CALCULUS ON MANIFOLDS

1. Riemannian manifolds

Recall that for any smooth manifold M, $\dim M = n$, the union $TM = \bigcup_{a\in M} T_aM$, called the *tangent bundle*, is itself a smooth manifold, $\dim TM = 2n$.

Example 1. Prove this and construct an atlas of charts on TM.

Recall that each bilinear form $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ can be made into a quadratic form $Q\colon \mathbb{R}^n \to \mathbb{R}$ by restricting on the diagonal, $Q(v) = B(v,v)$. Conversely, any quadratic form Q can be *polarized* to a *symmetric* bilinear form $B(u,v) = \frac{1}{2}(Q(u+v)-Q(u)-Q(v))$. It is positive definite, if $Q(v) \geqslant 0$ and $Q(v) = 0 \iff v = 0$.

1.1. **Riemannian manifolds: definitions.**

Definition 1. A Riemannian metric on a manifold M is a smooth function $g\colon TM \to \mathbb{R}_+$ which, when restricted on any tangent space T_aM, is a positive definite quadratic form. A Riemannian manifold is a manifold equipped with a specific Riemannian metric.

Example 2. The Euclidean space $\mathbb{R}^n$ equipped with the same standard quadratic form $g(v) = \langle v,v\rangle = \sum_{i=1}^n v_i^2$ for any $v = (v_1,\dots,v_n) \in T_a\mathbb{R}^n \simeq \mathbb{R}^n$.

The vast source of examples is provided by the *induced* structures: if M is a Riemannian manifold and $N \subseteq M$ is a submanifold. Then $TN \subseteq TM$, and the restriction of g on TN is a Riemannian metric making N into a Riemannian manifold.

Thus any smooth submanifold of $\mathbb{R}^n$, say, the unit sphere $\mathbb{S}^{n-1}$ inherits the structure of a Riemannian manifold.

The same manifold can be equipped with different Riemannian metrics. For instance, the torus obtained by rotation of a circle around an axis in $\mathbb{R}^3$ is embedded in $\mathbb{R}^3$ in the natural way and so inherits the Riemannian structure.

On the other hand, if we consider the quotient $\mathbb{T}^2_{a,b} = \mathbb{R}^2/a\mathbb{Z} + b\mathbb{Z}$ for a pair of linear independent vectors $a,b \in \mathbb{R}^2$, then it is naturally equipped by the Riemann metric in $\mathbb{R}^2$, since all translations preserve the scalar product. These are called *flat* tori.

Definition 2. Two Riemannian manifolds (M,g) and (N,h) are *isometric*, if there exists a diffeomorphism F between them, which transforms one Riemannian metric into another.

If $F\colon (M,g) \to (N,h)$ is a differentiable injective[1] map such $F^*h = g$, then we say that M is isometrically embedded in N.

1.2. **What can be done with a Riemannian metric?** First, it allows to identify $\mathscr{X}(M)$ with $\Omega^1(M)$: a vector field X and a 1-form ξ are dual, if for any other vector field $Y \in \mathscr{X}(M)$ one has

$$g(X,Y) = \xi(Y).$$

Example 3. The vector field $\operatorname{grad} f$ is dual to the 1-form $\mathrm{d}f$.

A *huge* generalization of this example is the *Hodge star operator* making $\Omega^k(M)$ and $\Omega^{n-k}(M)$ into dual spaces. Will be discussed later.

Second, we can measure lengths of the smooth curves and angles between them: by definition, the length of a smooth curve $\gamma\colon [0,1] \to M$ is the integral

$$L(\gamma) = \int_0^1 \sqrt{g(\dot\gamma(t), \dot\gamma(t))}\,\mathrm{d}t > 0.$$

Note that the length *is not* given by the integral of a 1-form, so $L(\gamma) = L(-\gamma)$, but is independent of the parametrization of the curve. A similar expression is given by the *action*[2], the integral

$$A(\gamma) = \int_0^1 g(\dot\gamma(t), \dot\gamma(t))\,\mathrm{d}t = \int_\gamma \xi_\gamma, \qquad \xi_\gamma = i_{\dot\gamma} g = g(\dot\gamma, \,\cdot\,) \in \Omega^1(M),$$

where ξ_γ is the 1-form dual to the velocity vector vector field $\dot\gamma$ defined along γ. The advantage of action is differentiability of the integrand, disadvantage is the explicit dependence on parametrization.

Having lengths, one can meaningfully talk about shortest curves connecting two given endpoints. Such curves are called geodesic curves. An smooth infinite curve is called geodesic, if it can be split into shortest curves (spell out the formal definition!).

Example 4. In the Euclidean space the shortest curves are line segments, and infinite geodesics are straight lines.

One can use simple geometric arguments to show that on the standard round sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ the shortest curves are arcs of the large circles (sections of the sphere by planes through the origin): it is enough to verify that such curves must be planar. The arc of length $> \pi$ is not shortest anymore (its complement is), and large circles (without the endpoints) are geodesics. There are no non-closed geodesics on the sphere!

There are no topological obstructions for a manifold to be Riemannian.

Theorem 1. *Any smooth manifold can be equipped with a Riemannian metric.*

---

[1]Why there cannot be isometric maps which *decrease* the dimension?

[2]Sometimes this integral is referred to as the energy and preceded by the coefficient $\frac{1}{2}$.

Proof. Let $\{U_i\}$ be a locally finite covering of M by charts $x_i\colon U_i \to \mathbb{R}^n$ and $1 = \sum_i \psi_i$ the subordinate partition of unity, $\mathscr{C}(M) \ni \psi \geqslant 0$, $\operatorname{supp} \psi_i \subseteq U_i$. In each chart we have metric g_i pulled back from $\mathbb{R}^n$. The sum $\sum \psi_i g_i$ yields a Riemannian metric which is nondegenerate.

Note that this argument would fail if instead of the positive definite quadratic form we were looking for a metric with a different signature! □

1.3. **Parallel transport and covariant derivation.** There is a rather large group of isometries of the Euclidean space $(\mathbb{R}^n, g)$, $g(x, v) = \langle v, v\rangle$, containing the subgroup (isomorphic to $\mathbb{R}^n$) of parallel translations.

The existence of the parallel transport allows one to calculate the derivative of a vector field Y along any smooth curve as follows: if the curve is given by the parametrization $\gamma : [0, 1] \to \mathbb{R}^n$, $t \mapsto x(t) = (x_1(t), \dots, x_n(t))$ and the vector field is given by its coordinate functions $(Y_1(x), \dots, Y_n(t))$, then we can take a composition $Y(\gamma(t))\colon [0, 1] \to \mathbb{R}^n$ and compute the velocity $\frac{\mathrm{d}}{\mathrm{d}\gamma} Y = \frac{\mathrm{d}}{\mathrm{d}t} Y(\gamma(t))$. By the chain rule, it can be expressed through the differential operator

$$\overline{\nabla}_X Y = (L_X Y_1, \dots, L_X Y_n), \qquad X, Y, \overline{\nabla}_X Y \in \mathscr{X}(\mathbb{R}^n), \tag{1}$$

as the derivation along the velocity vector $\dot\gamma$,

$$\frac{d}{\mathrm{d}\gamma} Y = \overline{\nabla}_{\dot\gamma} Y.$$

Looking at the operator $\overline{\nabla}$, we immediately see that it satisfies a number of properties:

(1) $\overline{\nabla}$ is $\mathbb{R}$-linear in both arguments X, Y,
(2) $\overline{\nabla}_X$ respects the Leibniz rule: $\overline{\nabla}_X(fY) = (\overline{\nabla}_X f)Y + f(\overline{\nabla}_X Y)$. There is only one way to interpret $\overline{\nabla}_X f$ for $f \in \mathscr{C}(M)$ as $L_X f$.
(3) $\overline{\nabla}_{fX} Y = f \overline{\nabla}_X Y$, that is, the operator $\overline{\nabla}$ acts on vector fields in the same way as L_X acts on functions (such behavior is called *tensorial*).
(4) Moreover, if $Y, Z \in \mathscr{X}(\mathbb{R}^n)$ are two vector fields, then the Leibniz rule holds for the scalar product $\langle \cdot, \cdot \rangle$:

$$\overline{\nabla}_X \langle Y, Z\rangle = L_X \langle Y, Z\rangle = \left\langle \overline{\nabla}_X Y, Z\right\rangle + \left\langle Y, \overline{\nabla}_X Z\right\rangle. \tag{2}$$

By definition, $\overline{\nabla}_X$ acts on vector fields in the same way as $X = L_X$ acts on scalar functions. In particular,

$$\overline{\nabla}_X Y - \overline{\nabla}_Y X = [X, Y]. \tag{3}$$

Definition 3. A differential operator $\nabla\colon \mathscr{X}(M) \times \mathscr{X}(M) \to \mathscr{X}(M)$ satisfying the properties (1)–(3) above, is called the *covariant derivation*. The image $\nabla_X Y$ is the *covariant derivative* of the vector field Y along the vector field X.

If γ is a (piecewise)-smooth curve, than the "vector field" Y defined only on the image of γ, is called *constant along* γ, if $\nabla_{\dot\gamma} Y = 0$. For any such field the value $Y(\gamma(1))$ is called the *parallel transport* of $Y(\gamma(0))$ along γ.

We say that ∇ defines a *connexion* on the manifold M. It is called *symmetric*, if (3) holds. If M is a Riemannian manifold and (2) holds, we say about the Riemannian connexion.

The property (2) means that the parallel transport in $\mathbb{R}^n$ is an isometry between $T_{\gamma(0)}\mathbb{R}^n$ and $T_{\gamma(1)}\mathbb{R}^n$.

1.4. **Induced covariant derivation.** We will show how the derivation $\overline{\nabla}_X Y$ can be "restricted" to vector fields tangent to a hypersurface $M \subseteq \mathbb{R}^n$, producing a vector field which is also tangent to M.

If $M^{n-1} \subset \mathbb{R}^n$ is a smooth hypersurface, then the parallel transport in $\mathbb{R}^n$ does not map tangent subspaces into themselves. On the infinitesimal level, the covariant derivative $\overline{\nabla}_X Y$ for two vector fields $X, Y \in \mathscr{X}(M)$ tangent to M is not necessarily tangent to M.

The required correction is both natural and minimal.

Definition 4. The covariant derivative $\nabla_X Y$ of a vector field $Y \in \mathscr{X}(M)$ along $X \in \mathscr{X}(M)$ is the orthonormal projection of $\overline{\nabla}_X Y \in \mathscr{X}(\mathbb{R}^n)$ onto $TM \subseteq T\mathbb{R}^n \simeq \mathbb{R}^n \times \mathbb{R}^n$. Of course, it depends on the hypersurface M.

A vector field $Y \in \mathscr{X}(M)$ is said to be parallel along a curve $\gamma : [0,1] \to M$, if $\nabla_{\dot\gamma} Y = 0$. This defines the linear operator between $T_{\gamma(0)}\mathbb{R}^n$ and $T_{\gamma(1)}\mathbb{R}^n$, called the *parallel transport*.

Example 5. Let $E_1, \dots, E_n \in \mathscr{X}(M)$ be (locally) linear independent vector fields on M, say, coming from the derivations $\partial_{x_1}, \dots, \partial_{x_n}$ in some chart. Then we can define n^3 functions on M by the conditions,

$$\nabla_{E_i} E_j = \sum_{k=1}^{n} \Gamma_{ij}^k E_k, \tag{4}$$

expanding the covariant derivatives in the basis $\{E_i\}$. The functions Γ_{ij}^k are called the *Christoffel symbols*. If the connexion ∇ is symmetric, then $\Gamma_{ij}^k = \Gamma_{ji}^k$.

Assume that the curve γ is given by its smooth parametric equation $t \mapsto \gamma(t)$. Then a vector field $Y = \sum y_i(t) E_i(t)$ is parallel along γ, if and only if $\nabla_{\dot\gamma} Y = 0$. Expanding this identity using the Leibniz rule, we obtain a system of linear homogeneous ODEs for the functions $y_i(t)$, which can be solved with any initial condition.

The fundamental question is as follows. Embedding of M into $\mathbb{R}^n$ induces the Riemannian metric on M. The parallel transport and covariant derivation are explicitly defined in terms of the embedding (which is much more information). Can one determine the result of the transport in terms of the induced metric only? The answer obtained by Gauss, the celebrated *Theorema Egregium*, is affirmative. We outline the steps towards explaining this result.

Let $M^{n-1} \subset \mathbb{R}^n$ be a smooth hypersurface and N a unit normal vector field on it:

$$\forall a \in M \ N(a) \in T_a\mathbb{R}^n, \quad \langle N(a), N(a)\rangle = 1, \quad \langle N(a), \cdot\rangle|_{T_aM} = 0.$$

The map $a \mapsto N(a)$ is sometimes called the Gauss map $M \to \mathbb{S}^{n-1} \subseteq \mathbb{R}^n$.

For any $X \in \mathscr{X}(M)$ the "ambient covariant derivative" $\overline{\nabla}_X N$ is tangent to M. Indeed, by the Leibniz rule

$$\left\langle \overline{\nabla}_X N, N \right\rangle = \tfrac{1}{2} L_X \langle N, N\rangle = 0, \qquad \text{hence} \quad N \perp \overline{\nabla}_X N \in \mathscr{X}(M).$$

Definition 5. The *Weingarten operator* is the linear operator $W_a \colon T_aM \to T_aM$ which sends a vector v into $\overline{\nabla}_v N$ (since $\overline{\nabla}$ is a covariant derivative, one can choose any $X \in \mathscr{X}(M)$ such that $X(a) = v$, and compute $\overline{\nabla}_X N$). For a vector field $X \in \mathscr{X}(M)$ we denote by WX the field $a \mapsto W_a X(a) \in T_aM$.

The Weingarten operator W_a is the Jacobian (derivative) of the Gauss map at a point $a \in M$.

Example 6. For the round unit sphere $\mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ the Gauss map is the identity, so the Weingarten operator is also identical.

Lemma 1. *The Weingarten operator is self-adjoint in the induced metric on T_aM.*

Proof. Let $X, Y \in \mathscr{X}(M)$ be any two vector fields tangent to M. We want to prove that

$$\langle WX, Y\rangle = \langle X, WY\rangle .$$

Without loss of generality we may assume that X, Y, N are defined in a neighborhood of M. Then, using twice the Leibniz rule (2) and the symmetry condition (3) of $\overline{\nabla}$, we conclude that

$$\begin{aligned}\langle WX, Y\rangle - \langle X, WY\rangle &= \left\langle \overline{\nabla}_X N, Y\right\rangle - \left\langle X, \overline{\nabla}_Y N\right\rangle \\ &= \left(\overline{\nabla}_X \langle Y, N\rangle - \left\langle \overline{\nabla}_X Y, N\right\rangle\right) - \left(\overline{\nabla}_Y \langle X, N\rangle - \left\langle N, \overline{\nabla}_Y X\right\rangle\right) \\ &= 0 - 0 - \left\langle N, \overline{\nabla}_X Y - \overline{\nabla}_Y X\right\rangle = -\langle N, [X, Y]\rangle .\end{aligned}$$

But the commutator of two vector fields X, Y tangent to M, is again tangent to M, hence the result is zero. □

Definition 6. The eigenvalues $\lambda_1(a), \dots, \lambda_{n-1}(a)$ of the Weingarten operator L_a are called the *principal curvatures* of the hypersurface. By construction, they are functions of the point on M. While the principal curvatures themselves may be non-smooth, their symmetric combinations (coefficients of the characteristic polynomial $\det(\lambda - L_a)$), are smooth. In particular, smooth are the determinant $K(a) = \prod_1^{n-1} \lambda_i(a)$, called the *Gaussian curvature*, and the trace $H(a) = \sum_1^{n-1} \lambda_i(a)$ called the *mean curvature* of M. The corresponding eigenspaces of T_pM are called *directions of curvature* (they are pairwise orthogonal for different eigenvalues).

Note that all these definitions are still non-intrinsic: they explicitly depend on the embedding of M into $\mathbb{R}^n$.

The meaning of the curvatures can be seen from the following construction. Let $X_i(a) \in T_aM$ be the eigenvector of $W = W_a$, corresponding to the eigenvalue $\lambda_i(a)$. Consider the 2-plane $\Pi_i(a)$ in $\mathbb{R}^n$ spanned by $N(a)$ and $X_i(a)$. Then the Gauss map restricted on the 1-dimensional subspace $\mathbb{R}X_i(a) \subset T_aM$ and its derivative can be instantly computed in terms of the osculating circle of the section $\Pi_i(a) \cap M$ as the inverse radius of this circle.

1.5. **The Gauss equation.** Using the Weingarten operator, one can easily express the induced covariant derivative as it was defined in Definition 4: to compute the orthogonal projection, one has to add to $\overline{\nabla}_X Y$ the normal vector N with a suitable coefficient,

$$\nabla_X Y = \overline{\nabla}_X Y - \left\langle \overline{\nabla}_X Y, N \right\rangle N = \overline{\nabla}_X Y + \left\langle Y, \overline{\nabla}_X N \right\rangle N = \overline{\nabla}_X Y + \langle WX, Y \rangle N,$$

since

$$0 = \overline{\nabla}_X \langle Y, N \rangle = \left\langle \overline{\nabla}_X Y, N \right\rangle + \left\langle Y, \overline{\nabla}_X N \right\rangle = \langle Y, WX \rangle .$$

1.6. **Curvature and the Codazzi–Mainardi equations.** We start with the obvious identity

$$\overline{\nabla}_X \overline{\nabla}_Y - \overline{\nabla}_Y \overline{\nabla}_X = \overline{\nabla}_{[X,Y]}$$

between first order differential operators: it is true when applied to functions, hence to vector fields in $\mathbb{R}^n$.

Substituting into this identity the Gauss identity $\nabla_X = \overline{\nabla}_X - \langle WX, \cdot \rangle N$, applying the result to a third vector field $Z \in \mathscr{X}(\mathbb{R}^n)$ and using several times the Leibniz rule and linearity of the Weingarten operator W, we can separate at the end the normal and tangential components of the result. These will yield us two identities:

$$\big(\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}\big) Z = \langle WY, Z \rangle WX - \langle WX, Z \rangle WY, \qquad (5)$$

and another identity valid for any Z, which implies that

$$\nabla_X(WY) - \nabla_Y(WX) = W[X,Y]. \qquad (6)$$

The identity (5) is very remarkable: it asserts that a certain differential operator that could apriori be of order 2 with respect to Z and of order 1 with respect to X, Y, is of order zero with respect to all arguments! In the classical language, it is a tensor, called the *curvature tensor*, usually considered as a multilinear scalar function of *four* vector arguments

$$\begin{aligned} R(X,Y,Z,V) &= \left\langle \big(\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}\big) Z, V \right\rangle \\ &= \left\langle \langle WY, Z \rangle WX - \langle WX, Z \rangle WY, V \right\rangle \\ &= \langle WY, Z \rangle \langle WX, V \rangle - \langle WX, Z \rangle \langle WY, V \rangle . \qquad (7) \end{aligned}$$

Note that the Weingarten operator is self-adjoint, which means that the curvature tensor has a very rich symmetry with respect to permutations of the arguments. The least obvious of these is the so called (the first) *Bianchi* *formula*, which is obtained by substituting the symmetry assumption $\nabla_X Y - \nabla_Y X = [X,Y]$ into the Jacobi identity $[X,[Y,Z]]+[Y,[X,Z]]+[Z,[X,Y]] = 0$.

Problem 1. Write down this identity explicitly.

For the case of 2-surfaces in $\mathbb{R}^3$, the geometric nature of this tensor can be described as follows. Let X, Y be two orthonormal vectors. Then

$$R(X,Y,X,Y) = \langle WX, Y\rangle \langle WY, X\rangle - \langle WY, X\rangle \langle WY, Y\rangle \\ = -\det W = -\lambda_1\lambda_2 = -R(X,Y,Y,X) \quad (8)$$

is the Gaussian curvature of the surface.

This formula is important since the left hand side of (5) is defined in terms of the intrinsic geometry of the surface M (the connexion ∇ and the Riemannian metric), while the Weingarten operator depends on the embedding of the surface M into $\mathbb{R}^3$.

1.7. **How unique is the Riemannian connexion?** There is only one covariant derivative which is compatible with a given Riemannian metric on a manifold. Recall that we conveniently denote the Lie derivation on functions as $\nabla_X = L_X = X$.

Theorem 2. *The covariant derivation ∇ on a Riemannian manifold, which is symmetric and preserves the Riemannian structure, i.e., $\forall X, Y, Z \in \mathscr{X}(M)$ satisfying the conditions*

$$\nabla_X Y - \nabla_Y X = [X,Y], \qquad \nabla_X \langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle,$$

is unique.

Remark 1. The compatibility condition means that the parallel transport along any curve is an isometry between the respective tangent spaces. Indeed, if both $X(t), Y(t)$ are parallel along γ, i.e., $\nabla_{\dot\gamma} X = \nabla_{\dot\gamma} Y = 0$, then

$$\langle X(1), Y(1)\rangle - \langle X(0), Y(0)\rangle = \int_0^1 \nabla_{\dot\gamma} \langle X(t), Y(t)\rangle \, \mathrm{d}t \\ = \int_0^1 \langle \nabla_{\dot\gamma} X, Y\rangle + \langle X, \nabla_{\dot\gamma} Y\rangle \, \mathrm{d}t = 0.$$

Proof of the Theorem. Let $X, Y, Z \in \mathscr{X}(M)$ be any three *commuting* vector fields. We have the following identities:

$$\begin{aligned} \nabla_X \langle Y, Z\rangle &= \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle, \\ \nabla_Y \langle X, Z\rangle &= \langle \nabla_Y X, Z\rangle + \langle X, \nabla_Y Z\rangle, \\ \nabla_Z \langle X, Y\rangle &= \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle \end{aligned} \quad (9)$$

Adding the first two and subtracting the third identity, we have

$$\begin{aligned}\nabla_X \langle Y, Z\rangle + \nabla_Y \langle X, Z\rangle - \nabla_Z \langle X, Y\rangle = \\ \langle \nabla_X Y, Z\rangle + \langle Y, [X, Z]\rangle + \langle \nabla_Y X, Z\rangle + \langle X, [Y, Z]\rangle \\ = 2\langle \nabla_X Y, Z\rangle - \langle Z, [X, Y]\rangle = 2\langle \nabla_X Y, Z\rangle. \quad (10)\end{aligned}$$

Note that the left hand side depends only on the Riemannian metric, while the right hand side involves the covariant derivative $\nabla_X Y$.

Let $E_1, \dots, E_n \in \mathscr{X}(M)$ be coordinate vector fields for any local coordinate system (commuting by definition), and apply the identity (10) to all triples E_i, E_j, E_k. As a result, for any pair i, j one gets an expression of the covariant derivative $\nabla_{E_i} E_j$ via its projections on each direction E_k in terms of the Riemannian metric. This determines the covariant derivative uniquely, if the Riemannian metric is non-degenerate (it always is by definition). □

This implies what Gauss called *Theorema Egregium* and is the "cartographer's nightmare": the surface of the Earth cannot be rendered isometrically on the flat paper. Indeed, the Gauss curvature of the sphere is positive (compute it for the sphere of radius $r > 0$), while that of the plane is zero. Yet the globus is perfect as a scaled image.