(R(X, Y)Z, W)$ *is alternating in* Z *and* W *and*
symmetry: $g(R(X, Y)Z, W) = g(R(Z, W)X, Y)$.

Proof. The Bianchi identity follows easily from the **symmetry** of ∇. For the antisymmetry it suffice to show that $g(R(X, Y)Z, Z) = 0$ for all X, Y, Z (apply this for fixed X and Y to W, Z and $W + Z$). By the **flatness of the metric** we have:

$$XY(g(Z, Z)) = X(2g(\nabla_Y Z, Z)) = 2g(\nabla_X \nabla_Y Z, Z) + 2g(\nabla_X Z, \nabla_Y Z).$$

Subtract from this the identity obtained by interchanging X and Y and the equation $[X, Y](g(Z, Z)) = 2g(\nabla_{[X,Y]} Z, Z)$ and find $0 = 2g(R(X, Y)Z, Z)$.

The last formula is a formal consequence of the other two. For if take the inner product of the Bianchi-identity for (X, Y, Z) with W and do the same for $(X, Y, W; Z)$, $(Z, W, X; -Y)$ and $(Z, W, Y; -X)$ and add, then the established anti-symmetry properties of the curvature form make all terms but two cancel and what remains is the asserted property. □

Sectional curvature. Since $g(R(X, Y)Z, W)$ is antisymmetric in the first two and the last two variables, we may regard this expression at $p \in M$ as a bilinear form $\wedge^2 T_p M \times \wedge^2 T_p M \to \mathbb{R}$ and the symmetry property shows that the latter is symmetric. Since a symmetric bilinear form is determined by the associated quadratic form, R_p is already given by the function $v \wedge w \in \wedge T_p M \mapsto g_p(R_p(v, w)v, w)$. If σ is a 2-dimensional subspace of $T_p M$, then we take for (v, w) an orthonormal basis of σ, and then $v \wedge w \in \wedge^2 T_p M$ only depends σ up to sign and hence $g(R(v, w)v, w)$ only depends on σ. We write $K(\sigma)$ for minus that expression: $K(\sigma) := -g(R(v, w)v, w)$ and call this the *sectional curvature* of M along σ. If (v, w) is an arbitrary basis of σ, then

(Sect. Curv.) $$K(\sigma) = -\frac{g(R(v, w)v, w)}{g(v, v)g(w, w) - g(v, w)^2}$$

In case $\dim M = 2$, we write $K(p)$ for $K(T_p M)$. In case M is a submanifold of $\mathbb{E}^3$ (with the metric inherited from $\mathbb{E}^3$), we shall identify $K(p)$ with the Gauß curvature of M. More about this in Chapter 5.

Remark 3.4. The construction of the Levi-Civita connection and the derivation of its properties did not use the fact that g_p is positive definite, but only that it is nondegenerate (recall that at some point we had to invert the matrix $(g_{k,l})$). This is important for general relativity theory, for then we deal with a 4-dimensional manifold (space-time), whose tangent bundle comes equipped with a nondegenerate quadratic form of signature $(1, 3)$.

Exercise 18. Prove that the holonomy group of the tangent bundle of a Riemannian manifold M at $p \in M$ (cf. Exercise 11) is a subgroup of the orthogonal group of $T_p M$.

Exercise 19. Prove that the sectional curvature of Euclidean space $\mathbb{E}^m$ is constant zero.

Exercise 20. Given $f : \mathbb{R}^2 \to (0, \infty)$ determine the de curvature of the metric $f\big((dx)^2 + (dy)^2\big)$ on $\mathbb{R}^2$.

Exercise 21. Let M be a Riemannian manifold whose Levi-Civita connection is flat. Prove that M is locally isometric to $\mathbb{E}^m$. (Hint: By assumption there exist locally a basis of flat vector fields $X_1, \dots X_m$. Check that $g(X_i, X_j)$ is locally constant, so that after a Gram–Schmidt orthonormalization process, we can assume that $g(X_i, X_j) = \delta_{i,j}$. Now observe that the Lie brackets between these vector fields all vanish and use this to show that locally these vector fields are the coordinate vector fields of a chart.)

Exercise 22. Prove that the curvature tensor of a Riemann manifold is completely determined by its sectional curvature.

Other curvature tensors. For every $p \in M$ and $v, w \in T_pM$, consider the linear map $e \in T_pM \mapsto R(e, v)w \in T_pM$. The trace of this endomorphism is denoted $\mathrm{Ric}_p(v, w)$. It is clearly a bilinear; is it also symmetric: If $(e_1, \dots, e_m)$ is an orthonormal basis of T_pM, then $\mathrm{Ric}_p(v, w) = \sum_i g(R(e_i, v)w, e_i)$, which according to Proposition 3.3 is indeed symmetric in v and w.

Now let $e \in T_pM$ be any vector of unit length and let $(e = e_1, \dots, e_m)$ extend this to an orthonormal basis. Then we see that

$$\mathrm{Ric}_p(e, e) = \sum_i g(R(e_i, e)e, e_i) = \sum_{i>1} K_p(\sigma_{1i}),$$

where $\sigma_{ij} = \mathbb{R}e_i + \mathbb{R}e_j$. So the right hand side is the sum of the sectional curvatures of the coordinate planes T_pM which contain $e = e_1$.

The Ricci form defines the self-adjoint map $\widetilde{\mathrm{Ric}}_p : T_pM \to T_pM$ characterized by the property that $g(\widetilde{\mathrm{Ric}}_p(v), w) = \mathrm{Ric}_p(v, w)$. So in terms of the above orthonormal basis $\widetilde{\mathrm{Ric}}_p(e_i) = \sum_j \mathrm{Ric}_p(e_i, e_j)e_j$. The trace of this endomorphism is called the *scalar curvature* $K(p)$ of M. This is twice the sum of the sectional curvatures of the coordinate planes in T_pM:

$$K(p) = \sum_j \mathrm{Ric}_p(e_j, e_j) = \sum_{i,j} g(R_p(e_i, e_j)e_j, e_i) = \sum_{i<j} 2K_p(\sigma_{ij}).$$

Remark 3.5. In general relativity theory an important role is played by a modification of the Ricci tensor, the *Einstein tensor*. This is the symmetric bilinear form on the tangent bundle of M defined by

$$G_p := \mathrm{Ric}_p - \tfrac{1}{2}K(p)g_p : T_pM \times T_pM \to \mathbb{R}.$$

It follows from the preceding that the unit vector $e \in T_pM$,

$$G_p(e, e) = -\sum_{1<j<k} K_p(\sigma_{jk}),$$

which is the sum of the sectional curvatures of the coordinate planes perpendicular to e. It should be thought of as minus the scalar curvature attached to the hyperplane perpendicular to e. (The minus sign is not present in general relativity theory: here g is a metric of signature $(+, -, -, -)$ and so if e is a timelike velocity vector whose orthogonal complement is necessarily negative definite, then $G_p(e, e)$ represents scalar curvature of a spatial hypersurface at p.)

Exercise 23. Prove that for a 3-dimensional Riemann manifold, the Ricci form determines the Gauß curvature.

Exercise 24. Let be given a 3-dimensional connected Riemann manifold M with the property that in every $p \in M$, Ric_p is proportional to g_p (such a Riemann manifold is called a *Einstein manifold*). Prove that this factor of proportionality is constant on M. (Hint: use the Bianchi identity.)

Exercise 25 (Exercise 8 continued). In problem 8 assume that $(a, b) = (0, \infty)$. Find the differential equation that f must satisfy in order that S_f has constant curvature -1.

4. Geodesics

In this chapter M is a *connected* Riemannian manifold with metric g.

The geodesic differential equation. Let $\gamma : I \to M$ be an arc, then for any $t \in I$, we have defined the velocity vector $\dot\gamma(t) \in T_{\gamma(t)}M$. We may regard this as defining a section of the pull-back of the tangent bundle, $\gamma^*TM \to I$. If we apply $\nabla_{\frac{d}{dt}}$ to this section we get another section of $\gamma^*TM \to I$, which is reasonably denoted $\nabla_{\dot\gamma}\dot\gamma$. Its value in t is called the *acceleration* of γ at t (so in contrast to the velocity, this notion requires the Levi-Civita connection). If the acceleration is constant zero, then we call γ a *geodesic*. From the identity

$$\frac{d}{dt} g(\dot\gamma, \dot\gamma) = 2g(\nabla_{\dot\gamma}\dot\gamma, \dot\gamma)$$

we deduce that for a geodesic the absolute velocity $\|\dot\gamma\|$ is constant. So the arc length function is then affine-linear in t. Since an affine-linear reparameterization $t \mapsto \gamma(at + b)$ multiplies the acceleration of γ by a^2, we may reparameterize a geodesic by arc length and still have a geodesic.

Let (U, κ) be chart for M and denote the associated coordinate vector fields $\partial_1, \dots, \partial_m$. Assume γ takes values in U and put $\gamma^i := \gamma^*\kappa^i : I \to \mathbb{R}$. Then $\dot\gamma(t) = \sum_i \dot\gamma^i(t)\partial_i$ and hence

$$\begin{aligned}\nabla_{\dot\gamma}\dot\gamma &= \sum_i \dot\gamma^i \nabla_{\dot\gamma}\partial_i + \sum_i \ddot\gamma^i\partial_i \\ &= \sum_{i,j,k} \gamma^*(\Gamma^k_{i,j})\dot\gamma^i\dot\gamma^j\partial_k + \sum_k \ddot\gamma^k\partial_k.\end{aligned}$$

So the acceleration vanishes if and only if $(\gamma^i)_i$ satisfies the system of second order differential equations

$$\text{(geod. eqn.)} \qquad \ddot\gamma^k = -\sum_{i,j,k} \gamma^*(\Gamma^k_{i,j})\dot\gamma^i\dot\gamma^j, \; k = 1, \dots, m.$$

The exponential map. The above system has for a given set of initial conditions specifying $(\gamma^i(t_0))_i$ and $(\dot\gamma^i(t_0))_i$ exactly one (local) solution, which in fact depends differentiably on these initial conditions. This leads to:

Lemma 4.1. *There is a neighborhood* Ω *of the zero section of* TM *in* TM *and a map* $(t, v) \in [0, 1] \times \Omega \to \gamma_v(t) \in M$ *with the following properties:*

(i) $\gamma_v : [0, 1] \to M$ *is the unique geodesic with* $\dot\gamma_v(0) = v$
(ii) *if* $v \in \Omega$ *and* $s \in [0, 1]$*, then* $sv \in \Omega$ *and we have* $\gamma_{sv}(t) = \gamma_v(st)$.

Proof. The preceding discussion shows that there exists a neighborhood W of $\{0\} \times TM$ in $\mathbb{R} \times TM$ on which the geodesic flow $(t, v) \in W \mapsto \gamma_v(t)$ is defined (γ_v is the geodesic with initial condition $\dot\gamma_v(0) = v$). We take W such that it meets every line $\mathbb{R} \times \{v\}$ in an interval and assume W maximal for this property. We observed that for every $s \in [0, 1]$, $t \mapsto \gamma_v(st)$ is a geodesic as well. Since its velocity at 0 equals $s\dot\gamma(0) = sv$, we must have $\gamma_v(st) = \gamma_{sv}(t)$. In particular, $(t, v) \in W$ implies $[0, 1] \times \{t'v\} \subset W$ for all $t' \in [0, t]$. It follows that the set of $v \in TM$ with $[0, 1] \times \{v\} \subset W$ is an open part of TM that contains the zero section. This is our Ω. □

Definition 4.2. If in Lemma 4.1 we take Ω maximal for the listed properties, then $v \in \Omega \mapsto \gamma_v(1)$ is called the *exponential map* of M. We denote it by exp and we write $\exp_p$ for its restriction to $\Omega_p := \Omega \cap T_pM$.

Lemma 4.3. *Let* $p \in M$*. Then the derivative of* $\exp_p$ *in* $0 \in T_pM$ *is, when regarded as a linear map from* $T_0T_pM \cong T_pM$ *to itself, the identity. Moreover there exists a neighborhood* U *of* p *in* M *and an* $\varepsilon > 0$ *such that*

(i) *for every* $q \in U$*,* $\exp_q$ *maps the open* ε*-ball in* T_qM *diffeomorphically onto an open subset of* M *which contains* U *and*
(ii) *for every pair* $q, q' \in U$ *there is precisely one geodesic parameterized by arc length* $\gamma_{q,q'}$ *of length* $< \varepsilon$ *which connects* q *with* q'.

Proof. The first assertion follows from the fact that for $v \in T_pM$ and small t we have $\exp(tv) = \gamma_{tv}(1) = \gamma_v(t)$, and hence the t-derivative of this map in 0 is $\dot\gamma_v(0) = v$.

In particular, $D_0 \exp|T_pM$ is nonsingular and so (i) follows from the implicit function theorem applied to exp.

To prove (ii), we note that (i) asserts the existence and uniqueness of a $v \in T_qM$ of length $\delta < \varepsilon$ with the property that $\exp_q(v) = q'$. So the arc $\gamma : [0, \delta] \to M$, $\gamma(t) = \exp_q(tv)$, is a geodesic parameterized by arc length from q to q'. As $\exp_q$ maps the open ε-ball in T_qM diffeomorphically onto an open subset containing U, this geodesic is the only one such with domain an interval of length $< \varepsilon$. □

We will show that in the situation of Lemma 4.3, $\gamma_{q,q'}$ is the arc of shortest length connecting q with q', in other words, the distance between q and q' is the arc length of $\gamma_{q,q'}$. From 4.4 till 4.6, U and ε are as in the preceding lemma.

The following lemma describes the metric near $q \in M$ in terms of 'polar coordinates'.

Lemma 4.4 (Gauß lemma). *Given $q \in U$, let S be the unit sphere in T_qM and let $F : [0, \varepsilon) \times S \to M$, be defined by $F(r, v) = \exp(rv)$. Then the pulled back metric F^*g has the form*

$$F^*g = (dr)^2 + r^2 h_r.$$

Here h_r is a metric on S depending on r.

Proof. Since the rays are parameterized by arc length, the restriction of F^*g to a ray is $(dr)^2$. We prove that F^*g does not contain mixed terms. Notice that F defines a diffeomorphism from $(0, \varepsilon) \times S$ onto a punctured neighborhood of q. Let R be radial vector field that is the image of the $\partial/\partial r$ under this diffeomorphism and for any vector field V on S denote by $\tilde{V}$ the image of $(0, V)$ under this diffeomorphism. We must show that $g(R, \tilde{V}) = 0$.

Since $\partial/\partial r$ and V commute, so do R and $\tilde{V}$. We show that $g(R, \tilde{V})$ is independent of r. Since the Levi-Civita connection is compatible with the metric, the derivative of $g(R, \tilde{V})$ with respect to R is computed by

$$R(g(R, \tilde{V})) = g(\nabla_R R, \tilde{V}) + g(R, \nabla_R \tilde{V}).$$

The vector field $\nabla_R R$ appearing in the first term is at $F(r, s)$ the acceleration vector of the geodesic γ_s in r, and hence vanishes. The second term is by the **symmetry** of ∇ equal to $g(R, \nabla_{\tilde{V}} R) = \frac{1}{2}\tilde{V} g(R, R) = \frac{1}{2}\tilde{V}(1) = 0$. So the value of $g(R, \tilde{V})$ in $F(r, s)$ only depends on s. If we restrict both vector fields to a radial geodesic (by fixing s) and let r tend to zero, then it follows from the fact that $D_0 \exp_q$ is the identity (Lemma 4.3) that R resp. $\tilde{V}$ tend to s resp. V_s. Since $V_s \in T_sS$ we have $g(s, V_s) = 0$. So $g(R, \tilde{V}) = 0$. □

Remark 4.5. One can show that in the Gauß lemma we have $h_r = h_0 + O(r^2)$, where h_0 is the metric S inherits from T_q. From that one can deduce that if we pull back the metric g along the exponential map $\exp_q$ to T_qM it is, at the origin of T_qM up to second order terms equal to the flat metric on T_qM defined by g_q. In other words, if $e_1, \dots, e_m$ is an orthonormal basis for T_qM, then in terms of the chart κ at q thus obtained (so a local inverse of $\exp_q$ at q followed by the identification $T_qM \cong \mathbb{E}^m$ defined by the basis), we have $g_{ij} = \delta_{ij} + O(\|\kappa\|^2)$ and the Christoffel symbols for this chart all vanish in q.

Corollary 4.6. *For any pair $q, q' \in U$, the arc length of $\gamma_{q,q'}$ equals the distance $d(q, q')$. Moreover, $\gamma_{q,q'}$ is essentially unique for this property: every piecewise differentiable arc ω from q to q' of arc length $d(q, q')$ is equal to $\gamma_{q,q'}$ after reparameterizing the latter by a piecewise differentiable function.*

Proof. Since q' in the image of F, we can write q' as $F(r, s) = \exp_q(rs)$. Let $\alpha : [0, c] \to M$ be a piecewise differentiable arc from q to q'. We must show that the arc length of α is $\geq r$ and that if equality holds, the arc is obtained by reparameterizing the geodesic ray $\gamma_s|[0,r]$. Let $0 < c' \le c$ be the first value of the parameter for which α hits the image of the sphere of radius r. So $\alpha([0,c']) \subset F([0,r] \times S)$ and $\alpha(c') \in F(\{r\} \times S)$. It is clear that $\alpha(t) \neq p$ for $t > 0$ and so for $t \in (0,c']$, $\alpha(t)$ can be written as $F(r(t), s(t))$. The Gauß lemma shows that for every $t \in (0,c']$ where α is differentiable, $\|\dot\alpha(t)\|^2 \ge \|\dot r(t)\|^2$ with equality holding only if $\dot s = 0$, i.e., when s is constant. We thus obtain the inequality

$$L(\alpha|[0,c']) \ge \int_0^{c'} |\dot r(t)| dt \ge |r(c') - r(0)| = r,$$

which is strict unless $r(t)$ is monotonous and s is constant. □

Theorem 4.7. *Every piecewise differentiable arc in* M *whose arc length equals the distance between its end points traverses a geodesic.*

Proof. Indeed, the previous corollary says that such an arc traverses in each of its points an open piece of a geodesic. □

Remark 4.8. Geodesics are encountered in various places in physics such as in mechanics, optics and general relativity. Suppose for instance given an ideal mechanical system consisting of a point particle that is forced to move in a surface in $\mathbb{R}^3$. If no forces act, and no friction is present, then the particle traverses a geodesic in M. The same is true for a ray of light forced to stay in M by making the latter of very thin glass with mirroring surfaces.

Completeness. It is not always possible to realize the distance between two points $q, q' \in M$ by (the arc length of) a piecewise differentiable arc. A simple example is obtained by removing a point p from a given Riemannian manifold M: take $q, q' \in U - \{p\}$, where U is the neighborhood of p which appears in Lemma 4.3, chosen in such a manner that p lies on the image of $\gamma_{q,q'}$. Notice that the Riemann manifold $M' := M - \{q\}$ is not complete (a Cauchy sequence in M' with limit q has no limit in M'. And $\exp_{q'}$ is not defined on all of $T_{q'}M$ (if regarded as the exponential map for M'. These deficiencies are not independent, witness the following theorem, due to Hopf-Rinow.

Theorem 4.9 (Hopf-Rinow)**.** *The following properties are equivalent:*

(i) *there exists a* $p \in M$ *such that* $\exp_p$ *is defined on all of* T_pM*,*
(ii) *the exponential map is defined on all of* TM*,*
(iii) M *is complete as a metric space,*

and if one of these is satisfied, then every pair of points of M *can be joined by a geodesic whose arc length equals the distance between them.*

This theorem is in fact the union of the following assertions.

Lemma 4.10. *If for some* $p \in M$ *the exponential map* $\exp_p$ *is defined on the whole tangent space* T_pM*, then every point of* M *can be joined with* p *by a geodesic whose arc length is its distance to* p.

Proof. Let $U \ni p$ and $\varepsilon > 0$ be as in Lemma 4.3. Let $q \in M$ have distance $d > 0$ to p. Choose $0 < \delta < \varepsilon$ such that q is not contained in the image under $\exp_p$ of the closed δ-ball B_δ in T_pM. Every continuous arc from p to q will hit the boundary sphere $S_\delta := \exp_p(\partial B_\delta)$. According to Corollary 4.6 every point of S_δ has distance δ to p. Since S_δ is compact, its distance to q is the distance of some $q' \in S$ to q, so that

$$\delta + d(q', q) = d. \tag{1_δ}$$

The properties of U guarantee the existence of a geodesic γ from p to q' of arc length δ. We suppose γ parameterized by arc length. By assumption, γ is as a geodesic defined on all of $\mathbb{R}$. The lemma follows if we show that $\gamma(d) = q$. In fact, we prove that for all $0 \le t \le d$ we have

$$t + d(\gamma(t), q) = d. \tag{1_t}$$

The set of $t \in [0, d]$ for which eqn. (1_t) holds, is closed for both members of (1_t) are continuous in t. Let t' be its maximal element. If $t' = d$, we are done. If $t' < d$, we produce a contradiction as follows. Let $p' := \gamma(t')$. Then according to $(1_{t'})$ the distance between p' and q equals $d - t' > 0$. Applying the preceding to the pair p', q (instead of p', q) we obtain a $q'' \in M$ and a geodesic γ' from p' to q'' of arc length $0 < \delta' < d - t'$ such that

$$\delta' + d(q'', q) = d - t'. \tag{2}$$

The geodesics $\gamma|[0, t']$ and γ' are the pieces of a piecewise differentiable arc (parameterized by arc length) α from p to q'' of total length $t' + \delta'$. According to (2), the latter length equals $d - d(q'', q)$. The triangle inequality (applied to p, q'', q) shows that the distance from p to q'' is at least that much, and so α realizes the distance. But then Theorem 4.7 tells us that α must be a geodesic. So $\gamma'(t) = \gamma(t' + t)$. This contradicts the maximal character of t'. □

Corollary 4.11. *Under the assumptions of the above lemma, every bounded closed subset of* M *is compact. In particular,* M *is a complete metric space.*

Proof. Let $K \subset M$ be closed and bounded. Then there exists an $r > 0$ such that every point of K has distance $\le r$ to p. The preceding lemma says that K is contained in the image under $\exp_p$ of the closed ball B in T_pM of radius r. Since B is compact, so is $\exp_p(B)$. Hence K is compact as it is a closed subset of $\exp_p(B)$.

The last assertion indeed follows from the first, for if K is the image of a Cauchy sequence without limit, then K must be bounded, discrete and infinite. This K is therefore bounded, closed, but not compact. This is a contradiction. □

The proof of the Hopf-Rinow theorem is finished by the following converse:

Proposition 4.12. *If* M *is complete as a metric space, then the exponential map is defined on all of* TM *(and so by Lemma 4.10 the distance between any two points of* M *is realized by a geodesic).*

Proof. Suppose the domain D of exp is strictly smaller than TM, so that there exist a $v \in TM$ of length 1 and a $r > 0$ such that $tv \in D$ for $0 \le t < r$, but $rv \notin D$ (D is open). The geodesic $\gamma_v : [0, r) \to M$ is parameterized by arc length and we have that for $0 \le t \le t' < r$,

$$d(\gamma_v(t), \gamma_v(t')) \le \int_t^{t'} \|\dot{\gamma}_v(\tau)\| d\tau = t' - t.$$

So the image of a Cauchy sequence in $[0, r)$ is one in M. If we make the former converge to r, than the latter has a limit $q \in M$. If we apply Lemma 4.1 to q, then we find that the exponential map can be extended to a neighborhood of rv and we arrive at a contradiction. □

Geodesic submanifolds. If N is a submanifold of M and a geodesic γ of M happens to lie in N, then γ is also a geodesic for N (with respect to its induced Riemann metric). This follows from the fact that γ is locally length minimizing in M: it is then also so for the induced metric in N and is hence a geodesic of N. The converse need not hold, of course: a submanifold of $\mathbb{E}^m$ which does not contain a straight line has no geodesic in common with $\mathbb{E}^m$.

Definition 4.13. We say that a submanifold N of M is *totally geodesic* if every geodesic of N (relative to the induced metric) is also one of M.

Totally geodesic submanifolds of dimension one are ample: the image of any geodesic without self-intersection is one. Interesting higher dimensional examples are sometimes obtained by means of

Lemma 4.14. *A submanifold* N *of* M *is totally geodesically embedded if it is open in the fixed point set of an isometry* $I : M \to M$.

Proof. If $v \in TN$ and γ is the geodesic of M with initial condition $\dot{\gamma}(0) = v$, then $I\gamma$ is also a geodesic with the same initial condition (for $(I\gamma)\dot{}(0) = dI(\dot{\gamma}(0)) = dI(v) = v$) and hence equal to γ. So γ then lies in N and is therefore a geodesic of N. □

Remark 4.15. It can be shown that the fixed point set of an isometry is always a submanifold. (This is definitely not true for a diffeomorphism: any closed subset of a compact manifold can be realized as the fixed point of a self-diffeomorphism.)

Examples 4.16. (i) Every affine-linear subspace $V \subset \mathbb{E}^m$ is totally geodesic in $\mathbb{E}^m$. This is clear.

(ii) Let $V \subset \mathbb{E}^{m+1}$ be a linear subspace. Then the orthogonal reflection in V restricts to an isometry of $\mathbb{S}^m$ which has the sphere $V \cap \mathbb{S}^m$ as fixed point set. So $V \cap \mathbb{S}^m$ is totally geodesic in $\mathbb{S}^m$. In particular (by taking $\dim V = 2$), we find that every great circle in $\mathbb{S}^m$ is the image of a geodesic. Since every nonzero tangent vector in $T_p\mathbb{S}^m$ lies in a 2-dimensional subspace of $\mathbb{E}^{m+1}$, we get in this way all the geodesics of $\mathbb{S}^m$.

(iii) We take up again Example 1.3-iii. Let $b(x, y)$ be the symmetric bilinear form defined by

$$x^1y^1 + \cdots x^my^m - x^{m+1}y^{m+1}$$

and let $\mathbb{H}^m$ be the hyperbolic space defined by $b(x, x) = -1,\ x_{m+1} > 0$, or equivalently, $x^{m+1} = \sqrt{(x^1)^2 + \cdots + (x^m)^2 + 1}$. If V is a linear subspace of $\mathbb{R}^{m+1}$ which meets $\mathbb{H}^m$, then the set $V^\perp$ of $x \in \mathbb{R}^{m+1}$ with $b(x, y) = 0$ for all $y \in V$ is positive definite and we have $\mathbb{R}^{m+1} = V \oplus V^\perp$. The transformation $\tilde{I}$ which is the identity on V en minus the identity on $V^\perp$ preserves the form b. It also leaves the nonempty set $V \cap \mathbb{H}^m$ pointwise fixed. Hence $\tilde{I}$ leaves $\mathbb{H}^m$ invariant and restricts to an isometry of $\mathbb{H}^m$. We conclude that $\mathbb{H}^m \cap V$ is totally geodesic. As before we find that every geodesic of $\mathbb{H}^m$ is the intersection of $\mathbb{H}^m$ with a 2p-lane in $\mathbb{R}^{m+1}$.

Exercise 26. Let $E \subset \mathbb{E}^{m+1}$ be the m-dimensional ellipsoid defined by the equation $\sum_i (x^i/a^i)^2 = 1$, $(a^i > 0)$. Prove that the intersection of E with a subspace spanned by basis vectors of $\mathbb{E}^{m+1}$ is totally geodesic in E. Conclude that we have $\binom{m+1}{2}$ closed geodesics on E. (But for $m = 2$ there many more: a theorem of Hilbert asserts that the union of their images is dense in E.)

Exercise 27. Let M be a connected Riemann manifold and I an isometry of M onto itself which fixes a point p and is the identity inT_pM. Prove that I is the identity. (Hint: use the exponential map.)

Exercise 28. Use the previous exercise to prove that every isometry of $\mathbb{S}^m$ is the restriction of an orthogonal transformation in $\mathbb{E}^{m+1}$.

Prove similarly that an isometry of $\mathbb{H}^m$ is induced by a linear transformation of $\mathbb{R}^{m+1}$ which preserves the symmetric bilinear form b and the component $\mathbb{H}^m$ of $b(x, x) = -1$.

Exercise 29. Determine the (images of the) closed geodesics of the flat torus $(\mathbb{S}^1)^2$ (see Example 1.3-(iv)).

Exercise 30. Let p and q be points of a complete Riemann manifold M. We say that two continuous arcs $\gamma_0, \gamma_1 : [a, b] \to M$ from p to q are *homotopic* if they can be connected by a family of such arcs $\{\gamma_s : [a, b] \to M\}_{0 \le s \le 1}$ with the property that the associated map $[a, b] \times [0, 1] \to M$ is continuous. Prove that every differentiable arc from p to q is homotopic to a geodesic.

5. The Second Fundamental form

In this section $(\tilde{M}, \tilde{g})$ stands for a Riemann manifold and $\tilde{\nabla}$ for its Levi-Civita connection. Furthermore, $M \subset \tilde{M}$ is a submanifold endowed with the induced Riemann metric g and ∇ is the Levi-Civita connection of (M, g).

Second Fundamental form. The tangent bundle of $\tilde{M}$ restricted to M is the direct sum of τ_M and the orthogonal complement $\tau_M^\perp$ of τ_M in $\tau_{\tilde{M}}$. The connection $\tilde{\nabla}$ on $\tau_{\tilde{M}}$ pulls back (here: restricts) to one on $\tau_{\tilde{M}}|M = \tau_M \oplus \tau_M^\perp$, which we continue to denote by $\tilde{\nabla}$. A comparison of this with the Levi-Civita connection on the summand τ_M leads to the notion of 'second fundamental form'. (So baptized by Gauß, whose terminology has stuck—to him the first fundamental form was the Riemann metric, but that has not stuck.)

Proposition 5.1. *If X and Y are vector fields on an open subset of M, then $\nabla_X Y$ is the τ_M-component van $\tilde{\nabla}_X Y$.*

Proof. Let $\pi : \tau_{\tilde{M}}|M \to \tau_M$ be the (orthogonal) projection. We verify that $\pi(\tilde{\nabla}_X Y)$ obeys the properties that caracterize the Levi-Civita connection. In view of the local nature of the issue, we may assume that any local vector field X on M extends to a vector field $\tilde{X}$ on an open subset of $\tilde{M}$. Then indeed:

$$\pi(\tilde{\nabla}_X Y) - \pi(\tilde{\nabla}_Y X) = \pi((\tilde{\nabla}_{\tilde{X}} \tilde{Y} - \tilde{\nabla}_{\tilde{Y}} \tilde{X})|M) = \pi([\tilde{X}, \tilde{Y}]|M) = [X, Y],$$

and

$$\begin{aligned} g(\pi(\tilde{\nabla}_X Y), Z) + g(Y, \pi(\tilde{\nabla}_X Z)) &= \tilde{g}(\tilde{\nabla}_X Y, Z) + \tilde{g}(Y, \tilde{\nabla}_X Z) \\ &= \tilde{g}(\tilde{\nabla}_{\tilde{X}} \tilde{Y}, \tilde{Z})|M + \tilde{g}(\tilde{Y}, \tilde{\nabla}_{\tilde{X}} \tilde{Z})|M = (\tilde{X}\tilde{g}(\tilde{Y}, \tilde{Z}))|M = Xg(Y, Z). \quad \square \end{aligned}$$

Let us write $H(X, Y)$ for the $\tau_M^\perp$-component of $\tilde{\nabla}_X Y$, so that

$$\tilde{\nabla}_X Y = \nabla_X Y + H(X, Y).$$

Proposition-definition 5.2. *The value of $H(X, Y)$ in $p \in M$ only depends on X_p and Y_p so that H defines a symmetric vector bundle homomorphism $H : \tau_M \otimes \tau_M \to \tau_M^\perp$. This homomorphism is called the* second fundamental form *of M in $\tilde{M}$.*

Proof. We have

$$H(X, Y) - H(Y, X) = (\tilde{\nabla}_X Y - \tilde{\nabla}_Y X) - (\nabla_X Y - \nabla_Y X) = [X, Y] - [X, Y] = 0.$$

so that H is indeed symmetric. For fixed Y, the values of $\nabla_X Y$ and $\tilde{\nabla}_X Y$ in p only depend on X_p, and so the same is true for $H(X, Y)$. Because of the symmetry this remains true if we interchange X and Y. $\square$

The second fundamental form H measures to what extent M fails to be totally geodesic in $\tilde{M}$:

Proposition 5.3. *The submanifold M is totally geodesic in $\tilde{M}$ if and only if its second fundamental form H is identically zero.*

Proof. Let $v \in T_p M$ be an arbitrary tangent vector and let γ be the geodesic of M with initial condition $\dot{\gamma}(0) = v$. Then $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$ and so $\tilde{\nabla}_{\dot{\gamma}} \dot{\gamma} = \nabla_{\dot{\gamma}} \dot{\gamma} + H(\dot{\gamma}, \dot{\gamma}) = H(\dot{\gamma}, \dot{\gamma})$. Hence γ is a geodesic in $\tilde{M}$ if and only if $H(\dot{\gamma}, \dot{\gamma}) \equiv 0$. The proposition follows from this. $\square$

With the help of the second fundamental form we can express the curvature of M in that of $\tilde{M}$:

Proposition 5.4. *Any four vector fields* X, Y, Z, W *on an open subset* M *satisfy the so-called Gauß equation*

$$g(R(X \wedge Y)Z, W) = \tilde{g}(\tilde{R}(X \wedge Y)Z, W) + \tilde{g}(H(X, W), H(Y, Z)) - \tilde{g}(H(X, Z), H(Y, W)).$$

Proof. It follows from $\nabla_Y Z = \tilde{\nabla}_Y Z - H(Y, Z)$ that $\nabla_X \nabla_Y Z = \pi \tilde{\nabla}_X \nabla_Y Z = \pi(\tilde{\nabla}_X \tilde{\nabla}_Y Z - \tilde{\nabla}_X H(Y, Z))$ and so

$$g(\nabla_X \nabla_Y Z, W) = \tilde{g}(\tilde{\nabla}_X \tilde{\nabla}_Y Z, W) - \tilde{g}(\tilde{\nabla}_X H(Y, Z), W).$$

Differentiation of the identity $\tilde{g}(H(Y, Z), W) = 0$ with respect to X yields

$$\tilde{g}(\tilde{\nabla}_X H(Y, Z), W) + \tilde{g}(H(Y, Z), \tilde{\nabla}_X W) = 0.$$

Since the normal component of $\tilde{\nabla}_X W$ equals $H(X, W)$ by definition, the second term equals $\tilde{g}(H(Y, Z), H(X, W))$. Feeding this into the displayed identity above gives

$$g(\nabla_X \nabla_Y Z, W) = \tilde{g}(\tilde{\nabla}_X \tilde{\nabla}_Y Z, W) + \tilde{g}(H(Y, Z), H(X, W)).$$

If we subtract from this the equality obtained by exchanging the roles of X and Y and use that $g(\nabla_{[X,Y]} Z, W) = \tilde{g}(\tilde{\nabla}_{[X,Y]} Z, W)$, we obtain the claimed identity. ☐

Curvature of a submanifold. Let σ be a plane in T_pM, spanned by $v, w \in T_pM$, say. It follows from Proposition 5.4 that the values of the sectional curvatures $\tilde{K}(\sigma)$ of $\tilde{M}$ and $K(\sigma)$ of M on this plane are related by

$$(\dagger) \qquad K(\sigma) = \tilde{K}(\sigma) + \tilde{g}(H(v, v), H(w, w)) - \tilde{g}(H(v, w), H(v, w)),$$

where $\{v, w\}$ is an orthonormal basis of σ. In particular, $K(\sigma) = \tilde{K}(\sigma)$ if M is totally geodesic in $\tilde{M}$.

We can also use this to understand the sectional curvature of $\tilde{M}$ on any plane $\sigma \subset T_p\tilde{M}$ geometrically. For this we no longer assume given M beforehand, but choose for M the image under exp of a small neighborhood of the origin of σ in σ. In general, M fails to be totally geodesic, but the argument in Proposition 5.3 does show that H is zero on $T_pM \times T_pM$. This implies that $\tilde{K}(\sigma) = K(\sigma) = K_M(p)$. Thus $\tilde{K}(\sigma)$ acquires an interpretation of the curvature of 2-dimensional submanifold.

Gauß derived the identity $(\dagger)$ for a surface in $\mathbb{E}^3$ in order to compute the curvature of that surface that is now named after him (notice that then $\tilde{K} = 0$). We can do this more generally for an hypersurface M in $\mathbb{E}^{m+1}$ as follows. For this we suppose given for every $p \in M$ a unit vector $N(p) \in \mathbb{S}^m$ such that $N(p)$ is perpendicular to T_pM, when displaced to p) which depends continuously on p. This makes the map $N : M \to \mathbb{S}^m$ in fact differentiable. Notice that the tangent spaces T_pM and $T_{N(p)}\mathbb{S}^m$ may be identified (after displacing them to the origin both are the orthogonal complement of $N(p)$). Thus the derivative $D_pN : T_pM \to T_{N(p)}\mathbb{S}^m$ can be understood as an endomorphism of T_pM.

Lemma 5.5. *If* h *denotes the symmetric bilinear form on* TM *defined by*

$$h_p(v,w) = \langle H(v,w), N(p)\rangle, \quad v,w \in T_pM,$$

i.e., the coefficient of H *on* N*, then*

$$h_p(v,w) = -\langle D_pN(v), w\rangle = -\langle v, D_pN(w)\rangle, \quad v,w \in T_pM.$$

In particular, D_pN *is a self-adjoint endomorphism of* T_pM.

If $\sigma \subset T_pM$ *is 2-plane and* $v,w \in T_pM$ *is an orthonormal basis of it, then*

$$K(\sigma) = h(v,v)h(w,w) - h(v,w)^2. \qquad (\ddagger)$$

Proof. Let V en W be vector fields on a neighborhood of p in M which take in p the value v resp. w. It follows from $\langle V,N\rangle = 0$ that

$$0 = W\langle V,N\rangle = \langle \tilde{\nabla}_W V, N\rangle + \langle V, \tilde{\nabla}_W N\rangle = h(V,W) + \langle V, \tilde{\nabla}_W N\rangle.$$

If we write $N = \sum_i N^i \partial_i$, then by our formula for the covariant derivative in $\mathbb{E}^{m+1}$, $\tilde{\nabla}_W N = \sum_i W(N^i)\partial_i$. Notice that we can write the latter as $DN(W)$. This proves that $h(V,W) = -\langle V, DN(W)\rangle$. If we combine this with the symmetry of h, we obtain the first assertion. The last formula follows from Equation (†). □

In a standard linear algebra course it is shown that if $(V, \langle\ ,\ \rangle)$ is a finite dimensional real inner product space, then every self-adjoint map $T : V \to V$ (i.e., for which $\langle v, Tw\rangle$ is symmetric in v and w) has an orthonormal basis of eigen vectors.

If this is applied to D_pN acting in T_pM, then we obtain an *orthonormal* basis $(e_1, \dots, e_m)$ of T_pM such that $D_pN(e_i) = k_ie_i$ $(i = 1, \dots, m)$ for certain $k_i \in \mathbb{R}$. So $h(e_i, e_j) = -k_i\delta_{ij}$. The coefficients k_i's are called the *principal curvatures* of M at p. These are *extrinsic* (curvature) invariants (by which we mean that they a priori depend on the way M is embedded in $\mathbb{E}^{m+1}$). But the products h_ih_j, $1 \le i < j \le m$, are *intrinsic* in the sense that they only depend on Riemann metric induced on M.

For $m = 2$ this follows immediately from Formula (‡): k_1k_2 is the Gauß curvature of M at p. When Gauß discovered this fact, he was so enchanted by it that he refered to this as the *Theorema Egregium* (the perfect theorem). He used the (more concrete) description of h that entered in Section 0 for $m = 2$: after a motion in $\mathbb{E}^{m+1}$ we can assume that $p = 0$ and $N_p = e_{m+1}$, so that M is there given at p as the graph of a differentiable function $\phi = \phi(x^1, \dots, x^m)$, whose Taylor development at 0 begins with a quadratic part, and after another orthogonal transformation in the hyperplane perpendicular to e_{m+1} we can also arrange that this quadratic part has no mixed terms:

$$\phi(x^1, \dots, x^m) = \sum_{i=1}^m -\tfrac{1}{2}k_i(x^i)^2 + \text{higher order terms}.$$

Our discussion in Section 0 shows that $D_pN_i(e_i) = k_ie_i$ and we conclude that $h_p(v, w) = \sum_{i=1}^m -k_iv^iw^i$. So the principal curvatures of M at p are $k_1, \dots, k_m$. This discussion also shows that if $v \in T_pM$ is of length 1, then $|h_p(v, v)|$ is the absolute curvature of the curve that M cuts out in the plane through p parallel to the span of v and the normal of M at p.

Examples 5.6. (i) Let $\mathbb{S}_r$ be the sphere of radius $r > 0$ a in $\mathbb{E}^{m+1}$ centered at 0. Then $N(x) = x/\|x\| = x/r$ and hence $\partial_jN^i|_{\mathbb{S}_r} = r^{-1}\delta_i^j - r^{-3}x^ix^j$. It follows that in $p = (p^1, \dots, p^{m+1})$ we have for $v, w \in T_p\mathbb{S}_r$ that

$$\begin{aligned} h(v, w) &= -\sum_i r^{-1}v^iw^i + \sum_{i,j} r^{-3}p^ip^jv^iw^j \\ &= -r^{-1}\langle v, w\rangle + r^{-3}\langle v, p\rangle\langle w, p\rangle = -r^{-1}\langle v, w\rangle. \end{aligned}$$

This shows that the sectional curvature of $\mathbb{S}_r$ takes on any plane in its tangent bundle the value r^{-2}.
(ii) Consider for $r > 0$ the cylinder C_r in $\mathbb{E}^3$ defined by the equation $x^2+y^2 = r^2$. It is easy to verify that the principal curvatures in evert point equal 0 en r^{-1} and so the Gauß curvature vanishes identically. This is easily seen directly: 'developing' the cylinder yields a local isometry with $\mathbb{E}^2$.

Exercise 31. Let M be a surface in $\mathbb{E}^3$ and $N : M \to \mathbb{S}^2$ a map which assigns to $p \in M$ a unit vector normal to T_pM and $K : M \to \mathbb{R}$ the Gauß curvature. Denote by μ_M the 2-form on M which takes the value 1 on any basis (v, w) of T_pM for which $(N(p), v, w)$ is an oriented orthonormal basis of $\mathbb{E}^3$ and let $\mu_{\mathbb{S}^2}$ be similarly defined. Prove that $N^*(\mu_{\mathbb{S}^2}) = K\mu_M$.

Exercise 32. Let M be a hypersurface in $\mathbb{E}^{m+1}$ and $p \in M$.
(a) Compute Ric_p in terms of an orthonormal basis on which the second fundamental form assumes the diagonal form.
(b) Suppose that Ric_p is proportional to the metric g_p. Prove that the sectional curvature on all planes of T_pM is constant.

6. The Gauss-Bonnet theorem

In this section M is an oriented Riemann manifold of dimension 2.

A remarkable differential. We begin with describing the Levi-Civita connection and the Gauß curvature in terms of a pair of orthonormal vector fields. Let X_1 and X_2 be vector fields on an open subset U of M with the property that they define in any $p \in U$ an oriented orthonormal basis of T_pM. Such a pair always exists locally: take an oriented basis of vector fields and then apply the Gram-Schmidt orthonormalization process. Denote by ω^1, ω^2 the corresponding dual basis of differentials: $\omega^i(X_j) = \delta^i_j$. So the 2-form $\mu_M := \omega^1 \wedge \omega^2$ is the area element of M.

Lemma 6.1. *Put* $c_i := \omega^i([X_1, X_2])$ *and put* $\alpha := c_1\omega^1 + c_2\omega^2$. *Then the Levi-Civita connection is given by :*

$$\nabla_X(X_1) = -\alpha(X)X_2, \quad \nabla_X(X_2) = \alpha(X)X_1.$$

In other words, its connection form relative to the basis (X_1, X_2) *is* $\left(\begin{smallmatrix} 0 & \alpha \\ -\alpha & 0 \end{smallmatrix}\right)$.

Proof. We only check the formula for $\nabla_X(X_1)$, the other is similar. Applying Formula (†) in the proof of Theorem 3.1 to the triple (X_i, X_1, X_k) gives

$$\omega^k(\nabla_{X_i} X_1) = \tfrac{1}{2}\left(\omega^k([X_i, X_1]) - \omega^1([X_i, X_k]) - \omega^i([X_1, X_k])\right).$$

This is clearly zero for $k = 1$. It remains to show that it equals $\alpha(X_i) = -c_i$ for $k = 2$. This is straightforward. For instance, if $i = 1$, then

$$\omega^2(\nabla_{X_1} X_1) = \tfrac{1}{2}\left(\omega^2([X_1, X_1]) - \omega^1([X_1, X_2]) - \omega^1([X_1, X_2])\right) = -c_1. \qquad \square$$

Corollary 6.2. *If* K *denotes the Gauß curvature, then* $K\mu_M = d\alpha$.

Proof. So this is saying that $K(X_1, X_2) = d\alpha(X_1, X_2)$. If X, Y are vector fields on U, then

$$\nabla_X \nabla_Y X_1 = \nabla_X(-\alpha(Y)X_2) = -\alpha(X)\alpha(Y)X_1 - X(\alpha(Y))X_2.$$

If we subtract from this the identity we get from interchanging X and Y and the identity $\nabla_{[X,Y]}X_1 = -\alpha([X, Y])X_2$, then we find

$$\begin{aligned} R(X, Y)X_1 &= -X(\alpha(Y))X_2 + Y(\alpha(X))X_2 + \alpha([X, Y])X_2 = \\ &= -\left(X(\alpha(Y)) - Y(\alpha(X)) - \alpha([X, Y])\right)X_2 = -d\alpha(X, Y)X_2 \end{aligned}$$

It follows that $K = -g(R(X_1, X_2)X_1, X_2)) = d\alpha(X_1, X_2)$. $\square$

Remark 6.3. The differential α depends on the orthonormal basis (X_1, X_2) and hence is not intrinsic to M. Still one can show that there is a canonically defined 'universal' differential $\tilde{\alpha}$ defined on the whole 3-manifold T_1M of unit tangent vectors of M which enjoys the following two properties (i) $d\tilde{\alpha} = \pi_1^*(K\mu_M)$, where $\pi_1 : T_1M \to M$ is the projection, and (ii) if X_1 is a unit vector field on an open $U \subset M$, and we regard X_1 as a map $U \to T_1U$ (so as a local section of π_1), then the pull-back $X_1^*(\tilde{\alpha})$ gives us just the α introduced above. Notice that property (i) already implies Corollary 6.2.

This approach (and the point of view that underlies it) is very fertile when going after higher dimensional generalizations.

Let $\gamma : [a, b] \to M$ be a differentiable arc parameterized by arc length: $\|\dot{\gamma}\| = 1$. The identity $0 = \frac{d}{dt} g(\dot{\gamma}, \dot{\gamma}) = 2g(\nabla_{\dot{\gamma}}\dot{\gamma}, \dot{\gamma})$ shows that the acceleration $\nabla_{\dot{\gamma}}\dot{\gamma}$ is always perpendicular to the velocity $\dot{\gamma}$. Let us denote by E_t the unique tangent vector in $T_{\gamma(t)}M$ for which $(\dot{\gamma}, E_t)$ is an oriented orthonormal basis. So $\nabla_{\dot{\gamma}}\dot{\gamma} = \rho(t)E_t$ for some scalar $\rho(t)$. For obvious reasons, $\rho(t)$

is called the *curvature* of γ at t. It is identically zero precisely when γ is a geodesic.

Now suppose that γ maps to U and denote by $\phi(t)$ the angle from $X_1(\gamma(t))$ to $\dot\gamma$ (measured counterclockwise relative the orientation). This angle is given modulo 2π, but its derivative $\dot\phi(t)$ is well-defined.

Lemma 6.4. *We have* $\gamma^*(\alpha) = (-\rho + \dot\phi)dt$.

Proof. This amounts to the equality $\alpha(\dot\gamma) = -\rho + \dot\phi$. We check this by computing the covariant derivative of $\dot\gamma = \cos\phi.\gamma^*X_1 + \sin\phi.\gamma^*X_2$ with respect to t. According to the Leibniz rule and Lemma 6.1 we have

$$\begin{aligned}\nabla_{\dot\gamma}\dot\gamma &= \nabla_{\dot\gamma}(\cos\phi.\gamma^*X_1 + \sin\phi.\gamma^*X_2)\\ &= \alpha(\dot\gamma)(\cos\phi. - \gamma^*X_2 + \sin\phi.\gamma^*X_1) + (-\sin\phi.\gamma^*X_1 + \cos\phi.\gamma^*X_2)\dot\phi\\ &= (-\alpha(\dot\gamma) + \dot\phi)E.\end{aligned}$$

If we take the inner product with E the left hand side becomes ρ. □

On winding numbers. Let $\gamma : [0, L] \to M$ be a piecewise differentiable loop parameterized by arc length (so $\gamma(L) = \gamma(0)$). Let $0 < t_1 < \cdots < t_N < L$ be the (finitely many) points of possible nondifferentiability of γ; we assume for simplicity that there is no kink at $t = 0$ so that $\dot\gamma(0) = \dot\gamma(L)$. Denote by $\Delta(t_i) \in (-\pi, \pi)$ the exterior angle of the kink at t_i:

$$\Delta(t_i) := \angle\big(\lim_{t\downarrow t_i}\dot\gamma(t), \lim_{t\uparrow t_i}\dot\gamma(t)\big)$$

(angles are always measured relative the given orientation and if $v, v' \in T_pM$ are nonzero, then $\angle(v, v')$ denotes the angle that we need to add to v' in order to reach v; the notation $\angle^{v}_{v'}$ would depict the situation better but is impractical). We call $\Delta_\gamma := \sum_i \Delta(t_i)$ the *total kink* of γ.

Let be given a nowhere zero *continuous* (not necessarily differentiable) vector field Y defined on an open subset of M. The reason we are happy with continuity is that we only use Y as a tool to measure angles with. So if p is in the domain of Y , then for any nonzero $v \in T_pM$ we have an angle $\angle(v, Y_p) \in \mathbb{R}/2\pi\mathbb{Z}$ defined.

We suppose the domain of Y contains the image of γ and we let $\phi(t) \in \mathbb{R}$ be essentially $\angle(\dot\gamma(t), Y(\gamma(t))$; precisely, let $\phi : [0, L] \setminus \{t_1, \ldots, t_N\} \to \mathbb{R}$ be defined by

(i) ϕ is continuous and equals the said angle modulo 2π
(ii) $\lim_{t\downarrow t_i}\phi(t) - \lim_{t\uparrow t_i}\phi(t) = \Delta(t_i)$ $(i = 1, \ldots, N)$.

Then $\phi(L) - \phi(0)$ will an integral multiple of 2π. The integral factor $(2\pi)^{-1}(\phi(L)-\phi(0))$ is called the *winding number* of Y along γ and we denote it by $w_\gamma(Y)$. Notice that

(windingno. formula) $$\int_0^L \dot\phi\, dt + \Delta_\gamma = 2\pi w_\gamma(Y).$$

This formula shows that the winding number is unaffected if γ is reparameterized by taking another point of departure: if $\gamma' : \mathbb{R} \to M$ is defined by $\gamma'(t) = \gamma(t + a)$ for some a, then $w_{\gamma'|[0,L]}(Y) = w_\gamma(Y)$. So if γ is without self-intersection, then the terms in the above equation only depend on the oriented image of γ; if the latter is denoted C, then we write Δ_C and $w_C(X)$ for them. Orientation reversal of γ makes both change sign.

If $\{Y_s\}_{0\le s\le 1}$ is a continuous family of nowhere continuouszero vector fields on a neighborhood of γ, then the angle $\phi_s(t)$ is continuous as a function of (s,t) and so $w_\gamma(Y_s)$ is continuous in s. On the other hand, this function is integer valued and so it must be constant in s. For similar reasons, w_γ is independent of the Riemann metric : if g_1 is another Riemann metric, then $g_s := (1-s)g|V + sg_1$ is a continuous family of such and hence the associated function $\phi_s(t)$ is continuous in (s,t).

Gauß-Bonnet I. We are now sufficiently equipped to prove the core part of the Gauß-Bonnet theorem.

Proposition 6.5. *Let* G *be a compact subset of* M *whose boundary* ∂G *is a union of piecewise differentiable loops. Orient* ∂G *as the boundary of* G *according to the standard convention (so that traversing it in the positive direction keeps* G *on the left). If* Y *is a nowhere zero* continuous *vector field on an open subset of* M *which contains* G*, then*

$$\int_G K\mu_M + \int_{\partial G} \rho\, dt + \Delta_{\partial G} = 2\pi w_{\partial G}(Y).$$

(Here dt *is the length form on the smooth part of* ∂G*. Since we allow* ∂G *to be disconnected,* $\Delta_{\partial G}$ *and* $w_{\partial G}$ *are to be understood as the sum of the contributions of its connected components.)*

Proof. The left hand side of the identity to be proven is independent of Y. The right hand side does not change if we first replace Y by a unit vector field X_1: we first pick a *differentiable* vector field Y_1 in such a manner that Y_1 nowhere points in direction opposite to Y and then let $X_1 := Y_1/\|Y_1\|$: now $\{(1-t)Y + tX_1\}_{0\le t\le 1}$ is a continuous family of nowhere zero vector fields and so the lefthand side has the same value for Y replaced by X_1.

So it is enough to prove the identity for X_1. Let X_2 be the unique vector field on the domain of Y such that (X_1, X_2) is there an oriented orthonormal basis. Then the preceding applies and we find

$$\int_G K\mu_M \overset{(6.2)}{=} \int_G d\alpha \overset{\text{Stokes}}{=\!=} \int_{\partial G} \alpha \overset{(6.4)}{=} \int_{\partial G} (-\rho + \dot\phi)dt.$$

The proposition now follows by invoking the winding number formula. □

Let us take the case when G is a *geodesic triangle*, i.e., the diffeomorphic image of a compact subset of $\mathbb{R}^2$ such that ∂G consists of three geodesic segments. Since the curvature of ∂G is now identically zero, that term does not enter in formula of Proposition 6.5. We may compute the winding number of ∂G using the Euclidean metric of $\mathbb{R}^2$ pulled back to G: then both integrals in Proposition 6.5 become zero and we find that $2\pi w_{\partial G}(X_1)$ equals the sum of the exterior angles of a triangle, that is 2π, which shows that $w_{\partial G}(X_1) = 1$. Feeding this into the formula, we see that the integral of K over G equals $2\pi - \Delta_{\partial G}$. If α, β, γ stand for the interior angles of G, then the last expression is known as the *angular excess* $\alpha + \beta + \gamma - \pi$. We thus obtain:

Proposition 6.6 (Gauß-Bonnet for a geodesic triangle)**.** *The angular excess of a geodesic triangle* G *in* M *equals* $\int_G K\mu_M$.

So the Gauß curvature in $p \in M$ has the same sign as the angular excess of a tiny geodesic triangle at p. Of special interest are the Riemannian surface of constant Gauß curvature $K \neq 0$, for instance the sphere $\mathbb{S}^2$ ($K > 0$) and the hyperbolic plane $\mathbb{H}^2$ ($K < 0$). Then the absolute value of the angular excess is proportional to the area of the triangle. It is easy to derive from this that the geodesic triangle is determined by $K \neq 0$ and its interior angles. This is of course not so for Euclidean triangles (for which $K = 0$), since these can be similar without being congruent.

Index of a planar vector field. Let Y be a continuous vector field on a differentiable 2-manifold N whose zero set $Z(Y)$ consists of isolated points. (Without proof we mention that 'in general' the points of $Z(Y)$ are isolated.)

We are going to define for every $p \in N$ an integer, the *index* of Y. Let (U, κ) be a chart at p with $\kappa(p) = 0$ and Y nowhere zero on $U - \{p\}$. Choose $R > 0$ so small that $\kappa(U)$ contains the closed disk D of radius R centered at 0 and parameterize ∂D by angle: $\gamma(t) := (R\cos t, R\sin t) \in \partial D$. Let $\alpha(t) \in \mathbb{R}$ be the angle from e_1 to $(\kappa_* Y)_{\gamma(t)}$ (we assume α continuous). Since α is defined modulo $2\pi\mathbb{Z}$ we have that $\alpha(2\pi) - \alpha(0) = 2\pi I$ for some $I \in \mathbb{Z}$. The left hand side is continuous in R and so I is independent of R. In fact:

Lemma-definition 6.7. *The integer* I *is independent of* κ*; we call it the* index *of* Y *at* p *and denote it by* $I_p(Y)$.

Proof. Suppose κ' is another chart as above and denote by $I(\kappa')$ the associated integer. If we can connect κ with κ' by a continuous family of such charts $\{\kappa_t\}_{0\le t\le 1}$, then $I(\kappa_t)$ is continuous in t, hence constant and we may conclude that $I(\kappa) = I(\kappa')$. We claim this is possible in case the nonsingular 2×2-matrix $\sigma := D_0(\kappa'\kappa^{-1})$ has positive determinant. Then the identity matrix 1_2 can be connected in $GL(2,\mathbb{R})$ with σ by a continuous family $\{\sigma_t\}_{0\le t\le 1}$. Hence $\sigma_t^{-1}\kappa'$ is a family of charts connecting κ' with $\kappa'' := \sigma^{-1}\kappa'$ and so $I(\kappa') = I(\kappa'')$. Notice that $D_0\kappa''\kappa^{-1} = 1_2$, or equivalently, $D_p\kappa = D_p\kappa''$. So $\kappa_t := (1-t)\kappa + t\kappa''$ also has the property that $D_p\kappa_t = D_p\kappa$ and the inverse function theorem shows that this is a continuous family of charts connecting κ with κ''. It follows that $I(\kappa) = I(\kappa'')$.

It remains to do a single case for which $\det(\sigma)$ is negative. For that we take $\kappa' = (\kappa^1, -\kappa^2)$. Then a corresponding angle function α' is given by

$\alpha'(t) = -\alpha(2\pi - t)$ and we find that $2\pi I(\kappa') = \alpha'(2\pi) - \alpha'(0) = -\alpha(0) + \alpha(2\pi) = 2\pi I(\kappa)$ so that we still have $I(\kappa') = I(\kappa)$. □

Gauß-Bonnet II. Returning to M, suppose that on a neighborhood U of p we are given continuous vector fields X and Y with X nowhere zero and Y nonzero on $U - \{p\}$. Let $\gamma : [0, 2\pi] \to M$ parameterize a small circle C around p (compatible with the orientation and relative some chart at p). If $\beta(t)$ denotes the angle from $X_{\gamma(t)}$ to $\dot{\gamma}(t)$, then $\beta(2\pi) = \beta(0) + 2\pi$. It follows that

$$\text{(index equality)} \qquad I_p(Y) = w_C(X) - w_C(Y) = 1 - w_C(Y).$$

Theorem 6.8 (Gauß-Bonnet)**.** *Let* M *be a compact oriented surface . If* Y *is a continuous vector field on* M *with isolated zeroes only, then*

$$\frac{1}{2\pi}\int_M K\mu_M = \sum_{p\in\Sigma(Y)} I_p(Y).$$

Proof. Choose for every zero p of Y a closed neighborhood D_p diffeomorphic to a disk, and such that $D_p \cap D_q = \emptyset$ if $p \neq q$. The orientation of M defines one for D_p and hence one for ∂D_p. The complement G in M of the union of the interiors of the D_p's is an oriented manifold met boundary $\partial G = \cup(-\partial D_p)$, where the minus sign indicates that the orientation is opposite to the one we just fixed. According to 6.5 we have

$$\int_G K\mu_M = -\int_{\partial G} \rho\, dt + 2\pi w_{\partial G}(Y).$$

If we choose a nowhere zero vector field X_p on a neighborhood of D_p, then again by 6.5:

$$\int_{D_p} K = -\int_{\partial D_p} \rho\, dt + 2\pi w_{\partial D_p}(X_p).$$

Adding these identities yields

$$\begin{aligned}
\int_M K\mu_M &= \sum_{p\in\Sigma(Y)} 2\pi\big(w_{-\partial D_p}(Y) + w_{\partial D_p}(X_p)\big) \\
&= \sum_{p\in\Sigma(Y)} 2\pi\big(-w_{\partial D_p}(Y) + w_{\partial D_p}(X_p)\big) \\
&= \sum_{p\in\Sigma(Y)} 2\pi I_p(Y) \quad \text{by the index equality.}
\end{aligned}$$

(The ρ-terms canceled out because of the difference in orientation.) □

The Gauß-Bonnet theorem is remarkable for at least two reasons. First of all because it tells us that the integral of the Gauß curvature over M is an integral multiple of 2π *that is independent of the metric*. And secondly its implication that the index sum of a vector field with only isolated zeroes is a constant *that is independent of the vector field*. This constant is called the

Euler characteristic of M; we denote it by $e(M)$. If that constant is nonzero, the tangent bundle of M cannot be trivial.

Euler introduced his characteristic differently: his point of departure was a *triangulation* of M, i.e., a paving of M by (the diffeomorphic images of) solid triangles that meet in edges or vertices. (It can be shown that a triangulation always exists.) Given a triangulation of M, then we can construct a *continuous* vector field Y on M by letting its restriction to any triangle as below:

So Y has zeroes in the vertices (of index $+1$), in de midpoints of edges (of index -1), and in the barycenters of the trangles (of index $+1$), so that

$$e(M) = \sum_p I_p(Y) = \#(\text{vertices}) - \#(\text{edges}) + \#(\text{triangles}).$$

A compact connected orientable surface is diffeomorphic to the surface of a body with $g \geq 0$ 'handles'. Now one easily verifies (for instance by triangulating) that the Euler characteristic of such a surface equals $2 - 2g$. Hence g is determined by M; we call it the *genus* of M. So if M has trivial tangent bundle, then its Euler characteristic vanishes, its genus therefore equals 1, and so M must be diffeomorphic to torus (which has a trivial tangent bundle indeed).

Corollary 6.9. *(i) There is no metric on S^2 with curvature ≤ 0 everywhere.*
(ii) The curvature of any metric on a torus changes sign or is constant zero.
(iii) A compact connected oriented surface of genus ≥ 2 has no metric whose curvature is ≥ 0.

Proof. This follows from the identity $\frac{1}{2\pi}\int_M K\mu_M = 2 - 2g$. □

Remarks 6.10. We say that two metrics g, g' on a Riemann manifold N are *conformally equivalent* if $g' = f.g$ for some differentiable function $f : N \to (0, \infty)$. In other words, g and g' define the same notion of angle. One can show that any metric on a connected compact surface, is conformally equivalent to one with constant curvature $1, 0$, of -1, the sign being the sign of the Euler characteristic. We get in the oriented case respectively

(1) the sphere $\mathbb{S}^2$,
(0) a flat torus $\mathbb{E}^2/(\mathbb{Z}v + \mathbb{Z}w)$, where (v, w) is a basis for $\mathbb{E}^2$,

(−1) a Riemannian surface isometric to $\Gamma\backslash\mathbb{H}^2$, where Γ is a discrete group of isometries of $\mathbb{H}^2$ which acts freely.

A conformal equivalence class (= a notion of angle) on a compact connected surface plus an orientation of that surface is the same thing as a 'complex structure' on M: M may be then considered as a complex manifold of complex dimension one. Such an object is called, perhaps somewhat confusingly, a *Riemann surface*. From that point of view (1) corresponds to the Riemann sphere and (0) to an elliptic curve. Riemann surfaces are mathematical gems—not only give they rise to deep and beautiful theorems, but at the same time they are quite fundamental.

Exercise 33. State and prove a common generalization of Proposition 6.5 and Theorem 6.8.

Exercise 34. Let (M, g) be a compact oriented Riemannian surface with boundary and let Y be a vector field on M that is tangent to ∂M and has only isolated zeroes, none of them on ∂M. Prove that

$$\int_M K\mu_M + \int_{\partial M} \rho\, dt = \sum_{p\in\Sigma(Y)} I_p(Y).$$

Conclude that the left hand side is independent of g and the right hand side is independent of Y.

Prove that for any triangulation of M, the members of the above identity are also equal to the Euler characteristic $e(M) = \#(\text{vertices}) - \#(\text{edges}) + \#(\text{triangles})$.

References

[1] S.S. Chern: *Differentiable manifolds*, Typescript, Instituto de Fysica e mattemática (1959).

[2] J.M. Lee: *Riemannian Manifolds An Introduction to Curvature*, Graduate Texts in Mathematics, Vol. 176, Springer (1997) Softcover ISBN: 978-0-387-98322-6.

[3] K. Nomizu: *Lie groups and differential geometry*, The Mathematical Society of Japan (1956).

[4] S. Kobayashi en K. Nomizu: *Foundations of differential geometry I, II*, Interscience Tracts, 15, Wiley (1963, 1968).

[5] W. Klingenberg: *Riemannian Geometry*, Studies in Mathematics 1, De Gruyter, (1982).

[6] M. Spivak: *A comprehensive introduction to differential geometry*, Typescript in five volumes, Publish or perish (1970...).

[7] D. Gromoll, W. Klingenberg and W. Meyer: *Riemannsche Geometrie im Großen*, Lecture Notes in Mathematics 55, Springer.

[8] S. Helgason: *Differential Geometry and Symmetric Spaces*, Pure and Applied Mathematics XII, Academic Press (1962).

[9] J.W. Milnor: *Morse Theory*, Annals of Math. Studies, Princeton UP 51 (1969).

The first three items are concise introductions that go well with these notes, (4)-(6) are more thorough and (7)-(9) emphasize particular aspects.