GUOWU MENG

CALCULUS III

PUBLISHER OF THIS BOOK

PUBLISHED BY PUBLISHER OF THIS BOOK

TUFTE-LATEX.GOOGLECODE.COM

First printing, February 2025

Contents

Dedicated to God who used beautiful mathematics

to create our universe.

Introduction

This book serves as an introductory text to modern differential geometry. Despite its focus on advanced mathematical concepts, the title "Calculus III" was chosen deliberately.

Calculus is the study of functions. In Calculus I, we examine real functions of one variable, while in Calculus II, we study vector-valued functions of several variables. These functions are defined on open connected sets of Euclidean spaces, with variables providing a smooth global parametrization for the domain. In Calculus III, the functions we study are more general in two ways. Firstly, a function may be defined in a space that locally resembles a Euclidean space. Secondly, a function may take value in a family of vector spaces parameterized by the domain of the function. Tensor fields, or more generally, sections of vector bundles, are the functions studied in Calculus III.

A space that locally resembles a Euclidean space is called a manifold. Unlike open sets of Euclidean spaces, a manifold can be a highly complex geometric object. As such, Calculus III is a part of geometry. Furthermore, the extension of calculus in this generality necessitates thinking without coordinates.

In summary, Calculus III is a significant expansion of our mathematical horizon in five interrelated areas: space, geometry, functions, differentiation, and integration.

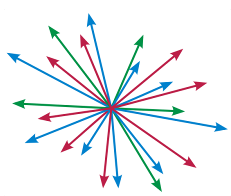
> "It is by logic that we prove, but by intuition that we discover."
>
> Henri Poincaré

§1 Review of Linear Algebra

The purpose of this section is to review the essential parts of Linear Algebra that are used in geometry, assuming all vector spaces are real vector spaces.

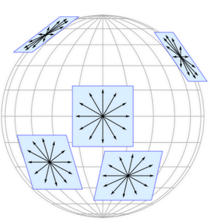
Examples of Vector Spaces

The notion of a vector space is abstracted from the collection of vectors on a plane with the same tail.



courtesy of https://commons.wikimedia.org/wiki/File:Vector_space_illust_rotate.svg

A typical example of a vector space on the geometric side is the tangent space of a smooth space at a point.



courtesy of an internet source

Vector spaces are also called linear spaces when emphasizing the algebraic side. If V is a vector space and X is a non-empty set, the set of maps from set X to vector space V, denoted by $\mathrm{Map}(X, V)$, is a vector space with a linear structure where for any $\alpha_1, \alpha_2 \in \mathrm{Map}(X, V)$ and any scalars c_1 and c_2, we have

Here, ":=" means "is defined to be," and $\cdot$ means scalar multiplication in V.

$$(c_1\alpha_1 + c_2\alpha_2)(x) := c_1 \cdot \alpha_1(x) + c_2 \cdot \alpha_2(x).$$

Here, the linear combination on $\mathrm{Map}(X, V)$ is defined point-wise.

Any non-empty subset of a vector space V that is closed under linear combination is a vector space and referred to as a **linear subspace** of V. Vector spaces often arise as subspaces, especially as subspaces of vector spaces of the form $\mathrm{Map}(X, V)$. As a result, one can quickly conclude that the kernel and image of linear maps, the span of a set of vectors, and $\mathrm{Hom}(V_1, V_2)$ (the set of linear maps from vector space V_1 to vector space V_2) are all vector spaces. In particular, the set of linear maps from V to $\mathbb{R}$ ($= \mathbb{R}^1$), denoted by V^*, is a vector space — the dual vector space of V.

If we call elements in V vectors, then elements in V^* are called co-vectors.

If $X = \{t^n \mid n = 0, 1, \ldots\}$, we have the natural vector space identification $\mathrm{Map}(X, \mathbb{R}) \equiv \mathbb{R}[[t]]$ — the vector space of real formal power series in t. The set of real polynomials in t, denoted by $\mathbb{R}[t]$, is a linear subspace of $\mathbb{R}[[t]]$.

For any two vector spaces V_1 and V_2, the Cartesian product $V_1 \times V_2$ comes with a natural linear structure:

$$c(\underline{u}_1, \underline{u}_2) + d(\underline{v}_1, \underline{v}_2) := (c\underline{u}_1 + d\underline{v}_1, c\underline{u}_2 + d\underline{v}_2),$$

thus is a vector space. The categorical product of vector spaces V_1 and V_2 means the linear space $V_1 \times V_2$ above together with the linear projections $p_1\colon V_1 \times V_2 \to V_1$ and $p_2\colon V_1 \times V_2 \to V_2$.

Let $\iota_1\colon V_1 \to V_1 \times V_2$ ($\iota_2\colon V_2 \to V_1 \times V_2$ resp.) be the linear map such that $\iota_1(x) = (x, 0)$ ($\iota_2(y) = (0, y)$ resp.). The categorical coproduct of vector spaces V_1 and V_2 means the linear space $V_1 \times V_2$ above together with the two linear injections above. In linear algebra, the categorical coproduct of V_1 and V_2 is called the direct sum of V_1 and V_2, and is often denoted by $V_1 \oplus V_2$. One can easily see that

$$p_i \iota_j = \delta_{ij} \quad \text{and} \quad \iota_1 p_1 + \iota_2 p_2 = 1.$$

Schematically, we have

$$V_1 \underset{p_1}{\overset{\iota_1}{\rightleftarrows}} V_1 \times V_2 \underset{\iota_2}{\overset{p_2}{\rightleftarrows}} V_2$$

From now on, we will only consider finite-dimensional real vector spaces unless stated otherwise.

Basis/trivialization

Let V be a real vector space of dimension $n > 0$. For example, V could be $\mathbb{R}^n$, i.e., the real vector space consisting of column matrices with n real entries. Elements in V shall be denoted by $\underline{u}$, $\underline{v}$, etc., and elements in $\mathbb{R}^n$ shall be denoted by $\vec{u}$, $\vec{v}$, etc. It is a fact that V is linearly equivalent to $\mathbb{R}^n$, though not naturally. We write this fact as

$$V \cong \mathbb{R}^n, \quad \text{but} \quad V \not\equiv \mathbb{R}^n.$$

We shall call $\mathbb{R}^n$ the model real vector space of dimension n.

The standard basis of $\mathbb{R}^n$ is $(\vec{e}_1, \ldots, \vec{e}_n)$, here $\vec{e}_i$ is the column matrix whose i-th entry is 1 and all other entries are zero. For example,

$$\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

A basis $\mathcal{B} = (\underline{v}_1, \ldots, \underline{v}_n)$ of V defines a coordinate map

$$[\ \ ]_{\mathcal{B}}: \quad V \to \mathbb{R}^n,$$

which is *the* linear equivalence that sends $\underline{v}_i$ to $\vec{e}_i$, i.e., sends the basis $\mathcal{B}$ of V to the standard basis of $\mathbb{R}^n$. It is easy to see that the map $\mathcal{B} \mapsto [\ \ ]_{\mathcal{B}}$ defines a natural bijection between the set of bases of V and the set of *trivializations* of V (i.e., linear equivalences from V onto $\mathbb{R}^n$). In other words, a trivialization $\equiv$ a basis.

A linear equivalence means an invertible linear transformation.

Exercise 1. *Verify the last statement.*

`Hint: it suffices to construct the inverse for the map` $\mathcal{B} \mapsto [\ \ ]_{\mathcal{B}}$.

We shall call matrix A in commutative diagram

$$\begin{array}{ccc} V_1 & \xrightarrow{T} & V_2 \\ {\scriptstyle [\,]_{\mathcal{B}_1}}\big\downarrow & & \big\downarrow{\scriptstyle [\,]_{\mathcal{B}_2}} \\ \mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^m \end{array}$$

the matrix representation of the linear map $T\colon V_1 \to V_2$ with respect to basis $\mathcal{B}_1$ for V_1 and basis $\mathcal{B}_2$ for V_2. Therefore, if we denote by $\vec{x}$ the matrix representation of $\underline{x}$ in V_1, by $\vec{y}$ the matrix representation of $\underline{y}$ in V_2, then equation $\underline{y} = T\underline{x}$ is represented by equation $\vec{y} = A\vec{x}$.

The simplest matrix representation of T, called the canonical form of T, is a matrix of the following block form

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

where I_r is the identity matrix of order r— the rank of T. Note that, using canonical representation can make computations or arguments very easy. As a practice of using canonical representation, please show that if T is injective, then T has a left inverse, i.e., there is a linear map $S\colon V_2 \to V_1$ such that $ST = 1_{V_1}$ — the identity map on V_1. Dually, if T is surjective, then T has a right inverse.

Let $\hat{e}^i\colon \mathbb{R}^n \to \mathbb{R}$ be the linear map that picks up the i-th entry, i.e., the linear map such that

$$\hat{e}^i(\vec{e}_j) = \delta^i_j.$$

Let $\mathcal{B} = (\underline{v}_1, \ldots, \underline{v}_n)$ be a basis of V and $\hat{v}^i = \hat{e}^i \circ [\ \ ]_{\mathcal{B}}$. Schematically we have

$$\begin{tikzcd} V \arrow[rr, "{[\ ]_{\mathcal{B}}}"] \arrow[dr, "\hat{v}^i"'] & & \mathbb{R}^n \arrow[dl, "\hat{e}^i"] \\ & \mathbb{R} & \end{tikzcd}$$

It is easy to see that

$$\hat{v}^i(\underline{v}_j) = \delta^i_j\,, \tag{1}$$

which can be viewed as a system of equations that defines $(\hat{v}^1, \ldots, \hat{v}^n)$ uniquely.

We claim that $(\hat{v}^1, \ldots, \hat{v}^n)$ is a basis of V^*. This basis for V^*, denoted by $\mathcal{B}^*$, is referred to as the corresponing dual basis. As a result we have $\dim V = \dim V^*$.

Proof. We just need to show that, for any $\alpha \in V^*$, equation

$$x_1\hat{v}^1 + \cdots + x_n\hat{v}^n = \alpha$$

has a unique solution. Indeed, by paring the equation with $\underline{v}_i$ with the help of identity (1), we arrive at equation $x_i = \alpha(\underline{v}_i)$. This proves the uniqueness of solution.

As for the existence of solution, we need to verify that identity $\alpha(\underline{v}_1)\hat{v}^1 + \cdots + \alpha(\underline{v}_n)\hat{v}^n = \alpha$ holds. Since both sides of the identity are linear maps, this amounts to verifying that both sides have the same value at each basis vector $\underline{v}_i$, which is the case. □

Remark. With respect to the basis $(\underline{v}_1, \ldots, \underline{v}_n)$, a vector $\underline{x}$ in V can be represented by a column vector $\vec{x}$ with the i-th entry x^i equal to $\hat{v}^i(\underline{x})$. Similarly, with respect to the corresponding dual basis $(\hat{v}^1, \ldots, \hat{v}^n)$, a vector α in V^* can be represented by a row vector $\underset{\cdot}{\alpha}$ with the i-th entry α_i equal to $\alpha(\underline{v}_i)$. Then $\alpha(\underline{x}) = \underset{\cdot}{\alpha}\vec{x}$ — the matrix multiplication of $\underset{\cdot}{\alpha}$ with $\vec{v}$, i.e., $\alpha(\underline{x}) = \alpha_1 x^1 + \cdots + \alpha_n x^n$.

In other words, $\underline{x} = \hat{v}^1(\underline{x})\underline{v}_1 + \cdots + \hat{v}^n(\underline{x})\underline{v}_n$.

We always identify a number with a square matrix of order 1. With this in mind, the identity

$$\det A^T = \det A$$

can be remembered in the usual way: the two operations —taking determinant and taking transpose — commute with each other.

Multilinear Maps, Pairing

For each integer $i = 1, 2, 3$, we let V_i be a finite dimensional real vector space. A map

$$\lambda: \quad V_1 \times V_2 \to V_3$$

is called bilinear if it is linear in each of the two arguments. An equivalent reformulation for the bilinear map λ is the linear map $\lambda_\natural\colon V_1 \to \mathrm{Hom}(V_2, V_3)$ that maps $\underline{u}_1$ to $\lambda_\natural(\underline{u}_1) : \underline{u}_2 \mapsto \lambda(\underline{u}_1, \underline{u}_2)$.

In general a bilinear map $\omega\colon V_1 \times V_2 \to \mathbb{R}$ is called non-degenerate if its equivalent reformulation $\omega_\natural\colon V_1 \to V_2^*$ is a linear equivalence.

So, a necessary condition for ω to be non-degenerate is that $\dim V_1 = \dim V_2^* = \dim V_2$

Upon choosing a basis $(\underline{u}_1, \ldots, \underline{u}_n)$ for V_1 and a basis $(\underline{v}_1, \ldots, \underline{v}_n)$ for V_2 (thus a basis $(\hat{v}^1, \ldots, \hat{v}^n)$ for V_2^*), ω is represented by matrix $\underset{\sim}{\omega} := [\omega(\underline{u}_i, \underline{v}_j)]$ in the sense that $\omega(\underline{x}, \underline{y}) = \vec{x}^T \underset{\sim}{\omega} \vec{y}$ and $\omega_\natural$ is represented by matrix $\underset{\sim}{\omega}^T$. Therefore, ω is non-degenerate if and only if its matrix representation $\underset{\sim}{\omega}$ is an invertible matrix.

The pairing

$$\begin{array}{rrcl} \langle\, , \rangle : & V^* \times V & \to & \mathbb{R} \\ & (\alpha, \underline{u}) & \mapsto & \alpha(\underline{u}) \end{array} \tag{2}$$

between elements in V^* and elements in V is non-degenerate because its equivalent reformulation $\langle\, , \rangle_\natural$ is the identity map on V^*. Then, an argument based on matrix representation shows that the pairing

$$\begin{array}{rrcl} \langle\, , \rangle' : & V \times V^* & \to & \mathbb{R} \\ & (\underline{u}, \alpha) & \mapsto & \alpha(\underline{u}) \end{array}$$

is non-degenerate as well, so its equivalent reformulation $\langle\, , \rangle'_\natural$, typically written as

$$\iota : \quad V \to V^{**}$$

is a linear equivalence. Note that, for a finite dimensional vector space V, we have

$$V \equiv V^{**}, \quad V \cong V^* \quad \text{but} \quad V \not\equiv V^*.$$

The **double** of V, denoted by $D(V)$, is defined to be $V \oplus V^*$ — the direct sum of V with its dual vector space V^*. Although $V \not\equiv V^*$, $D(V) \equiv D(V^*)$ in the sense that $(\underline{u}, \alpha) \equiv (\alpha, \iota(\underline{u}))$.

Exercise 2. *Let V be real vector space of dimension $n > 0$ and V^* be its dual. Suppose that there are n elements $\underline{v}_1, \ldots, \underline{v}_n$ in V and n elements $\alpha^1, \ldots, \alpha^n$ in V^* such that $\langle \alpha^i, \underline{v}_j \rangle = \delta^i_j$. Show that $(\underline{v}_1, \ldots, \underline{v}_n)$ is a basis of V and $(\alpha^1, \ldots, \alpha^n)$ is a basis of V^*. (Each shall be called the dual basis of the other.)*

`Hint: to check` $(\alpha^1, \ldots, \alpha^n)$ `is a basis of` V^*`, it suffices to show that` $\alpha^1, \ldots, \alpha^n$ `are linearly independent.`

Note that $(\hat{e}^1, \ldots, \hat{e}^n)$ is the basis of $(\mathbb{R}^n)^*$ and it is dual to the standard basis of $\mathbb{R}^n$. We shall identify $(\mathbb{R}^n)^*$ with the vector space consisting of row matrices with n entries so that the pairing $\langle\, , \rangle$: $(\mathbb{R}^n)^* \times \mathbb{R}^n \to \mathbb{R}$ is just the matrix multiplication. It is useful to keep the following table in mind.

	Abstract Side	Concrete Side
vector space	V	$\mathbb{R}^n$
basis	$(\underline{v}_1, \ldots, \underline{v}_n)$	$(\vec{e}_1, \ldots, \vec{e}_n)$
dual space	V^*	$(\mathbb{R}^n)^*$
dual basis	$(\hat{v}^1, \ldots, \hat{v}^n)$	$(\hat{e}^1, \ldots, \hat{e}^n)$
pairing	$\langle \hat{v}^i, \underline{v}_j \rangle = \delta^i_j$	$\langle \hat{e}^i, \vec{e}_j \rangle = \delta^i_j$

Exercise 3. *Let $T\colon V_1 \to V_2$ be a linear map and $\mathcal{B}_i$ be a basis of vector space V_i. Denote by $\mathcal{B}_i^*$ the dual basis of $\mathcal{B}_i$ for vector space V_i^*, by $T^*\colon V_2^* \to V_1^*$ the dual map of T, i.e., the map that sends α to $\alpha \circ T$.*

Let A be the matrix representation of T with respect to bases $\mathcal{B}_1$, $\mathcal{B}_2$ and A^* be the matrix representation of T^* with respect to bases $\mathcal{B}_2^*$, $\mathcal{B}_1^*$. Show that A^* and A are transposes of each other.

Hint: It is helpful to start with commutative diagrams

This hint shows you a general trick in linear algebra: go to the matrix representation side!

$$\begin{tikzcd} V_1 \arrow[r, "T"] \arrow[d, "{[\,]_{\mathcal{B}_1}}"'] & V_2 \arrow[d, "{[\,]_{\mathcal{B}_2}}"] \\ \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} V_1^* \arrow[d, "{[\,]_{\mathcal{B}_1^*}}"'] & V_2^* \arrow[l, "T^*"'] \arrow[d, "{[\,]_{\mathcal{B}_2^*}}"] \\ \mathbb{R}^n & \mathbb{R}^m \arrow[l, "A^*"'] \end{tikzcd}$$

and then observe that the defining equation $\langle \alpha, T(\underline{u}) \rangle = \langle T^*(\alpha), \underline{u} \rangle$ is represented by the defining equation

$$\vec{u} \cdot A\vec{v} = A^*\vec{u} \cdot \vec{v} \quad \text{or} \quad \vec{u} \cdot A\vec{v} = \vec{v} \cdot A^*\vec{u}$$

Take $\vec{u} = \vec{e}_i$ and $\vec{v} = \vec{e}_j$, then the identity says that the (i, j)-entry of A is equal to the (j, i)-entry of A^*.

Two-forms

Let V be a real vector space. A linear map $\alpha\colon V \to \mathbb{R}$ is called a one-form on V. So the dual space V^* is the vector space of one-forms on V.

A bilinear map $\omega\colon V \times V \to \mathbb{R}$ is called a two-form on V. In case $\omega(\underline{u}, \underline{v}) = -\omega(\underline{v}, \underline{u})$ for any $\underline{u}$, $\underline{v}$ in V, we say that ω is a skew-symmetric two-form on V. In case $\omega(\underline{u}, \underline{v}) = \omega(\underline{v}, \underline{u})$ for any $\underline{u}$, $\underline{v}$ in V, we say that ω is a symmetric two-form on V. Note that, an inner product on V is nothing but a positive-definite symmetric two-form on V.

If ω is a symmetric two-form, then the map $V \to \mathbb{R}$ which maps $\underline{u}$ to $Q(\underline{u}) := \omega(\underline{u}, \underline{u})$ is a quadratic form on V. Actually, quadratic forms on V are defined this way, i.e., a quadratic form on V is a real function that sends $\underline{u}$ to $\omega(\underline{u}, \underline{u})$ for some symmetric two-form ω on

V. In case $\underline{v}_i$ is a basis of V and we write $\underline{u} = u^i \underline{v}_i$, then $Q(\underline{u}) = \omega(\underline{v}_i, \underline{v}_j) u^i u^j$ is a quadratic homogeneous polynomial in coordinate variables u^i. Here we have used Einstein convention for sum over indices.

Example 1. *Let $V = \mathbb{R}^n$ and A be a real symmetric matrix of order n. Then $Q(\vec{x}) := \vec{x} \cdot A\vec{x}$ is a quadratic form on $\mathbb{R}^n$, $\omega(\vec{x}, \vec{y}) := \vec{x} \cdot A\vec{y}$ is a symmetric 2-from on $\mathbb{R}^n$, moreover, Q and ω determine each other.*

Consider a vector space V with a basis $(\underline{v}_1, \ldots, \underline{v}_n)$. Let $(\hat{v}^1, \ldots, \hat{v}^n)$ be the corresponding dual basis of V^*. Suppose that ω is a symmetric two-form on V. Then, if we write $\underline{u} = u^i \underline{v}_i$ and $\underline{w} = w^i \underline{v}_i$, then $\omega(\underline{u}, \underline{w}) = u^i w^j \omega_{ij}$ with $\omega_{ij} := \omega(\underline{v}_i, \underline{v}_j)$, and it is in this sense we say that the symmetric matrix $[\omega_{ij}]$ is the matrix representation of the symmetric two-form ω with respect to basis $(\underline{v}_1, \ldots, \underline{v}_n)$. In short, $\omega(\underline{u}, \underline{w}) = \vec{u} \cdot [\omega_{ij}]\vec{w}$.

One can show by a computation that a symmetric 2-form ω on V can be written as $\omega = \omega_{ij}\, \hat{v}^i \hat{v}^j$. Here, $\hat{v}^i \hat{v}^j$ is the symmetric 2-form on V that sends $(\underline{x}, \underline{y})$ to $\frac{1}{2}(\hat{v}^i(\underline{x})\hat{v}^j(\underline{y}) + \hat{v}^i(\underline{y})\hat{v}^j(\underline{x}))$ and is called the symmetric tensor product of $\hat{v}^i$ with $\hat{v}^j$.

Remark 1. In general, for any two one-forms α and β on V, we can form the tensor product, symmetric tensor product, and anti-symmetric tensor product of α with β. Denoted by $\alpha \otimes \beta$, $\alpha\beta$, and $\alpha \wedge \beta$ respectively, they are all two-forms on V, and are defined as follows:

1. $\alpha \otimes \beta(\underline{u}, \underline{v}) := \alpha(\underline{u})\beta(\underline{v})$.
2. $\alpha\beta := \frac{1}{2}(\alpha \otimes \beta + \beta \otimes \alpha)$, i.e., $\alpha\beta(\underline{u}, \underline{v}) = \frac{1}{2}\left(\alpha(\underline{u})\beta(\underline{v}) + \alpha(\underline{v})\beta(\underline{u})\right)$.
3. $\alpha \wedge \beta := \frac{1}{2}(\alpha \otimes \beta - \beta \otimes \alpha)$, i.e., $\alpha \wedge \beta(\underline{u}, \underline{v}) = \frac{1}{2}\left(\alpha(\underline{u})\beta(\underline{v}) - \alpha(\underline{v})\beta(\underline{u})\right)$.

Note that $\alpha\beta = \beta\alpha$, $\alpha \wedge \beta = -\beta \wedge \alpha$, and $\alpha \otimes \beta = \alpha\beta + \alpha \wedge \beta$.

Suppose ω is further assumed to be non-degenerate, i.e. its equivalent formulation $\omega_\natural \colon V \to V^*$ is assumed to be an invertible linear map, then we can transport the symmetric two-form ω on V to a symmetric two-form ω^{-1} on V^* via equation

$$\omega(\underline{u}, \underline{v}) = \omega^{-1}(\omega_\natural(\underline{u}), \omega_\natural(\underline{v})).$$

Let $\omega^{ij} := \omega^{-1}(\hat{v}^i, \hat{v}^j)$, then $\omega^{-1} = \omega^{ij}\underline{v}_i\underline{v}_j$. Here, $v_i v_j$ is the symmetric tensor product of v_i with v_j, viewed as elements in V^{**}. We claim that matrix $[\omega^{ij}]$ and matrix $[\omega_{ij}]$ are inverse of each other.

Exercise 4. *Prove this claim.*

Hint: If we go to representation of ω by the symmetric invertible matrix A (thus $\omega_\natural$ by A as well) and ω^{-1} by matrix B, then the defining equation for ω^{-1} becomes this defining equation for B:

$$\vec{u} \cdot A\vec{v} = A\vec{u} \cdot B(A\vec{v}).$$

Pseudo-inner Product

By definition, an inner product on a real vector space V is a positive-definite symmetric 2-form on V. The positivity condition implies that any inner product must be non-degenerate. In general, a non-degenerate symmetric 2-form on V is called a pseudo-inner product on V, and a non-degenerate skew-symmetric 2-form on V is called a symplectic form on V.

For example, on $\mathbb{R}^n$, the quadratic form

$$Q_k(\vec{u}) := u_1^2 + \cdots + u_k^2 - u_{k+1}^2 \cdots - u_n^2$$

corresponds to a pseudo-inner product on $\mathbb{R}^n$ with signature $(k, n - k)$. Note that Q_n actually corresponds to an inner product.

A real vector space with an inner product is called a Euclidean vector space, a real vector space with a pseudo-inner product is called a pseudo-Euclidean vector space, a real vector space with a symplectic form is called a symplectic space.

It is a fact that ω is an inner product on $V \Leftrightarrow$ the matrix representation $[\omega_{ij}]$ of ω is positive-definite $\Leftrightarrow$ the leading principal minors of $[\omega_{ij}]$ are all positive, i.e., the determinant of the upper-left k by k sub-matrix of $[\omega_{ij}]$ is positive for all $1 \le k \le n$.

Exercise 5. *Continuing the discussion, we let ω_1 be a non-degenerate symmetric two-forms on vector space V and T be an endomorphism on V. Suppose that T is self-adjoint with respect to ω_1, this means that, for any $\underline{u}, \underline{v}$ in V, we have*

$$\omega_1(\underline{u}, T\underline{v}) = \omega_1(T\underline{u}, \underline{v}).$$

Let ω_2 be this two-form on V: $\omega_2(\underline{u}, \underline{v}) = \omega_1(\underline{u}, T\underline{v})$. Show that

1) ω_2 is a symmetric two-form on V.

2) Writing $\omega_1 = g_{ij}\,\hat{v}^i\hat{v}^j$ and $\omega_2 = h_{ij}\,\hat{v}^i\hat{v}^j$, show that

$$\det T = \frac{\det[h_{ij}]}{\det[g_{ij}]}.$$

Hint: representing ω_1 by invertible symmetric matrix g, ω_2 by matrix h, T by matrix A. Then $\omega_1(\underline{u}, T\underline{v}) = \omega_1(T\underline{u}, \underline{v})$ becomes $\vec{u} \cdot gA\vec{v} = A\vec{u} \cdot g\vec{v}$, and $\omega_2(\underline{u}, \underline{v}) = \omega_1(\underline{u}, T\underline{v})$ becomes $\vec{u} \cdot h\vec{v} = \vec{u} \cdot gA\vec{v}$. Then we have $gA = A^T g$ and $h = gA$.

Symplectic Form

In case ω is a symplectic form on V we have $\omega = \omega_{ij}\, \hat{v}^i \wedge \hat{v}^j$ for a unique real anti-symmetric matrix $[\omega_{ij}]$. Here, $\omega_{ij} = \omega(\underline{v}_i, \underline{v}_j)$, and $\hat{v}^i \wedge \hat{v}^j$ is the anti-symmetric 2-form on V that sends $(\underline{x}, \underline{y})$ to $\frac{1}{2}(\hat{v}^i(\underline{x})\hat{v}^j(\underline{y}) - \hat{v}^i(\underline{y})\hat{v}^j(\underline{x}))$ and is called the anti-symmetric tensor product of $\hat{v}^i$ with $\hat{v}^j$. It is a fact that n must be an even integer, and the basis $(\hat{v}^1, \cdots, \hat{v}^n)$ can be chosen such that

This basis is called a symplectic basis.

$$[\omega_{ij}] = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

where 0 and I are respectively the zero matrix and the identity matrix of order $n/2$.

Exercise 6. *Show that the dimension of a symplectic space must be even. Show also that the pairing in Eq. (2) yields a symplectic form on $D(V)$:*

This exercise says that the double of V, $D(V)$, is a canonical symplectic space.

$$\omega((\underline{u}_1, \alpha_1), (\underline{u}_2, \alpha_2)) = \langle \alpha_1, \underline{u}_2 \rangle - \langle \alpha_2, \underline{u}_1 \rangle.$$

Hint: let $(\underline{v}_1, \ldots, \underline{v}_n)$ be a basis for V and $(\hat{v}^1, \ldots, \hat{v}^n)$ be the dual basis for V^*. Then, with respect to the basis

$$(\underline{v}_1, \ldots, \underline{v}_n, \hat{v}^1, \ldots, \hat{v}^n)$$

of $D(V)$, ω is represented by the block matrix $\begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}$.

<u>Fact</u>: Up to equivalences, there is only one

(i) vector space,
(ii) Euclidean vector space,
(iii) pseudo-Euclidean vector space of a given signature,
(iv) symplectic space

of a given dimension. (The dimension must be even for symplectic space)

Orientation

Let V be a real vector space of dimension $n > 0$, and $\mathcal{B}_V$ be the set of bases of V. Suppose that $\mathcal{B}_1$ and $\mathcal{B}_2$ are two bases of V, then the coordinate map $[\ ]_{\mathcal{B}_i}\colon V \to \mathbb{R}^n$ is a linear equivalence for each i. We say $\mathcal{B}_1 \sim \mathcal{B}_2$ if the determinant of the linear equivalence

$$[\ ]_{\mathcal{B}_1} \circ ([\ ]_{\mathcal{B}_2})^{-1} : \mathbb{R}^n \to \mathbb{R}^n$$

is positive. Schematically we have

$$\begin{tikzcd} & V \arrow[dl, "{[\ ]_{\mathcal{B}_2}}"'] \arrow[dr, "{[\ ]_{\mathcal{B}_1}}"] & \\ \mathbb{R}^n \arrow[rr, "A"] & & \mathbb{R}^n \end{tikzcd} \qquad \det A > 0$$

It is clear that $\sim$ is an equivalence relation on $\mathcal{B}_V$.

One can see that there are *exactly two* equivalence classes on $\mathcal{B}_V$, each one of which is called an orientation on V. So any real vector space V admits exactly two orientations. By definition, a real vector space V with an orientation is called an oriented (real) vector space.

Up to equivalences, there is only one oriented real vector space of a given dimension. For example, on $\mathbb{R}^3$, the standard basis $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ represents the standard orientation and basis $(\vec{e}_2, \vec{e}_1, \vec{e}_3)$ represents the opposite (i.e., the other) orientation. However, the linear equivalence $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ such that $T(\vec{e}_1) = \vec{e}_2$, $T(\vec{e}_2) = \vec{e}_1$ and $T(\vec{e}_3) = \vec{e}_3$ identifies the oriented vector space $(\mathbb{R}^3, [(\vec{e}_1, \vec{e}_2, \vec{e}_3)])$ with the oriented vector space $(\mathbb{R}^3, [(\vec{e}_2, \vec{e}_1, \vec{e}_3)])$.

Because of the natural identification of $\mathcal{B}_V$ with $\mathcal{B}_{V^*}$ that identifies a basis with its dual basis, Exercise 3 implies that the orientation on V and the orientation on V^* can be naturally identified with each other.

Let us use $o(V)$ to denote the set of orientations on V. We can summarize the natural identifications we have encountered as follows:

$$V \equiv V^{**}, \quad D(V) \equiv D(V^*), \quad \mathcal{B}_V \equiv \mathcal{B}_{V^*}, \quad o(V) \equiv o(V^*).$$

List of Symbols

V	linear space or vector space
$\mathbb{R}^n$	the real linear space of column vectors of n entries.
	It is also called the model real linear space of dimension n
V^*	the dual (linear) space of V
V^{**}	the double dual space of linear space V, i.e., $(V^*)^*$
$D(V)$	the double of linear space V. By definition, $D(V) = V \oplus V^*$
$\mathrm{Map}(X, V)$	the collection of set maps from set X to linear space V. It is a linear space
$\mathrm{Hom}(V_1, V_2)$	the collection of linear maps from linear space V_1 to linear space V_2.
	It is a linear subspace of $\mathrm{Map}(V_1, V_2)$
$\mathbb{R}[[t]]$	the linear space of real power series in real variable t
$\mathbb{R}[t]$	the linear space of real polynomials in real variable t. It is a linear subspace of $\mathbb{R}[[t]]$
$\equiv$	naturally equivalent to
$\cong$	equivalent to
$\underline{u}, \underline{v}$, etc.	vectors in abstract linear spaces
$\vec{u}, \vec{v}$, etc.	vectors in model linear spaces
$T\underline{u}$	evaluation of linear map T at vector $\underline{u}$, i.e., $T(\underline{u})$
TS	the composition of map T with map S
c, d, etc.	scalars, i.e., real numbers
α, β, etc.	one-forms
$\alpha \wedge \beta$	the wedge product of one-form α with one-form β
$(\underline{v}_1, \ldots, \underline{v}_n)$	a basis of V
$(\hat{v}^1, \ldots, \hat{v}^n)$	the corresponding dual basis of V^*
$(\vec{e}_1, \ldots, \vec{e}_n)$	the standard basis of $\mathbb{R}^n$
$\mathcal{B}, \mathcal{B}^*$	a basis of V and its corresponding dual basis for V^*
$\mathcal{B}_V$	the set of bases on linear space V
$o(V)$	the set of orientations on real linear space V
$[\ ,\ ]_{\mathcal{B}}$	the coordinate map w.r.t. basis $\mathcal{B}$
$\langle\ ,\ \rangle$	either the paring of co-vectors with vectors
	or the inner product on a real linear space
A^T	the transpose of matrix A

§2 Review of Affine Space

The spaces occurring in Calculus III are certain subsets of affine spaces. Intuitively an affine space is just a line, or a plane, or their analogues in dimensions other than 1 and 2. Formally we have

In terms of group-theoretical language, an affine space is a principal G-space where the group G is a vector space for which the group multiplication is the vector addition.

Definition 1 (Affine Space). *Let V be a vector space. An affine space modelled on V is a non-empty set $\mathbb{A}$ together with a map*

$$\begin{array}{rccc} T: & \mathbb{A}\times V & \to & \mathbb{A} \\ & (p, \underline{u}) & \mapsto & p+\underline{u} \end{array} \tag{3}$$

such that

1) $(p+\underline{u})+\underline{v} = p+(\underline{u}+\underline{v})$ *for any point* $p \in \mathbb{A}$ *and any vectors* $\underline{u}$, $\underline{v}$ *in* V.

2) $p+\underline{0} = p$ *for any point* $p \in \mathbb{A}$. *(here* $\underline{0}$ *is the zero vector in* V*)*

3) *For any two points* p, q *in* $\mathbb{A}$, *there IS a UNIQUE vector* $\underline{u}$ *in* V *such that* $p+\underline{u} = q$.

Remark 2. The point $p+\underline{u}$ is called the translation of point p by the vector $\underline{u}$. For a good reason, the unique vector $\underline{u}$ in condition 3) shall be denoted by $q-p$. If $\dim V = n$, we say that $\mathbb{A}$ is an affine space of dimension n. An affine space of dimension 0 is just a point, an affine space of dimension 1 is called an affine line or simply line, an affine space of dimension 2 is called an affine plane or simply plane. If V is a real (complex resp.) vector space we say $\mathbb{A}$ is a real (complex resp.) affine space. In this book we are only interested in real affine spaces.

Remark 3. Affine spaces are intimately related to vector spaces. We shall see later that a fundamental trick in understanding affine space is to turn problems in affine spaces into the corresponding problems in linear algebra.

Example 2 (Model Affine Space)**.** *We take V to be $\mathbb{R}^n$, $\mathbb{A}$ to be*

$$\mathbb{A}^n_{\mathbb{R}} := \{(x^1, \ldots, x^n) \mid x^i \in \mathbb{R}\},$$

and the translation map T to be the one such that

$$(x^1, \ldots, x^n) + \begin{bmatrix} u^1 \\ \vdots \\ u^n \end{bmatrix} = (x^1 + u^1, \ldots, x^n + u^n).$$

We shall denote by 0 the point $(0, \cdots, 0) \in \mathbb{A}^n_{\mathbb{R}}$. For any point $x \in \mathbb{A}^n_{\mathbb{R}}$, we shall denote by $\vec{x}$ the column vector $x - 0$. The correspondence $x \leftrightarrow \vec{x}$ is called the point-vector correspondence.

$\mathbb{A}^n_{\mathbb{R}}$ is called the model affine space of dimension n. It is also called the n-dimensional coordinate space. In many textbooks, it is also denoted by $\mathbb{R}^n$.

Example 3 (Categorical Product of Affine Spaces)**.** *For $i = 1, 2$, we let $\mathbb{A}_i$ be an affine space modelled on vector space V_i. Then the categorical product $\mathbb{A}_1 \times \mathbb{A}_2$ is an affine space modelled on vector space $V_1 \times V_2$, having the Cartesian product of $\mathbb{A}_1$ with $\mathbb{A}_2$ as its underlying space, together with the projections onto the two factors.*

Example 4 (Vector Spaces as Affine Spaces)**.** *A vector space V is an affine space modelled on itself, with the translation map*

$$\begin{array}{rccc} T: & V \times V & \to & V \\ & (\underline{v}, \underline{u}) & \mapsto & \underline{v} + \underline{u} \end{array}$$

When we view a vector space V as an affine space, we rewrite it as V_{aff}.

Affine Combination and Affine Map

The affine combination of $(k+1)$ points $p_0, \ldots, p_k$ in $\mathbb{A}$ weighted by real numbers $c^0, \ldots, c^k$ with $\sum_{i=0}^k c^i = 1$, denoted by

$$c^0 p_0 + \cdots + c^k p_k \quad \text{or simply} \quad c^i p_i,$$

is defined as follows:

$$c^i p_i := q + \sum_i c^i (p_i - q). \tag{4}$$

The affine combination can be viewed as the center of charge of the $(k+1)$ points.

Please draw a picture to help you understand this formula. The reference point enables us to turn affine combination in affine space into linear combination in linear algebra. The fundamental trick in Remark 3 has been used here!

Here, q is any (reference or observation) point in $\mathbb{A}$. This definition makes sense because of

Exercise 7. *Show that*[1] *, if $\sum_i c_i = 1$, then, for any two reference points $q, q' \in \mathbb{A}$, we have*

$$q + \sum_i c^i (p_i - q) = q' + \sum_i c^i (p_i - q').$$

Show also that, in $\mathbb{A}^n_{\mathbb{R}}$, we have formula

$$t(x^1, \ldots, x^n) + (1-t)(y^1, \ldots, y^n) = (tx^1 + (1-t)y^1, \ldots, tx^n + (1-t)y^n)$$

for affine combination.

[1]

Proof. Write $q = q' + \underline{u}$. Since

$$\begin{aligned} q' + (\underline{u} + (p_i - q)) &= (q' + \underline{u}) + (p_i - q) \\ &= q + (p_i - q) = p_i, \end{aligned}$$

we have $\underline{u} + (p_i - q) = p_i - q'$. Then

$$\begin{aligned} q + \sum_i c^i (p_i - q) &= (q' + \underline{u}) + \sum_i c^i (p_i - q) \\ &= q' + (\underline{u} + \sum_i c^i (p_i - q)) \\ &= q' + (\sum_i c_i \underline{u} + \sum_i c^i (p_i - q)) \\ &= q' + \sum_i c_i (\underline{u} + (p_i - q)) \\ &= q' + \sum_i c_i (p_i - q'). \end{aligned}$$

$\square$

Exercise 8. *Draw two points p and q on a plane. Then draw the following points on the plane:*

$$0p + 1q, \quad \frac{1}{2}p + \frac{1}{2}q, \quad \frac{1}{3}p + \frac{2}{3}q, \quad 2p + (-1)q.$$

Warning: the notions of addition and scalar multiplication are not valid for affine spaces!

Let $\mathbb{A}_1$ and $\mathbb{A}_2$ be two affine spaces. A map $F\colon \mathbb{A}_1 \to \mathbb{A}_2$ is called an affine map if F respects the affine combinations, i.e. F commutes with taking affine combination: $F(c^i p_i) = c^i F(p_i)$. In case F is a bijection and affine, F^{-1} must be an affine map; in this case we say that F is an affine equivalence and we write $\mathbb{A}_1 \cong \mathbb{A}_2$.

Example 5. *If $\phi\colon V \to W$ is a linear map, then $\phi\colon V_{aff} \to W_{aff}$ is an affine map. In particular,*

$$V \cong W \implies V_{aff} \cong W_{aff}.$$

I.e., a linear equivalence is an affine equivalence between the underlying affine spaces.

Affine Independence

We say that points $p_0, \ldots, p_k$ in $\mathbb{A}$ are affinely independent if vectors $p_1 - p_0, \ldots.\ p_k - p_0$ in V are linearly independent. This definition makes sense because

$$\begin{gathered}\{p_i - p_0 \mid i \neq 0\} \text{ is a linearly independent set} \\ \Updownarrow \\ \text{for any fixed } j,\ \{p_i - p_j \mid i \neq j\} \text{ is a linearly independent set.}\end{gathered}$$

Exercise 9. *Prove this last statement.*

`Hint:` we may assume that $j = 1$. Since $p_i - p_1 = (p_i - p_0) - (p_1 - p_0)$, we have

$$[p_0 - p_1, p_2 - p_1, \ldots, p_k - p_1] = [p_1 - p_0, p_2 - p_0, \ldots, p_k - p_0]A$$

where the matrix A is the invertible block matrix $\begin{bmatrix} -1 & a \\ 0 & I \end{bmatrix}$ for which I is the identity matrix of order $k-1$, $a = [-1, \cdots, -1]$ and 0 is the column matrix zero.

Affine Span

The set of all affine combinations of points $p_0, \ldots, p_k$ is called the affine span of points $p_0, \ldots, p_k$. It is denoted by $\operatorname{span}\{p_0, \ldots, p_k\}$.

One can check that the affine span of a finite set of points in $\mathbb{A}$ is an affine subspace of $\mathbb{A}$. We say that the set of points $\{p_0, \ldots, p_k\}$ is a spanning set for $\mathbb{A}$ if $\operatorname{span}\{p_0, \ldots, p_k\} = \mathbb{A}$. A spanning set for $\mathbb{A}$ is called a minimal spanning set if any of its proper subset is not a spanning set for $\mathbb{A}$. It is a fact that a minimal spanning set is an affinely independent spanning set.

A non-empty subset of $\mathbb{A}$ is called an **affine subspace** of $\mathbb{A}$ if it is closed under affine combination.

A co-dimension one affine subspace of an affine space $\mathbb{A}$ is called a hyperplane of $\mathbb{A}$. So, a hyperplane of $\mathbb{A}^2_{\mathbb{R}}$ is a line, a hyperplane of $\mathbb{A}^3_{\mathbb{R}}$ is a plane.

Pointed Affine Spaces are Vector Spaces of the same dimension

We have remarked that a vector space V can be viewed as an affine space in a canonical way: the affine combinations are just those linear combinations for which the sum of weights is equal to 1. On the other hand, a pointed affine space $(\mathbb{A}, p)$ — an affine space $\mathbb{A}$ together with a point p in $\mathbb{A}$ — is a vector space. Indeed, for any two points q_1 and q_2 in $\mathbb{A}$ and any two real numbers c^1 and c^2, the linear combination $c^1q_1 + c^2q_2$ in vector space $(\mathbb{A}, p)$ is defined to be

$$p + \left(c^1(q_1 - p) + c^2(q_2 - p)\right),$$

i.e., the translation of point p by vector $c^1(q_1 - p) + c^2(q_2 - p)$. The fact that the above formula for $c^1q_1 + c^2q_2$ indeed defines a linear structure can be seen from the next paragraph.

Let $\mathbb{A}$ be an affine space modelled on vector space V, and $p \in \mathbb{A}$. Then the map

$$\begin{array}{rcl} \psi_p : V & \to & (\mathbb{A}, p) \\ \underline{u} & \mapsto & p + \underline{u} \end{array}$$

is a linear equivalence. Indeed, property 3) in Definition 1 says that ψ_p is bijection; since

$$\begin{array}{rcl} \psi_p(c^i\underline{u}_i) & = & p + c^i\underline{u}_i \quad \text{definition of } \psi_p \\ & = & c^i(p + \underline{u}_i) \quad \text{definition of linear combination on } (\mathbb{A}, p) \\ & = & c^i\psi_p(\underline{u}_i) \quad \text{definition of } \psi_p \end{array}$$

ψ_p is linear. **Remark**: this actually proves that $(\mathbb{A}, p)$ is a linear space of dimension equal to $\dim \mathbb{A}$.

Since ψ_p is a linear equivalence, it is also an affine equivalence from V_{aff} onto $\mathbb{A} = (\mathbb{A}, p)_{aff}$. We claim that the assignment $p \mapsto \psi_p$ is an injective map from $\mathbb{A}$ into the set of affine equivalence of V_{aff} with $\mathbb{A}$, but not a surjective map.

Exercise 10. *Prove this claim.*

Recall that, when a vector space V is viewed as an affine space, we denote it by V_{aff}. Then,

$$\begin{array}{rcll} V & = & (V_{aff}, \underline{0}) & \text{as vector space} \\ \mathbb{A} & = & (\mathbb{A}, p)_{aff} & \text{as affine space} \end{array}$$

So, *an affine space is just a vector space with the distinguished element (i.e., the zero vector) forgotten, and if you put the distinguished element back, you get back to the vector space.*

In particular, an affine equivalence is a linear equivalence between pointed affine spaces.

Exercise 11. *Let $F\colon \mathbb{A} \to \mathbb{B}$ be a map and $p \in \mathbb{A}$. Show that F is an affine map if and only if F is a linear map from $(\mathbb{A}, p)$ to $(\mathbb{B}, F(p))$.*

Exercise 12. *Let $p_0, \ldots, p_k$ be points in $\mathbb{A}$ and $\mathbb{B}$ be the affine span of these points. Show that the vector space $(\mathbb{B}, p_0)$ is the linear span of the vectors $p_1, \ldots, p_k$. Then conclude that a minimal spanning set is an affinely independent spanning set.*

The uniqueness of affine spaces

Let $n \geq 0$ be an integer. It is a fact that, up to linear equivalences, there is a unique real vector space of dimension n. Consequently, up to affine equivalences, there is a unique real affine space of dimension n. Indeed, if $\mathbb{A}$ and $\mathbb{A}'$ are affine spaces of equal dimension, then, for any point p in $\mathbb{A}$ and any point p' in $\mathbb{A}'$, because they have the same dimension, vector spaces $(\mathbb{A}, p)$ and $(\mathbb{A}', p')$ are linearly equivalent, then their underlying affine spaces $\mathbb{A}$ and $\mathbb{A}'$ are affine equivalent. In particular,

If $\mathbb{A}$ is affine space of dimension n, then $\mathbb{A} \cong \mathbb{A}^n_{\mathbb{R}}$, but not canonically.

An affine equivalence of $\mathbb{A}$ with the model affine space of dimension $\dim \mathbb{A}$ is called a trivialization of $\mathbb{A}$.

Affine Frame

An affine frame of an affine space $\mathbb{A}$ is an ordered minimal affine spanning set for $\mathbb{A}$. The affine frame $(0, e_1, \ldots, e_n)$ of $\mathbb{A}^n_{\mathbb{R}}$ is called the standard affine frame of $\mathbb{A}^n_{\mathbb{R}}$. Here $e_i = 0 + \vec{e}_i$.

In case the affine space is the vector space V and $(\underline{v}_1, \ldots, \underline{v}_n)$ is a basis on V, then $(\underline{0}, \underline{v}_1, \ldots, \underline{v}_n)$ is an affine frame of the affine space V_{aff}.

Exercise 13. *Prove this last statement.*

Since affine equivalences send an affine frame to an affine frame, if $\mathbb{A}$ is an affine space modelled on vector space V, p is a point in $\mathbb{A}$, and $(\underline{v}_1, \ldots, \underline{v}_n)$ is a basis on V, then, using affine equivalence

$$\psi_p : \ V_{aff} \to \mathbb{A}$$

that sends $\underline{v}$ to $p + \underline{v}$, we conclude that $(p, p + \underline{v}_1, \ldots, p + \underline{v}_n)$ is an affine frame of $\mathbb{A}$.

Exercise 14. *Let $(p_0, p_1, \ldots, p_n)$ be an affine frame of $\mathbb{A}$, and $p \in \mathbb{A}$.*

1) Show that

$$p = \lambda^i p_i$$

for a unique choice of real numbers $\lambda^0, \ldots, \lambda^n$ such that $\sum_{i=0}^n \lambda^i = 1$.

These numbers λ^i are called the barycentric coordinates of p.

2) Show that the corresponding affine coordinate map, i.e., the map from $\mathbb{A}$ to $\mathbb{A}^n_{\mathbb{R}}$ which sends p to $(\lambda^1, \ldots, \lambda^n)$, is an affine equivalence.

3) Show that the assignment of the affine coordinate map to an affine frame is a bijection from set of affine frames of $\mathbb{A}$ to the set of affine equivalences from $\mathbb{A}$ onto $\mathbb{A}^n_{\mathbb{R}}$.

Recall that an affine equivalence from $\mathbb{A}$ to $\mathbb{A}^n_{\mathbb{R}}$ is called a trivialization of $\mathbb{A}$. So an affine frame of $\mathbb{A}$ and a trivialization of $\mathbb{A}$ are the same thing.

Tangent Space and Cotangent Space

A vector in the affine space $\mathbb{A}$ is just an arrow in $\mathbb{A}$, i.e., a line segment in $\mathbb{A}$ together with a direction. If the arrow has head point q and tail point p, it is denoted by $\overrightarrow{pq}$ and is called a tangent vector of $\mathbb{A}$ at point p. The set of all tangent vectors of $\mathbb{A}$ at point p, denoted by $T_p\mathbb{A}$, is a vector space, called the tangent space of $\mathbb{A}$ at point p. This is clearly true intuitively. To see it precisely, we note that vector $\overrightarrow{pq}$ can be recorded as (p, q) or $(p, q - p)$. Therefore, we can write

The line segment with end points p and q, denoted by $\overline{pq}$, is the subset $\{tp + (1-t)q \mid t \in [0, 1]\}$

$$T_p\mathbb{A} = \{(p, \underline{u}) \mid \underline{u} \in V\} = \{p\} \times V.$$

With this understood, one can write down the linear structure on $T_p\mathbb{A}$ as follows:

$$c_1(p, \underline{u}_1) + c_2(p, \underline{u}_2) = (p, c_1\underline{u}_1 + c_2\underline{u}_2).$$

We claim that $T_p\mathbb{A}$ and $(\mathbb{A}, p)$ are naturally equivalent, so we write

$$T_p\mathbb{A} \equiv (\mathbb{A}, p).$$

Indeed, since the natural map ϕ_p that sends $(p, \underline{u})$ to $p + \underline{u}$ is clearly a bijection, one just needs to verify that

$$\phi_p(c_1(p, \underline{u}_1) + c_2(p, \underline{u}_2)) = c_1\phi_p(p, \underline{u}_1) + c_2\phi_p(p, \underline{u}_2),$$

i.e., $p + (c_1\underline{u}_1 + c_2\underline{u}_2) = c_1(p + \underline{u}_1) + c_2(p + \underline{u}_2)$, which is just the definition of affine combination.

By definition, the dual vector space of $T_p\mathbb{A}$, denoted by $T_p^*\mathbb{A}$, is called the cotangent space of $\mathbb{A}$ at point p, and an element of $T_p^*\mathbb{A}$, i.e., a linear map from $T_p\mathbb{A}$ to $\mathbb{R}$, is called a cotangent vector of $\mathbb{A}$ at point p. *Unlike tangent vectors, cotangent vectors are not intuitive objects; however, they appear naturally in calculus.*

We shall see in the next chapter that, for a real smooth function defined on an open neighborhood of p in $\mathbb{A}$, the coordinate-free version of its first order partial derivatives at point p is a cotangent vector of $\mathbb{A}$ at point p.

Orientation

An orientation on an affine space is a *continuous* assignment of an orientation to each of its tangent spaces. It is a fact that an affine space admits exactly two orientations.

Affine Complement

Two affine subspaces of $\mathbb{A}$, say $\mathbb{A}_1$ and $\mathbb{A}_2$, are called affine complement of each other if they intersect at a single point p and the vector space identity $(\mathbb{A}, p) = (\mathbb{A}_1, p) \oplus (\mathbb{A}_2, p)$ holds, i.e., for any point q in $\mathbb{A}$, there is a unique point q_1 in $\mathbb{A}_1$ and a unique point q_2 in $\mathbb{A}_2$ such that q is equal to the affine combination $(-1)p + 1q_1 + 1q_2$, i.e. the point $p + ((q_1 - p) + (q_2 - p))$. In case this happens we write $\mathbb{A} = \mathbb{A}_1 \oplus \mathbb{A}_2$. As an example, we take $\mathbb{A}$ to be $\mathbb{A}^3_{\mathbb{R}}$, and

$$\mathbb{A}_1 = \{(x_1, x_2, 0) \mid x_1, x_2 \in \mathbb{R}\},\ \mathbb{A}_2 = \{(0, 0, x_3) \mid x_3 \in \mathbb{R}\}.$$

In case $q = p$, we have $q_1 = q_2 = q = p$ because $p = (-1)p + 1p + 1p$, which corresponds to identity $0 = 0 + 0$ in linear algebra. Please draw a picture on a plane to visualize it.

For simplicity, people prefer to write $(-1)p + 1q_1 + 1q_2$ as (q_1, q_2).

As another example, if $\mathbb{A}_1$ and $\mathbb{A}_2$ are affine spaces, and $p_i \in \mathbb{A}_i$, then affine subspaces $\mathbb{A}_1 \times \{p_2\}$ and $\{p_1\} \times \mathbb{A}_2$ of the categorical product $\mathbb{A}_1 \times \mathbb{A}_2$ are affine complement of each other, i.e.,

$$\mathbb{A}_1 \times \mathbb{A}_2 = \mathbb{A}_1 \times \{p_2\} \oplus \{p_1\} \times \mathbb{A}_2.$$

Euclidean Space

A Euclidean space is an affine space together with a Euclidean structure, i.e., a translation-invariant assignment of inner product to each tangent space: for any points p, q_1, q_2 in the affine space, if we use p', q_1' and q_2' to denote their respective translations by an arbitrary vector $\underline{u}$, then

$$\langle \overrightarrow{pq_1}, \overrightarrow{pq_2} \rangle = \langle \overrightarrow{p'q_1'}, \overrightarrow{p'q_2'} \rangle.$$

Here we use $\langle\, , \rangle$ to denote the inner product on each tangent space.

We use $\mathbb{E}^n$ to denote the Euclidean space for which the underlying affine space is $\mathbb{A}^n_{\mathbb{R}}$ and the Euclidean structure is the standard one:

$$\langle (p, \vec{u}), (p, \vec{v}) \rangle = \vec{u} \cdot \vec{v}, \quad \text{i.e. the dot product of } \vec{u} \text{ with } \vec{v}.$$

$\mathbb{E}^n$ is called the model Euclidean space of dimension n. In most textbooks, it is also denoted by $\mathbb{R}^n$.

Remark 4. An affine subspace of a Euclidean space $\mathbb{E}$ is a Euclidean space with the inherited Euclidean structure, and it is called a **Euclidean subspace** of $\mathbb{E}$.

Remark 5. Suppose that $\mathbb{A}$ is an affine space modelled on vector space V, and $p \in \mathbb{A}$. Then the natural map

$$\begin{array}{ccc} T_p\mathbb{A} & \to & V \\ (p, \underline{u}) & \mapsto & \underline{u} \end{array}$$

is a linear equivalence. Suppose that an inner product on V has been chosen, then the above natural identification of V with $T_p\mathbb{A}$ produces an inner product on $T_p\mathbb{A}$ for each point p, which is invariant under the translation of p. So an inner product on V gives us a Euclidean structure on $\mathbb{A}$. In particular this says a Euclidean structure always exists on an affine space. In fact, all Euclidean structures on $\mathbb{A}$ can be obtained this way.

In general, if $\mathbb{E}$ is a Euclidean space of dimension n, then $\mathbb{E}$ is affine isometric to $\mathbb{E}^n$, this means that there is an affine equivalence ϕ: $\mathbb{E} \to \mathbb{E}^n$ such that, for any points p, q_1, q_2 in $\mathbb{E}$, if we let $p' = \phi(p)$, $q_1' = \phi(q_1)$ and $q_2' = \phi(q_2)$, then

To put it differently, up to equivalences, there is only one Euclidean space for any given dimension.

$$\langle \overrightarrow{pq_1}, \overrightarrow{pq_2} \rangle = \langle \overrightarrow{p'q_1'}, \overrightarrow{p'q_2'} \rangle.$$

In other words, ϕ is an affine equivalence that respects the Euclidean structures.

By definition, a symmetry (or automorphism) of the Euclidean space $\mathbb{E}$ is an affine equivalence ϕ: $\mathbb{E} \to \mathbb{E}$ which respects the Euclidean structure on $\mathbb{E}$. It is clear that the reflection about any hyperplane in $\mathbb{E}$ is a symmetry of $\mathbb{E}$. A theorem of E. Cartan says the symmetries of this kind are the basic building blocks for the symmetries of Euclidean spaces: any symmetry of a Euclidean space is the composition of some reflections, and any symmetry that preserves orientation is the composition of an even number of reflections.

By definition, a rigid motion of $\mathbb{E}$ is a symmetry of $\mathbb{E}$ which respects any choosen orientation on $\mathbb{E}$.

Exercise 15 (rigid motion)**.** *Let ϕ: $\mathbb{E}^n \to \mathbb{E}^n$ be a rigid motion. Show that there is a rotation matrix A of order n such that*

$$\phi(x + \vec{u}) = \phi(x) + A\vec{u} \quad \textit{for any } x \in \mathbb{E}^n \textit{ and any } \vec{u} \in \mathbb{R}^n.$$

`Hint:` ϕ`:` $(\mathbb{E}^n, x) \to (\mathbb{E}^n, \phi(x))$ `is a linear map, i.e.,` $\phi(x + \vec{u}) = \phi(x) + A\vec{u}$ `for a matrix` A`.`

Since $x = 0 + \vec{x}$, we have

$$\phi(x) = \phi(0) + A\vec{x} = (0 + \overrightarrow{\phi(0)}) + A\vec{x} = (0 + A\vec{x}) + \overrightarrow{\phi(0)},$$

we see that a rigid motion of $\mathbb{E}^n$ is always a rotation about point 0 (in fact this point can be any point) followed by a translation.

Remark 6. In a Euclidean space $\mathbb{E}$, the length of a line segment

$$\overline{AB} := \{tA + (1-t)B \mid 0 \le t \le 1\}$$

is

$$|\overline{AB}| := \sqrt{\left\langle \overrightarrow{AB}, \overrightarrow{AB} \right\rangle}$$

and the angle $\angle ABC$ is

$$\arccos\left(\frac{\left\langle \overrightarrow{BA}, \overrightarrow{BC} \right\rangle}{|\overline{BA}| \cdot |\overline{BC}|}\right).$$

Exercise 16. *Consider two parallel lines $ax + by + c_1 = 0$ and $ax + by + c_2 = 0$ inside $\mathbb{E}^2$. Find the distance between them. What about the distance between two parallel affine planes in $\mathbb{E}^3$?*

`Hint: use a rigid motion of` $\mathbb{E}^2$ `to reduce the problem to the special case in which` $b = 0$. `Answer:` $\frac{|c_1 - c_2|}{\sqrt{a^2+b^2}}$.

Exercise 17. *Let $p = (x_0, y_0)$ be a point in $\mathbb{E}^2$ and L be the line $ax + by + c = 0$ inside $\mathbb{E}^2$. Find the distance between p and L. What about the distance between a point in $\mathbb{E}^n$ and a hyperplane in $\mathbb{E}^n$?*

`Hint: point` p `is on the parallel line` $ax + by + c' = 0$ `where` $c' = -(ax_0 + by_0)$. `Then use result from the previous exercise problem. Answer:` $\frac{|ax_0+by_0+c|}{\sqrt{a^2+b^2}}$.

One can use a Euclidean structure on an affine space to introduce the notion of open set and show that this notion is independent of the choice of the Euclidean structure.

Since the tangent space is a concept determined by the local data, and locally an open set U of $\mathbb{A}$ looks like $\mathbb{A}$, so the tangent space of U at a point $p \in U$, denoted by T_pU, is the same as $T_p\mathbb{A}$. I.e.,

This would be obvious if we view a tangent vector at point p as the velocity vector of a moving particle at the moment it arrives at point p.

$$T_pU = T_p\mathbb{A}.$$

An element of T_pU is called a tangent vector of U at point p. The dual of T_pU, denoted by T_p^*U, is called the cotangent space of U at point p. An element of T_p^*U is called the cotangent vector of U at p.

In case the affine space $\mathbb{A}$ is modelled on vector space V, the total tangent space $TU := \cup_{p\in U} T_pU$ and the total cotangent space $T^*U := \cup_{p\in U} T_p^*U$ can be described as follows:

$$TU = U \times V, \quad T^*U = U \times V^*.$$

In particular, we have $T\mathbb{A} = \mathbb{A} \times V$, $T^*\mathbb{A} = \mathbb{A} \times V^*$, which are affine spaces modelled on vector spaces $V \oplus V$ and $V \oplus V^*$ respectively. Then both TU and T^*U are open sets of an affine space: TU is an open set of $T\mathbb{A}$ and T^*U is an open set of $T^*\mathbb{A}$.

Summary

A vector space is a non-empty set together with a linear structure, i.e., a way to form linear combination. An affine space is a non-empty set together with an affine structure, i.e., a way to form affine combination. A Euclidean vector space is a vector space together with an inner product. A Euclidean space is an affine space together with a Euclidean structure. Up to equivalence, there is only one vector space, affine space, Euclidean vector space and Euclidean space of dimension n. The model spaces are $\mathbb{R}^n$, $\mathbb{A}^n_{\mathbb{R}}$, $\mathbb{R}^n$, $\mathbb{E}^n$ respectively, and their symmetry groups are $\mathrm{GL}(n,\mathbb{R})$, $\mathrm{Aff}(n,\mathbb{R}) = \mathbb{R}^n \rtimes \mathrm{GL}(n,\mathbb{R})$, $\mathrm{O}(n)$, $\mathrm{E}(n) = \mathbb{R}^n \rtimes \mathrm{O}(n)$ respectively. An equivalence with the model space is called a trivialization and it amounts to a basis, an affine frame, an orthonormal basis and an orthonormal frame respectively.

A vector space is an affine space canonically, an affine space is a topological space canonically, an Euclidean space is a metric space canonically. A linear map is an affine map, an affine map is a continuous map. A pointed affine space is a vector space canonically and an affine map between pointed affine spaces is a linear map.

Let us end the discussion so far with a table that compares affine spaces with vector spaces.

	Affine Space	Vector Space
generic notation	$\mathbb{A}$	V
model space	$\mathbb{A}^n_{\mathbb{R}}$	$\mathbb{R}^n$
morphism	affine map	linear map
isomorphism	affine equivalence	linear equivalence
elements	points	vectors
algebraic structure	affine combination	linear combination
geometric structure	Euclidean structure	inner product
equivalence with a model means	affine frame $(p_0, p_1, \ldots, p_n)$	basis $(\underline{v}_1, \ldots, \underline{v}_n)$
subobject	affine subspace	linear subspace
	affine independence	linear independence
	span	span

Appendix: Extra examples and Exercises

Example 6. *Describe the solution set X (as a subset of $\mathbb{E}^2$) of quadratic equation*

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

where real numbers a, b, c are not all equal to zero.

Solution. We shall use rigid motions of $\mathbb{E}^2$ to simplify the analysis. First of all, up to rotation about point $(0,0)$, we may assume [2] that $b = 0$. Let us further divide the analysis into various cases.

[2] Here we use the two facts: 1) a quadratic form is represented by a real symmetric matrix, 2) a real symmetric matrix can be diagonalized by a rotation matrix.

Case I. Coefficients a, c are all nonzero. Up to translation we may assume that $d = e = 0$.

for example, $ax^2 + dx = a(x + \frac{d}{2a})^2 - \frac{d^2}{4a}$.

Case $f \neq 0$. After dividing by f, we end up with equation of the form

$$ax^2 + cy^2 = 1$$

where a, c are all nonzero. Further analysis can be divided into a few sub-cases:

Sub-case I_1. a, c are all negative. X is an empty set.

Sub-case I_2. a, c are all positive. Up to rotation, X is the solution set of

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

where $a \geq b > 0$, i.e., an ellipse.

Sub-case I_3. a, c have opposite sign. Up to rotation, X is the solution set of

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

where $a, b > 0$, i.e., a hyperbola.

Case $f = 0$. Then we end up with equation of the form

$$ax^2 + cy^2 = 0$$

where a, c are all nonzero. Further analysis can be divided into two sub-cases:

Sub-case I_4. $ac > 0$. X is a point.

Sub-case I_5. $ac < 0$. Then X is the solution set of

$$y^2 = (ax)^2$$

where $a > 0$, i.e., two intersecting lines.

Case II. Only one of coefficients a, c is zero. We may assume that $c = 0$, then we may further assume that $d = 0$, so we may assume (after dividing by a) that the equation is of the form

$$x^2 + ey + f = 0.$$

Further analysis can be divided into two sub-cases:

Sub-case II_1. $e \neq 0$. Then, up to a translation plus a possible rotation, X is the solution set of

$$y = ax^2$$

where $a > 0$, i.e., a parabola.

Sub-case II_2. $e = 0$ and $f < 0$. Then X is the solution set of

$$x^2 = a^2$$

where $a > 0$, i.e., two parallel lines.

Sub-case II_3. $e = 0$ and $f = 0$. Then X is the solution set of

$$x^2 = 0,$$

i.e., a line.

Sub-case II_4. $e = 0$ and $f > 0$. X is an empty set.

In summary, as a subset of $\mathbb{E}^2$, the solution set of a real quadratic equation in two variables is one of the followings:

(1) the empty set,
(2) a point,
(3) a line,
(4) two parallel lines (with any non-zero distance),
(5) two intersecting lines (with any angle θ such that $\sin\theta \neq 0$),
(6) a parabola (of any shape),
(7) an ellipse (of any shape),
(8) a hyperbola (of any shape).

Sets (1) and (2) are not curves, sets (3) and (4) are degenerate quadratic curves, sets (5) through (8) are non-degenerate quadratic curves. Here *non-degenerate* means not degenerate, and *degenerate* means that, modulo an affine transformation of $\mathbb{E}^2$, the degree 2 algebraic equation that defines the quadratic curve contains fewer than 2 variables. ◀

Remark 7. In the above example, the result is a classification of curves of degree 2 modulo rigid motions. If we classify these curves modulo affine equivalences, the result is even simpler: As a subset of $\mathbb{A}^2_{\mathbb{R}}$, curves of degree 2 modulo affine equivalences is one of the followings:

(1) the empty set,
(2) the point $(0,0)$,
(3) the line $y = 0$,
(4) the two parallel lines $y = 1$ and $y = -1$,
(5) the two intersecting lines $y = 0$ and $x = 0$,
(6) the parabola $y = x^2$,
(7) the ellipse $x^2 + y^2 = 1$,
(8) the hyperbola $x^2 - y^2 = 1$.

Exercise 18. *Describe, modulo affine equivalences, the solution set (as a subset of $\mathbb{A}^3_{\mathbb{R}}$) of a real quadratic equation in three variables.*

Hint: The empty set, a point, and a line should be in the classification list, though they are not surfaces. The cylinder over the curves (3) - (8) listed in the preceding remark should be in the classification list, and they are called *degenerate* quadrics because, up to affine transformations, the quadratic equation contains fewer than 3 variables. The non-degenerated quadrics are the following six:

$$x^2+y^2+z^2=1,\quad x^2+y^2-z^2=1,\quad x^2+y^2-z^2=-1,\quad x^2+y^2-z^2=0,\quad z=x^2+y^2,\quad z=x^2-y^2.$$

Altogether there are $3+6+6=15$ of them, 12 of which are quadrics -- surfaces defined by quadratic equations.

Exercise 19. *Let p_i, q_j be points in affine space $\mathbb{A}$ and c^i, d^j, α and β be real numbers. Here $0 \le i \le m$ and $0 \le j \le n$. Suppose that $\sum_{i=0}^m c^i = \sum_{j=0}^n d^j = \alpha+\beta = 1$.*

(a) Show that $\alpha(c^0p_0+\cdots+c^mp_m)+\beta(d^0q_0+\cdots+d^nq_n)$ is equal to

$$(\alpha c^0)p_0+\cdots+(\alpha c^m)p_m+(\beta d^0)q_0+\cdots+(\beta d^n)q_n.$$

(b) Show that the three medians of a triangle always meet at a common point.

Hint: one may assume that $\mathbb{A}=\mathbb{A}^n_{\mathbb{R}}$. Then we can write p_i as $(p_i^1,\ldots,p_i^n)$ and q_j as $(q_j^1,\ldots,q_j^n)$. Recall that affine combinations in $\mathbb{A}^n_{\mathbb{R}}$ are done component-wisely.

Exercise 20. *Let $\mathbb{A}$ be an affine space, $\mathbb{B}$ be an affine subspace of $\mathbb{A}$, and p be a point in $\mathbb{B}$.*

1) Show that the pointed space $(\mathbb{B}, p)$ is linear subspace of the vector space $(\mathbb{A}, p)$. So $(\mathbb{B}, p)$ is a vector space as well.

2) Show that $\mathbb{B}$ is an affine space modelled on vector space $(\mathbb{B}, p)$. This says that affine subspaces are affine spaces.

3) Let $T\colon V \to W$ be a linear map, $\underline{w}$ be a vector in W such that $T^{-1}(\underline{w}) := \{\underline{v}\in V \mid T(\underline{v})=\underline{w}\}$ is a nonempty set. Show that $T^{-1}(\underline{w})$ is an affine subspace of the affine space V (i.e., V_{aff}), but not a linear subspace of the vector space V unless $\underline{w}=\underline{0}$. This says that taking inverse image of a linear map will get us out of the realm of vector spaces.

4) Let $F\colon \mathbb{A}_1 \to \mathbb{A}_2$ be an affine map, y be a point in $\mathbb{A}_2$ such that $F^{-1}(y)$ is a nonempty set. Show that $T^{-1}(y)$ is an affine subspace of the affine space $\mathbb{A}_1$.

5) Suppose that $\mathbb{A}_1$ and $\mathbb{A}_2$ are affine subspaces of $\mathbb{A}$, and $\mathbb{A}_1 \subseteq \mathbb{A}_2$. Show that $\dim\mathbb{A}_1 \le \dim\mathbb{A}_2$ with equality holds if and only if $\mathbb{A}_1=\mathbb{A}_2$.

Hint to part 5): pick a point p in $\mathbb{A}_1$ so that one can turn the problem into a problem in linear algebra.

Exercise 21. *Let $\mathbb{A}_i$ be an affine space modelled on vector space V_i and $F\colon \mathbb{A}_1 \to \mathbb{A}_2$ be an affine map. Show that*

1) There is a linear map $T\colon V_1 \to V_2$ such that

$$F(p+\underline{u}) = F(p)+T\underline{u} \quad \text{for any } p\in\mathbb{A}_1 \text{ and any } \underline{u}\in V_1.$$

2) The graph Γ_F of F is an affine subspace of the affine space $\mathbb{A}_1 \times \mathbb{A}_2$.

Recall that the product of two affine spaces is an affine space.

3) The map from $\mathbb{A}_1$ to Γ_F that maps $p \in \mathbb{A}_1$ to $(p, F(p))$ is an affine equivalence.

`Hint: Without loss of generality (WLOG) we may assume that` $\mathbb{A}_i$ `are model affine spaces (say` $\mathbb{A}_1 = \mathbb{A}^n_{\mathbb{R}}$ `and` $\mathbb{A}_2 = \mathbb{A}^m_{\mathbb{R}}$`). Then` V_1 `and` V_2 `are model linear spaces` $\mathbb{R}^n$ `and` $\mathbb{R}^m$ `respectively, and` $p + \vec{u}$ `is the affine combination`

$$(1 - \sum_i u^i)p + u^1(p + \vec{e}_1) + \cdots + u^n(p + \vec{e}_n).$$

`Let` A `be the` $m \times n$`-matrix whose` j`-th column is the column vector` $F(p + \vec{e}_j) - F(p)$`. Show that` $F(p + \vec{u}) = F(p) + A\vec{x}$`.`

Part 3) in the last exercise says that, for a given affine map, its domain and its graph are naturally affine equivalent.

Exercise 22. *Let $F\colon \mathbb{A}_1 \to \mathbb{A}_2$ be an affine map. Show that any two of the following three statements being true implies the third is true.*

1) F is one-to-one,

2) F is onto,

3) $\dim \mathbb{A}_1 = \dim \mathbb{A}_2$.

`Hint: pick a point` p_1 `in` $\mathbb{A}_1$ `and let` $p_2 = F(p_1)$`. Then` F `is a linear map from` $(\mathbb{A}_1, p_1)$ `to` $(\mathbb{A}_2, p_2)$`.`

Exercise 23. *Let $F\colon \mathbb{A}_1 \to \mathbb{A}_2$ be an affine map. Show that*

1) F is one-to-one $\implies \dim \mathbb{A}_1 \le \dim \mathbb{A}_2$,

2) F is onto $\implies \dim \mathbb{A}_1 \ge \dim \mathbb{A}_2$,

3) For any affine subspace $\mathbb{B}$ of $\mathbb{A}_1$,

$$F(\mathbb{B}) = \{F(p) \mid p \in \mathbb{B}\}$$

is an affine subspace of $\mathbb{A}_2$ and $\dim F(\mathbb{B}) \le \dim \mathbb{B}$,

4) For any affine subspace $\mathbb{C}$ of $\mathbb{A}_2$,

$$F^{-1}(\mathbb{C}) = \{p \in \mathbb{A}_1 \mid F(p) \in \mathbb{C}\}$$

is an affine subspace of $\mathbb{A}_1$.

`Hint: by picking a point in an affine space, the problems can be turned into problems in linear algebra.`

List of Symbols

$\mathbb{A}, \mathbb{B}, etc.$	affine spaces
V_{aff}	the underlying affine space of the linear space V
$\mathbb{A}^n_{\mathbb{R}}$	the affine space of n-tuples of real numbers. It is called the model real affine space of dimension n, also called the coordinate space of dimension n
$\mathbb{E}^n$	the model Euclidean space of dimension n
p, q, etc.	points in abstract affine spaces
x, y, etc.	points in model affine spaces
$\vec{x}$	the vector in $\mathbb{R}^n$ such that $0 + \vec{x} = x$ or $\vec{x} = x - 0$
$\overline{pq}$	the line segment with end points p and q in an affine space
$\lvert\overline{pq}\rvert$	the length of the line segment $\overline{pq}$
$\overrightarrow{pq}$	the vector from point p to point q in an affine space
U	an open set of an affine space
T_pU	the tangent space of U at point p
$(\mathbb{A}, p)$	an pointed affine space. It is a linear space and $(\mathbb{A}, p) \equiv T_p\mathbb{A}$
TU	the total tangent space of U
T^*_pU	the cotangent space of U at point p
T^*U	the total cotangent space of U
$\mathrm{Aut}(X)$	the group of symmetries of X or the automorphism group of X
$\mathrm{GL}(V)$	the group of symmetries of the linear space V, i.e., $\mathrm{Aut}(V)$
$\mathrm{GL}_+(V)$	the group of symmetries of the oriented linear space V
$\mathrm{O}(V)$	the group of symmetries of the Euclidean vector space V
$\mathrm{SO}(V)$	the group of symmetries of the oriented Euclidean vector space V
$\mathrm{Aff}(\mathbb{A})$	the group of symmetries of the affine space $\mathbb{A}$, i.e., $\mathrm{Aut}(\mathbb{A})$
$\mathrm{GL}(n, \mathbb{R})$	the group of invertible square matrices of order n, naturally equivalent to $\mathrm{GL}(\mathbb{R}^n)$
$\mathrm{Aff}(n, \mathbb{R})$	the semi-direct product $\mathbb{R}^n \rtimes \mathrm{GL}(n, \mathbb{R})$, naturally equivalent to $\mathrm{Aff}(\mathbb{A}^n_{\mathbb{R}})$
$\mathrm{E}(n)$	the semi-direct product $\mathbb{R}^n \rtimes \mathrm{O}(n)$, naturally equivalent to $\mathrm{Aut}(\mathbb{E}^n)$

§3 Review of Differential Calculus

We review differential calculus here. The main message is this: differentiation is geometric. In modern language, differentiation is a functor.

The ordinary textbook approach to differential calculus is similar to the matrix approach to linear algebra. In both cases we take the quick route at the expanse of not showing the essence of the subject.

Let U be an open set of an affine space $\mathbb{A}$ modelled on vector space V, $f\colon U \to \mathbb{R}$ be a map, and $p \in U$. Nothing is lost if we assume that $\mathbb{A}$ is a model affine space, say $\mathbb{A}^n_{\mathbb{R}}$.

Assuming that f is smooth, which means that *the partial derivative functions of f to any order exist and are continuous*. Note that polynomial functions, the natural exponential function and natural logarithmic function, the sine and cosine functions are all smooth, so are functions made out of smooth functions via arithmetic operations and composition. For example, if x, y, z are the standard coordinate functions on $\mathbb{A}^3_{\mathbb{R}}$, then functions

If $\alpha > 0$, we define $\alpha^\beta = e^{\beta \ln \alpha}$. So $\sqrt{x}$ and $x^{\frac{1}{2}}$ are different functions. By definition, $\sqrt{\ }$ is the inverse of the map from $[0, \infty)$ to $[0, \infty)$ that sends x to x^2. Similarly, arcsin is the inverse of the map from $[-\frac{\pi}{2}, \frac{\pi}{2}]$ to $[-1, 1]$ that sends x to $\sin x$.

$$x^2 + y^2 - z^2, \quad \sin(x + xy + xyz), \quad \ln(x^2 + y^2 + z^2 + 1), \quad (x^2 + 1)^{x+y+z}$$

are all smooth functions on $\mathbb{A}^3_{\mathbb{R}}$. Hereafter we are only interested in smooth functions.

Calculus I

In Calculus I we study functions of one variable, i.e., functions $f\colon I \to \mathbb{R}$ or $\mathbb{R}^n$, where I is a non-empty open interval, say (a, b) with $-\infty \le a < b \le \infty$. We usually write function f this way:

$$y = f(x)$$

where x is called the *independent variable* and y is called the *dependent variable*. The key concept here is the derivative of f, which is written in various notations:

$$\frac{\mathrm{d}y}{\mathrm{d}x}, \quad \frac{\mathrm{d}f}{\mathrm{d}x}, \quad f', \quad \dot{f}, \quad f^{(1)}.$$

The derivative f' of function f is a function whose value at x_0 is defined to be the limit of a difference quotient:

$$\frac{\mathrm{d}y}{\mathrm{d}x} := \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

More precisely,

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x_0) := \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Note that, in practical situation derivative $f'(x_0)$ has the meaning such as the slope of the graph Γ_f at point $(x_0, f(x_0))$, the rate of change of f at x_0, the velocity at time x_0, etc.

The most efficient way of computing derivatives is to use rules for differentiation:

1. $f' = 0$ if f is a constant function, i.e., independent of x.
2. The map $f \mapsto f'$ is an additive operator, i.e., $(f+g)' = f' + g'$.
3. Product Rule: $(fg)' = f'g + fg'$.
4. Chain Rule: $(f \circ g)'(x) = f'(g(x))g'(x)$, i.e., $(f \circ g)' = (f' \circ g)\, g'$.

The product rule says that the derivative of a product is the sum of product, one product for each factor, e.g.,

$$(fgh)' = f'gh + fg'h + fgh'.$$

The chain rule says that the derivative of a composition of functions is the product of derivatives, one for each function in the composition, e.g.,

$$(f \circ g \circ h)' = (f' \circ g \circ h)(g' \circ h)h'.$$

Calculus II

In Calculus II we study functions $f \colon U \to W$, where U is a non-empty open set of $\mathbb{E}^n$ and W is a non-empty open set of $\mathbb{E}^m$. We write f as $y = f(x)$ in compact notation or $y^i = f^i(x^1, \ldots, x^n)$. Here f^i is the i-th component of f, i.e., $f = (f^1, \ldots, f^m)$. Note that $x = (x^1, \ldots, x^n)$ and $y = (y^1, \ldots, y^m)$, so we may write $y^i = f^i(x)$ as well. Sometime we write

$$\begin{cases} y^1 & = & f^1(x^1, \ldots, x^n) \\ \vdots & & \\ y^m & = & f^m(x^1, \ldots, x^n) \end{cases}$$

In this case the analogue of the derivative in Calculus I is the Jacobean matrix of f:

$$Jf := \begin{bmatrix} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^n} \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial x^1} & \cdots & \frac{\partial f^m}{\partial x^n} \end{bmatrix} \quad \text{or} \quad Jf = \left[\frac{\partial f^i}{\partial x^j}\right] \quad \text{in short.}$$

Here are other notations: $\frac{\partial y}{\partial x}$, $\frac{\partial(y^1,\ldots,y^m)}{\partial(x^1,\ldots,x^n)}$. Note: In case $m = n = 1$, Jf is nothing but the derivative of f.

We can also write Jf in the column form:

$$Jf = \left[\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n}\right]$$

where $\frac{\partial f}{\partial x^i}$, a $\mathbb{R}^m$-valued function on U, is the derivative of $f(x^1, \ldots, x^n)$ with respect to variable x^i with all other independent variables frozen. So

$$\begin{aligned} \frac{\partial f}{\partial x^i}(x) &= \lim_{h\to 0} \frac{f(\cdots, x^i + h, \cdots) - f(\cdots, x^i, \cdots)}{h} \\ &= \lim_{h\to 0} \frac{f(x + h\vec{e}_i) - f(x)}{h} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x + t\vec{e}_i) \end{aligned} \tag{5}$$

We call $\frac{\partial f}{\partial x^i}$ the partial derivative of f along the i-th direction. A small generalisation is the directional derivative of f along direction $\vec{n}$ (a unit vector in $\mathbb{R}^n$):

$$\frac{\partial f}{\partial \vec{n}}(x) := \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x + t\vec{n}).$$

Here is a further generalization: for any $\vec{u} \in \mathbb{R}^n$, let

Chain Rule: The composition $f \circ g$ of smooth maps f and g between open sets of model affine spaces is smooth and

$$\frac{\partial (f\circ g)^i}{\partial x^j}(x) = \frac{\partial f^i}{\partial y^k}(g(x))\frac{\partial g^k}{\partial x^j}(x)$$

$$\begin{aligned} f_{\vec{u}}(x) &:= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x + t\vec{u}) \\ &= (f\circ g)'(0) \quad \text{where } g(t) = x + t\vec{u} \\ &= f'(g(0))g'(0) \quad \text{chain rule} \\ &= Jf(x)\vec{u} \\ &= u^i \frac{\partial f}{\partial x^i}(x). \end{aligned}$$

I.e., $f_{\vec{u}} = Jf\,\vec{u}$. Then we know that $f_{\vec{u}}$ depends on $\vec{u}$ linearly. Note that $f_{\vec{n}} = \frac{\partial f}{\partial \vec{n}}$ and $f_{\vec{e}_i} = \frac{\partial f}{\partial x^i}$.

Exercise 24. *Let $f = x^2 + y^2 - z^2$ be the real smooth function on $\mathbb{A}^3_{\mathbb{R}}$. ($x = x^1$, $y = x^2$, $z = x^3$) Compute $f_{\vec{0}}$, $f_{\vec{e}_1}$, $f_{\vec{e}_1+2\vec{e}_2}$, $f_{\vec{e}_1+2\vec{e}_2+3\vec{e}_3}$.*

Please note that the discussion above makes sense for functions on non-empty open sets of any affine space, with $\vec{u}$ being written as $\underline{u}$. For example, $f_{\underline{u}}$ depends on $\underline{u}$ linearly. In calculus II, we also introduce higher order derivatives such as

$$\frac{\partial^2 f}{\partial x^i \partial x^j} := \frac{\partial}{\partial x^i}\left(\frac{\partial f}{\partial x^j}\right) = (f_{\vec{e}_j})_{\vec{e}_i} =: f_{\vec{e}_i \vec{e}_j}. \quad \text{In general, } f_{\vec{u}\vec{v}} := (f_{\vec{v}})_{\vec{u}}.$$

Other notations:

$$\frac{\partial f}{\partial x^i}, \quad f_{x^i}, \quad \partial_{x^i} f, \quad \partial_i f \quad \text{all mean the same}$$

$$\frac{\partial^2 f}{\partial x^i \partial x^j}, \quad f_{x^i x^j}, \quad \partial_{x^i}\partial_{x^j} f, \quad \partial_i \partial_j f \quad \text{all mean the same}$$

$$\frac{\partial^2 f}{\partial x^i \partial x^i}, \quad \frac{\partial^2 f}{\partial (x^i)^2}, \quad \partial^2_{x^i} f, \quad \partial^2_i f \quad \text{all mean the same}$$

Since $\frac{\partial}{\partial x^i}$ is really $\frac{\mathrm{d}f}{\mathrm{d}x^i}$ in calculus I, $\frac{\partial}{\partial x^i}$ satisfies the following rules:

1. $\frac{\partial f}{\partial x^i} = 0$ if f is independent of x^i.
2. $\frac{\partial}{\partial x^i}$ is additive: $\frac{\partial}{\partial x^i}(f+g) = \frac{\partial f}{\partial x^i} + \frac{\partial g}{\partial x^i}$.
3. Product Rule: $\frac{\partial}{\partial x^i}(f \cdot g) = \frac{\partial f}{\partial x^i} \cdot g + f \cdot \frac{\partial g}{\partial x^i}$. Here $\cdot$ can be any product which is linear in each of its factors. For example, if $\vec{A}$, $\vec{B}$, $\vec{C}$ are $\mathbb{R}^3$-valued functions of x, y, z, then

$$\frac{\partial}{\partial x}(\vec{A} \times \vec{B}) = \frac{\partial \vec{A}}{\partial x} \times \vec{B} + A \times \frac{\partial \vec{B}}{\partial x}.$$

$$\frac{\partial}{\partial x}(\vec{A} \cdot (\vec{B} \times \vec{C})) = \frac{\partial \vec{A}}{\partial x} \cdot (\vec{B} \times \vec{C}) + \vec{A} \cdot \left(\frac{\partial \vec{B}}{\partial x} \times \vec{C}\right) + \vec{A} \cdot \left(\vec{B} \times \frac{\partial \vec{C}}{\partial x}\right).$$

4. Chain Rule: Given the sequence of smooth maps $U_1 \xrightarrow{g} U_2 \xrightarrow{f} U_3$, we have $f \circ g$ smooth and $J(f \circ g)_p = Jf_{g(p)} \cdot Jg_p$ for all $p \in U_1$. Here $\cdot$ is the matrix multiplication and Jg_p denote the value of Jg at point p, etc.

Recall that matrices represent linear maps and matrix multiplication represents composition of linear maps. So the chain rule stated above should have an intrinsic formulation. Indeed, this is the case: Let $p_i \in U_i$ with $p_2 = g(p_1)$ and $p_3 = f(p_2)$, then we have commutative digrams

$$\begin{tikzcd}
T_{p_1}U_1 \arrow[r, "T_{p_1}g"] \arrow[d, equal] & T_{p_2}U_2 \arrow[d, equal] \\
\{p_1\} \times \mathbb{R}^{n_1} \arrow[d, equal] & \{p_2\} \times \mathbb{R}^{n_2} \arrow[d, equal] \\
\mathbb{R}^{n_1} \arrow[r, "Jg_{p_1}"] & \mathbb{R}^{n_2}
\end{tikzcd}$$

and

$$\begin{tikzcd} & T_{p_2}U_2 \arrow[dr, "T_{p_2}f"] & \\ T_{p_1}U_1 \arrow[ur, "T_{p_1}g"] \arrow[rr, "T_{p_1}(g\circ f)"] & & T_{p_3}U_3 \end{tikzcd}$$

Hereafter the further review of Calculus I & II shall be done in a more intrinsic way.

Total Differential

Continuing the discussion above, we observe that the assignment of $f_{\underline{u}}(p)$ to tangent vector $(p, \underline{u})$ is a cotangent vector of U at point p. This cotangent vector, denoted by df_p, is called the differential of f at point p. By definition, the map

$$\begin{array}{rccc} \mathrm{d}f: & U & \to & T^*U \\ & p & \mapsto & \mathrm{d}f_p \end{array}$$

is the total differential of f. Please note that the equation

$$f_{\underline{u}}(p) = \langle \mathrm{d}f_p, (p, \underline{u})\rangle$$

defines $\mathrm{d}f_p$ uniquely and is thus said to be a defining equation for $\mathrm{d}f_p$.

Let E_i be the vector field on $\mathbb{A}^n_{\mathbb{R}}$ whose value at point p is the tangent vector $(p, \vec{e}_i)$. For the sanity of notations, the restriction of E_i to any open set of $\mathbb{A}^n_{\mathbb{R}}$ is also denoted by E_i. Then, for any smooth function f on a non-empty open set U of $\mathbb{A}^n_{\mathbb{R}}$, we have

$$\frac{\partial f}{\partial x^i}(p) = f_{\vec{e}^i}(p) = \langle \mathrm{d}f_p, E_i(p)\rangle \quad \text{or simply} \quad \frac{\partial f}{\partial x^i} = \langle \mathrm{d}f, E_i\rangle$$

if we hide the variable point p. Thus the total differential $\mathrm{d}f$ *determines* the partial derivatives $\frac{\partial f}{\partial x^i}$.

Since x^j is the j-th standard coordinate function on $\mathbb{A}^n_{\mathbb{R}}$, we have

Again, for the sanity of notations, the restriction of x^j to U is denoted by x^j as well. The identity on the left should be viewed point-wisely in case you are confused.

$$\langle \mathrm{d}x^j, E_i\rangle = \frac{\partial x^j}{\partial x^i} = \delta^j_i, \text{which is a constant function on } U.$$

This identity says that $(E_1, \ldots, E_n)$ is a tangent frame and $(\mathrm{d}x^1, \ldots, \mathrm{d}x^n)$ is the corresponding dual cotangent frame; consequently, we have

A tangent frame on U means a smooth map that sends point p in U to a basis for the tangent space T_pU. One can have a similar definition for cotangent frame.

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

Indeed, since $(\mathrm{d}x^1, \ldots, \mathrm{d}x^n)$ is a cotangent frame on U, we must have $\mathrm{d}f = f_i\,\mathrm{d}x^i$ for some unique choice of real functions f_i on U; then, paring with E_j, we have $\langle \mathrm{d}f, E_j\rangle = \langle f_i\mathrm{d}x^i, E_j\rangle$ or $\frac{\partial f}{\partial x^j} = f_j$.

If you get stuck about this point, please think point-wisely

The identity $\mathrm{d}f = \frac{\partial f}{\partial x^i}\,\mathrm{d}x^i$ implies that the partial derivatives $\frac{\partial f}{\partial x^i}$ *determine* the total differential $\mathrm{d}f$.

Remark 8. Since the first order partial derivatives and the total differential determine each other, the total differential is just another version of the first order partial derivatives. Being coordinate-free, the total differential is said to be the *geometric* or *intrinsic* version of the first order partial derivatives. Put differently, we say that the row-matrix valued function $[\partial_1 f, \ldots, \partial_n f]$ is the representation of $\mathrm{d}f$ with respect to coordinate functions x^i. Here $\partial_i f$ is a simpler notation for $\frac{\partial f}{\partial x^i}$.

Remark 9. In the geometric version, differentiation is given by the operator d: $f \mapsto \mathrm{d}f$. While f is a smooth real function, $\mathrm{d}f$, though smooth, is no longer a real function, so the set of smooth real functions is not complete with respect to the intrinsic way of taking derivatives. By a completion process with respect to applying operator d, we discover the missing functions: differential forms. More on this in later chapters.

Exercise 25. *Let f and g be two real smooth functions on an open set U of an affine space $\mathbb{A}$. Let $\underline{u}$ be any vector in the vector space on which $\mathbb{A}$ is modelled. Please prove the following:*

1) **Additivity Law***: $(f+g)_{\underline{u}} = f_{\underline{u}} + g_{\underline{u}}$, thus $\mathrm{d}(f+g) = \mathrm{d}f + \mathrm{d}g$.*

2) **Product Rule***: $(fg)_{\underline{u}} = g\, f_{\underline{u}} + f\, g_{\underline{u}}$, thus $\mathrm{d}(fg) = f\,\mathrm{d}g + g\,\mathrm{d}f$.*

Hint: You may assume that $\mathbb{A} = \mathbb{A}^n_{\mathbb{R}}$ and $\underline{u} = \vec{u} = [u^1, \cdots, u^n]^T$. You may also use the identity $\mathrm{d}f = \frac{\partial f}{\partial x^i}\,\mathrm{d}x^i$.

Smooth Map and its Tangent

For $i = 1, 2$ we let U_i be an open set of the affine space $\mathbb{A}_i$, modelled on the vector space V_i. Suppose that a map

$$F: \quad U_1 \to U_2$$

is smooth, which means that *the partial derivatives of F to any order exist and continuous*. Just as before, the first order partial derivative $F_{\underline{u}}$ of F is a map that sends p to $\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} F(p + t\underline{u})$ — a vector in V_2.

Such a map F shall be called calculus-smooth map hereafter.

Exercise 26. *Assume that each $\mathbb{A}_i$ is a model affine space. Show that*

$$F_{\vec{u}}(p) = JF_p \cdot \vec{u}$$

where JF_p is the Jacobian matrix of F at point p and $\cdot$ is the matrix multiplication.

Hint: $F^i_{\vec{u}}(p) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} F^i(p + t\underline{u})$ and then apply chain rule to arrive at identity $F^i_{\vec{u}}(p) = \frac{\partial F^i}{\partial x^j}(p) u^j$.

Exercise 27. *Let U_i be a non-empty open set of $\mathbb{A}_i$ and $f\colon U_1 \to U_2$ be a map. We use $\iota\colon U_2 \to \mathbb{A}_2$ to denote the inclusion map and let $\hat{f} = \iota \circ f$.*

(a) Show that f is smooth $\iff$ $\hat{f}$ is smooth. `Hint:` $f_{\underline{u}} = \hat{f}_{\underline{u}}$

(b) Show that $\hat{f}$ is smooth $\iff$ $l \circ \hat{f}$ is smooth for each affine map l: $\mathbb{A}_2 \to \mathbb{R}$.

`Hint: You may assume that`

`0) affine spaces` $\mathbb{A}_i$ `are model affine spaces and let` $\vec{f} = \hat{f} - 0$, `here` 0 `is` c_0 `--- the constant map with constant value 0, and` l^i `be the affine map that sends` $(y^1, \ldots, y^m)$ `to` y^i,

`1)` $\hat{f}$ `is smooth` $\iff$ $\vec{f}$ `is smooth,`

`2)` $\vec{f} = (l^i \circ \hat{f}) c_{\vec{e}_i}$

`3) constant maps and affine maps are smooth maps,`

`4) composition, addition, and multiplication of smooth maps are smooth maps.`

Remark: The result in this exercise implies that a map between open sets of (model) affine spaces is smooth if and only if its component functions are all smooth.

It is clear that the inclusion map $\iota : U \to \mathrm{A}$ is smooth, and affine maps are smooth. Also, if $f\colon U \to \mathbb{R}$ is smooth, then $\mathrm{d}f\colon U \to T^*U$ is smooth.

Exercise 28. *Prove the last statement.*

`Hint: one may take` U `to be an open set of` $\mathbb{A}^n_{\mathbb{R}}$. `Then` $\mathrm{d}f$ `sends` p `to` Jf_p, `so, it is a map with components` $\frac{\partial f}{\partial x^1}$, `...,` $\frac{\partial f}{\partial x^n}$. `You should use Exercise 27.`

Just as before, the higher order partial derivatives, say $F_{\underline{uv}} := (F_{\underline{v}})_{\underline{u}}$, can be defined iteratively. Unlike F which is map from U_1 to U_2, the partial derivative of F of any order are maps from U_1 to V_2 — the real vector space on which $\mathbb{A}_2$ is modelled. The Taylor series of F around point $p \in U_1$ is

$$F(p + \underline{u}) \sim F(p) + \frac{1}{1!} F_{\underline{u}}(p) + \frac{1}{2!} F_{\underline{uu}}(p) + \cdots$$

The 1st order Taylor polynomial of F around point p is the affine map

$$\begin{array}{rccc} L_p F: & \mathbb{A}_1 & \to & \mathbb{A}_2 \\ & p + \underline{u} & \mapsto & F(p) + F_{\underline{u}}(p) \end{array}$$

Since $F_{\underline{u}}(p)$ is linear in $\underline{u}$, part (1) in Exercise 21 says that $L_p F$ is an affine map.

which is a linear map from $(\mathbb{A}_1, p)$ to $(\mathbb{A}_2, F(p))$ as well. Recall that $T_p U_1$ is naturally identified with $(\mathbb{A}_1, p)$ and $T_{F(p)} U_2$ is naturally identified with $(\mathbb{A}_2, F(p))$, this linear map — referred to as the tangent of F at p — is then rewritten as follows:

$$\begin{array}{rccc} T_p F: & T_p U_1 & \to & T_{F(p)} U_2 \\ & (p, \underline{u}) & \mapsto & (F(p), F_{\underline{u}}(p)) \end{array}$$

Of course, $L_p F$ and $T_p F$ determine each other. Later we shall see that $T_p F$ is a better version because it works in more general setting.

By assembling these linear maps together, we arrive at the tangent map of F:

$$\begin{array}{rccc} TF: & TU_1 & \to & TU_2 \\ & (p, \underline{u}) & \mapsto & (F(p), F_{\underline{u}}(p)) \end{array}$$

Exercise 29. *Suppose that $F\colon U_1 \to U_2$ is smooth. Show that $TF\colon TU_1 \to TU_2$ is smooth.*

`Hint: Use Exercise 27.`

Chain Rule

Here comes a very important theorem in calculus.

This is the geometric version of the chain rule in Calculus II. Being able to capture the essence of the chain rule, this version is much easier to remember: T commutes with composition.

Theorem 1 (Chain Rule)**.** *Suppose that $F\colon U_1 \to U_2$ and $G\colon U_2 \to U_3$ are smooth maps between open sets of affine spaces. Then the composed map $GF\colon U_1 \to U_3$ is smooth as well; moreover*

$$T(GF) = TG\,TF.$$

In other words, T preserves the commutativity for diagrams of smooth maps:

$$\begin{tikzcd}
& U_2 \arrow[dr, "G"] & & & & TU_2 \arrow[dr, "TG"] & \\
U_1 \arrow[ur, "F"] \arrow[rr, "H"'] & & U_3 & \Longrightarrow & TU_1 \arrow[ur, "TF"] \arrow[rr, "TH"'] & & U_3
\end{tikzcd}$$

It is also clear that T maps an identity map to an identity map.

Let $\mathcal{O}$ be the category whose morphisms are smooth maps between open sets of affine spaces, then we can summarize everything by saying that T is a functor from $\mathcal{O}$ to $\mathcal{O}$.

Exercise 30. *Let U_i be an open set of $\mathbb{A}_{\mathbb{R}}^{n_i}$ and $F\colon U_1 \to U_2$ and $G\colon U_2 \to U_3$ be smooth maps. Show that Theorem 1 is equivalent to saying that for any $p \in U_1$, we have*

$$J(GF)_p = JG_{F(p)} \cdot JF_p.$$

Here $\cdot$ is the matrix multiplication.

Exercise 31. *Let $F\colon U_1 \to U_2$ be a diffeomorphism, i.e., a smooth map with a smooth map $G\colon U_2 \to U_1$ such that $F \circ G = 1_{U_2}$ and $G \circ F = 1_{U_1}$. Using the fact that T is a functor to show that TF is a diffeomorphism and then T_pF is a linear equivalence for each point $p \in U_1$.*

Exercise 32. *Let F and G: $\mathbb{A}^n_{\mathbb{R}} \to \mathbb{A}^n_{\mathbb{R}}$ be polynomial maps. Show that, if F and G are inverse of each other, then $\det(JF)$ is a non-zero constant.*

Exercise 33. *Let U_1 and U_2 be non-empty open sets of affine space $\mathbb{A}$ and $p \in U_1 \subset U_2$. Then $T_pU_1 = T_pU_2 = T_p\mathbb{A}$. Show that, if ι: $U_1 \to U_2$ is the inclusion map, then $T_p\iota$ is the identity map on $T_p\mathbb{A}$.*

`Hint: you may assume that` $\mathbb{A} = \mathbb{A}^n_{\mathbb{R}}$.

Remark. If f is a calculus-smooth map, i.e., a smooth map between open sets of affine spaces, then $T_pf = T_p\hat{f}$. However, Tf and $T\hat{f}$ may not be equal because they might have different codomains.

Inverse Function Theorem and Implicit Function Theorem

Let F: $U_1 \to U_2$ be a calculus-smooth map and $p_1 \in U_1$ and $p_2 = f(p_1)$. By definition, a local version of F around point p_1 is a calculus-smooth map

$$\begin{array}{rccc} F_{loc}: & (U_1)_{loc} & \to & (U_2)_{loc} \\ & x & \mapsto & F(x) \end{array}$$

for which $(U_i)_{loc}$ is an open subset of U_i that contains p_i.

For the sanity of notation, we may simply rewrite F_{loc} as F. In general notation X_{loc} denotes an open neighborhood of a point in the topological space X.

We say that F is a diffeomorphism if there is a smooth map G: $U_2 \to U_1$ such that $FG = 1_{U_2}$ and $GF = 1_{U_1}$ In terms of categorical language we say that F is an isomorphism in the category $\mathcal{O}$ of calculus-smooth maps if it has an inverse G. Applying the tangent functor T, we conclude that F is a diffeomorphism $\implies$ TF is a diffeomorphism, in particular, the linear map $T_{p_1}F$ must be a bijection, thus a linear equivalence (or isomorphism).

In general 1_X denotes the identity map on X: $x \mapsto x$.

We say that F is a local diffeomorphism (or local isomorphism) around point p_1 if a local version of F around point p_1 is a diffeomorphism. It is clear that if F is a local diffeomorphism around point p_1, then $T_{p_1}f$ is a linear equivalence. In short, *local isomorphism* implies *infinitesimal isomorphism* and this is really an easy corollary of the Chain Rule. A nice fact is that the converse of the preceding statement — *Inverse Function Theorem*— is also true. In summary, we have

$$\textit{locally invertible} \iff \textit{infinitesimally invertible}.$$

Inverse Function Theorem is a very useful theorem because it enables us to draw interesting but difficult conclusion (local isomorphism) by verifying something much easier (infinitesimal isomorphism). One of its interesting corollaries is the Implicit Function Theorem. To describe it, we start with a set-up: Let F be a calculus-smooth map with p_1 in the domain of F and $p_2 = F(p_1)$. Let $\mathbb{A}'_1 = (\ker T_{p_1}F)_{aff}$ and $\mathbb{A}''_1 = (\ker T_{p_1}F)^{\perp}_{aff}$, then $\mathbb{A}'_1 \cap \mathbb{A}''_1 = \{p_1\}$ and affine spaces $\mathbb{A}'_1$ and $\mathbb{A}_1{}''$ are affine complements of each other in the affine space $\mathbb{A}_1$ — the affine space on which F is locally defined around point p_1.

If W is a linear subspace of V, we use $W^{\perp}$ to denote a linear complement of W in V, i.e., a linear subspace of V such that $V = W \oplus W^{\perp}$. Here $V = T_{p_1}\mathbb{A}_1$ and $W = \ker T_{p_1}F$. Recall that W_{aff} denotes the underlying affine space of the linear space W.

Recalling that the affine combination $(-1)p_1 + 1p + 1q$ with $p \in \mathbb{A}'_1$ and $q \in \mathbb{A}''_1$ is denoted by (p, q), we can now state **Implicit**

Function Theorem (geometric version): *Let F be a calculus-smooth map with $F(p_1) = p_2$. Suppose that $T_{p_1}F$ is surjective, then equation $F(x) = p_2$ implicitly defines a calculus-smooth map $f\colon (\mathbb{A}'_1)_{loc} \to (\mathbb{A}''_1)_{loc}$ such that*
1) $(\mathbb{A}'_1)_{loc} \cap (\mathbb{A}''_1)_{loc} = \{p_1\}$, 2) $f(p_1) = p_1$, 3) $F(p, f(p)) = p_2$.

Note that, in most places, for example here, Implicit Function Theorem is formulated in terms of coordinates, i.e., analytically. Note also that, *Inverse Function Theorem* $\iff$ *Implicit Function Theorem*.

Function f here is *germ-unique* in the sense that any two are equal locally, i.e., have a common local version. Note that, linear space $\ker T_{p_1}F$ ($(\ker T_{p_1}F)^\perp$ resp.) is the tangent space (the normal space resp.) of the graph of f at point p_1.
As an example, equation $x^2 + y^2 = 1$ defines an implicit function $q = f(p)$ around point $p_1 := (0, 1)$. In terms of coordinates $x = p - p_1$ and $y = q - p_1$, we have $y = \sqrt{1 - x^2}$. Therefore, $f(p) = p_1 + \sqrt{1 - (p - p_1)^2}$. In this case $F(x, y) = x^2 + y^2$, so $T_{p_1}F\colon T_{p_1}\mathbb{E}^2 \to T_{p_2}\mathbb{E}^1$ is the linear map that sends $(p_1, [u, v]^T])$ to $(1, [2v])$. Please draw a picture to visualize the theorem for this example. What about $p_1 = (-1, 0)$?

List of Symbols

f	a smooth real function on an open set U of the model Euclidean space $\mathbb{E}^n$
x^i	the i-the coordinate function on $\mathbb{E}^n$
$f_{\vec{u}}$	defined to be smooth function $u^i \frac{\partial f}{\partial x^i}$
$\mathrm{d}f_p$	the differential of f at point p, which is a cotangent vector of U at point p that sends tangent vector $(p, \vec{u})$ to real number $Jf_p \cdot \vec{u}$, i.e., $f_{\vec{u}}(p)$
$\mathrm{d}f$	the total differential of f, which is a cotangent-vector-valued smooth function on U
Jf_p	the Jacobean matrix of f at point p. By definition, it is the row matrix $\left[\frac{\partial f}{\partial x^1}(p), \ldots, \frac{\partial f}{\partial x^n}(p)\right]$.
$F : U_1 \to U_2$	a smooth map between open sets of model Euclidean spaces
F^i	the i-th component of F
$F_{\vec{u}}$	defined to be the vector-valued smooth function $u^i \frac{\partial F}{\partial x^i}$
JF_p	the Jacobean matrix of F at point p. By definition, its (i, j)-entry is $\frac{\partial F^i}{\partial x^j}(p)$
$T_pF : T_pU_1 \to T_{F(p)}U_2$	the tangent of F at point p. By definition, it is the linear map that sends $(p, \vec{u})$ to $(F(p), JF_p \cdot \vec{u})$
$TF : TU_1 \to TU_2$	the tangent of F, also a smooth map
$\mathrm{d}F_p$	the differential of F at point p, which is the linear map that sends tangent vector $(p, \vec{u})$ to column vector $JF_p \cdot \vec{u}$
$\mathcal{O}$	the category of smooth maps between open sets of affine spaces (not necessarily the model affine spaces)
$T : \mathcal{O} \to \mathcal{O}$	the tangent functor. It is an endofunctor on $\mathcal{O}$
X_{loc}	an open neighborhood of a point in the topological space X

§4 Curve

The purpose here is to introduce regular curves in Euclidean spaces, say in $\mathbb{E}^2$ or $\mathbb{E}^3$.

We all have an intuitive understanding about curves since childhood, but when it comes to studying them mathematically, we must have a good definition to start with. Any curve, assuming it is nice enough, admits a smooth parametrization, at least locally. So our first step is to introduce the notion of parametrized smooth curve. Let $n \geq 2$ be an integer.

Definition 2. *A parametrized smooth curve in $\mathbb{E}^n$ is just a smooth map α: $I \to \mathbb{E}^n$ from an open interval I into $\mathbb{E}^n$.*

The underlying curve of α — the image of α — is called the trace of α. For example,

$$\alpha(t) = (\cos t, \sin t, t)$$

is a parametrized smooth curve in $\mathbb{E}^3$ whose trace is a helix. Put it differently, we say α above is a smooth parametrization of the helix.

An interesting class of curves is the class of regular curves. We say that a parametrized smooth curve α: $I \to \mathbb{E}^n$ is regular (or 1-regular) if

The word "regular" and "1-regular" are interchangeable in this book.

$$\alpha'(t) := \lim_{h \to 0} \frac{1}{h}\left(\alpha(t+h) - \alpha(t)\right)$$

is a nonzero vector in $\mathbb{R}^n$ for each $t \in I$. Note that $\alpha' = \alpha_{[1]}$ if you still remember the notation $F_{\underline{u}}$. By definition, a regular curve in $\mathbb{E}^n$ is the trace of a regular parametrized smooth curve. For example, the helix above is a regular curve.

For any parametrized smooth curve α, we let

It is helpful to view a parametrized smooth curve as the position function of a moving particle and its trace the trajectory of the moving particle.

$$\dot{\alpha}(t) := (\alpha(t), \alpha'(t)), \quad \ddot{\alpha}(t) := (\alpha(t), \alpha''(t)).$$

Being referred to as the velocity vector and acceleration vector at time t respectively, $\dot{\alpha}(t)$ and $\ddot{\alpha}(t)$ are both tangent vectors of $\mathbb{E}^n$ at point $\alpha(t)$, and their easy versions are column vectors $\alpha'(t)$ and $\alpha''(t)$ respectively.

Remark 10. A curve in $\mathbb{E}^n$ is called a plane curve if it lies inside an affine subspace of dimension two. A line inside $\mathbb{E}^n$ is a regular plane curve. Ellipses, parabolas, and one branch of any hyperbola are all regular plane curves.

Remark 11. A regular curve can have a non regular parametrization or even a non smooth parametrization. For example, the diagonal line $\Delta = \{(x, x) \mid x \in \mathbb{R}\}$ in $\mathbb{E}^2$ is a regular curve, but it has a non regular parametrization $\alpha(t) = (t^3, t^3)$ as well as a non smooth parametrization $\beta(t) = (\sqrt[3]{t}, \sqrt[3]{t})$.

The graph of a smooth map of one-variable is always a regular curve. To see this, we let $F\colon I \to \mathbb{E}^{n-1}$ be a smooth map, then it is easy to see that its graph

$$\Gamma_F = \{(t, F(t)) \mid t \in I\} \subset \mathbb{E}^n$$

is a regular curve in $\mathbb{E}^n$.

A nice fact about a regular parametrized smooth curve α is the existence of its tangent line: the tangent line at (the point parametrized by) t_0 is the line

$$\{\alpha(t_0) + (t - t_0)\alpha'(t_0) \mid t \in \mathbb{R}\}$$

together with the distinguished point $\alpha(t_0)$, i.e., the one-dimensional linear subspace $\operatorname{span}\{\dot{\alpha}(t_0)\}$ of $T_{\alpha(t_0)}\mathbb{E}^n$. If the trace C of this regular parametrized smooth curve α does not cross itself at point $p := \alpha(t_0)$, then this pointed line is called the tangent line of the regular curve C at point p. In case p is a self crossing point of the regular curve C, the tangent line of C at p is not defined.

The graph of a smooth function of one-variable never crosses itself.

A nice fact about regular curves is that they always admit an arc-length parametrization, i.e., mileage parametrization.

Theorem 2 (Arc-length Parametrization)**.** *Let C be a regular curve, then C admits an arc-length parametrization, i.e., C is the trace of a regular parameterized smooth curve α such that $|\alpha'| = 1$ everywhere.*

Proof. Since C is a regular curve, it must be the trace of a parameterized smooth curve $\beta\colon J \to \mathbb{E}^n$ such that $|\beta'| \neq 0$ everywhere. Let $s = g(t) := \int_{t_0}^{t} |\beta'(\tau)|\, d\tau$. Since $g' = |\beta'| > 0$ on J, g is smooth and has a smooth inverse $t = f(s)$ defined on an open interval I. Then Chain Rule implies that

$$f'(s) = \frac{1}{g'(f(s))} = \frac{1}{|\beta'(f(s))|}.$$

Let $\alpha = \beta \circ f$. Then $|\alpha'(s)| = |\beta'(f(s))| \cdot f'(s) = 1$. So α is an arc-length parameterization for regular curve C. □

Let $J = (c, d)$ with $c < d$. For any integer $n > \frac{2}{d-c}$, since g is a continuous and increasing real function on J, the closed interval $[g(c + \frac{1}{n}), g(d - \frac{1}{n})]$ must be a subset of the image $\operatorname{Im} g$. Then $\operatorname{Im} g = I := (a, b)$ where $a = \lim_{n\to\infty} g(c + \frac{1}{n})$ and $b = \lim_{n\to\infty} g(d - \frac{1}{n})$.

Remark 12. One could view a parametrized smooth curve $\beta(t)$ as the position function of a moving particle and the underlying curve as the trajectory. With this in mind, t is the time and $s(t)$ is the distance the particle travels from time t_0 to time t, and the word "regular" means the velocity is never zero.

Frenet-Serret frame

Suppose that $\alpha\colon I \to \mathbb{E}^2$ is a regular parameterized smooth curve. By a reparametrization if necessary, we can assume that $|\alpha'| = 1$ on I.

At $s \in I$, we have the unit tangent vector $\mathbf{T}(s) := \dot{\alpha}(s)$. After the counterclockwise rotation of $90°$, $\mathbf{T}(s)$ is turned into a unit normal vector $\mathbf{N}(s)$. Then $(\mathbf{T}(s), \mathbf{N}(s))$ is an orthonormal basis of $T_{\alpha(s)}\mathbb{E}^2$, which shall be refereed to as the orthonormal frame of α at s. Therefore, as s moves inside I, we get a canonical moving frame $(\mathbf{T}(s), \mathbf{N}(s))$ along the regular parametrized smooth curve α.

We assume the standard orientation on $\mathbb{E}^2$, so the orthonormal basis thus obtained is an oriented basis as well.

In general we say that a parameterized smooth curve $\alpha\colon I \to \mathbb{E}^n$ is k-regular if

$$\alpha'(t), \alpha''(t), \ldots, \alpha^{(k)}(t)$$

are linearly independent for any t in the domain of α.

Suppose that $\alpha\colon I \to \mathbb{E}^3$ is a regular parameterized smooth curve such that $|\alpha'| = 1$ on I. Now the normal space is two-dimensional, to get a canonical moving frame, we must further assume that $|\alpha''| \neq 0$ on I, i.e, α is 2-regular. As before, at $s \in I$, we have the unit tangent vector $\mathbf{T}(s) := \dot{\alpha}(s)$. Next, we introduce

Because $\alpha' \cdot \alpha'' = 0$.

$$\mathbf{N} := \frac{\ddot{\alpha}}{|\ddot{\alpha}|}, \quad \mathbf{B} := \mathbf{T} \times \mathbf{N}.$$

Note that, $(\mathbf{T}(s), \mathbf{N}(s), \mathbf{B}(s))$ is an orthonormal basis of $T_{\alpha(s)}\mathbb{E}^3$, which shall be refereed to as the orthonormal frame of α at s. Therefore, as s moves inside I, we get a canonical moving frame $(\mathbf{T}(s), \mathbf{N}(s), \mathbf{B}(s))$ along the 2-regular parametrized smooth curve α, i.e., the Frenet-Serret frame. Here $\mathbf{N}(s)$ is called the unit normal vector at s and $\mathbf{B}(s)$ is called the unit binormal vector at s. The two orthonormal vectors $\mathbf{T}(s)$, $\mathbf{N}(s)$ span the osculating plane of the smooth parametrized curve at s.

Writing $\mathbf{T}(s) = (\alpha(s), \vec{t}(s))$, $\mathbf{N}(s) = (\alpha(s), \vec{n}(s))$ and $\mathbf{B}(s) = (\alpha(s), \vec{b}(s))$, we can introduce

$$\dot{\mathbf{T}}(s) = (\alpha(s), \vec{t}\,'(s)), \quad \dot{\mathbf{N}}(s) = (\alpha(s), \vec{n}\,'(s)), \quad \dot{\mathbf{B}}(s) = (\alpha(s), \vec{b}\,'(s)).$$

A basic fact about the Frenet-Serret frame is the Frenet-Serret formulas: there are smooth functions $\kappa > 0$ and τ on I such that

$$\begin{cases} \dot{\mathbf{T}} &= & & \kappa\mathbf{N} & \\ \dot{\mathbf{N}} &= & -\kappa\mathbf{T} & + & \tau\mathbf{B} \\ \dot{\mathbf{B}} &= & & -\tau\mathbf{N} & \end{cases} \tag{6}$$

Proof. Let $A = [\vec{t}, \vec{n}, \vec{b}]$. Since $\vec{t}(s), \vec{n}(s), \vec{b}(s)$ are orthonormal, $A(s)$ is an orthogonal matrix, i.e., $A^T A = I$. Let $\omega = A^T A'$, then $\omega(s)$ is an anti-symmetric matrix for any s, so we can write

That is because $\omega + \omega^T = (A^T A)' = I' = 0$.

$$\omega = \begin{bmatrix} 0 & -\kappa & -\lambda \\ \kappa & 0 & -\tau \\ \lambda & \tau & 0 \end{bmatrix}.$$

In view of the fact that $\omega = A^T A'$ and $A^{-1} = A^T$, we have $A' = A\omega$, i.e.,

$$[\vec{t}', \vec{n}', \vec{b}'] = [\kappa\vec{n} + \lambda\vec{b}, -\kappa\vec{t} + \tau\vec{b}, -\lambda\vec{t} - \tau\vec{n}].$$

In particular $\vec{t}' = \kappa\vec{n} + \lambda\vec{b}$. By the definition of $\vec{n}$, $\vec{t}' = |\vec{t}'|\vec{n}$, so $\lambda = 0$ and $\kappa = |\vec{t}'|$. The rest is clear. □

Remark 13. Functions κ and τ are respectively called the curvature functions and the torsion function of the 2-regular parametrized curve α in $\mathbb{E}^3$. For the 2-regular plane curves, such as conics, the torsion function is identically zero. If the trace C of α does not cross itself, then κ ($|\tau|$ resp.) is called the curvature (torsion resp.) function of the 2-regular curve C.

Remark 14. The two smooth functions $\kappa > 0$ and τ determine a 2-regular smooth arc-length parametrized space curve uniquely up to a rigid motion of $\mathbb{E}^3$ and a shifting of the arc-length parameter s. This is a consequence of the existence and uniqueness theorem in ODE.

Remark 15. For 1-regular smooth arc-length parametrized curve α in $\mathbb{E}^2$, we would have

$$\begin{cases} \dot{\mathbf{T}} &= \kappa\mathbf{N} \\ \dot{\mathbf{N}} &= -\kappa\mathbf{T} \end{cases} \tag{7}$$

However, κ — curvature function of α — is not necessarily positive. If the trace C of α does not cross itself, then $|\kappa|$ is called curvature function of the 1-regular curve C.

Exercise 34. *Let C be the regular plane curve $y = f(x)$, where f is a smooth function defined on an open interval I. Show that the curvature of C at point $(x, f(x))$ is*

$$\frac{|f''(x)|}{(1 + (f'(x))^2)^{3/2}}.$$

Please use this formula to compute the curvature function on ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

`Hint:` $\kappa = \left|\frac{\mathrm{d}^2}{\mathrm{d}s^2}\alpha(t(s))\right|$ `and` $\frac{\mathrm{d}s}{\mathrm{d}t} = |\alpha'|$. `You need chain rule for help.`

Exercise 35. *Let C be a regular curve in $\mathbb{E}^3$ with $\alpha\colon \mathbb{R} \to \mathbb{E}^3$ being a regular smooth parametrization of C. Suppose that $\alpha' \times \alpha'' = \vec{0}$ on $\mathbb{R}$, show that C is a line.*

`Hint: Assume the arclength parametrization for` α `and then conclude that` $\alpha'' = \vec{0}$.

Exercise 36. *Let C be a 2-regular curve in $\mathbb{E}^3$ with $\alpha\colon I \to \mathbb{E}^3$ being a 2-regular smooth parametrization of C. Suppose that $(\alpha' \times \alpha'') \cdot \alpha''' = 0$ on I, show that C is a plane curve, i.e., curve C is a subset of an affine plane.*

`Hint: Assume the arclength parametrization for` α `and then applying Frenet-Serret formulas.`

Local canonical representation of 2-regular space curves

The purpose here is to examine the local structure of parametrized 2-regular space curves.

Let $\alpha\colon I \to \mathbb{E}^3$ be an arclength parameterized 2-regular smooth curve, and $0 \in I$. Recall that we write $\mathbf{T}(s) = (\alpha(s), \vec{t}(s))$, $\mathbf{N}(s) = (\alpha(s), \vec{n}(s))$ and $\mathbf{B}(s) = (\alpha(s), \vec{b}(s))$. Then

$$\begin{aligned}\alpha(s) &= \alpha(0) + \alpha'(0)s + \frac{1}{2}\alpha''(0)s^2 + \frac{1}{6}\alpha'''(0)s^3 + \cdots \\ &= \alpha(0) + \vec{t}(0)s + \frac{1}{2}\kappa(0)\vec{n}(0)s^2 + \frac{1}{6}\alpha'''(0)s^3 + \cdots\end{aligned}$$

Since $\alpha''' = \kappa'\vec{n} + \kappa\vec{n}' = \kappa'\vec{n} + \kappa(-\kappa\vec{t} + \tau\vec{b}) = -\kappa^2\vec{t} + \kappa'\vec{n} + \kappa\tau\vec{b}$. Up to a rigid motion, we may assume $\alpha(0) = 0$ and $(\vec{t}(0), \vec{n}(0), \vec{b}(0)) = (\vec{i}, \vec{j}, \vec{k})$, then, with $\kappa(0)$ rewritten as κ, $\kappa'(0)$ rewritten as κ', and $\tau(0)$ rewritten as τ, we have

$$\begin{aligned}\alpha(s) &= 0 + \vec{i}s + \frac{1}{2}\kappa\vec{j}s^2 + \frac{1}{6}(-\kappa^2\vec{i} + \kappa'\vec{j} + \kappa\tau\vec{k})s^3 + \cdots \\ &= (s - \frac{\kappa^2}{6}s^3, \frac{\kappa}{2}s^2 + \frac{\kappa'}{6}s^3, \frac{\kappa\tau}{6}s^3) + o(s^3)\end{aligned}$$

Note that $x = s - \frac{\kappa^2}{6}s^3 + o(s^3)$ is invertible near $s = 0$, due to the fact that $x'(0) = 1 \neq 0$. In fact, it is not hard to see that $s = x + \frac{\kappa^2}{6}x^3 + o(x^3)$. Then

$$\alpha(s(x)) = (x, \frac{\kappa}{2}x^2 + \frac{\kappa'}{6}x^3 + o(x^3), \frac{\kappa\tau}{6}x^3 + o(x^3)).$$

It is then clear that, up to a rigid motion, the piece of curve near $s = 0$ is just the graph of map $F\colon I \to \mathbb{E}^2$ near $x = 0$, where

$$F(x) = (\frac{\kappa}{2}x^2 + \frac{\kappa'}{6}x^3, \frac{\kappa\tau}{6}x^3) + o(x^3). \tag{8}$$

This representation (i.e., $x \mapsto (x, F(x))$ with F in the above form) of a 2-regular parametrized curve α is called the local canonical representation of α around $s = 0$.

Remark 16. The function F above is a smooth map from an open set of the tangent line to the normal space.

Remark 17. Given a formal real function

$$y = b_2 x + b_3 x^2 + \cdots$$

where x is a formal variable and b_2, b_3, ..., are real numbers and $b_2 \neq 0$. Then the formal inverse can be derived iteratively as follows:

$$\begin{aligned}
x &= b_2^{-1}(y - \sum_{k>2} b_k x^{k-1}) \\
&= b_2^{-1} y - b_2^{-1} b_3 x^2 - b_2^{-1} \sum_{k>3} b_k x^{k-1} \\
&= b_2^{-1} y - b_2^{-1} b_3 x^2|_{x=b_2^{-1}y+\cdots} - b_2^{-1} \sum_{k>3} b_k x^k \\
&= b_2^{-1} y - (b_2^{-1})^3 b_3 y^2 + \cdots \\
&= b_2^{-1} y - b_2^{-1}(b_3 x^2 + b_4 x^3)|_{x=b_2^{-1}y-(b_2^{-1})^3 b_3 y^2+\cdots} - b_2^{-1} \sum_{k>4} b_k x^{k-1} \\
&= b_2^{-1} y - (b_2^{-1})^3 b_3 y^2 + \left(2(b_2^{-1})^5 b_3^2 - (b_2^{-1})^4 b_4\right) y^3 + \cdots \\
&\vdots
\end{aligned}$$

In fact each term in this formula can be expressed in terms of certain easily remembered Feynman diagrams at the tree level. Although it is not hard to prove this formula directly, you may find an interesting way to do it by reading this paper.

Remark 18. By definition, a circle of radius R has curvature equal to $\frac{1}{R}$ everywhere. For the circle $x^2 + (y - R)^2 = R^2$, its tangent parabola (a.k.a. 2nd order approximation) at point $(0, 0)$ is

In the original equation, y^2 is deleted because its order is higher than quadratic.

$$x^2 - 2Ry = 0, \quad \text{or} \quad y = \frac{1}{2}\frac{1}{R}x^2.$$

Therefore, the curvature is an information that can be extracted from the tangent parabola.

In general, the curvature of a regular parameterized smooth curve α in $\mathbb{E}^N$ at a parameter point t_0 is an information that can be extracted from the tangent parabola:

$$\alpha(t_0) + \alpha'(t_0)(t - t_0) + \frac{1}{2}\alpha''(t_0)(t - t_0)^2$$

at the parameter point t_0. That is because this tangent parabola is also the tangent parabola of the tangent circle whose curvature is by definition the curvature of α at the the parameter point t_0.

Exercise 37. *Let $\alpha\colon I \to \mathbb{E}^N$ be a regular parameterized smooth curve in $\mathbb{E}^N$ ($N \geq 2$) and $0 \in I$. Let $\vec{v} = \alpha'(0)$, $\vec{a} = \alpha''(0)$, and κ be the curvature of α at 0. Please derive a formula that expresses the curvature κ as a function of the velocity $\vec{v}$ and the acceleration $\vec{a}$.*

`Hint: Try to turn Taylor expansion` $p + t\vec{v} + \frac{1}{2}t^2\vec{a} + o(t^2)$ `(where` $p = \alpha(0)$`) into the canonical form:`

$$p + s\vec{t} + \frac{1}{2}s^2\kappa\vec{n} + o(s^2)$$

`where` s `is a smooth function of` t`,` $\vec{t}$ `and` $\vec{n}$ `are orthonormal vectors in` $\mathbb{R}^N$`. Then` $\kappa = 2 \cdot \lim_{s\to 0} \frac{|\frac{1}{2}s^2\kappa\vec{n}+o(s^2)|}{|s\vec{t}+o(s^2)|^2}$`. Anwser:` $\frac{\sqrt{|\vec{a}|^2\,|\vec{v}|^2-(\vec{a}\cdot\vec{v})^2}}{|\vec{v}|^3}$`. One can guess this answer by a dimensional analysis plus the observations: 1)` $\kappa = 0$ `if` $\vec{a}$ `and` $\vec{v}$ `are parallel, 2)` $\kappa = \frac{|\vec{a}|}{|\vec{v}|^2}$ `for the circular motion in which` $\vec{a}$ `and` $\vec{v}$ `are perpendicular.`

Remark 19. For the regular plane curve, up to a rigid motion, the piece of curve near $s = s_0$ is just the graph of map $f\colon I \to \mathbb{R}$ near $x = 0$, where

$$f(x) = \frac{\kappa}{2}x^2 + o(x^2). \tag{9}$$

Exercise 38. *Continuing Exercise 34, find the local canonical representation of the curve $2x^2 - y^2 = 1$ at point $(1,1)$.*

Example 7. *Find the curvature and torsion of the graph of $F(x) = (\cos x, (\sin x)^3)$ at point $(0,1,0)$.*

Solution. Let $\beta(x) = (x, F(x))$. Around $x = 0$, we have the Taylor expansion

$$\begin{aligned}
\beta(x) &= (x, 1 - \frac{x^2}{2} + o(x^3), x^3 + o(x^3)) \\
&= (x, -\frac{x^2}{2} + o(x^3), x^3 + o(x^3)) \quad \text{modulo a translation} \\
&= (x, \frac{x^2}{2} + o(x^3), -x^3 + o(x^3)) \quad \text{module a rotation of } 180^\circ \text{ of the } yz \text{ plane}
\end{aligned}$$

So, up to a rigid motion, locally the piece of curve near $x = 0$ is just the graph of the map $F(x) = (\frac{x^2}{2}, -x^3) + o(x^3)$ near $x = 0$. By comparing with Eq. (8), this implies the curvature and torsion at $x = 0$ are 1 and $|-6| = 6$ respectively. ◀

Example 8. *Show that the cusp $x^3 = y^2$ is not a regular curve.*

Solution. Proof by contradiction. Suppose that it is regular, i.e., there is a regular parametrized smooth curve $\alpha(t) = (x(t), y(t))$ such that

$$x(t)^3 = y(t)^2. \tag{10}$$

We may assume that $\alpha(0) = (0,0)$, then we have

$$x(t) = at + o(t), \quad y(t) = bt + o(t)$$

for some real numbers a and b, so Eq. (10) becomes

$$(at + o(t))^3 = (bt + o(t))^2.$$

Dividing this identity by t^2 and then taking limit as $t \to 0$, we arrive at identity $0 = b^2$, so $b = 0$, hence we have

$$x(t) = at + o(t), \quad y(t) = ct^2 + o(t^2).$$

for some real number c. Plugging these into Eq. (10), we arrive at identity

$$(at + o(t))^3 = (ct^2 + o(t^2))^2.$$

Dividing this identity by t^3 and then taking limit as $t \to 0$, we get $a^3 = 0$, so $a = 0$.

Then $\alpha'(0) = [a, b]^T = \vec{0}$, contradicting to the assumption that α is a regular parametrized smooth curve. So the cusp $x^3 = y^2$ cannot be a regular curve. ◀

List of Symbols

I	a non-empty open interval
α	a smooth map from I to $\mathbb{E}^n$, which can be viewed as either the position function of a moving particle or a smooth parametrized curve
α'	the derivative of α, which is a $\mathbb{R}^n$-valued smooth function on I
$\dot{\alpha}$	defined to be (α, α') and is called the velocity function, a tangent-vector-valued function on I
$\ddot{\alpha}$	defined to be (α, α'') and is called the acceleration function, a tangent-vector-valued function on I
Γ_F	the graph of F, a subset of $\mathbb{E}^n$ for $F\colon I \to \mathbb{E}^{n-1}$
C	a regular curve, i.e., a one dimensional geometric object that admits a regular smooth parametrization α
$(\mathbf{T}, \mathbf{N}, \mathbf{B})$	the Frenet-Serret (orthonormal) frame function of an oriented 2-regular smooth parametrized space curve α. It is a smooth function on I
$(\vec{t}, \vec{n}, \vec{b})$	the "easy version" of the Frenet-Serret frame function: $\mathbf{T} = (\alpha, \vec{t})$, etc.
$\dot{\mathbf{T}}, \dot{\mathbf{N}}, \dot{\mathbf{B}}$	defined to be $(\alpha, \vec{t}')$, $(\alpha, \vec{n}')$, $(\alpha, \vec{b}')$ respectively
κ	the curvature function of a 1-regular curve, either for parametrized (a function on I) or for unparameterized (a function on C, multivalued at each self-crossing point)
τ	the torsion function of a 2-regular space curve, either for parametrized (a function on I) or for unparameterized (a function on C, multivalued at each self-crossing point)

§5 Manifold

The Chain Rule stated in Theorem 1 says that there is a canonical endofunctor T on the category of smooth maps between open sets of affine spaces. It turns out that this theorem is valid for the larger category of smooth maps between spaces that locally look like open sets of affine spaces, i.e., the category of smooth maps between manifolds. For technical convenience, it is better to start with the even bigger category of smooth maps between subsets of affine spaces.

Smooth Map

In general, if A and B are two subsets of affine spaces, we say that A looks like B, or more formally A is diffeomorphic to B, if there is a *bijective* map $f\colon A \to B$ such that both f and f^{-1} are both smooth. We shall write $A \cong B$ if such a map f — referred to as a **diffeomorphism** or **smooth equivalence** from A onto B — exists; furthermore, in case $a \in A$, $b \in B$ and $f(a) = b$, we shall write $(A, a) \cong (B, b)$ — a diffeomorphism between *pointed spaces*. We shall write $(A, a) \cong_{locally} (B, b)$ if $(A_{loc}, a) \cong (B_{loc}, b)$. We shall see later that both $\cong$ and $\cong_{locally}$ are equivalence relations.

In general (X_{loc}, x) is a pointed space with X_{loc} being an open neighborhood of the point x in the topological space X.

Let me explain the meaning of the phrase "map $f\colon A \to B$ is smooth". Suppose that A is a subset of affine space $\mathbb{A}$ and B is a subset of affine space $\mathbb{B}$, then we say that f is smooth if there is a calculus-smooth map $\tilde{f}\colon U \to V$ that extends f in the following sense: U is an open neighborhood of A in $\mathbb{A}$, V is an open neighborhood of B in $\mathbb{B}$, and $\tilde{f}(a) = f(a)$ for any $a \in A$. This can be summarized in the commutative square

Recall that a smooth map between open sets of affine spaces is called a **calculus-smooth map** in this book.

We use the Greek letter ι exclusively for any inclusion map.

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \downarrow{\scriptstyle\iota} & & \downarrow{\scriptstyle\iota} \\ U & \xrightarrow{\tilde{f}} & V \end{array} \tag{11}$$

In short, we say a map is smooth if it can be extended to a calculus-smooth map. Note that if f is a map between opens sets of affine spaces, then f is a calculus-smooth map if and only if it is smooth in this (generalized) sense.

Example 9. *Let A and B be subsets of affine spaces. Then the projection $A \times B \to A$ is a smooth map. Indeed, one can take $\tilde{f}$ to be the projection map $p_{\mathbb{A}}$: $\mathbb{A} \times \mathbb{B} \to \mathbb{A}$.*

Example 10. *Any inclusion map ι: $A \to B$ is a smooth map, in particular the identity map on a subset A of an affine space is a smooth map. Indeed, one can take $\tilde{f}$ to be the identity map $1_{\mathbb{A}}$ on $\mathbb{A}$ provided that A, B are subsets of $\mathbb{A}$.*

Exercise 39. *Show that the composition of two smooth maps is a smooth map and then conclude that smooth maps between subsets of affine spaces form a category.*

`Hint: Start with`

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dd, "\iota"] & B \arrow[r, "g"] \arrow[d, "\iota"] & C \arrow[d, "\iota"] \\ & U \arrow[r, "\tilde{g}"] \arrow[d, "\iota"] & \mathbb{C} \\ V \arrow[r, "\tilde{f}"] & \mathbb{B} & \end{tikzcd}$$

`then one can extend` gf `to a calculus-smooth map on the open neighborhood` $\tilde{f}^{-1}(U)$ `of` A`.`

Exercise 40. *For $i = 1$ and 2, we let f_i: $A \to B_i$ be a set map. Show that $f_1 \times f_2$: $A \to B_1 \times B_2$ is smooth if and only if the components f_1 and f_2 are all smooth.*

Exercise 41. *Assume f: $A \to B$ is a smooth map between subsets of affine spaces. Show that $\Gamma_f \cong A$, with the map $(a, f(a)) \to a$ being a diffeomorphism. In other words, for a smooth map, its graph looks like its domain.*

`Hint: use result from the previous exercise.`

Suppose that A is a subset of affine space $\mathbb{A}$ and B is a subset of affine space $\mathbb{B}$, and f: $A \to B$ is a set map. Let $\bar{f}$: $A \to \operatorname{Im} f$ and $\hat{f}$: $A \to \mathbb{B}$ denote the set maps such that $\bar{f}(a) = f(a) = \hat{f}(a)$ for any $a \in A$. We summarize this by commutative diagram

$$\begin{tikzcd} & \mathbb{B} \\ A \arrow[ur, "\hat{f}"] \arrow[r, "f"] \arrow[dr, "\bar{f}"'] & B \arrow[u, "\iota"'] \\ & \operatorname{Im} f \arrow[u, "\iota"'] \end{tikzcd}$$

We shall call $\hat{f}$ and $\bar{f}$ the hat version and the bar version of f respectively. We say that f is an imbedding if $\bar{f}$ is an diffeomorphism.

For the sanity of notations or sloppiness, in many textbooks, these three maps are all denoted by f. This common practice causes no confusions to anyone who can read from context.

Exercise 42. *For the three essentially the same maps f, $\bar{f}$ and $\hat{f}$, show that if one is smooth then the other two are smooth.*

Remark 20. If $f\colon A \to B$ is a map, then f is smooth if and only if f is smooth locally (i.e., any point of A admits an open neighborhood U in A such that $f \circ \iota\colon U \to B$ is smooth.) One direction is easy, the other direction requires help from the smooth partitions of unity.

Velocity Space

Let X be a subset of an affine space $\mathbb{A}$ and p be a point of X. By definition, a **velocity vector** of X at point p is $\dot{\alpha}(0)$ where α is a smooth map from an open interval containing 0 into X such that $\alpha(0) = p$. Recall that $\dot{\alpha}(0) = (\alpha(0), \alpha'(0))$ and

This map α is called a smooth parametrized curve in X

$$\alpha'(0) := \lim_{t \to 0} \frac{\alpha(t) - \alpha(0)}{t} = \hat{\alpha}'(0).$$

So a velocity vector of X is a tangent vector of the ambient affine space $\mathbb{A}$.

The set of all velocity vectors of X at point p, denoted by $V_p X$, is called the **velocity space** of X at point p. Although a subset of $T_p\mathbb{A}$, $V_p X$ may not be a linear subspace of $T_p\mathbb{A}$.

Example 11. *Let $X \subset \mathbb{E}^2$ be the solution set of equation $xy = 0$, i.e., the union of x-axis with the y-axis. Then $V_{(0,0)}X$ is not a linear subspace of $T_{(0,0)}\mathbb{E}^2$.*

Proof. Let $\alpha(t) = (x(t), y(t))$ be a smooth parametrized curve in X with $\alpha(0) = (0,0)$. Write

$$\alpha'(0) = \begin{bmatrix} a \\ b \end{bmatrix}.$$

Then

$$x(t) = at + o(t), \quad y(t) = bt + o(t).$$

Since $\mathrm{Im}(\alpha(t)) \subset X$, we have $x(t)y(t) = 0$, i.e.,

$$abt^2 + o(t^2) = 0.$$

Divide by t^2 and then take the limit as $t \to 0$, we have $ab = 0$. So we must have either $a = 0$ or $b = 0$.

On the other hand, if α is the smooth parametrized curve in X of the form $t \to (at, 0)$ or $t \to (0, bt)$, then $\alpha'(0)$ is

$$\begin{bmatrix} a \\ 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 \\ b \end{bmatrix}.$$

Therefore, the set of velocity vectors of X at point $(0,0)$ is the union of

$$\operatorname{span}\left\{\left((0,0), \begin{bmatrix}1\\0\end{bmatrix}\right)\right\}$$

with

$$\operatorname{span}\left\{\left((0,0), \begin{bmatrix}0\\1\end{bmatrix}\right)\right\}.$$

Then $V_{(0,0)}X$ is not a linear subspace of $T_{(0,0)}\mathbb{E}^2$. □

It is clear that if X_{loc} is an open neighborhood a point p in X, then, $V_p X_{loc} = V_p X$. It is also easy to see that, for open sets of affine spaces, velocity vector and tangent vector are the same concept, and velocity space and tangent space are the same linear space.

The set of all velocity vectors of X, denoted by VX, is called the **total velocity space** of X. Since $V_p X$ is a subset $T_p\mathbb{A}$, VX is a subset of $T\mathbb{A}$.

Velocity Functor

The assignment of VX to X extends to an endo-functor V on the category of smooth maps between subsets of affine spaces. This means that, to each morphism $f\colon X \to Y$ in this category, there assigns a morphism $Vf\colon VX \to VY$ in this category such that V preserves identity and composition, that is, $V1_X = 1_{VX}$ for any object X and $V(f \circ g) = Vf \circ Vg$ for any two composable morphisms f and g.

The morphism Vf is defined naturally:

$$\dot{\alpha}(0) \mapsto \dot{\overline{f_*\alpha}}(0).$$

I.e., it sends the velocity vector represented by a smooth parametrised curve α in X to the velocity vector represented by $f_*\alpha := f \circ \alpha$, which is a smooth parametrised curve in Y.

It is obvious that V preserves identity and composition if we can check that Vf is well-defined (i.e., $\dot{\overline{f_*\alpha}}(0)$ depends only on f and $\dot{\alpha}(0)$) and is smooth as well. The guided exercise below provides such a checking.

Exercise 43. *Let $\tilde{f}\colon \mathbb{A}_{loc} \to \mathbb{B}$ be a calculus-smooth extension of f.*

1) $\dot{\overline{f_*\alpha}}(0) = T_{\alpha(0)}\tilde{f}(\dot{\alpha}(0))$ and then conclude that $\dot{\overline{f_*\alpha}}(0)$ depends only on f and $\dot{\alpha}(0)$. `Hint:` $(f_*\alpha)'(0) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}\tilde{f}(\alpha(t))$.

2) $Vf = Tf$ if f is a calculus-smooth map.

3) $T\tilde{f}$ is a calculus-smooth extension of Vf and then conclude that Vf is smooth.

For any point p in X, it is clear that Vf maps V_pX into $V_{f(p)}Y$, so we have a set map $V_pf\colon V_pX \to V_{f(p)}Y$. Since V_pX may not be a linear space, in general, V_pf is not a linear map. However, we claim that V_pf is **pre-linear**, i.e., V_pf respects linear combination in this weak sense: for any two vectors $\underline{u}$ and $\underline{v}$ in V_pX and any two real numbers a and b, if $a\underline{u} + b\underline{v}$ is in V_pX, then

We shall call V_pX a **pre-linear space** in the sense that it is a subset of a linear space that contains the zero vector.

$$V_pf(a\underline{u} + b\underline{v}) = aV_pf(\underline{u}) + bV_pf(\underline{v}).$$

That is because $V_pf(\underline{u}) = T_p\tilde{f}(\underline{u})$ for any calculus-smooth extension $\tilde{f}$ of f.

Exercise 44. *Show that, if f is a diffeomorphism, then V_pf is a pre-linear equivalence, i.e., both bijective and pre-linear.* `Hint:` V `is a functor.`

Exercise 45. *Let $p \in X_{loc}$ and $\iota\colon X_{loc} \to X$ be the inclusion. Show that $V_p\iota = 1_{V_pX}$.*

Exercise 46. *Let X be a (non-empty) subset of affine space $\mathbb{A}$. Show that the projection map $\nu\colon VX \to X$ that sends velocity vector $\dot{\alpha}(0)$ to point $\alpha(0)$ is a smooth map.* `Hint: Consider the commutative square`

$$\begin{array}{ccc} VX & \xrightarrow{\nu} & X \\ \downarrow{\scriptstyle\iota} & & \downarrow{\scriptstyle\iota} \\ V\mathbb{A} & \xrightarrow{\nu} & \mathbb{A} \end{array}$$

We shall call Vf the velocity map of f and V_pf the velocity map of f at point p. For convenience, we introduce the **easy version** of the velocity space of X at point p, which is denoted by VX_p. By definition,

$$VX_p := \{\alpha'(0) \mid \dot{\alpha}(0) \in V_pX\}.$$

Then we have the easy version of the velocity map of f at point p:

$$Vf_p : VX_p \to VY_{f(p)}$$

which sends $\alpha'(0)$ to $(f_*\alpha)'(0)$.

Exercise 47. *Let X and ∂X be the following subsets of $\mathbb{E}^2$:*

$$X = \{(x,y) \in \mathbb{E}^2 \mid y \geq 0\}, \quad \partial X = \{(x,y) \in X \mid y = 0\}.$$

Let $p \in \partial X$, q be a point of X away from ∂X.

1) Describe VX_p and VX_q as subsets of the linear space $\mathbb{R}^2$ and then conclude that there is no pre-linear equivalence between them.

2) Show that there is no diffeomorphism from X onto X that sends q to p.

Exercise 48. *Let X and Y be the following subsets of $\mathbb{E}^2$:*

$$X = \{(x,y) \mid y \geq 0\}, \quad Y = \{(x,y) \mid x \geq 0, y \geq 0\}$$

Let $p = (0, 0)$.

1) Describe both VX_p *and* VY_p *as subsets of* $\mathbb{R}^2$.

2) Show that there is no diffeomorphism from X *onto* Y.

3) Show that a square and a disk are never diffeomorphic to each other.

Definition of Manifold

We have observed the following facts:

(a) on the category of calculus-smooth maps, we have the tangent endofunctor $T\colon f \mapsto Tf$. Moreover, Tf is fiberwise linear, which means that $T_p f$ is linear for each point p in the domain of f.

(b) on the category of smooth maps between subsets of affine spaces, we have the velocity endofunctor $V\colon f \mapsto Vf$. Moreover, Vf is fiberwise pre-linear, which means that $V_p f$ is pre-linear for each point p in the domain of f.

(c) V is an extension of T, that is, $Vf = Tf$ for any calculus-smooth map f.

While the extension of functor T to functor V is natural, the fiberwise linearity of the image morphism Vf is lost in the extension, that is because $V_p f$ may not be linear. However, there is an intermediate category which is closed under the velocity functor V, moreover, the fiberwise linearity of the image morphism is kept. Because of that, when we work with this intermediate category, we prefer to call the velocity functor the tangent functor, velocity vector the tangent vector, velocity space the tangent space, and so on.

The objects in this intermediate category are called **manifolds**. In particular this means that open sets of affine spaces are manifolds.

Definition 3. *A subset* X *of an affine space is called an* n*-manifold if any point of* X *has an open neighborhood diffeomorphic to an open set of* $\mathbb{E}^n$. *Equivalently,*

$$\forall p \in X, \quad (X, p) \cong_{locally} (\mathbb{E}^n, 0)$$

In this definition, the manifold X is said to be modelled on $\mathbb{E}^n$, or we say that $\mathbb{E}^n$ is a local model for the n-manifold X. Note that, the empty set is an n-manifold for any integer $n \geq 0$, and is called the **empty manifold**.

Remark 21. Manifolds are locally represented by open sets of affine spaces, so smooth maps between manifolds are locally represented by calculus-smooth maps. More precisely, if $f\colon X \to Y$ is a smooth map from n-manifold X to m-manifold Y, then, for any point $p \in X$, one can form a commutative square

In the diagram, the vertical arrows are diffeomorphisms and they provide local representations for manifolds, the dotted arrow is uniquely determined by the other three arrows in the diagram.

$$\begin{tikzcd} (X_{loc}, p) \arrow[r, "f"] & (Y_{loc}, f(p)) \\ (\mathbb{E}^n_{loc}, 0) \arrow[u, "\cong"] \arrow[r, dotted, "\tilde{f}"] & (\mathbb{E}^m_{loc}, 0) \arrow[u, "\cong"] \end{tikzcd}$$

in which the calculus-smooth map $\tilde{f}$ is called a **local calculus representation** of f around point p.

By design, the notions of tangent space and tangent map are valid for manifolds. Indeed, if X is an n-manifold, then, for any point $p \in X$, since $(X_{loc}, p) \cong (\mathbb{E}^n_{loc}, 0)$, by applying the functor V, we conclude that the pre-linear space V_pX is pre-linear equivalent to the linear space $T_0\mathbb{E}^n$, so V_pX is a linear space of dimension n. Because of this fact, the velocity space V_pX shall be renamed as the tangent space of X at point p and be rewritten as T_pX.

To see that the category of smooth maps between manifolds is closed under the velocity functor, it remains to show that the total velocity space of a manifold is a manifold. A simple proof is outlined in the exercise below.

Exercise 49. *Assume that A and B are subsets of $\mathbb{A}$. Show that*

1) If A is an open set of B, then VA is an open set of VB. `Hint:` VA `is the pre-image of` A `under the smooth map` $\nu\colon\ VB \to B$ `in Exercise 46`.

2) If X is an n-manifold, then VX is a $2n$-manifold.

In case $f\colon X \to Y$ is a smooth map between manifolds, we shall rewrite $Vf\colon VX \to VY$ as $Tf\colon TX \to TY$. For $p \in X$, the easy version of $T_pf\colon T_pX \to T_{f(p)}Y$ shall be written as $Tf_p\colon TX_p \to TY_{f(p)}$.

Remark 22. Since smooth maps between manifolds can be represented locally by calculus-smooth maps, we can restate Inverse Function Theorem as follows: for smooth maps between manifolds,

$$\textit{local invertibility} \iff \textit{infinitesimal invertibility}.$$

In the following exercise, you are invited to use Inverse Function Theorem to prove the local canonical form of a smooth map about a point in two special cases.

Exercise 50. *Let $f\colon X \to Y$ be a smooth map from n-manifold X to m-manifold Y and $p \in X$. Show that, locally around point p, function f can be represented by 1) the standard injection*

$$((\mathbb{E}^n)_{loc}, 0) \quad \to \quad ((\mathbb{E}^m)_{loc}, 0)$$

$$x \mapsto (x, 0)$$

if $T_p f$ is injective, 2) the standard projection

$$\begin{array}{rcl} ((\mathbb{E}^n)_{loc}, 0) & \to & ((\mathbb{E}^m)_{loc}, 0) \\ (x, y) & \mapsto & y \end{array}$$

if $T_p f$ is surjective.

`Remark: The two results 1) and 2) above are respectively called the injective and surjective form of Inverse Function Theorem.`

Exercise 51. *Show that a non-empty subset X of an affine space $\mathbb{A}$ is an n-manifold if locally it is always the graph of a calculus-smooth map whose domain is an open set of an n-dimensional affine subspace of $\mathbb{A}$.*

`Hint: use the result in Exercise 41.`

Example 12. *The circle S^1: $x^2 + y^2 = 1$ is a 1-manifold inside $\mathbb{E}^2$.*

Proof. Let $p = (x_0, y_0) \in S^1$. We need to verify that, locally around point p, S^1 is the graph of a smooth function on an open interval. Since either $x_0 \neq 0$ or $y_0 \neq 0$, we can divide the task into four cases: $y_0 > 0$, $y_0 < 0$, $x_0 > 0$, $x_0 < 0$.

In case $y_0 < 0$,

$$p \in S^1 \cap \{(x, y) \in \mathbb{E}^2 \mid y < 0\} = \Gamma_f$$

where $f(x) = -\sqrt{1 - x^2}$ with $|x| < 1$.

The other cases can be shown similarly . $\square$

If you employ symmetry, you just need to verify the condition in one of the four cases.

Ellipses, parabolas and hyperbolas are 1-manifolds as well. The trace of α: $\mathbb{R} \to \mathbb{E}^2$, where α maps $t \in \mathbb{R}$ to $(\cos t, \sin t \cos t) \in \mathbb{E}^2$, though a regular curve, is not a 1-manifold because it is a figure 8.

One can check quadrics

$$x^2 + y^2 + z^2 = 1,\ x^2 + y^2 - z^2 = 1,\ x^2 + y^2 - z^2 = -1,\ z = x^2 + y^2,\ z = x^2 - y^2$$

in $\mathbb{E}^3$ are 2-manifolds, but quadric $x^2 + y^2 - z^2 = 0$ is not a 2-manifold.

Exercise 52. *Continuing Exercise 48, show that X is not a manifold.*

Exercise 53. *Show that the cusp $y^2 = x^3$ is not a manifold.*

Local Canonical Form

For any point p in an n-manifold $X \subseteq \mathbb{E}^N$, if we let $N_p X$ denote the normal space of X at point p, i.e., the orthogonal complement of $T_p X$ in $T_p \mathbb{E}^N$, then we have the decomposition

$$\mathbb{E}^N = (T_p X)_{aff} \oplus (N_p X)_{aff}$$

so that any point $q \in \mathbb{E}^N$ admits a unique decomposition $q = (q_t, q_n)$ with $q_t \in (T_pX)_{aff}$ and $q_n \in (N_pX)_{aff}$, thus a resulting decomposition $\iota = (\iota_t, \iota_n)$ for the inclusion map $\iota\colon X \to \mathbb{E}^N$.

Note that point (q_t, q_n) is the affine combination $(-1)p + 1q_t + 1q_n$, i.e., the point $p + ((q_t - p) + (q_n - p))$.

One can check easily that $T_p\iota_t = 1_{T_pX}$, so, by Inverse Function Theorem, $\iota_t\colon X \to (T_pX)_{aff}$ is a local diffeomorphism around point p. Then, if we let $f = \iota_n \circ (\iota_t)^{-1}_{loc}$ — a calculus-smooth map, then $\Gamma_f = X_{loc}$ — the domain of the smoothly invertible map $(\iota_t)_{loc}$, so the graph Γ_f is an open neighborhood of p in X. In short, we have shown that locally any non-empty manifold is the graph of a calculus-smooth map. Combining with the result in Exercise 51, we have shown that *a non-empty subset of an affine space is a manifold if and only if locally it is the graph of a calculus-smooth map.*

That is because $q \in X_{loc} \iff q_n = f(q_t)$.

There are many ways of representing a manifold locally as the graph of a calculus-smooth map f, and the one produced in the last paragraph is called the **local canonical form** and it is germ-unique, i.e., the canonical map (or function) f is germ-unique in the sense that any two have a common local version around point p in $(T_pX)_{aff}$. Analytically, up to a rigid motion of $\mathbb{E}^N = \mathbb{E}^n \times \mathbb{E}^{N-n}$, this canonical map is of the form

$$y = 0 + Q(x) + o(x^2)$$

where each component of $Q(x)$ is a quadratic form in $x \in \mathbb{E}^n$. Furthermore, in case X is a hypersurface in $\mathbb{E}^N$ (i.e., $\dim X = N - 1$), by a further rotation, we arrive at the canonical function

$$y = \frac{1}{2}\sum_i \kappa_i (x^i)^2 + o(x^2)$$

for some real numbers $\kappa_1, \ldots, \kappa_{\dim X}$, which are refereed to as the principal curvatures of the hypersurface X at point p.

Example 13. *Find the local canonical form of the surface $x^2 + y^2 - z^2 = 1$ at point $p = (1, 1, 1)$.*

Solution. Ideas: apply a rigid motion of $\mathbb{E}^3$ to move point p to coordinate origin$(0,0,0)$ and the tangent plane at point p to the xy-coordinate plane.

Please recall this **handy fact**: let ϕ be a rigid motion of $\mathbb{E}^3$, then ϕ maps surface $f = 0$ to surface $f \circ \phi^{-1} = 0$. That is because, $f(p) = 0$ if and only if $(f \circ \phi^{-1})(\phi(p)) = 0$.

Step 1. Applying the translation of $\mathbb{E}^3$ that turns point $(1, 1, 1)$ into $(0, 0, 0)$, the surface $x^2 + y^2 - z^2 = 1$ is tuned into surface $(x + 1)^2 + (y + 1)^2 - (z + 1)^2 = 1$ or $2(x + y - z) + x^2 + y^2 - z^2 = 0$, i.e., the surface

$$\vec{n} \cdot \vec{r} + \frac{x^2 + y^2 - z^2}{2\sqrt{3}} = 0 \tag{12}$$

where

$$\vec{n} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \quad \vec{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

So we have the tangent plane $\vec{n} \cdot \vec{r} = 0$.

Step 2. Let us first find a rotation $\mathbb{E}^3$ about point $(0,0,0)$:

$$\vec{r} \mapsto A\vec{r},$$

that turns the tangent plane $\vec{n} \cdot \vec{r} = 0$ into plane $z = 0$, i.e., find a rotation matrix A such that $\vec{n} \cdot A^{-1}\vec{r} = z$ or $\vec{n} = A^{-1}\vec{e}_3$. One possible choice is

$$A^{-1} = \begin{bmatrix} -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{bmatrix}, \text{ so } A = (A^{-1})^T = \begin{bmatrix} -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{bmatrix}.$$

Step 3. Applying the rotation of $\mathbb{E}^3$ found in step 2, surface (12) is tuned into surface

Note that $x^2 + y^2 - z^2 = (x^2 + y^2 + z^2) - 2z^2$, $x^2 + y^2 + z^2$ is invariant under any rotation, but $z = \vec{e}_3 \cdot \vec{r}$ is turned into $\vec{e}_3 \cdot A^{-1}\vec{r} = A\vec{e}_3 \cdot \vec{r} = \frac{1}{\sqrt{6}}x + \frac{1}{\sqrt{2}}y - \frac{1}{\sqrt{3}}z$.

$$z + \frac{x^2 + y^2 + z^2 - 2(\frac{1}{\sqrt{6}}x + \frac{1}{\sqrt{2}}y - \frac{1}{\sqrt{3}}z)^2}{2\sqrt{3}} = 0$$

or

$$z + \frac{\frac{2}{3}x^2 - \frac{2}{\sqrt{3}}xy + \frac{z^2}{3} + 2\sqrt{2}(\frac{y}{\sqrt{3}} + \frac{x}{3})z}{2\sqrt{3}} = 0.$$

Then

$$z = -\frac{\frac{2}{3}x^2 - \frac{2}{\sqrt{3}}xy}{2\sqrt{3}} + o(x^2 + y^2) = -\frac{1}{3\sqrt{3}}(x^2 - \sqrt{3}xy) + o(x^2 + y^2)$$

By a further rotation around the z-axis, we arrive at equation

Note that, the quadratic form $x^2 - \sqrt{3}xy$ is represented by the real symmetric matrix $\begin{bmatrix} 1 & -\sqrt{3}/2 \\ -\sqrt{3}/2 & 0 \end{bmatrix}$ which is similar to matrix $\begin{bmatrix} 3/2 & 0 \\ 0 & -1/2 \end{bmatrix}$ that represents the quadratic form $\frac{3}{2}x^2 - \frac{1}{2}y^2$.

$$z = -\frac{1}{6\sqrt{3}}(3x^2 - y^2) + o(x^2 + y^2).$$

In summary, up to rigid motion, locally around point $p = (1,1,1)$ the surface $x^2 + y^2 - z^2 = 1$ is the graph of smooth function of the form

$$z = -\frac{1}{2\sqrt{3}}x^2 + \frac{1}{6\sqrt{3}}y^2 + o(x^2 + y^2).$$

◀

Local Coordinate System

Let X be an n-manifold inside $\mathbb{E}^N$, $p \in X$. Then there is an open set U of $\mathbb{E}^n$ and an imbedding

$$\phi : \quad U \to X$$

such that $\phi(U)$ is an open neighborhood of p in X.

Recall that ϕ is an imbedding means that ϕ is a diffeomorphism from its domain onto its image, i.e., $\bar{\phi}$ is a diffeomorphism.

On $\mathrm{Im}(\phi)$ we have n smooth real functions: $x^1, \ldots, x^n$. By definition,

$$x^i : \quad \mathrm{Im}(\phi) \quad \to \quad \mathbb{R}$$

$$q \quad \mapsto \quad \text{the } i\text{-th component of } \bar{\phi}^{-1}(q) \in \mathbb{E}^n \qquad (13)$$

We call $(x^1, \ldots, x^n)$, or simply x^i, a local coordinate system on X around point p, $(\mathrm{Im}(\phi), x^1, \ldots, x^n)$, or simply $(\mathrm{Im}(\phi), x^i)$, a coordinate chart on X and $\mathrm{Im}(\phi)$ the corresponding coordinate patch. We shall call ϕ a *local parametrization* of X around point p.

Example 14 (spherical coordinates)**.** *Let $X \subset \mathbb{E}^3$ be the unit sphere $x^2 + y^2 + z^2 = 1$ inside $\mathbb{E}^3$, $U \subset \mathbb{E}^2$ be the open region $(0, \pi) \times (0, 2\pi)$ of $\mathbb{E}^2$. Then*

$$\begin{aligned} \phi : U &\to X \\ (\vartheta, \varphi) &\mapsto (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta) \end{aligned} \qquad (14)$$

is an imbedding onto an open set of X, so we have a local coordinate system (ϑ, φ) with $\vartheta(x, y, z) = \arccos z$ and $e^{\mathrm{i}\varphi(x,y,z)} = \frac{x+y\mathrm{i}}{\sqrt{1-z^2}}$.

`Note: this local coordinate system is called the spherical coordinate system. It is *almost global* in the sense that the corresponding coordinate patch is an open dense subset of the whole manifold.`

Question: What is the local parametrisation for $\mathbb{E}^2$ that corresponds to the polar coordinate system on $\mathbb{E}^2$?

Total Differential

Just as for open sets of affine spaces, for a manifold X and a point p in X, a cotangent vector of X at point p is a linear map from the tangent space T_pX into $\mathbb{R}$ and the set of all cotangent vectors of X at point p is called the cotangent space of X at point p and is denoted by T_p^*X.

In Calculus II, the total differential of a smooth real function is introduced as a formal expression. Here we shall see that it is a function whose value at point p is a cotangent vector at point p. To be more precise, let f be a real smooth function on a manifold X, and p be a point of X, then the **differential of f at point** p, denoted by $\mathrm{d}f_p$, is the cotangent vector of X at point p that sends tangent vector $(p, \vec{u})$ to the real number $Tf_p(\vec{u})$. Schematically we have

$$\begin{tikzcd} T_pX \arrow[r, "T_pf"] \arrow[dr, "\mathrm{d}f_p"'] & T_{f(p)}\mathbb{R} \arrow[d, leftrightarrow] \\ & \mathbb{R} \end{tikzcd}$$

The map $\mathrm{d}f\colon X \to T^*X$ that maps $p \in X$ to $\mathrm{d}f_p$ is called the **total differential of** f. One can show that T^*X is a $2n$-manifold. Note that $\mathrm{d}f$ is a differential one-form on X.

By definition, a smooth map $\alpha\colon X \to T^*X$ with $\alpha(p) \in T_p^*X$ for any $p \in X$ is called a **differential one-form** on X.

The notions of differential at a point and total differential make sense even for the smooth map $F\colon X \to \mathbb{E}^m$. But then $\mathrm{d}F_p$ is a $\mathbb{R}^m$-valued cotangent vector of X at point p and $\mathrm{d}F$ is a $\mathbb{R}^m$-valued differential one-form.

The 1st fundamental form

Let X be an n-manifold inside a Euclidean space $\mathbb{E}^N$. For any point $p \in X$, we let I_p be the following positive-definite symmetric two-form on T_pX:

$$\mathrm{I}_p((p,\vec{u}),(p,\vec{v})) = \vec{u}\cdot\vec{v}, \quad \text{for } (p,\vec{u}) \text{ and } (p,\vec{v}) \text{ in } T_pX. \tag{15}$$

We use I to denote the map $p \to \mathrm{I}_p$. This is a map from X to $S^2(T^*X)$ — the space of symmetric two forms on the tangent spaces of X.

One can show that $S^2(T^*X)$ is a manifold of dimension $n + \frac{n(n+1)}{2}$, and $\mathrm{I}\colon X \to S^2(T^*X)$ is a smooth map.

By definition, I is called the 1st fundamental form on X. It is clear that the 1st fundamental form on X is nothing but the restriction of the Euclidean structure on $\mathbb{E}^N$ to X. In other words, (X, I) is a Riemannian submanifold of the Euclidean space $\mathbb{E}^N$.

Remark 23. Let X be an n-manifold inside a Euclidean space $\mathbb{E}^N$ and $\iota\colon X \to \mathbb{E}^N$ be the inclusion map. Then $\mathrm{d}\iota$ is an $\mathbb{R}^N$-valued differential one-form. Combining the dot product for vectors in $\mathbb{R}^N$ with the pointwise symmetric tensor product for one-forms on each tangent space of X, we arrive at $\mathrm{d}\iota\cdot\mathrm{d}\iota$, which is the smooth map that sends point p in X to the symmetric 2-form $\mathrm{d}\iota_p\cdot\mathrm{d}\iota_p$ on the tangent space T_pX: $((p,\vec{u}),(p,\vec{v})) \mapsto \mathrm{d}\iota_p(p,\vec{u})\cdot\mathrm{d}\iota_p(p,\vec{v})$. We claim that

$$\mathrm{I} = \mathrm{d}\iota\cdot\mathrm{d}\iota. \tag{16}$$

Exercise 54. *Show that $\mathrm{d}\iota_p(p,\vec{u}) = \vec{u}$. Then prove the above claim.*

Manifold with Structure

A Euclidean space is an example of manifold with structure (i.e., euclidean structure). A more general concept is *Riemannian manifold* — manifold with a Riemannian structure. The first fundamental form is an example of Riemannian structure. By definition, a Riemannian structure on a manifold M is a smooth assignment of inner product to each of tangent space of M, i.e., a smooth map $g\colon M \to S^2(T^*M)$ such that $g(p)$ is an inner product on T_pM for each point p in M. Similarly, one can introduce the concepts of pseudo-Riemannian structure and pseudo-Riemannian manifold.

By definition, an orientation on M is a smooth assignment of orientation to each tangent space of M, and a manifold together with an orientation is called an **oriented manifold**. A manifold may or may not have an orientation and is called **orientable** if it admits an orientation, and **non-orientable** otherwise. Möbius strip and the real projective plane are non-orientable, but Lie groups are always orientable.

There are many other types of structure on manifolds, such as *spin structure, almost complex structure, symplectic structure, complex structure, Kähler structure*, ..., and the corresponding manifolds with structure are respectively called *spin manifold, almost complex manifold, symplectic manifold, complex manifold, Kähler manifold*, ...

Let us conclude this chapter with the following sequence of inclusion functors:

$$\begin{array}{c}
\text{The category of vector spaces and linear maps} \\
\downarrow \\
\text{The category of affine spaces and affine maps} \\
\downarrow \\
\text{The category } \mathcal{O} \text{ of open sets of affine spaces and smooth maps} \\
\downarrow \\
\text{The category } \mathcal{M} \text{ of manifolds and smooth maps} \\
\downarrow \\
\text{The category } \mathcal{S} \text{ of subsets of affine spaces and smooth maps}
\end{array}$$

and a commutative diagram

$$\begin{tikzcd}
\mathcal{O} \arrow[r, hook] \arrow[d, "T"'] & \mathcal{M} \arrow[r, hook] \arrow[d, "T"] & \mathcal{S} \arrow[d, "V"] \\
\mathcal{O} \arrow[r, hook] & \mathcal{M} \arrow[r, hook] & \mathcal{S}
\end{tikzcd}$$

of functors.

List of Symbols

$\cong$	diffeomorphic to
$\cong_{locally}$	locally diffeomorphic to
ι	an inclusion map
$f : A \to B$	a map between subsets of affine spaces
$\hat{f}$	the hat version of f. It is f, but with the target enlarged to the affine space that B sits in
$\bar{f}$	the bar version of f. It is f, but with the target shrunk to the image of f
$\tilde{f}$	a calculus-smooth map that extends the map f between subsets of affine spaces or a local calculus representation of the smooth map f between manifolds.
f_*	the pushforward by f, so $f_*\alpha = f \circ \alpha$
X	a subset of an affine space
V_pX	the velocity space of X at point p
VX_p	the easy version of V_pX. So $V_pX = \{p\} \times VX_p$
VX	the total velocity space of X
Vf	the velocity map of smooth map f
V_pf	the velocity map of smooth map f at point p. It is a pre-linear map
Vf_p	the easy version of V_pf
I	the first fundamental form of X
I_p	the first fundamental form of X at point p
$\phi : U \to X$	a local parametrization of X
x^i	a locally coordinate system
$\mathcal{S}$	the category of smooth maps between subsets of affine spaces
V	the velocity functor, which is an endofunctor on $\mathcal{S}$
$\mathcal{M}$	the category of smooth maps between (smooth) manifolds
T	the tangent functor, which is an endofunctor on $\mathcal{M}$

§6 Gauss-Pontryagin Map

For any manifold inside a Euclidean space, there are two canonical maps from the manifold, one is the inclusion map into the ambient Euclidean space, the other is the Pontryagin map that we shall introduce here.

Let X be an n-manifold inside the Euclidean space $\mathbb{E}^N$. Recall that, for each point $p \in X$, the easy version of the tangent space of X at p, denoted by TX_p, is a linear subspace of $\mathbb{R}^N$ such that $T_pX = \{p\} \times TX_p$. Let NX_p be the orthogonal complement of TX_p in $\mathbb{R}^N$, then NX_p is the easy version of the normal space of X at p. The assignment $p \to NX_p$ defines a map from X to the set of linear subspaces of $\mathbb{R}^N$ of dimension $k := N - n$. This set is denoted by $\mathrm{Gr}_k(\mathbb{R}^N)$ and is called the Grassmannian of k-dimensional linear subspaces of $\mathbb{R}^N$. We shall show in the appendix that Grassmannians are manifolds.

Unless said otherwise, we always assume the standard orientation on $\mathbb{R}^N$. Then, since $\mathbb{R}^N = TM_p \oplus NM_p$, we conclude that the orientations on TM_p and the orientations on NM_p correspond to each other.

Pontryagin Map

Let X be an n-manifold inside the Euclidean space $\mathbb{E}^N$ and $k = N - n$. Then the natural map

$$\begin{array}{cccc} P: & X & \to & \mathrm{Gr}_k(\mathbb{R}^N) \\ & p & \mapsto & NX_p \end{array} \tag{17}$$

is called the Pontryagin map on X. One can show that P is a smooth map.

Denote by $\widetilde{\mathrm{Gr}}_k(\mathbb{R}^N)$ the set of oriented k-dimensional linear subspaces of $\mathbb{R}^N$. One can show that $\widetilde{\mathrm{Gr}}_k(\mathbb{R}^N)$ is a manifold, and the orientation-forgetting 2-to-1 covering map

$$\pi : \widetilde{\mathrm{Gr}}_k(\mathbb{R}^N) \to \mathrm{Gr}_k(\mathbb{R}^N)$$

is smooth.

By definition, a Gauss-Pontryagin map on X is a smooth map $\tilde{P}$: $X \to \widetilde{\mathrm{Gr}}_k(\mathbb{R}^N)$ such that $\pi \circ \tilde{P} = P$. Schematically we have

$$\begin{tikzcd} & \widetilde{\mathrm{Gr}}_k(\mathbb{R}^N) \arrow[d, "\pi"] \\ X \arrow[ur, dashed, "\tilde{P}"] \arrow[r, "P"] & \mathrm{Gr}_k(\mathbb{R}^N) \end{tikzcd}$$

Note that $\tilde{P}$ (called a *lifting* of P) may not exist, moreover, a Gauss-Pontryagin map on X is nothing but an orientation on X. Note also that $\widetilde{\mathrm{Gr}}_1(\mathbb{R}^N)$ can be naturally identified with S^{N-1} and $\mathrm{Gr}_1(\mathbb{R}^N)$ is the real projective space of dimension $N-1$, customarily written as $\mathbb{R}P^{N-1}$.

Gauss Map and Orientation

Let M be an n-manifold inside the Euclidean space $\mathbb{E}^{n+1}$. (Such a manifold is called a hypersurface in $\mathbb{E}^{n+1}$.) Suppose that M is orientable, this means that there is a smooth map

If M is a compact hypersurface in an affine space, then M is orientable. Here is a proof of this fact.

$$g: \quad M \to \mathrm{S}^n := \{x \in \mathbb{E}^{n+1} \mid \vec{x} \cdot \vec{x} = 1\}$$

such that for any $p \in M$, $\overrightarrow{g(p)}$ (i.e. $g(p) - 0$) spans NM_p — the easy version of normal space of M at p. This map g — a special case of Gauss-Pontryagin map — is called the Gauss map on hypersurface M.

As an example, if $M = \mathrm{S}^n$, then the identity map 1_M: $x \to x$ is a Gauss map, and the map -1_M: $x \to -x$ is another Gauss map. There are no other Gauss maps. This simply reflects the fact that S^n admits exactly two orientations: oriented outward and oriented inward. Of course these two orientations are opposite of each other.

It is easy to see that $TM_p = T\mathrm{S}^{\mathrm{n}}{}_{g(p)}$. That is because both are the orthogonal complement of NM_p in $\mathbb{R}^{n+1}$. Therefore, Tg_p is an endomorphism on TM_p. Recall that an endomorphism T on a Euclidean vector space is called self-adjoint if $\langle T\underline{u}, \underline{v}\rangle = \langle \underline{u}, T\underline{v}\rangle$ for any two vectors $\underline{u}, \underline{v}$, i.e., the two-form that sends $(\underline{u}, \underline{v})$ to $\langle T\underline{u}, \underline{v}\rangle$ is symmetric. In case A is the real square matrix that represents T with respect to an orthonormal basis, then, T is a self-adjoint $\iff$ A is symmetric.

Theorem 3. *Tg_p: $TM_p \to TM_p$ is self-adjoint with respect to the inner product on TM_p.*

Proof. We may assume that $p = (0, \ldots, 0)$ and M is the graph of function

$$x_{n+1} = f(x_1, \ldots, x_n) = \frac{1}{2}\sum_i \kappa_i x_i^2 + o(x^2)$$

near point p. Here κ_i are some fixed real numbers, called principal curvatures of M at point p, and their arithmetic mean is called the mean curvature of the hypersurface at point p. Then, up to a sign, locally the Gauss map g has the following calculus-smooth extension:

Recall that $(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + \cdots$

$$\tilde{g} = 0 + \frac{1}{\sqrt{1+\sum_i (\frac{\partial f}{\partial x_i})^2}} \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \\ -1 \end{bmatrix} = 0 + \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \\ -1 \end{bmatrix} + \cdots$$

Here $\cdots$ means a function that starts with the quadratic term. Then

$$\left.\frac{\partial \tilde{g}_i}{\partial x_j}\right|_p = \frac{\partial^2 f}{\partial x_j \partial x_i}(0,\ldots,0) = \delta_{ij}\kappa_i \tag{18}$$

if $i, j \leq n$ and is zero otherwise. Note that $TM_p = \text{span}\{\vec{e}_1,\ldots,\vec{e}_n\}$, so, we have $Tg_p(\vec{e}_i) = J\tilde{g}_p \cdot \vec{e}_i = \kappa_i\vec{e}_i$ if $i \leq n$. Then, for any $\vec{u}$, $\vec{v}$ in TM_p,

$$\vec{u} \cdot Tg_p(\vec{v}) = \sum_{i=1}^{n} \kappa_i u^i v^i$$

is symmetric in $\vec{u} := \sum_{i=1}^n u^i\vec{e}_i$ and $\vec{v} := \sum_{i=1}^n v^i\vec{e}_i$. Therefore Tg_p is self-adjoint. □

Remark 24. From the proof above we can see that the endomorphism Tg_p can be represented by a diagonal matrix with the principal curvatures at point p as its diagonal entries. Then $\det(Tg_p)$ is the product of the principal curvatures at point p.

The 2nd fundamental form

Recall that the 1st fundamental form I is defined for any manifold inside a *Euclidean* space. In comparison, the second fundamental form II is defined only for oriented hypersurfaces in a Euclidean space.

Let X be an oriented hypersurface of $\mathbb{E}^{n+1}$ and $g\colon X \to S^n$ be the corresponding Gauss map. For each point $p \in X$, we let II_p be the following two-form on T_pX:

$$\mathrm{II}_p((p,\vec{u}),(p,\vec{v})) = \vec{u} \cdot Tg_p(\vec{v}).$$

Due to Theorem 3, II_p is a symmetric two-form on T_pX. We use II to denote the map $p \to \mathrm{II}_p$. This is a smooth map from X to $S^2(T^*X)$ — the space of symmetric two-forms on the tangent spaces of X. By definition, II is called the 2nd fundamental form on X. Since I is non-degenerate, II and Tg determine each other, as one can see from commutative diagram [3]

[3] Here we have abused the notation a bit for the southeast arrow.

$$\begin{tikzcd} & T_pX \arrow[dl, "(\mathrm{II}_p)_\natural"'] \arrow[dr, "T_pg"] & \\ T_p^*X & & T_pX \arrow[ll, "\cong", "(\mathrm{I}_p)_\natural"'] \end{tikzcd}$$

Locally a hypersurface is always orientable, so $\det(Tg_p)$, being the product of principal curvatures, is a well-defined for any even-dimensional hypersurface, irrespective of the existence of a Gauss map. That is because principal curvatures are well-defined up to permutations and a global sign. In any case, the number

$$K_p := \begin{cases} \det(Tg_p) & \text{if } n \text{ is even} \\ |\det(Tg_p)| & \text{if } n \text{ is odd} \end{cases}$$

on any hypersurface is well defined, irrespective of the existence of a Gauss map. Let us call this number the Gauss curvature of the hypersurface at point p.

Exercise 55. *Let C be a non-self crossing regular plane curve, and $p \in C$. Then K_p is equal to the value of $|\kappa|$ at point p in Remark 15.*

Remark 25. For an oriented hypersurface X in $\mathbb{E}^{n+1}$, we let

$$\iota:\ X \to \mathbb{E}^{n+1} \quad \text{and} \quad \hat{g}:\ X \to \mathbb{E}^{n+1}$$

be the inclusion map and the hat version of the Gauss map respectively. Then $\mathrm{d}\iota$ and $\mathrm{d}\hat{g}$ are $\mathbb{R}^{n+1}$-valued differential one-forms. Combining the dot product for vectors in $\mathbb{R}^{n+1}$ with the pointwise symmetric tensor product for one-forms on each tangent space of X, we arrive at $\mathrm{d}\iota \cdot \mathrm{d}\hat{g}$. We claim that

$$\mathrm{II} = \mathrm{d}\iota \cdot \mathrm{d}\hat{g}. \tag{19}$$

Exercise 56. *Please prove this claim, i.e., for any two tangent vectors $(p, \vec{u})$ and $(p, \vec{v})$ on X, we have*

$$\mathrm{II}_p((p,\vec{u}),(p,\vec{v})) = \frac{1}{2}\left(\mathrm{d}\iota_p(p,\vec{u}) \cdot \mathrm{d}\hat{g}_p(p,\vec{v}) + \mathrm{d}\iota_p(p,\vec{v}) \cdot \mathrm{d}\hat{g}_p(p,\vec{u})\right).$$

Appendix 1: A useful theorem

Theorem 4 (Regular Value Theorem)**.** *Let f be a smooth real function on an open set U of $\mathbb{E}^{n+1}$. Suppose that $X := f^{-1}(0)$ is non-empty and $\mathrm{d}f_p \neq 0$ for any $p \in X$. Then X is an orientable hypersurface in $\mathbb{E}^{n+1}$.*

Proof. We use $\operatorname{grad} f$ to denote the gradient vector field of f and ∇f to denote its easy version. So, if $p \in U$, we have

$$\nabla f(p) = (Jf_p)^T, \quad \operatorname{grad} f(p) = (p, \nabla f(p)).$$

Suppose that $p \in X$. Up to a rigid motion, we may assume $p = 0$ and $\nabla f(p) = |\nabla f(p)|\vec{e}_{n+1}$, then the Taylor expansion of f around p is

$$\begin{aligned} f(x_1, \cdots, x_{n+1}) &= f(p) + \sum_{i=1}^{n+1} \frac{\partial f}{\partial x_i}(p)x_i + \frac{1}{2!}\sum_{1\le i,j\le n+1} \frac{\partial^2 f}{\partial x_i \partial x_j}(p)x_i x_j + o(x^2) \\ &= |\nabla f(p)|x_{n+1} + \frac{1}{2!}\sum_{1\le i,j\le n+1} \frac{\partial^2 f}{\partial x_i \partial x_j}(p)x_i x_j + o(x^2). \end{aligned}$$

Therefore, locally around point p, X is the graph of smooth function

$$x_{n+1} = -\frac{1}{2|\nabla f(p)|}\sum_{1\le i,j\le n} \frac{\partial^2 f}{\partial x_i \partial x_j}(p)x_i x_j + o(x_1^2 + \cdots + x_n^2).$$

So X is a manifold inside $\mathbb{E}^{n+1}$ with co-dimension one. It is easy to see that

$$g = 0 + \frac{\nabla f}{|\nabla f|}$$

is a Gauss map on X. So X is an orientable hyper-surface in $\mathbb{E}^{n+1}$.

□

Continuing the discussion in the proof above, up to a rotation about the x_{n+1}-axis, we arrive at the local canonical form:

$$x_{n+1} = \frac{1}{2}\sum_{i=1}^{n} \kappa_i x_i^2 + o(x_1^2 + \cdots + x_n^2).$$

Considering the commutative diagram

$$\begin{array}{ccc} X & \xrightarrow{g} & S^n \\ \downarrow \iota & & \downarrow \iota \\ U_{loc} & \xrightarrow[\tilde{g}]{} & \mathbb{E}^{n+1} \end{array}$$

where $U_{loc} := \{df \neq 0\}$ and $\tilde{g}$ is the map that takes the assigning rule for g, if we linearise it at point p, we arrive at the following commutative diagram:

$$\begin{array}{ccc} TX_p & \xrightarrow{Tg_p} & TX_p \\ \downarrow \iota & & \downarrow \iota \\ \mathbb{R}^{n+1} & \xrightarrow[A|_p]{} & \mathbb{R}^{n+1} \end{array}$$

with

$$A := J\tilde{g} = \pi \frac{Hf}{|\nabla f|} \tag{20}$$

where

$$\pi = I - \frac{\nabla f \, (\nabla f)^T}{|\nabla f|^2}$$

represents the orthogonal projection of $\mathbb{R}^{n+1}$ onto NX_p and Hf is the Hessian-matrix of f whose (i,j)-entry is $\frac{\partial^2 f}{\partial x^i \partial x^j}$. Therefore, for $\vec{u}, \vec{v}$ in TX_p, we have

$$\begin{aligned} \vec{u} \cdot Tg_p(\vec{v}) &= \vec{u} \cdot A|_p \vec{v} \\ &= \vec{u} \cdot (A\pi)|_p \vec{v} \quad \text{because } \pi|_p(\vec{v}) = \vec{v} \\ &= (A\pi)|_p \vec{u} \cdot \vec{v} \quad \text{because } (A\pi)|_p \text{ is a symmetric matrix of order } (n+1) \end{aligned}$$

Therefore Tg_p is self-adjoint. Since locally X is the graph of a smooth function, so locally $X = f^{-1}(0)$ for some real smooth function f with $\mathrm{d}f$ nowhere vanishing, so this gives another proof of Theorem 3.

Proposition 1. *Let f be a smooth real function on an open set U of $\mathbb{E}^3$. Suppose that $X := f^{-1}(0)$ is non-empty and $\mathrm{d}f_p \neq 0$ for any $p \in X$. Then X is a hypersurface in $\mathbb{E}^3$ with Gauss curvature function*

$$K = \frac{1}{2}((\operatorname{tr} A)^2 - \operatorname{tr} A^2). \tag{21}$$

and principal curvature functions

κ_1 and κ_2 are defined only up to order and global sign.

$$\kappa_1 = \frac{\operatorname{tr} A + \sqrt{2\operatorname{tr} A^2 - (\operatorname{tr} A)^2}}{2}, \quad \kappa_2 = \frac{\operatorname{tr} A - \sqrt{2\operatorname{tr} A^2 - (\operatorname{tr} A)^2}}{2} \tag{22}$$

Exercise 57. *Prove this proposition. What if $\mathbb{E}^3$ in the theorem is replaced by $\mathbb{E}^2$?*

`Hint:` Let $T\colon V \to V$ be a linear map. Suppose that $V = V_1 \oplus V_2$ with $\dim V_1 > 0$, and $T = \begin{bmatrix} T_1 & 0 \\ 0 & 0 \end{bmatrix}$ with respect to the above decomposition of V. Then $\operatorname{tr} T_1 = \operatorname{tr} T$, $\operatorname{tr} T_1^2 = \operatorname{tr} T^2$, and $\det T_1 = \frac{1}{2}((\operatorname{tr} T_1)^2 - \operatorname{tr} T_1^2) = \frac{1}{2}((\operatorname{tr} T)^2 - \operatorname{tr} T^2)$ if $\dim V_1 = 2$.

One can use the above formulae for principal curvatures to find the local canonical representation for X.

Example 15. *Let X be the quadric $x^2 + y^2 - z^2 = 1$. Find 1) the Gauss curvature function K, 2) the local canonical representation for X at point $p = (x_0, y_0, z_0)$.*

Solution. Let $f = \frac{1}{2}(x^2 + y^2 - z^2 - 1)$. Then $X = f^{-1}(0)$ and

$$\nabla f = \begin{bmatrix} x \\ y \\ -z \end{bmatrix}, \quad Hf = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

So

$$\pi = \left(I - \frac{1}{x^2+y^2+z^2} \begin{bmatrix} x^2 & xy & -xz \\ xy & y^2 & -yz \\ -xz & -yz & z^2 \end{bmatrix} \right) = \frac{1}{x^2+y^2+z^2} \begin{bmatrix} y^2+z^2 & -xy & xz \\ -xy & x^2+z^2 & yz \\ xz & yz & x^2+y^2 \end{bmatrix}$$

and

$$A = \frac{1}{(x^2+y^2+z^2)^{3/2}} \begin{bmatrix} y^2+z^2 & -xy & -xz \\ -xy & x^2+z^2 & -yz \\ xz & yz & -x^2-y^2 \end{bmatrix}.$$

Then

$$\operatorname{tr} A = \frac{2z^2}{(x^2+y^2+z^2)^{3/2}}, \quad \operatorname{tr} A^2 = \frac{2}{(x^2+y^2+z^2)^3}((x^2+y^2)^2+z^4).$$

By using the formulae in the above theorem and equation $x^2+y^2-z^2=1$, we have

$$K = -\frac{1}{(1+2z^2)^2}, \quad \kappa_\pm = \frac{z^2 \pm (1+z^2)}{(1+2z^2)^{3/2}}.$$

Therefore the Gauss curvature is everywhere negative, and up to a rigid motion, locally around point (x_0, y_0, z_0), X is the graph of a function $z = f(x, y)$ where

$$f(x,y) = \frac{1}{2(1+2z_0^2)^{1/2}} x^2 - \frac{1}{2(1+2z_0^2)^{3/2}} y^2 + o(x^2+y^2).$$

In particular,

$$f(x,y) = \frac{1}{2\sqrt{3}} x^2 - \frac{1}{6\sqrt{3}} y^2 + o(x^2+y^2)$$

if $(x_0, y_0, z_0) = (1, 1, 1)$. Cf. Example 13. ◀

Appendix 2: Grassmannians are manifolds

The purpose of this appendix is to prove

Proposition 2. *The Grassmannian* $\mathrm{Gr}_k(\mathbb{R}^n)$ *is a manifold of dimension* $k(n-k)$*.*

Proof. Step 1: Show that $\mathrm{Gr}_k(\mathbb{R}^n)$ is a subset of an affine space.

To see this, we note that a k-dimensional linear subspace of $\mathbb{R}^n$ can be uniquely specified by the orthogonal projection of $\mathbb{R}^n$ onto it, which is in turn specified by a rank-k real symmetric matrix A of order n such that $A^2 = A$. So, if we let $\mathrm{H}_n(\mathbb{R})$ be the set of real symmetric matrices of order n — a real vector space (hence a real affine space) of dimension $\frac{1}{2}n(n+1)$, then

$$\mathrm{Gr}_k(\mathbb{R}^n) \equiv \{A \in \mathrm{H}_n(\mathbb{R}) \mid A^2 = A, \quad r(A) = k\} \subset \mathrm{H}_n(\mathbb{R}).$$

We shall view $\mathrm{H}_n(\mathbb{R})$ as a Euclidean space with the translation-invariant inner product at point A being this one:

$$\langle (A, Z), (A, W) \rangle = \operatorname{tr}(ZW).$$

Step 2: Show that locally $\mathrm{Gr}_k(\mathbb{R}^n)$ is the graph of a smooth map.

To see this, we let A_0 be a point in $\mathrm{Gr}_k(\mathbb{R}^n)$. Up to a rotation of $\mathbb{R}^n$, we may assume

$$A_0 = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}$$

where I_k is the identity matrix of order k. We shall find a local canonical representation of $\mathrm{Gr}_k(\mathbb{R}^n)$ as a graph of smooth map.

Let V_1 be the subspace of $\mathrm{H}_n(\mathbb{R})$ consisting of matrices of the form

$$\begin{bmatrix} 0 & X \\ X^T & 0 \end{bmatrix}$$

Here V_1 is the solution set of the linearized version of equation $A^2 = A$ at point $A = A_0$:

$$A_0 Z + Z A_0 = Z.$$

and V_2 be the orthogonal complement of V_1 in $\mathrm{H}_n(\mathbb{R})$, i.e., V_2 is the subspace of $\mathrm{H}_n(\mathbb{R})$ consisting of matrices of the form

Vector space V_1 (V_2 resp.) would be the easy version of the tangent (normal resp.) space of $\mathrm{Gr}_k(\mathbb{R}^n)$ at point A_0.

$$\begin{bmatrix} Y & 0 \\ 0 & Z \end{bmatrix}$$

where Y and Z are symmetric real matrices of order k and $n-k$ respectively.

Let U be an open neighborhood of 0 in V_1 such that $\left(\frac{I_k+\sqrt{I_k-4XX^T}}{2}\right)^{-1}$ makes sense for any

$$\begin{bmatrix} 0 & X \\ X^T & 0 \end{bmatrix}$$

in U. Let the affine subspaces $\mathbb{A}_1$ be $A_0 + V_1$ and its affine complement be $\mathbb{A}_2 := A_0 + V_2$. Further more, we let the open set of $\mathbb{A}_1$ be $A_0 + U$. Then, by solving equation

If $Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \in \mathrm{H}_n(\mathbb{R})$ with Z_{11} being invertible and $r(Z) = r(Z_{11})$, then $Z_{22} = Z_{21}(Z_{11})^{-1}Z_{12}$.

$$\begin{bmatrix} Z_{11} & X \\ X^T & X^T Z_{11}^{-1} X \end{bmatrix}^2 = \begin{bmatrix} Z_{11} & X \\ X^T & X^T Z_{11}^{-1} X \end{bmatrix}$$

where $Z_{11} = I_k + o(X)$, we get $Z_{11} = \frac{I_k+\sqrt{I_k-4XX^T}}{2}$. So, locally around A_0, $\mathrm{Gr}_k(\mathbb{R}^n)$ is the graph of smooth map $f\colon A_0 + U \to \mathbb{A}_2$ where

$$f\left(\begin{bmatrix} I_k & X \\ X^T & 0 \end{bmatrix}\right) = \begin{bmatrix} \frac{I_k+\sqrt{I_k-4XX^T}}{2} & 0 \\ 0 & X^T\left(\frac{I_k+\sqrt{I_k-4XX^T}}{2}\right)^{-1} X \end{bmatrix}.$$

Combining the two steps above we conclude that $\mathrm{Gr}_k(\mathbb{R}^n)$ is a manifold with dimension equal to $\dim V_1 = k(n-k)$. $\square$

Exercise 58. *Show that* $\mathrm{Gr}_k(\mathbb{R}^n)$ *is naturally diffeomorphic to* $\mathrm{Gr}_{n-k}(\mathbb{R}^n)$.

Exercise 59. *Let* W *be a* k*-dimensional linear subspace of* $\mathbb{R}^n$ *and* $W^\perp$ *be its orthogonal complement in* $\mathbb{R}^n$. *Show that the easy version of the tagent space of* $\mathrm{Gr}_k(\mathbb{R}^n)$ *at point* W *can be naturally identified with* $\mathrm{Hom}(W, W^\perp)$ *— the linear space consisting of all linear maps from* W *to* $W^\perp$.

Exercise 60. *Show that the Pontryagin map on a manifold* $X \subset \mathbb{E}^N$ *is a smooth map.*

Hint: in view of Remark 20, it suffices to show that P is locally smooth. Also, if A is $n \times k$-matrix with rank k, then $A(A^TA)^{-1}A^T$ is the standard matrix for the orthogonal projection from $\mathbb{R}^n$ onto $\mathrm{Col}\, A$ --- the column space of A.

List of Symbols

TX_p	the easy version of the tangent space T_pX
NX_p	the easy version of the normal space N_pX
$\mathrm{Gr}_k(\mathbb{R}^N)$	the Grassmannian of k-dim. linear subspaces of $\mathbb{R}^N$
$\widetilde{\mathrm{Gr}}_k(\mathbb{R}^N)$	the Grassmannian of k-dim. linear oriented subspaces of $\mathbb{R}^N$
P	the Pontryagin map
$\tilde{P}$	the Gauss-Pontryagin map
g	the Gauss map
II	the second fundamental form of an oriented hyper-surface
K	the Gauss curvature function of a hyper-surface
κ_i	the principal curvatures of a hyper-surface at a point
∇f	the easy version of the gradient vector field of the real smooth function f

§7 Local Representations and Local Computations

Locally a manifold can be represented by open sets of an affine space, so, if we study manifolds *locally*, we fall back to Calculus II. To facilitate the local study of manifolds, we collect some revevant definitions, notations and basic facts here.

Local tangent frame and local cotangent frame

Let M be an n-manifold inside $\mathbb{E}^N$, $p \in M$. Let $\phi\colon U \to M$ be a local parametrization of M around point p and x^i the resulting local coordinate system on M. On U we have n standard (smooth) vector fields $E_1, \ldots, E_n$:

$$E_i(x) = (x, \vec{e}_i) \quad \text{for any } x \in U$$

The diffeomorphism $\bar{\phi}$ maps the vector field E_i on U to a vector field on manifold $\mathrm{Im}(\phi)$, which shall be denoted by $\frac{\partial}{\partial x^i}$. So, for any $x \in U$, we have

$$\frac{\partial}{\partial x^i}\Big|_{\phi(x)} := T\bar{\phi}(E_i(x)) = T\hat{\phi}(E_i(x)) = (\phi(x), J\hat{\phi}_x \cdot \vec{e}_i).$$

So the easy version of $\frac{\partial}{\partial x^i}$ is the smooth map

$$\hat{\phi}_{\vec{e}_i} \circ \bar{\phi}^{-1}: \quad \mathrm{Im}(\phi) \to \mathbb{R}^N.$$

For simplicity, one may write $\frac{\partial}{\partial x^i}$ as ∂_{x^i} or simply ∂_i.

Remark 26. The local vector field ∂_i on M is a smooth map from $M_{loc} := \mathrm{Im}(\phi)$ to TM_{loc} that sends point q to tangent vector $(q, \hat{\phi}_{\vec{e}_i}(\bar{\phi}^{-1}(q)))$, i.e.,

$$\partial_i = (1, \hat{\phi}_{\vec{e}_i} \circ \bar{\phi}^{-1}) \tag{23}$$

Please read this formula from context because the right hand side is a smooth map from M_{loc} to $M_{loc} \times \mathbb{R}^N$ rather than its subset TM_{loc}.

Now, if f is a real smooth function on M, then the total differential $\mathrm{d}f$ is a differential one-form on M, i.e., a smooth assignment of a cotangent vector at p to each point p in M. So, on $\mathrm{Im}(\phi)$, one can form the pairing of $\mathrm{d}f$ with ∂_{x^i} to yield a real smooth function on the coordinate patch $\mathrm{Im}(\phi)$. Let us denote this function by $\frac{\partial f}{\partial x^i}$ or $\partial_{x^i} f$ or $\partial_i f$, then, by definition,

$$\frac{\partial f}{\partial x^i} := \langle \mathrm{d}f, \partial_{x^i} \rangle$$

which is the easy version of $Tf(\partial_{x^i})$. Since $T(f \circ \phi) = Tf \circ T\phi$, we have

$$\left.\frac{\partial f}{\partial x^i}\right|_{\phi(x)} = \langle \mathrm{d}(f \circ \phi)_x, E_i(x) \rangle = (f \circ \phi)_{\vec{e}_i}(x).$$

In other word, $\frac{\partial f}{\partial x^i} = (f \circ \phi)_{\vec{e}_i} \circ \bar{\phi}^{-1}$ or

The calculus-smooth function $f \circ \phi$ is called the *local representation* of f with respect to the coordinate system x^i.

$\left.\frac{\partial f}{\partial x^i}\right|_q$ is the i-th partial derivative of the calculus-smooth function $f \circ \phi$ at point $\bar{\phi}^{-1}(q)$.

(It is in this sense notation $\frac{\partial f}{\partial x^i}$ is justified.) In particular, we have $\frac{\partial x^i}{\partial x^j} = \delta^i_j$.

On coordinate patch $\mathrm{Im}(\phi)$, we also have n differential one-forms: $\mathrm{d}x^1, \ldots, \mathrm{d}x^n$. Since $\frac{\partial x^i}{\partial x^j} = \delta^i_j$, we have

$$\langle \mathrm{d}x^i, \partial_{x^j} \rangle = \delta^i_j. \tag{24}$$

Consequently $(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n})$ is a local tangent frame on M and $(\mathrm{d}x^1, \ldots, \mathrm{d}x^n)$ is a local cotangent frame on M, with respect to the local coordinate system x^i. Moreover, we have

Please read from context. Here $\mathrm{d}f$ really means $\mathrm{d}f|_{\mathrm{Im}(\phi)}$. For the sanity of notations, people prefer to write $\mathrm{d}f|_{\mathrm{Im}(\phi)}$ as $\mathrm{d}f$.

$$\mathrm{d}f = \frac{\partial f}{\partial x^i}\, \mathrm{d}x^i. \tag{25}$$

Remark 27. The total differential $\mathrm{d}f$ is the coordinate-free version of 1st order partial derivatives of f.

Local representation of fundamental forms

The first fundamental form I on M identifies a cotangent vector with a tangent vector, thus a differential one-form with a vector field, in particular the total differential $\mathrm{d}f$ with the gradient vector field $\mathrm{grad} f$, via the defining equation

$$\langle \mathrm{d}f, X \rangle = \mathrm{I}(\mathrm{grad} f, X) \quad \text{for any vector field } X.$$

On $\mathrm{Im}(\phi)$ we let

$$g_{ij} := \mathrm{I}(\partial_i, \partial_j) = (\hat{\phi}_{\vec{e}_i} \cdot \hat{\phi}_{\vec{e}_j}) \circ \bar{\phi}^{-1}.$$

Then the symmetric matrix function $[g_{ij}]$, being a positive-definite matrix at each point, has a pointwise matrix inverse. Let us denote this point-wise matrix inverse by $[g^{ij}]$. Then, solving the defining equation via a local computation, we have

Since matrix $[g_{ij}|p]$ is the matrix representation of the inner product I_p with respect to basis $(\partial_1|_p, \ldots, \partial_n|_p)$, it is positive-definite.

$$\mathrm{grad} f = g^{ij}\, \partial_i f\, \partial_j \quad \text{on } \mathrm{Im}(\phi).$$

In view of the fact that $\mathrm{I} = \mathrm{d}\iota \cdot \mathrm{d}\iota$, we claim that

$$\mathrm{I} = g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j \quad \text{on } \mathrm{Im}(\phi). \tag{26}$$

Here, $\mathrm{d}x^i\,\mathrm{d}x^j$ is the point-wise symmetric tensor product of one forms, so it maps a pair of vector fields (X, Y) to local function $\frac{1}{2}(\langle \mathrm{d}x^i, X\rangle\langle \mathrm{d}x^j, Y\rangle + \langle \mathrm{d}x^j, X\rangle\langle \mathrm{d}x^i, Y\rangle)$. We call g_{ij} the local representation of I with respect to the local coordinate system x^i.

Remark 28. For any smooth map $F\colon M \to \mathbb{E}^N$, we have

$$\mathrm{d}F = \partial_i F\,\mathrm{d}x^i,$$

where $\partial_i F$ as well as $\partial_i\partial_j F$ and so on (which are all $\mathbb{R}^N$-valued local functions) can be computed by the following rules:

1. Represent F locally by the calculus-smooth function $F \circ \phi$.

2. Compute the partial derivatives for the local representation $F \circ \phi$ (just as we did in Calculus II), then we have $(F \circ \phi)_{\vec{e}_i}$ as well as $(F \circ \phi)_{\vec{e}_i\vec{e}_j} := ((F \circ \phi)_{\vec{e}_j})_{\vec{e}_i}$ and so on.

3. Translate the computed result from $\mathbb{E}^n_{loc} := \mathrm{domain}\,\phi$ to $M_{loc} := \mathrm{Im}\,\phi$ via $\bar{\phi}$, then we have

$$\partial_i F = (F \circ \phi)_{\vec{e}_i} \circ \bar{\phi}^{-1}, \quad \partial_i\partial_j F = (F \circ \phi)_{\vec{e}_i\vec{e}_j} \circ \bar{\phi}^{-1}, \quad \text{etc.}$$

For example, $\partial_i\iota = \hat{\phi}_{\vec{e}_i} \circ \bar{\phi}^{-1}$. So,

$$\partial_i\iota \text{ is the easy version of } \partial_i.$$

Exercise 61. *Verify claim (26) in the above.*

Exercise 62. *Recall that* $\mathrm{II} = \mathrm{d}\iota \cdot \mathrm{d}\hat{g}$. *Show that, locally*

$$\mathrm{II} = -\vec{g} \cdot (\hat{\phi}_{\vec{e}_i\vec{e}_j} \circ \bar{\phi}^{-1})\,\mathrm{d}x^i\,\mathrm{d}x^j.$$

Hint: $\vec{g} \cdot \partial_i\iota = 0$ because the (easy version of) normal vector is orthogonal to (the easy version of) tangent vector ∂_i. Then $\partial_j\vec{g} \cdot \partial_i\iota = -\vec{g} \cdot \partial_j\partial_i\iota$.

Exercise 63. *Show that the third fundamental form* $\mathrm{III} := \mathrm{d}\hat{g} \cdot \mathrm{d}\hat{g}$ *is completely determined by the first two fundamental forms.*

Hint: 1) Since $\hat{g} = 0 + \vec{g}$, we have $\partial_i\hat{g} = \partial_i\vec{g}$; 2) Since $\vec{g} \cdot \vec{g} = 1$, we have $\vec{g} \cdot \partial_i\vec{g} = 0$, i.e., $\partial_i\vec{g}$ is a vector field on the coordinate patch, so $\partial_i\vec{g} = c_i^j\partial_j\iota$ for a unique set of local functions c_i^j. Then $\mathrm{III} = c_i^l c_j^m g_{lm}\,\mathrm{d}x^i\mathrm{d}x^j$. So all we need to show is that functions c_i^j are determined by the first two fundamental forms. For that, we need to solve equation $\partial_i\vec{g} = c_i^j\partial_j\iota$ for c_i^j by taking (pointwise) dot product with $\partial_k\iota$. Note: $g_{jk} = \partial_j\iota \cdot \partial_k\iota$ and $\vec{g} \cdot \partial_k\iota = 0$.

Remark 29. Let M be an oriented hypersurface in $\mathbb{E}^3$ and with (u, v) as a local coordinate system. We write

$\mathrm{d}u^2$ means $\mathrm{d}u\,\mathrm{d}u$, etc.

$$\mathrm{I} = E\,\mathrm{d}u^2 + 2F\,\mathrm{d}u\,\mathrm{d}v + G\,\mathrm{d}v^2, \quad \mathrm{II} = e\,\mathrm{d}u^2 + 2f\,\mathrm{d}u\,\mathrm{d}v + g\,\mathrm{d}v^2.$$

Then Exercise 5 yields formula

$$K = \frac{eg - f^2}{EG - F^2}$$

for the Gauss curvature function on the coordinate patch with coordinate functions u and v.

Example 16. *This is a continuation of Example 14.*

Question 1: Please describe the local tangent frame $(\partial_\vartheta, \partial_\varphi)$ as well as the local cotangent frame $(\mathrm{d}\vartheta, \mathrm{d}\varphi)$ associated with the local coordinate system (ϑ, φ).

Question 2: Let f be the real smooth function on X whose value at point $p = (x, y, z)$ is z. Please expand the total differential $\mathrm{d}f$ in terms of $\mathrm{d}\vartheta$, $\mathrm{d}\varphi$ and the gradient vector field $\mathrm{grad} f$ in terms of ∂_ϑ, ∂_φ.

Question 3: Express the 1st fundamental form in terms of $\mathrm{d}\vartheta$ and $\mathrm{d}\varphi$.

Solution. 1) Recall that $\phi(\vartheta, \varphi) = (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta)$ and $\partial_\vartheta = (1, \partial_\vartheta\iota)$, $\partial_\varphi = (1, \partial_\varphi\iota)$, with

Here, following the custom, we rewrite the local coordinate functions x^1 and x^2 as ϑ and φ respectively.

$$\partial_\vartheta\iota = \hat{\phi}_{\vec{e}_1} \circ \bar{\phi}^{-1} = \begin{bmatrix} \cos\vartheta\cos\varphi \\ \cos\vartheta\sin\varphi \\ -\sin\vartheta \end{bmatrix}, \quad \partial_\varphi\iota = \hat{\phi}_{\vec{e}_2} \circ \bar{\phi}^{-1} = \begin{bmatrix} -\sin\vartheta\sin\varphi \\ \sin\vartheta\cos\varphi \\ 0 \end{bmatrix}.$$

Since the local tangent frame $(\partial_\vartheta, \partial_\varphi)$ and the local cotangent frame $(\mathrm{d}\vartheta, \mathrm{d}\varphi)$ are dual of each other, we know that as one form on T_pS^2, $\mathrm{d}\vartheta_p$ and $\mathrm{d}\varphi_p$ map tangent vector

$$\left(p, \begin{bmatrix} u^1 \\ u^2 \\ u^3 \end{bmatrix}\right) \in T_pS^2$$

to real numbers

$$(u^1\cos\vartheta\cos\varphi + u^2\cos\vartheta\sin\varphi - u^3\sin\vartheta)|_p \quad \text{and} \quad \frac{-u^1\sin\varphi + u^2\cos\varphi}{\sin\vartheta}\Big|_p$$

respectively. This answer is obtained by solving the equation

$$\vec{u} = \langle \mathrm{d}\iota_p, (p, \vec{u})\rangle = \partial_\vartheta\iota|_p\langle \mathrm{d}\vartheta_p, (p, \vec{u})\rangle + \partial_\varphi\iota|_p\langle \mathrm{d}\varphi_p, (p, \vec{u})\rangle$$

by using the dot product with $\partial_\vartheta\iota|_p$ and $\partial_\varphi\iota|_p$. In other words, we have

$$\mathrm{d}\vartheta = (1, [\cos\vartheta\cos\varphi, \cos\vartheta\sin\varphi, -\sin\vartheta])$$

and

$$\mathrm{d}\varphi = \left(1, \frac{1}{\sin\vartheta}[-\sin\varphi, \cos\varphi, 0]\right).$$

2) Since $g_{ij} = \partial_i \iota \cdot \partial_j \iota$, we have

$$g_{\vartheta\vartheta} = 1, \quad g_{\vartheta\varphi} = 0, \quad g_{\varphi\varphi} = (\sin\vartheta)^2$$

and

$$g^{\vartheta\vartheta} = 1, \quad g^{\vartheta\varphi} = 0, \quad g^{\varphi\varphi} = \frac{1}{(\sin\vartheta)^2}.$$

Since $f = z = \cos\vartheta$, we have

$$\frac{\partial f}{\partial\vartheta} = -\sin\vartheta, \quad \frac{\partial f}{\partial\varphi} = 0$$

and

$$\mathrm{d}f = \frac{\partial f}{\partial\vartheta}\mathrm{d}\vartheta + \frac{\partial f}{\partial\varphi}\mathrm{d}\varphi = -\sin\vartheta\,\mathrm{d}\vartheta.$$

Since $\mathrm{grad} f = g^{ij}\,\partial_{x^i} f\,\partial_{x^j}$ with $x^1 = \vartheta$ and $x^2 = \varphi$, we have

$$\mathrm{grad} f = -\sin\vartheta\,\frac{\partial}{\partial\vartheta}.$$

3) Locally we have $\mathrm{I} = \mathrm{d}\vartheta^2 + (\sin\vartheta)^2\,\mathrm{d}\varphi^2$.

◀

Example 17. *Let M be the quadric $x^2 + y^2 - z^2 = -1$, (x, y) be the local coordinate system associated with imbedding*

$$\begin{aligned} \phi : \mathbb{E}^2 &\to M \\ (x, y) &\mapsto (x, y, -\sqrt{1 + x^2 + y^2}) \end{aligned} \tag{27}$$

1) Please express the first fundamental form in this local coordinate system.

2) Consider the real smooth function $f = xz$. Find $\mathrm{grad} f|_p$ for $p = (0, 0, -1)$.

Solution. 1) The easy version of ∂_x and ∂_y are

$$\partial_x \iota = \begin{bmatrix} 1 \\ 0 \\ -\frac{x}{\sqrt{1+x^2+y^2}} \end{bmatrix} \quad \text{and} \quad \partial_y \iota = \begin{bmatrix} 0 \\ 1 \\ -\frac{y}{\sqrt{1+x^2+y^2}} \end{bmatrix}.$$

So

$$[g_{ij}] = \frac{1}{1 + x^2 + y^2}\begin{bmatrix} 1 + 2x^2 + y^2 & xy \\ xy & 1 + x^2 + 2y^2 \end{bmatrix}.$$

and

$$[g^{ij}] = \frac{1}{1 + 2(x^2 + y^2)}\begin{bmatrix} 1 + x^2 + 2y^2 & -xy \\ -xy & 1 + 2x^2 + y^2 \end{bmatrix}.$$

Therefore

$$I = \frac{(1+2x^2+y^2)\,\mathrm{d}x^2 + 2xy\,\mathrm{d}x\,\mathrm{d}y + (1+x^2+2y^2)\,\mathrm{d}y^2}{1+x^2+y^2}.$$

2) Since $f = -x\sqrt{1-x^2-y^2} = -x + o(x^2+y^2)$ on the coordinate patch, we have

$$\partial_x f|_p = -1, \quad \partial_y f|_p = 0.$$

Also, $[g^{ij}]|_p = I$. Then

$$\mathrm{grad} f|_p = -\partial_x|_p = (p, -\vec{e}_1).$$

◀

Exercise 64. *Continuing the preceding example, let us assume that the orientation on M is given by normal vectors pointing towards the other hyperbolic sheet. Please express the second fundamental form* II *in this local coordinate system.*

`Answer:` $\vec{g} = \frac{1}{\sqrt{x^2+y^2+z^2}}[x, y, -z]^T$, $\partial_x\partial_y\iota = [0, 0, \frac{xy}{z^3}]^T$, `etc..`
`Then`

$$\mathrm{II} = \frac{1}{(1+x^2+y^2)\sqrt{1+2(x^2+y^2)}}\Big(-(1+y^2)\,\mathrm{d}x^2 + 2xy\,\mathrm{d}x\,\mathrm{d}y - (1+x^2)\,\mathrm{d}y^2\Big)$$

Formulae for local coordinate changes

Suppose that we have two local coordinate systems on n-manifold M, say x^i and $\tilde{x}^j$, and their corresponding coordinate patches overlap. Let W be the overlap, i.e., the common intersection of the two coordinate patches. Then, on the overlap W, we have

Please read from context. For example, $\mathrm{d}x^j$ is actually $\mathrm{d}(x^j|_W) = \mathrm{d}x^j|_W$. We prefer simplicity over precision here.

$$\mathrm{d}\tilde{x}^i = \frac{\partial \tilde{x}^i}{\partial x^j}\,\mathrm{d}x^j,$$

consequently we have

$$\frac{\partial \tilde{x}^i}{\partial x^j}\frac{\partial x^j}{\partial \tilde{x}^k} = \delta^i_k, \quad \partial_{x^i} = \frac{\partial \tilde{x}^j}{\partial x^i}\partial_{\tilde{x}^j} \quad \text{and} \quad \partial_{x^i} f = \frac{\partial \tilde{x}^j}{\partial x^i}\partial_{\tilde{x}^j} f.$$

Exercise 65. *Prove these identities .*

Suppose that X is a vector field on manifold M, α is a differential one-form on M, then there are unique real smooth functions X^i and α_j on the coordinate patch of the local coordinate system x^i such that, on this coordinate patch, we have

$$X = X^i\,\partial_{x^i} \quad \text{and} \quad \alpha = \alpha_i\,\mathrm{d}x^i.$$

Here, X^i (α_i resp.) is called the local representation of X (α resp.) with respect to the local coordinate system x^i. A simple computation shows that $\langle\alpha, X\rangle = \alpha_i X^i$ on this coordinate patch. Speaking of local representation, we know that $\partial_{x^i} f$ is the local representation $\mathrm{d}f$ and $g^{ij}\partial_{x^j} f$ is the local representation of $\mathrm{grad} f$.

Similarly, with local coordinate system $\tilde{x}^i$, we have

$$X = \tilde{X}^i\, \partial_{\tilde{x}^i} \quad \text{and} \quad \alpha = \tilde{\alpha}_i\, \mathrm{d}\tilde{x}^i.$$

It is then a simple exercise in linear algebra that, on the overlap W, we have

$$\tilde{X}^i = \frac{\partial \tilde{x}^i}{\partial x^j}\, X^j \quad \text{and} \quad \tilde{\alpha}_i = \frac{\partial x^j}{\partial \tilde{x}^i}\, \alpha_j.$$

Also, on the overlap W, we have

$$\tilde{g}_{ij} = \frac{\partial x^k}{\partial \tilde{x}^i}\frac{\partial x^l}{\partial \tilde{x}^j} g_{kl}, \quad \tilde{g}^{ij} = \frac{\partial \tilde{x}^i}{\partial x^k}\frac{\partial \tilde{x}^j}{\partial x^l} g^{kl}.$$

Exercise 66 (Rectangular Coordinate System)**.** *Let $M = \mathbb{E}^n$ and ϕ: $\mathbb{E}^n \to M$ be the identity map. Then the associated coordinate system x^i on M is just the standard coordinate system. I.e., if point p in M is $(q_1, \ldots, q_n)$, then $x^i(p) = q_i$.*

Show that the easy version of ∂_{x^i} is the constant vector-valued function with the constant value $\vec{e}_i \in \mathbb{R}^n$. Then conclude that the easy version of $\mathrm{d}x^i$ is the constant vector-valued function with the constant value $\hat{e}^i \in (\mathbb{R}^n)^*$. In other words, we have $\partial_{x^i}\big|_p = (p, \vec{e}_i)$ and $\mathrm{d}x^i\big|_p = (p, \hat{e}^i)$, or

$$\partial_{x^i} = (1, c_{\vec{e}_i}), \quad \mathrm{d}x^i = (1, c_{\hat{e}^i}).$$

Here c_v denotes the constant function with constant value v.

Show also that if f is a smooth real function on M, then $\partial_{x^i} f$ is precisely the i-th partial derivative function of f.

Finally, show that the first fundamental form is

$$\mathrm{I} = (\mathrm{d}x^1)^2 + \cdots + (\mathrm{d}x^n)^2.$$

Exercise 67 (Polar Coordinate System)**.** *Continuing the exercise above with $n = 3$. Rewrite x^1 as x, x^2 as y and x^3 as z. Let $U = (0, \infty) \times (0, \pi) \times (0, 2\pi)$. Then the polar coordinate system (r, ϑ, φ) on X is the local coordinate system associated with the local parametrization*

$$\begin{aligned} \psi : U &\to M \\ (r, \vartheta, \varphi) &\mapsto (r \sin\vartheta \cos\varphi, r \sin\vartheta \sin\varphi, r\cos\vartheta) \end{aligned} \tag{28}$$

1) Please express ∂_r, ∂_ϑ and ∂_φ in terms of ∂_x, ∂_y and ∂_z.

2) Please express the first fundamental form I on $\mathrm{Im}(\psi)$ in terms of $\mathrm{d}r$, $\mathrm{d}\vartheta$ and $\mathrm{d}\varphi$. `Answer:` $I = \mathrm{d}r^2 + r^2\mathrm{d}\vartheta^2 + r^2(\sin\vartheta)^2\mathrm{d}\varphi^2$.

3) Let f be a real smooth function on M, ∇f be the easy version of the gradient vector field $\mathrm{grad} f$. Please express ∇f in terms of polar coordinates.

Remark 30. On the standard euclidean space $\mathbb{E}^3$, the rectangular coordinate system is global. Although the polar coordinate system is local, it is nearly global in the sense that the coordinate patch is a dense and open subset of $\mathbb{E}^3$. In practice most people pretend the polar coordinate system to be global. That is OK, because, by using continuity argument, to show an identity on $\mathbb{E}^3$, it suffices to verify the identity on a dense and open subset.

List of Symbols

ϕ	a local representation of a manifold
x^i	the local coordinate system resulting from a local representation of a manifold
E_i	the i-th standard vector field on an open set of $\mathbb{E}^n$
$\partial_i, \partial_{x^i}, \frac{\partial}{\partial x^i}$	the i-th local vector field associated with a local coordinate system x^i
$\frac{\partial f}{\partial x^i}$	defined to be $\langle \mathrm{d}f, \partial_i \rangle$
g_{ij}	local representation of the first fundamental form with respect to a local coordinate system x^i
g^{ij}	the (i,j)-entry of the matrix inverse of the matrix $[g_{ij}]$
E, F, G	local representation of the first fundamental form of a surface in $\mathbb{E}^3$
e, f, g	local representation of the 2nd fundamental form of an oriented surface in $\mathbb{E}^3$

§8 Tensor and Tensor Field

Having extended the notion of space from open sets of affine spaces to manifolds, we are now ready to extend the notion of function from real functions to tensor fields. Tensor is a notion broader than real number, and tensor is to tensor field is what real number is to real smooth function.

Tensor

Throughout this section we assume that V is a real vector space of dimension $n \geq 1$ and V^* denotes the dual space of V. Recall that, once we have a basis $\mathcal{B}$ chosen for V, a vector $\underline{v}$ in V can be represented as a set of n real numbers v^i. Similarly, with respect to the dual basis $\mathcal{B}^*$ for V^*, an element α in V^* can be represented as a set of n real numbers α_i. Note that, in term of these representations, the pairing of α with $\underline{v}$, denoted by $\langle \alpha, \underline{v} \rangle$, is the real number $\alpha_i v^i$.

In the proceeding paragraph three types of tensor on V have appeared:

1. $\underline{v}$ — a tensor of type $(1, 0)$, represented by a set of real numbers v^i with the upper index i running from 1 to $\dim V$
2. α — a tensor of type $(0, 1)$, represented by a set of real numbers α_i with the lower index i running from 1 to $\dim V$
3. $\langle \alpha, \underline{v} \rangle$ — a tensor of type $(0, 0)$, a single real number. No index is needed

In general, a tensor on V of type (r, s) is represented by a set of real numbers $t^{i_1 \cdots i_r}_{j_1 \cdots j_s}$ with r upper indexes and s lower indexes (all running from 1 to $\dim V$), with the understanding that two sets of real numbers $t^{i_1 \cdots i_r}_{j_1 \cdots j_s}$ and $\tilde{t}^{\tilde{i}_1 \cdots \tilde{i}_r}_{\tilde{j}_1 \cdots \tilde{j}_s}$ (with respect to bases $\mathcal{B} = (\underline{v}_1, \ldots, \underline{v}_n)$ and $\tilde{\mathcal{B}} = (\underline{\tilde{v}}_1, \ldots, \underline{\tilde{v}}_n)$ respectively) represent the same tensor if they can be transformed into each other as follows: let $\underline{v}_k = \underline{\tilde{v}}_{\tilde{i}} a^{\tilde{i}}{}_k$ and $\underline{\tilde{v}}_{\tilde{j}} = \underline{v}_l b^l{}_{\tilde{j}}$, then

Note: matrices $[a^{\tilde{i}}{}_k]$ and $[b^i{}_{\tilde{k}}]$ are inverse of each other.

$$\tilde{t}^{\tilde{i}_1 \cdots \tilde{i}_r}_{\tilde{j}_1 \cdots \tilde{j}_s} = a^{\tilde{i}_1}{}_{k_1} \cdots a^{\tilde{i}_r}{}_{k_r} \, t^{k_1 \cdots k_r}_{l_1 \cdots l_s} \, b^{l_1}{}_{\tilde{j}_1} \cdots b^{l_s}{}_{\tilde{j}_s} \tag{29}$$

In other words, a tensor on V of type (r, s) is an equivalence class of the pair $(t^{i_1\cdots i_r}_{j_1\cdots j_s}, \mathcal{B})$ where $t^{i_1\cdots i_r}_{j_1\cdots j_s}$ is a set of real numbers and $\mathcal{B}$ is a basis for V. Note that the basis vector $\underline{v}_k$ in V is represented by δ^i_k and the basis vector $\hat{v}^k$ in V^* is represented by δ^k_i.

A tensor of type $(r,0)$ with $r > 0$ is called a contra-variant tensor of rank r, a tensor of type $(0,s)$ with $s > 0$ is called a covariant tensor of rank s, a tensor of type $(0,0)$ is also called scalar type.

The set of tensors on V of type (r,s) is a real vector space of dimension $(\dim V)^{r+s}$ and shall be denoted by $T^{r,s}V$. Note that $T^{0,0}V = \mathbb{R}$. Then the set of all tensors on V,

$$T^{\cdot,\cdot}V := \bigoplus_{r,s\geq 0} T^{r,s}V,$$

is a bi-graded real vector space, in fact, an associative bi-graded algebra with unity. Here the product is called the **tensor product** and is denoted by $\otimes$. Here is the formula for tensor product: Fixing a basis $\mathcal{B}$ for V, the tensor product of tensor x (represented by $(x^{i_1\cdots i_{r_1}}_{j_1\cdots j_{s_1}}, \mathcal{B})$) with tensor y (represented by $(y^{k_1\cdots k_{r_2}}_{l_1\cdots l_{s_2}}, \mathcal{B})$), denoted by $x \otimes y$, is the tensor represented by $(x^{i_1\cdots i_{r_1}}_{j_1\cdots j_{s_1}} y^{k_1\cdots k_{r_2}}_{l_1\cdots l_{s_2}}, \mathcal{B})$. We simply write this as

$$x^{i_1\cdots i_{r_1}}_{j_1\cdots j_{s_1}} \otimes y^{k_1\cdots k_{r_2}}_{l_1\cdots l_{s_2}} = x^{i_1\cdots i_{r_1}}_{j_1\cdots j_{s_1}} y^{k_1\cdots k_{r_2}}_{l_1\cdots l_{s_2}}.$$

Exercise 68. *Let $\mathcal{B} = (\underline{v}_1, \ldots, \underline{v}_n)$ be a basis for V. Show that $\{\underline{v}_i \otimes \underline{v}_j\}$ is a minimal spanning set for $T^{2,0}V$. Similarly, $\{\hat{v}^i \otimes \hat{v}^j\}$ is a minimal spanning set for $T^{0,2}V$.*

Hint: Fixing a basis $\mathcal{B}$ for V and then work with the resulting representation. The rank-2 contra-variant tensor $\underline{v}_i \otimes \underline{v}_j$ is represented by the set of real numbers $\delta^k_i \delta^l_j$. Let a rank-2 contra-variant tensor t be represented by the set of real numbers t^{kl}, then $t^{kl} = t^{ij}\delta^k_i\delta^l_j$ for all k and l, i.e., $t = t^{ij}\underline{v}_i \otimes \underline{v}_j$. Also, we need to check that equation $x^{ij}\underline{v}_i \otimes \underline{v}_j = 0$ (i.e. equations $x^{ij}\delta^k_i\delta^l_j = 0$ for all k and l) has only the trivial solution $x^{ij} = 0$ for all i and j.

For simplicity, we can hide the phrase "for all k and l" in writing.

Similarly, the rank-2 covariant tensor $\hat{v}^i \otimes \hat{v}^j$ is represented by the set of real numbers $\delta^i_k \delta^j_l$.

One can also see that the set of all covariant tensors on V and the set of all contravariant tensors on V are both associative graded algebra with unity.

Example 18. *The* **delta tensor** *δ is by definition the tensor of type $(1,1)$ that is represented by the Kronecker delta δ^i_j with respect to any basis of V. Does it make sense? Well, one can check that equality*

$$\delta^{\tilde{i}}_{\tilde{j}} = a^{\tilde{i}}{}_k\, \delta^k_l\, b^l{}_{\tilde{j}}$$

indeed holds.

A tensor on V of type $(1,1)$ is nothing but a linear map from V to V. Indeed, if t^i_j is the representation with respect to the basis $\mathcal{B} = (\underline{v}_1, \ldots, \underline{v}_n)$, then the linear map is the one that sends vector x^j to vector $t^i_j x^j$. In particular, the delta tensor on V is just the identity map on V.

A tensor on V of type (r,s), with representation $t^{i_1\cdots i_r}_{j_1\cdots j_s}$, can be naturally identified with a linear map from $T^{s,0}V$ to $T^{r,0}V$ this way: it sends $x^{j_1\cdots j_s}$ to $t^{i_1\cdots i_r}_{j_1\cdots j_s} x^{j_1\cdots j_s}$. More generally, a tensor on V of type (r_2+s_1, s_2+r_1) can be naturally identified with a linear map from $T^{r_1,s_1}V$ to $T^{r_2,s_2}V$ this way: it sends $x^{j_1\cdots j_{r_1}}_{i_1\cdots i_{s_1}}$ to $t^{k_1\cdots k_{r_2} i_1\cdots i_{s_1}}_{l_1\cdots l_{s_2} j_1\cdots j_{r_1}} x^{j_1\cdots j_{r_1}}_{i_1\cdots i_{s_1}}$. As an example, we have the natural identification

$$(T^{r,s}V)^* \equiv T^{s,r}V.$$

In particular we have $(T^{r,r}V)^* \equiv T^{r,r}V$. In case $r = 1$, this says that $\mathrm{End}(V) := \mathrm{Hom}(V,V)$ is naturally identified with its dual space. Under the identification $T^{1,1}V \equiv \mathrm{End}(V) \equiv (\mathrm{End}(V))^*$, we have the identification $\delta \equiv 1 \equiv \mathrm{tr}$. Here tr is the trace map:

$$x^j_i \mapsto \delta^i_j x^j_i = x^i_i$$

and 1 is that identity map on V: $v^i \mapsto \delta^i_j v^j = v^i$.

Remark 31. $T^{r,0}V = V^{\otimes r}$ — the tensor product of r copies of V and $T^{0,s}V = (V^*)^{\otimes s}$ — the tensor product of s copies of V^*. Also, $T^{r,s}V \equiv (V^*)^{\otimes s} \otimes V^{\otimes r} \equiv \mathrm{Hom}(V^{\otimes s}, V^{\otimes r})$; more generally

$$T^{s_1+r_2, s_2+r_1}V \equiv \mathrm{Hom}(T^{r_1,s_1}V, T^{r_2,s_2}V).$$

A contra-variant tensor $t^{i_1\cdots i_r}$ is called symmetric if

$$t^{i_{\sigma(1)}\cdots i_{\sigma(r)}} = t^{i_1\cdots i_r}$$

for any permutation σ and is called skew-symmetric if

$$t^{i_{\sigma(1)}\cdots i_{\sigma(r)}} = \mathrm{sign}(\sigma) t^{i_1\cdots i_r}$$

for any permutation σ. Here, $\mathrm{sign}(\sigma)$ is the sign of σ.

sign: $\Sigma_n \to \{\pm 1\}$ is a group homomorphism.

The set of contra-variant symmetric tensors on V of rank r, denoted by S^rV, is a vector space, and the set of all contra-variant symmetric tensors on V,

$$S^{\cdot}V := \bigoplus_{r\geq 0} S^rV,$$

is an associative and commutative graded algebra with unity. The product xy here — the symmetric tensor product of x in S^rV with y in S^sV— can be written as

$$(xy)^{i_1\cdots i_r i_{r+1}\cdots i_{r+s}} = \frac{1}{(r+s)!} \sum_{\sigma\in\Sigma_{r+s}} x^{i_{\sigma(1)}\cdots i_{\sigma(r)}} y^{i_{\sigma(r+1)}\cdots i_{\sigma(r+s)}}.$$

Once can show that

$$\sum t^r \dim S^r V = (1-t)^{-\dim V}.$$

A smart way to prove this identity is to observe that the graded algebra $S^\cdot V$ is isomorphic to the real polynomial algebra in n ($:= \dim V$) variables, which is in turn isomorphic to the tensor product of n copies of the real polynomial algebra $\mathbb{R}[x]$ in one variable x, so

$$\sum t^r \dim S^r V = \mathrm{tr}_{S^\cdot V} t^{\deg} = \left(\mathrm{tr}_{\mathbb{R}[x]} t^{\deg}\right)^n = (1-t)^{-n}.$$

Exercise 69. *Let $\mathcal{B} = (\underline{v}_1, \ldots, \underline{v}_n)$ be a basis for V. Show that $\{\underline{v}_i \underline{v}_j \mid 1 \le i \le j \le n\}$ is a minimal spanning set for $S^2 V$.*

The set of contra-variant skew-symmetric tensors on V of rank r, denoted by $\wedge^r V$, is a vector space, and the set of all contra-variant skew-symmetric tensors on V,

$$\wedge^\cdot V := \bigoplus \wedge^r V,$$

is an associative and *graded commutative* graded algebra with unity. The product $x \wedge y$ here — the skew-symmetric tensor product (also called wedge product) of x with y — can be written as

$$(x\wedge y)^{i_1\cdots i_r i_{r+1}\cdots i_{r+s}} = \frac{1}{(r+s)!}\sum_{\sigma\in\Sigma_n} \mathrm{sign}\,(\sigma) x^{i_{\sigma(1)}\cdots i_{\sigma(r)}} y^{i_{\sigma(r+1)}\cdots i_{\sigma(r+s)}}.$$

Once can show that

$$\sum t^r \dim \wedge^r V = (1+t)^{\dim V}.$$

Here is the law of graded commutativity for wedge product:

$$x \wedge y = (-1)^{|x||y|} y \wedge x$$

where $|x|$ denotes the rank of the homogeneous element x.

Exercise 70. *Let $\mathcal{B} = (\underline{v}_1, \ldots, \underline{v}_n)$ be a basis for V. Show that $\{\underline{v}_i \wedge \underline{v}_j \mid 1 \le i < j \le n\}$ is a minimal spanning set for $\wedge^2 V$.*

Similarly, one can talk about co-variant symmetric tensor as well as co-variant skew-symmetric tensor. We shall use $S_r V$ to denote the vector space of all co-variant symmetric tensors on V of rank r and $\wedge_r V$ to denote the vector space of all co-variant skew-symmetric tensors on V of rank r. Note that we have the identification

$$(S^r V)^* \equiv S_r V \equiv S^r V^*, \quad (\wedge^r V)^* \equiv \wedge_r V \equiv \wedge^r V^*.$$

Determinant Line and Orientation

Let $n = \dim V$. One can see easily that $\wedge^n V$ is an one-dimensional real vector space, moreover, if T is an endomorphism on V, then the induced endomorphism on $\wedge^n V$ is the scalar multiplication by $\det T$. Thus $\wedge^n V$ is called the **determinant line** of V and is rewritten as $\det V$. Concerning orientation, here is a basic identification:

$$o(V) \equiv o(V^*) \equiv o(\det V) \equiv o(\det V^*)$$

in the sense that

$$[(v_1, \ldots, v_n))] \equiv [(\hat{v}^1, \ldots, \hat{v}^n)] \equiv [v_1 \wedge \cdots \wedge v_n] \equiv [\hat{v}^1 \wedge \cdots \wedge \hat{v}^n].$$

These natural identifications enable us to describe orientations on V in *4 different ways.*

Tensor Field

Real smooth functions, vector fields, differential one-forms, the 1st and 2nd fundamental forms are all examples of tensor fields on manifolds.

Let M be a manifold and p be a point in M. A tensor on T_pM of type (r, s) is called a tensor of M of type (r, s) at point p. By definition, a **tensor field** t on M of type (r, s) is a smooth function on M whose value at any point p is a tensor of M of type (r, s) at point p. In a coordinate patch with coordinates x^i, the associated tangent frame ∂_{x^i} enables us to represent a tensor field t of type (r, s) by a set of real smooth functions $t^{i_1 \cdots i_r}_{j_1 \cdots j_s}$. So, on the overlap with another coordinate patch with coordinates $\tilde{x}^i$, we have

Recall that a vector field on M is a smooth function on M whose value at any point p is a tangent vector at point p

$$\tilde{t}^{i_1 \cdots i_r}_{j_1 \cdots j_s} = \frac{\partial \tilde{x}^{i_1}}{\partial x^{k_1}} \cdots \frac{\partial \tilde{x}^{i_r}}{\partial x^{k_r}} \frac{\partial x^{l_1}}{\partial \tilde{x}^{j_1}} \cdots \frac{\partial x^{l_s}}{\partial \tilde{x}^{j_s}} t^{k_1 \cdots k_r}_{l_1 \cdots l_s}.$$

This identity is complicated-looking. A short hand notation for it is $\tilde{t}^I_J = \frac{\partial \tilde{x}^I}{\partial x^K} \frac{\partial x^L}{\partial \tilde{x}^J} t^K_L$ where I, J, K and L are multi-indices (for example, $I = (i_1, \ldots, i_r)$) and

To remember the transformation rule easily, one may write $\tilde{t}^I_J = \frac{\partial \tilde{x}^I}{\partial x^K} t^K_L \frac{\partial x^L}{\partial \tilde{x}^J}$.

$$\frac{\partial \tilde{x}^I}{\partial x^K} = \frac{\partial \tilde{x}^{i_1}}{\partial x^{k_1}} \cdots \frac{\partial \tilde{x}^{i_r}}{\partial x^{k_r}}.$$

Then a real smooth function is a tensor field of type $(0, 0)$, a vector field is a tensor field of type $(1, 0)$, a differential one-form is a tensor field of type $(0, 1)$, the first fundamental form is a tensor field of type $(0, 2)$. A tensor field of type $(r, 0)$ is called a contra-variant tensor field of rank r and a tensor field of type $(0, s)$ is called a covariant tensor field of rank s. A covariant tensor field t is called symmetric if $t_I = t_{\sigma(I)}$ for any permutation σ. For example, the first fundamental form is symmetric because $g_{ij} = g_{ji}$. A covariant tensor field t is called skew-symmetric if $t_I = \operatorname{sign}(\sigma) t_{\sigma(I)}$ for any permutation σ. For example, on $\mathbb{E}^2$, $\omega := \mathrm{d}x^1 \wedge \mathrm{d}x^2$ is a skew-symmetric because $\omega_{12} = \frac{1}{2}$ and $\omega_{21} = -\frac{1}{2} = -\omega_{12}$.

Let M be an n-manifold. A skew-symmetric covariant tensor field α of rank k on M is called a differential k-form on M. Locally we write

$$\alpha = \alpha_I \, \mathrm{d}x^I$$

where I is the multi-index $(i_1, \ldots, i_k)$, so $\alpha_I = \alpha_{i_1,\ldots,i_k}$ and $\mathrm{d}x^I = \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$.

The set of differential k-forms on M is denoted by $\Omega^k(M)$. By convention, $\Omega^0(M) = C^\infty(M)$ — the ring of real smooth functions on M. It is an easy fact that $\Omega^k(M)$ is a module over the ring $C^\infty(M)$.

Note that $\Omega^k(M) = 0$ if $k > n$. By convention, we let $\Omega^k(M) = 0$ if $k < 0$. Let $\Omega^\cdot(M) = \bigoplus_k \Omega^k(M)$. Then $\Omega^\cdot(M)$ is a graded module over the ring $C^\infty(M)$, in fact, a graded algebra over $C^\infty(M)$. The $C^\infty(M)$-bilinear product in this algebra, denoted by $\wedge$, is called the wedge product. This product is associative, unital with the constant function 1 as its unit, and is graded commutative in the sense that

$$\alpha \wedge \beta = (-1)^{|\alpha||\beta|} \beta \wedge \alpha$$

for any two differential forms α and β. Here, $|\alpha|$ denotes the degree (i.e., rank) of α. Note that, in case f is a real smooth function, $f \wedge \alpha = \alpha \wedge f$ is the same as the scalar multiplication $f\alpha$.

A skew-symmetric contravariant tensor P of rank p on M is called a p-vector field on M, and locally we write

$$P = P^I \frac{\partial}{\partial x^I}$$

where I is the multi-index $(i_1, \ldots, i_k)$, so $P^I = P^{i_1,\ldots,i_k}$ and $\frac{\partial}{\partial x^I} = \frac{\partial}{\partial x^{i_1}} \wedge \cdots \wedge \frac{\partial}{\partial x^{i_p}}$.

The set of k-vector fields on M is denoted by $A_k(M)$. By convention, $A_0(M) = C^\infty(M)$ — the ring of real smooth functions on M, and $A_k(M) = 0$ if $k < 0$. It is an easy fact that $A_k(M)$ is a module over the ring $C^\infty(M)$. Also, the graded module

$$A.(M) := \bigoplus A_k(M)$$

is an associative, unital and graded commutative algebra over $C^\infty(M)$.

Remark 32. Denote by $\Gamma(T^{r,s}M)$ the space of type-(r,s) tensor fields on M, then $A_1(M) = \Gamma(T^{1,0}M)$, $\Omega^1(M) = \Gamma(T^{0,1}M)$, and both are modules over the commutative ring $C^\infty(M)$. It is a fact that $\Omega^1(M) \equiv \mathrm{Hom}_{C^\infty(M)}(A_1(M), C^\infty(M))$. More generally, we have

For example, $\mathrm{d}f \equiv$ the $C^\infty(M)$-linear map: $X \mapsto \langle \mathrm{d}f, X\rangle$.

$$\Gamma(T^{s_1+r_2,s_2+r_1}M) \equiv \mathrm{Hom}_{C^\infty(M)}(\Gamma(T^{r_1,s_1}M), \Gamma(T^{r_2,s_2}M)).$$

Having extended the notion of functions, we shall be ready to extend the notion of differentiation in the next few sections. There are three inter-related types:

i) the **exterior differentiation** of differential forms,

ii) the **Lie differentiation** of tensor fields,

iii) **covariant differentiation** of tensor fields. Note: the set of covariant differentiations is a real affine space of infinite dimension.

For details, please continue.

§9 The Exterior Differentiation

This chapter aims to extend partial differentiation from real smooth functions (i.e., differential 0-forms) to all differential forms.

Recall that, on a manifold M, taking partial derivatives turns a real smooth function f to its total differential $\mathrm{d}f$ — a differential one-form rather than a real smooth function. So the space of real smooth functions on M is not complete with respect to taking partial derivatives, i.e., applying the operator d. What would be the completion then? Here is the answer: $\Omega^\cdot(M)$ — the space of all differential forms on M.

Indeed, the operator d: $\Omega^0(M) \to \Omega^1(M)$ can be extended naturally to the whole $\Omega^\cdot(M)$. In terms of local coordinates x^i, we have the formulae

$$\mathrm{d} = \mathrm{d}x^i \wedge \partial_i. \tag{30}$$

For example, if f is a real smooth function, then, locally we have

$$\mathrm{d}f = \mathrm{d}x^i \wedge \partial_i f = \partial_i f \mathrm{d}x^i$$

which is indeed the total differential of f. In general, if $\alpha = \alpha_I \mathrm{d}x^I$, we have

$$\mathrm{d}\alpha = \mathrm{d}x^i \wedge (\partial_i \alpha_I)\mathrm{d}x^I = \partial_i \alpha_I \, \mathrm{d}x^i \wedge \mathrm{d}x^I.$$

By a straightforward computation, one can check that, on the overlap of two coordinate charts with coordinates x^i and $\tilde{x}^j$, we have

$$\partial_{x^i} \alpha_I \, \mathrm{d}x^i \wedge \mathrm{d}x^I = \partial_{\tilde{x}^j} \tilde{\alpha}_J \, \mathrm{d}\tilde{x}^i \wedge \mathrm{d}\tilde{x}^J,$$

so $\mathrm{d}\alpha$ is well-defined on the whole manifold M.

It is clear that

$$\mathrm{d}^2\alpha := \mathrm{d}(\mathrm{d}\alpha) = \partial_j \partial_i \alpha_{i_1 \cdots i_k} \, \mathrm{d}x^j \wedge \mathrm{d}x^i \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} = 0.$$

In other words,

$$\mathrm{d}^2 = 0. \tag{31}$$

de Rham complex

The **exterior derivative operator** d introduced above is a degree-one graded derivative operator on the graded algebra $\Omega^{\cdot}(M)$ of differential forms over the ring $C^\infty(M)$. This means that the operator d is $\mathbb{R}$-linear, turns a differential k-form into a differential $(k+1)$-form, and satisfies the *graded product rule*:

$$\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^{|\alpha|}\alpha \wedge \mathrm{d}\beta \tag{32}$$

as one can verify easily by a direct local computation.

Remark 33. The exterior derivative operator d is the *unique* degree-one graded derivative operator on the algebra of differential forms such that

0) d is natural transformation from the functor $\Omega^{\cdot}$ to functor $\Omega^{\cdot}$,
1) $\mathrm{d}f$ is the total differential form of f,
2) $\mathrm{d}^2 = 0$.

Please note that the map $\mathrm{d}\colon \Omega^0(M) \to \Omega^1(M)$ is not linear over $C^\infty(M)$, so it is not a tensor field of type $(0,1)$. In short, the exterior differentiation is not a tensor field.

Exercise 71. *Please prove the statement in the above remark.*

```
Hint: just do a local computation.
```

Remark 34. The differential complex

$$0 \to \Omega^0(M) \xrightarrow{\mathrm{d}} \Omega^1(M) \xrightarrow{\mathrm{d}} \cdots \xrightarrow{\mathrm{d}} \Omega^n(M) \to 0$$

is called the *de Rham complex* of the n-manifold M, and the resulting k-th cohomology space, denoted by $H^k_{dR}(M)$, is called the k-th **de Rham cohomolgy group** of M. By definition,

$$H^k_{dR}(M) = \frac{\ker(\Omega^k(M) \xrightarrow{\mathrm{d}} \Omega^{k+1}(M))}{\mathrm{Image}(\Omega^{k-1}(M) \xrightarrow{\mathrm{d}} \Omega^k(M))}.$$

In case M is a compact, connected, orientable manifold of dimension n, the Poincaré duality says that $H^k_{dR}(M) \cong H^{n-k}_{dR}(M)$, in particular, we have

$$H^n_{dR}(M) \cong H^0_{dR}(M) \cong \mathbb{R}.$$

Note: The wedge product on differential forms passes to de Rham cohomology classes so that

$$H^{\cdot}_{dR}(M) := \bigoplus H^k_{dR}(M)$$

becomes an associative, graded commutative ring with unity (in fact a graded algebra over $\mathbb{R}$) and is called the **de Rham cohomolgy algebra** of M. It is a fact that $H^k_{dR}(M) = H^k(M) \otimes_{\mathbb{Z}} \mathbb{R}$ — the tensor product of the integral cohomology group of M with $\mathbb{R}$ over $\mathbb{Z}$.

The pairing between differential forms and multi-vector fields

The pairing $\langle\,,\,\rangle$ of differential 1-forms with vector fields extends to the **pairing** of differential k-forms with k-vector fields. By convention, for functions f and g, we have the pairing $\langle f, g\rangle = fg$. Assume that $k > 0$, then the pairing is the $C^\infty(M) := C^\infty(M, \mathbb{R})$-bilinear map

$$\langle\,,\,\rangle : \quad \Omega^k(M) \times A_k(M) \to C^\infty(M) \tag{33}$$

such that

$$\langle \alpha_1 \wedge \cdots \wedge \alpha_k, X_1 \wedge \cdots X_k\rangle = \det[\langle \alpha_i, X_j\rangle].$$

Locally, we write $\alpha = \alpha_I\, \mathrm{d}x^I$, $X = X^I \frac{\partial}{\partial x^I}$, then

$$\langle \alpha, X\rangle = |\alpha|!\, \alpha_I X^I.$$

Interior product

It is a simple fact that the pairing above is non-degenerate, so the operator $X\wedge$ on multi-vector fields has its adjoint operator, ι_X, on differential forms. In case ω is a differential $(k+1)$-form and K is a k-vector field, then we have the defining equation

$$\langle \iota_X\omega, K\rangle = \langle \omega, X \wedge K\rangle \tag{34}$$

for the interior product operator ι_X. For example, if α is a differential one-form, then

$$\iota_X\alpha = \langle \alpha, X\rangle.$$

By convention $\iota_X f = 0$ for any real smooth function f.

Exercise 72. *Let X be a vector field and α, $\alpha_0, \ldots, \alpha_k$ be differential one-forms. Show that*

(i) $\iota_X\alpha = \langle \alpha, X\rangle$.

(ii) $\iota_X(\alpha_0 \wedge \cdots \wedge \alpha_k) = \sum_{i=0}^k (-1)^i \iota_X\alpha_i\ \alpha_0 \wedge \cdots \wedge \hat{\alpha}_i \wedge \cdots \wedge \alpha_k$ where $\hat{\alpha}_i$ means the term α_i is missing.

In general, if P is a p-vector field, then ι_P is a *degree $-p$ graded derivation* on the graded algebra $(\Omega^\bullet(M), \wedge)$ with the defining equation

$$\langle \iota_P\omega, K\rangle = \langle \omega, P \wedge K\rangle \tag{35}$$

Here we assume K is a k-vector field and ω is a differential form of degree $p + k$. In particular, we have $\iota_P\omega = \langle \omega, P\rangle$ if $|\omega| = |P|$.

§10 The Lie Differentiation

The Lie differentiation is a differentiation on tensor fields: for each vector field X, there is exactly one tensor-type preserving derivative operator, $\mathcal{L}_X$, on the vector space of tensor fields. In its authentic definition, the Lie differentiation of tensor fields on a manifold measures the infinitesimal change of tensor fields under *infinitesimal symmetries* of that manifold (i.e., *vector fields* on that manifold). However, for technical simplicity we shall introduce the Lie differentiation in an alternative way.

Lie Bracket

Let X and Y be two vector fields on manifold M, and x^i and $\tilde{x}^j$ be two coordinate systems with overlapped coordinate patches. Let us write $X = X^i\,\partial_{x^i}$, $X = \tilde{X}^i\,\partial_{\tilde{x}^i}$, $Y = Y^i\,\partial_{x^i}$, $Y = \tilde{Y}^i\,\partial_{\tilde{x}^i}$. Introduce local functions

$$Z^i := X^j\frac{\partial Y^i}{\partial x^j} - Y^j\frac{\partial X^i}{\partial x^j} \quad \text{and} \quad \tilde{Z}^i := \tilde{X}^j\frac{\partial \tilde{Y}^i}{\partial \tilde{x}^j} - \tilde{Y}^j\frac{\partial \tilde{X}^i}{\partial \tilde{x}^j}.$$

A simple computation shows that $\tilde{Z}^i = \frac{\partial \tilde{x}^i}{\partial x^j} Z^j$ on the overlap W. Therefore we have a unique vector field on M, denoted by $[X, Y]$, such that

$$[X, Y] = Z^i \partial_{x^i}$$

on the coordinate patch associated with the local coordinate system x^i. This vector field is called the Lie bracket of X with Y. It is obvious that

This is a quick but *uninspiring* way to introduce the Lie bracket of two vector fields.

$$[\partial_{x^i}, \partial_{x^j}] = 0.$$

Let X, Y and Z be vector fields, and f be a real smooth function. One can check easily that

Please note that the Lie bracket $[\,,\,]\colon A_1(M) \times A_1(M) \to A_1(M)$ is not bilinear over $C^\infty(M)$, so it is not a tensor field of type $(1, 2)$. In short, the Lie bracket is not a tensor field.

0) $[X, Y + Z] = [X, Y] + [X, Z]$,
1) $[X, Y] = -[Y, X]$,
2) $[X, fY] = \langle \mathrm{d}f, X\rangle Y + f[X, Y]$,
3) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.

Exercise 73. *Prove these identities and then conclude that the space of smooth vector fields on M is a real Lie algebra under the Lie bracket.*

Lie derivative

Let X be a vector field on M and f be a real smooth function on M. There is something called the Lie derivative of f with respect to X, denoted by $\mathcal{L}_X f$. Instead of telling you its original definition here, let me write down an identity

$$\mathcal{L}_X f = \langle \mathrm{d}f, X\rangle \tag{36}$$

which shall serve as an alternative definition for $\mathcal{L}_X f$. Obviously we have

$$\mathcal{L}_{\partial_{x^i}} f = \frac{\partial f}{\partial x^i}.$$

It is clear that
1) $\mathcal{L}_X(f+g) = \mathcal{L}_X f + \mathcal{L}_X g$ and $\mathcal{L}_X(fg) = f\,\mathcal{L}_X g + g\,\mathcal{L}_X f$,
2) $\mathcal{L}_{gX} f = g\,\mathcal{L}_X f$,
3) $(\mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X - \mathcal{L}_{[X,Y]})f = 0$.

Exercise 74. *Prove these identities.*

```
Note: identity 3) can be used as the defining equation
for Lie bracket.
```

Let Y be another vector field on M. There is also something called the Lie derivative of Y with respect to X, denoted by $\mathcal{L}_X Y$. Again, instead of telling you its original definition here, let me write down an identity

$$\mathcal{L}_X Y = [X, Y] \tag{37}$$

which shall serve as an alternative definition for $\mathcal{L}_X Y$. It is easy to verify that
1) $\mathcal{L}_X(Y+Z) = \mathcal{L}_X Y + \mathcal{L}_X Z$ and $\mathcal{L}_X(fY) = f\,\mathcal{L}_X Y + \mathcal{L}_X f\,Y$,
2) $\mathcal{L}_{gX} Y \neq g\mathcal{L}_X Y$ in general,
3)$\mathcal{L}_X([Y,Z]) = [\mathcal{L}_X Y, Z] + [Y, \mathcal{L}_X Z]$,

Exercise 75. *Prove these identities.*

Continuing Example 14, if we let M be the coordinate patch there, then

$$\mathcal{L}_{\partial_\vartheta} f = \frac{\partial f}{\partial \vartheta} = -\sin\vartheta, \quad \mathcal{L}_{\partial_\varphi} f = \frac{\partial f}{\partial \varphi} = 0, \quad \mathcal{L}_{\mathrm{grad} f} f = -\sin\vartheta \frac{\partial f}{\partial \vartheta} = (\sin\vartheta)^2$$

and

$$\mathcal{L}_{\partial_\vartheta} \mathrm{grad} f = -\cos\vartheta \frac{\partial}{\partial \vartheta} = -\mathcal{L}_{\mathrm{grad} f} \partial_\vartheta.$$

The Lie derivative operator $\mathcal{L}_X$ extends uniquely to tensor fields of any type. The key to the existence of this extension is the product rule for $C^\infty(M)$-multilinear product of tensor fields. In particular, this means that, if α is a differential one-form, Y is another vector field, then identity

$$\mathcal{L}_X(\langle\alpha, Y\rangle) = \langle\mathcal{L}_X\alpha, Y\rangle + \langle\alpha, \mathcal{L}_X Y\rangle,$$

required by the product rule, serves as the defining equation for $\mathcal{L}_X\alpha$. Then, by a local computation, one can find that, locally

$$\mathcal{L}_X\alpha = (X^k\partial_k\alpha_i + \alpha_k\partial_i X^k)\,\mathrm{d}x^i.$$

More generally, since the pairing of differential k-form ω with k-vector field K is $C^\infty(M)$-bilinear, we have

$$\mathcal{L}_X(\langle\omega, K\rangle) = \langle\mathcal{L}_X\omega, K\rangle + \langle\omega, \mathcal{L}_X K\rangle.$$

One can see this from the local formula $\langle\omega, K\rangle = k!\omega_I K^I$. As another example, the product $\mathrm{I}(Y, Z) = g_{ij}Y^iZ^j$ is $C^\infty(M)$-trilinear, so we have

$$\mathcal{L}_X(\mathrm{I}(Y, Z)) = \mathcal{L}_X\mathrm{I}(Y, Z) + \mathrm{I}(\mathcal{L}_X Y, Z) + \mathrm{I}(Y, \mathcal{L}_X Z).$$

Exercise 76. *Show that the local formula for $\mathcal{L}_X\alpha$ above makes sense, i.e., on the overlap of any two coordinate patches, the two local formulae that correspond to the two local coordinate systems always give the same answer.*

We say the vector field X is a Killing vector field of the Riemannian manifold (M, I) if $\mathcal{L}_X\mathrm{I} = 0$, i.e., X is an infinitesimal symmetry of the Riemannian manifold (M, I).

Exercise 77. *Locally the first fundamental form is customarily written as*

$$\mathrm{I} = g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j.$$

Writing $X = X^i\,\partial_i$ locally and let $X_i = g_{ij}X^j$, show that, locally

$$\mathcal{L}_X\mathrm{I} = (X_{i;j} + X_{j;i})\,\mathrm{d}x^i\,\mathrm{d}x^j.$$

Here, $X_{i;j} = \partial_j X_i - \Gamma^k_{ij}X_k$ with Γ^i_{jk} being the Christoffel symbols.

Exercise 78. *Assume the standard global rectangular coordinate system (x^1, x^2, x^3) on $\mathbb{E}^3$, show that the vector fields*

$$J_1 = -x^2\partial_3 + x^3\partial_2,\ J_2 = -x^3\partial_1 + x^1\partial_3,\ J_3 = -x^1\partial_2 + x^2\partial_1$$

are Killing vector fields on $\mathbb{E}^3$ and satisfy commutation relations

$$[J_1, J_2] = J_3,\ [J_2, J_3] = J_1,\ [J_3, J_1] = J_2.$$

Hint: Note that $[X, Y] = \mathcal{L}_X Y = -\mathcal{L}_Y X$, so you can use the identities $\mathcal{L}_{\partial_i}x^j = \delta^j_i$ and $\mathcal{L}_{\partial_i}\partial_j = 0$. For example, $\mathcal{L}_{J_1}x^1 = -\mathcal{L}_{x^2\partial_3}x^1 + \mathcal{L}_{x^3\partial_2}x^1 = -x^2\delta^1_3 + x^3\delta^1_2 = 0$, $\mathcal{L}_{J_1}\partial_3 = -\mathcal{L}_{\partial_3}J_1 = \mathcal{L}_{\partial_3}(x^2\partial_3) - \mathcal{L}_{\partial_3}(x^3\partial_2) = (\mathcal{L}_{\partial_3}x^2)\partial_3 - (\mathcal{L}_{\partial_3}x^3)\partial_2 = 0\cdot\partial_3 - 1\cdot\partial_2 = -\partial_2$.

Cartan's Formula

Here is a simple fact: if α is differential one form, then

$$\langle \mathrm{d}\alpha, X \wedge Y\rangle = \langle \mathcal{L}_X\alpha, Y\rangle - \langle \mathcal{L}_Y\alpha, X\rangle - \langle \alpha, \mathcal{L}_Y X\rangle \tag{38}$$

for any two vector fields X and Y. This fact can be checked by a local computation.

Formula (38) can be rewritten as

$$\langle \iota_X(\mathrm{d}\alpha), Y\rangle = \langle \mathcal{L}_X\alpha, Y\rangle - \mathcal{L}_Y(\langle \alpha, X\rangle) = \langle \mathcal{L}_X\alpha, Y\rangle - \langle \mathrm{d}\langle \alpha, X\rangle, Y\rangle$$

for any vector field Y, so it is equivalent to identity

$$(\iota_X \circ \mathrm{d})\alpha = \mathcal{L}_X\alpha - \mathrm{d}\langle \alpha, X\rangle = \mathcal{L}_X\alpha - (\mathrm{d} \circ \iota_X)\alpha$$

for any one-form α. Therefore, as operators on $\Omega^1(M)$, we have $\mathcal{L}_X = \mathrm{d} \circ \iota_X + \iota_X \circ \mathrm{d}$ for any vector field X. More generally we have

Theorem 5 (Cartan's Formula)**.** *As operators on differential forms,*

$$\mathcal{L}_X = \mathrm{d} \circ \iota_X + \iota_X \circ \mathrm{d} \tag{39}$$

for any vector field X, consequently, $\mathcal{L}_X \circ \mathrm{d} = \mathrm{d} \circ \mathcal{L}_X$.

Exercise 79. *Prove this theorem.*

Hint: it suffices to show that $\mathrm{d} \circ \iota_X + \iota_X \circ \mathrm{d}$ is derivative operator on differential forms such that it maps function f to function $\mathcal{L}_X f$ and differential one-form α to differential one-form $\mathcal{L}_X\alpha$. Why?

Remark 35. Cartan's Formula says that Lie differentiation $\mathcal{L}$ and exterior differentiation d determine each other.

Let us end this chapter with the following easy fact: The map that sends vector field X to derivative operator $\mathcal{L}_X$ is a Lie algebra homomorphism, i.e., a linear map from the real Lie algebra $(A_1(M), [\,,\,])$ to the real Lie algebra of derivations on the vector space of tensor fields such that

$$\mathcal{L}_{[X,Y]} = \mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_X \circ \mathcal{L}_Y.$$

This identity is obviously true because we know already it is true on functions and on vector fields.

Appendix: Integral curve, flow map, and Lie differentiation

The purpose of this appendix is to introduce the authentic definition of Lie differentiation. To do that, we need to introduce a very useful fact:

Any vector field X on a compact n-manifold M is an infinitesimal symmetry of M in the sense that its integration Φ_X is a one-parameter family of global symmetries of M. Here, the smooth map

$$\Phi_X : \mathbb{R} \times M \to M$$

is a group action of $\mathbb{R}$ on M such that, for any point $p \in M$, the parametrized curve in M, $c_p(t) := t \cdot p$ (i.e., $\Phi_X(t, p)$), is the *integral curve* of X through point p:

$$\begin{cases} \dot{c}_p(t) &= X(c_p(t)) \\ c_p(0) &= p \end{cases} \tag{40}$$

In other words,

$$\begin{cases} \partial_t \Phi_X(t, p) &= X(\Phi_X(t, p)) \\ \Phi_X(0, p) &= p \end{cases} \tag{41}$$

Remark 36. One might call Φ_X the "total" integral curve of X, instead, we call it the **flow map** for X.

In case M is not compact, the integration Φ_X of X still exists, but its domain is just an open neighborhood of $0 \times M$ in $\mathbb{R} \times M$ and is only *germ-unique*: any two such integrations always agree on the intersection of their domains. In general, the flow is incomplete, so the group action is incomplete in the sense that the “associativity" law: “for each point p in M, we have

$$t_1 \cdot (t_2 \cdot p) = (t_1 + t_2) \cdot p,"$$

holds only when t_1 and t_2 are sufficiently small.

In applications, we only need the flow information for sufficiently small time t, so it makes no difference whether M is compact or not, so one may pretend M is compact in these applications.

Since the action of $\mathbb{R}$ on M induces a tensor-type preserving $\mathbb{R}$-linear action on the space of tensor fields, it is natural to make

Definition 4. *For any vector field X on M and any tensor field T on M, the Lie derivative of T with respect to X, denoted by $\mathcal{L}_X T$, is*

$$\mathcal{L}_X T := (t \cdot T)'|_{t=0}. \tag{42}$$

Here $'$ stands for the derivative with respect to t, and $t \cdot T$ is the result of the induced linear action of t on T, referred to as the translation of T by t.

Let us write ϕ_t for symmetry $q \mapsto t \cdot q$, i.e., $\phi_t(q) = \Phi_X(t, q)$, then we can work out a few computations.

1. If T is a scalar type tensor field f, then $t \cdot f = \phi_t^* f := f \circ \phi_t$ or $(\phi_{-t})_? f := f \circ (\phi_{-t})^{-1}$, depending on whether we view f as a covariant tensor field or contra-variant tensor field. Then

$$\begin{aligned}
\mathcal{L}_X f|_p &= (\phi_t^* f(p))'(0) \\
&= (f(c_p(t))'(0) \quad \because c_p(t) = \phi_t(p) = \Phi_X(t, p) \\
&= \langle \mathrm{d}f_p, \dot{c}_p(0) \rangle \quad \text{chain rule} \\
&= \langle \mathrm{d}f_p, X(p) \rangle \quad \text{using Eq. for integral curve} \\
&= \langle \mathrm{d}f, X \rangle|_p.
\end{aligned}$$

In other words, $\mathcal{L}_X f = \langle \mathrm{d}f, X \rangle$.

2. If T is a vector field Y, then $t \cdot Y = (\phi_{-t})_? Y$, so

$$\begin{aligned}
\mathcal{L}_X Y|_p &= ((\phi_{-t})_? Y(p))'(0) \quad \text{definition of action of } t \text{ on contravariant tensor filed} \\
&= (T_{\phi_t(p)}\phi_{-t}(Y(\phi_t(p)))'(0) \quad \text{using definition of the push-forward } _?.
\end{aligned}$$

Since this is a local calculation, by going to local representations, without loss of generality, we may pretend that we are doing calculation in ordinary calculus, for example, we write Y as the column vector $\vec{Y}$ and $T\phi_{-t}$ as the the Jacobean matrix $J\phi_{-t}$. Then the computation continues as follows:

ϕ_0 is the identity map, so $J\phi_0$ is the identity matrix and $\frac{\partial^2 \phi_0^i}{\partial x^k \partial x^j} = 0$.

$$\begin{aligned}
(\mathcal{L}_X Y)^i|_p &= ((J\phi_{-t})^i_j|_{\phi_t(p)}\, Y^j(\phi_t(p)))'(0) \\
&= ((J\phi_{-t})^i_j|_{\phi_t(p)})'(0)\, Y^j(p) + (Y^i(\phi_t(p)))'(0) \quad \text{the product rule.}
\end{aligned}$$

To continue, on the one hand, the chain rule and the integral curve equation combined together imply that

$$(Y^i(\phi_t(p)))'(0) = (Y^j(c_p(t)))'(0) = \partial_j Y^j(p)(c_p^j)'(0) = (X^j \partial_j Y^i)(p);$$

on the other hand,

$$\begin{aligned}
((J\phi_{-t})^i_j|_{\phi_t(p)})'(0) &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \left(\frac{\partial \Phi_X^i}{\partial x^j}(-t, \Phi_X(t, p)) \right) \quad \text{definition of } \phi_t \\
&\overset{\text{chain rule}}{=} \left(\frac{\partial^2 \Phi_X^i}{\partial t \partial x^j}(-t, \Phi_X(t, p)) \cdot (-1) + \frac{\partial^2 \Phi_X^i}{\partial x^k \partial x^j}(-t, \Phi_X(t, p)) \frac{\partial \Phi_X^k}{\partial t}(t, p) \right) \Big|_{t=0} \\
&= -\frac{\partial X^i}{\partial x^j}(p) + \frac{\partial^2 \phi_0^i}{\partial x^k \partial x^j}(p) X^k(p) \quad \text{using flow equation} \\
&= -\frac{\partial X^i}{\partial x^j}(p) \quad \because \phi_0^i = x^i
\end{aligned}$$

Therefore,

$$(\mathcal{L}_X Y)^i|_p = \left(-\frac{\partial X^i}{\partial x^j} Y^j + \frac{\partial Y^i}{\partial x^j} X^j \right) \Big|_p$$

In other words, $\mathcal{L}_X Y = [X, Y]$.

§11 Covariant Differentiation and Curvature

The goal of this chapter is to introduce a kind of differentiation for tensor fields, in terms of which, the "directional derivative" of tensor fields at a point can be introduced. We shall see later that, this kind of differentiation — refereed to as **covariant differentiation** — exists in huge abundance.

A covariant differentiation ∇ on tensor fields is a *direct* generalization of partial differentiation in the sense that it sends a type-T tensor field to a type-T-tensor-valued differential one-form: $s \mapsto \mathrm{d}_\nabla s$. So, for any vector field X and a tensor field s, if we let

For example, if f is a scalar tensor field, i.e., a real smooth function, then $\mathrm{d}_\nabla f$ is equal to $\mathrm{d}f$, thus a differential one-form.

$$\nabla_X s := \langle \mathrm{d}_\nabla s, X\rangle,$$

then we can paraphrase the previous sentence this way: A covariant differentiation ∇ on tensor fields is a map that sends a vector field X to a tensor-type-preserving derivative operator ∇_X on the linear space of tensor fields, moreover, ∇_X is $C^\infty(M)$-linear in X.

Note that $\nabla_X f = \mathcal{L}_X f$ because $\mathrm{d}_\nabla f = \mathrm{d}f$

While the derivative operator ∇_X is $C^\infty(M)$-linear in X, the same is not true for the derivative operator $\mathcal{L}_X$, cf. Exercise 75. Note that, the "directional derivative" of tensor field s with respect to the tangent vector $\underline{u}$ at point p is the tensor $\langle \mathrm{d}_\nabla s\big|_p, \underline{u}\rangle$ of M at point p.

What is covariant differentiation then? By definition, a **covariant differentiation** ∇ on tensor fields is an assignment of the derivative operator ∇_X on the linear space of tensor fields on M to any vector field X on M such that the following *desired property*:

Please note that the map $\nabla\colon A_1(M)\times A_1(M)\to A_1(M)$ that sends (X, Y) to $\nabla_X Y$ is not bilinear over $C^\infty(M)$, so it is not a tensor field of type $(1,2)$. In short, a covariant differentiation is not a tensor field.

$$\nabla_X \text{ is } C^\infty(M)\text{-linear in } X,$$

holds. Obviously the Lie differentiation of tensor fields is not a covariant differentiation.

The set of covariant differentiations (of tensor fields) on M is an affine space modelled on the infinite dimensional real linear space of tensor fields of type $(1,2)$. Indeed, if ∇ is a covariant differentiation on M and

Please note that ϕ is a tensor field of type $(1,2)$ and can be represented locally by functions $\phi^k_{ij} := \langle \mathrm{d}x^k, \phi(\partial_{x^i}, \partial_{x^j})\rangle$

$$\phi:\ A_1(M)\times A_1(M)\to A_1(M)$$

is a $C^\infty(M)$-bilinear map — a tensor field of type $(1,2)$, then

$$\nabla' := \nabla + \phi,$$

being a covariant differentiation on M, is the translation of ∇ by ϕ. In case you are confused, please note that $\nabla'_X f = \nabla_X f = \mathcal{L}_X f$ and $\nabla'_X Y = \nabla_X Y + \phi(X,Y)$.

Torsion Tensor for a covariant differentiation

For any two vector fields X and Y, we let $T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$. A simple computation shows that $T(X,Y)$ is $C^\infty(M)$-bilinear in (X,Y), so T is a tensor of type $(1,2)$ — the torsion tensor of ∇. We say that ∇ is **torsion free** if its torsion tensor T is zero.

The canonical covariant differentiation on pseudo-Riemannian manifolds

Let M be an n-manifold and $0 \le k \le n$ be an integer. Suppose that

$$g: \quad M \to S^2(T^*M)$$

is a smooth map such that $g|_p$ for any point $p \in M$ is a non-degenerate quadratic form on T_pM with signature $(k, n-k)$, then g is called a pseudo-Riemannian metric on M and the pair (M,g) is called a pseudo-Riemannian manifold. We drop the prefix "pseudo" if $k = n$. Any Euclidean space is a Riemannian manifold. Any manifold M inside $\mathbb{E}^N$ is a Riemannian sub-manifold of $\mathbb{E}^N$ with the first fundamental form I being its Riemannian metric.

g is a covariant symmetric tensor field of rank 2

Locally we can write $g = g_{ij}\, dx^i\, dx^j$. Let $[g^{ij}] = [g_{ij}]^{-1}$. One can see that there is a unique smooth map

$$g^{-1}: \quad M \to S^2(TM)$$

such that locally

$$g^{-1} = g^{ij}\,\partial_{x^i}\,\partial_{x^j}.$$

Cf. Exercise 4.

Remark 37. The Minkowski space in special relativity is the pseudo-Riemannian manifold $(\mathbb{A}^4_{\mathbb{R}}, g)$ where g is the standard Lorentzian structure on $\mathbb{A}^4_{\mathbb{R}}$: at each point $p \in \mathbb{A}^4_{\mathbb{R}}$, we have

$$g|_p((p,\vec{u}),(p,\vec{v})) = \eta(\vec{u},\vec{v}).$$

Here, η is the standard Lorentzian inner product on $\mathbb{R}^4$, so

$$\eta(\vec{u},\vec{v}) = u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3 \quad \text{for} \quad \vec{u} = \begin{bmatrix} u_0\\ u_1\\ u_2\\ u_3 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} v_0\\ v_1\\ v_2\\ v_3 \end{bmatrix}.$$

Note that g is a translation-invariant assignment of Lorentzian inner product to each tangent space of $\mathbb{A}^4_{\mathbb{R}}$. Therefore, the Minkowski space is the analogue of the $4D$ Euclidean space $\mathbb{E}^4$.

If we write $\eta = \eta_{ij}\hat{e}^i\hat{e}^j$, then

$$[\eta_{ij}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

In terms of the standard affine coordinate x^i, we have

$$g = \eta_{ij}\, \mathrm{d}x^i\, \mathrm{d}x^i = \mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2.$$

Here $x^0 = t$, $x^1 = x$, $x^2 = y$ and $x^3 = z$.

Exercise 80. *Let $R > 0$ be a real number and X be the hypersurface*

$$t^2 - x^2 - y^2 - z^2 = R^2$$

in the Minkowski space $(\mathbb{A}^4_{\mathbb{R}}, g)$. Show that the restriction of $-g$ to X is a Riemannian metric on X. What if X is the hypersurface $t^2 - x^2 - y^2 - z^2 = -R^2$? What if X is the circular cone $t^2 - x^2 - y^2 - z^2 = 0$ with the cone point removed?

`Hint: To compute the restriction of the Minkowski metric` g `to the hyper-surface` $t^2 - x^2 - y^2 - z^2 = R^2$`, one starts with identity` $\mathrm{d}(t^2 - x^2 - y^2 - z^2) = \mathrm{d}R^2$`, i.e.,` $t\,\mathrm{d}t - x\,\mathrm{d}x - y\,\mathrm{d}y - z\,\mathrm{d}z = 0$.

Levi-Civita connection

Let (M, g) be a pseudo-Riemannian manifold, X and Y be vector fields on M. Then the covariant derivative of Y with respect to X, $\nabla_X Y$, is uniquely defined provided that it is torsion free and metric compatible. More formally we have (recall that $\nabla_X f = \mathcal{L}_X f$)

Theorem 6. *To each pair of vector fields (X, Y) on M, there assigns a unique vector field $\nabla_X Y$ such that*

1) $\nabla_X Y$ is $\mathbb{R}$-linear in X and Y,

2) For each real smooth function f on M we have

(i) $\nabla_X(fY) = \nabla_X f\, Y + f\nabla_X Y$,

(ii) $\nabla_{fX} Y = f\nabla_X Y$,

3) (torsion free) $\nabla_X Y - \nabla_Y X - [X, Y] = 0$,

4) (metric compatible) $\nabla_X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$.

Indeed, on a coordinate patch with local coordinates x^i, if we let

$$\Gamma^k_{ij} = \frac{1}{2} g^{kl}(\partial_i g_{lj} - \partial_l g_{ij} + \partial_j g_{il}), \tag{43}$$

Symbols Γ^k_{ij} are called Christoffel symbols. One also introduces symbol $\Gamma_{lij} := \frac{1}{2}(\partial_i g_{lj} - \partial_l g_{ij} + \partial_j g_{il})$.

then, a simple computation shows that $\nabla_{\partial_i}\partial_j = \Gamma^k_{ij}\partial_k$. This proves the uniqueness.

Let us write $X = X^i\partial_i$ and $Y = Y^i\partial_i$ locally. For simplicity we write $\partial_i f$ as $f_{,i}$. Let

$$Y^i{}_{;j} := \partial_j Y^i + \Gamma^i_{jk} Y^k = Y^i{}_{,j} + \Gamma^i_{jk} Y^k,$$

then we would have

$$\nabla_X Y = X^j Y^i{}_{;j}\,\partial_i$$

provided that the right-hand side of this equation is the local formula of a vector field, which is the case: a simple computation shows that

$$\tilde{X}^j \tilde{Y}^i{}_{;j} = \frac{\partial \tilde{x}^i}{\partial x^k} X^l Y^k{}_{;l}.$$

This proves the existence.

The $\mathbb{R}$-linearity and the product rule enable us to extend the covariant derivative operator ∇_X uniquely to tensor fields of any type. As a practice, let us see how the covariant differentiation of differential one-form α is defined. Due to the product rule, we must have the defining equation

$$\nabla_X(\langle \alpha, Y\rangle) = \langle \nabla_X \alpha, Y\rangle + \langle \alpha, \nabla_X Y\rangle$$

for $\nabla_X\alpha$. Locally if we write $\alpha = \alpha_i\, \mathrm{d}x^i$, then we claim that

$$\nabla_{\partial_j}\alpha = \alpha_{i;j}\,\mathrm{d}x^i$$

where

$$\alpha_{i;j} = \partial_j \alpha_i - \Gamma^k_{ij}\alpha_k = \alpha_{i,j} - \Gamma^k_{ij}\alpha_k.$$

Exercise 81. *Please prove this claim.*

One can also compute $\nabla_X g$ and get the answer $\nabla_X g = 0$. Indeed, it suffices to prove it in a coordinate patch with local coordinates x^i. Since $g = g_{ij}\, \mathrm{d}x^i\, \mathrm{d}x^j$, by the product rule, we have

$$\nabla_{\partial_k} g = g_{ij;k}\, \mathrm{d}x^i\, \mathrm{d}x^j \tag{44}$$

where

$$\begin{aligned} g_{ij;k} &= \partial_k g_{ij} - \Gamma^l_{ik} g_{lj} - \Gamma^l_{kj} g_{il} \\ &= \partial_k g_{ij} - \frac{1}{2}(\partial_i g_{jk} - \partial_j g_{ik} + \partial_k g_{ij}) - \frac{1}{2}(\partial_k g_{ij} - \partial_i g_{kj} + \partial_j g_{ki}) \\ &= 0. \end{aligned}$$

Then $\nabla_X g = X^k \nabla_{\partial_k} g = 0$.

Please note that the identity $\nabla_X g = 0$ is just a compact way of saying that ∇ is metric compatible. Indeed, for any vector fields X, Y and Z, we have

$$(\nabla_X g)(Y, Z) = \nabla_X(g(Y, Z)) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z).$$

So $\nabla_X g = 0$ iff ∇ is metric compatible.

Exercise 82. *Show that $\nabla_X g^{-1} = 0$.* `Hint:` $\nabla_{\partial_k} g^{-1} = g^{ij}{}_{;k}\, \partial_i \partial_j$.

A tensor field T on (M, g) is called covariantly constant if $\nabla_X T = 0$ for any vector field X on M. So metric tensor g and its inverse g^{-1} are covariantly constant.

Remark 38. The above unique covariant differentiation ∇ on tensor fields is called the Levi-Civita connection of the pseudo-Riemannian manifold (M, g).

Exercise 83. *Let (X, g) be an oriented n-dimensional pseudo-Riemannian manifold and x^i is an oriented local coordinate system in the sense that $\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$ represents the orientation of X on the coordinate patch. Writing $g = g_{ij}\, \mathrm{d}x^i \wedge \mathrm{d}x^j$, show that there is a differential n-form* vol *on X such that, locally*

$$\mathrm{vol} = \sqrt{|\det[g_{ij}]|}\, \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n.$$

Note: vol *is called the* **oriented volume form** *on X and it is covariantly constant.*

`Hint: Just need to show that if` $\tilde{x}^i$ `is another oriented local coordinate system, then on the overlap of the two coordinate patches (assuming non-empty), we have`

$$\sqrt{|\det[\tilde{g}_{ij}]|}\, \mathrm{d}\tilde{x}^1 \wedge \cdots \wedge \mathrm{d}\tilde{x}^n = \sqrt{|\det[g_{ij}]|}\, \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n.$$

Example 19 (Poincaré half-plane)**.** *Let M be the upper half plane $y > 0$ with Riemannian metric*

$$g = \frac{\mathrm{d}x^2 + \mathrm{d}y^2}{y^2}.$$

Here x and y are the standard global coordinates. For simplicity we let $x^1 = x$ and $x^2 = y$. Compute $\nabla_{\partial_i} \partial_j$.

This Remannian manifold is called the Poincaré half-plane.

Solution.

$$g_{11} = g_{22} = \frac{1}{y^2}, \quad g_{12} = g_{22} = 0.$$

So

$$g^{11} = g^{22} = y^2, \quad g^{12} = g^{22} = 0.$$

Let us compute the Christoffel symbols now.

$$\Gamma^1_{11} = g^{1i}\Gamma_{i11} = g^{11}\Gamma_{111} = 0,$$

$$
\begin{aligned}
\Gamma^1_{12} &= g^{1i}\Gamma_{i12} = g^{11}\Gamma_{112} = \frac{1}{2}g^{11}g_{11,2} = -\frac{1}{y},\\
\Gamma^1_{22} &= g^{1i}\Gamma_{i22} = g^{11}\Gamma_{122} = 0,\\
\Gamma^2_{11} &= g^{2i}\Gamma_{i11} = g^{22}\Gamma_{211} = -\frac{1}{2}g^{22}g_{11,2} = \frac{1}{y},\\
\Gamma^2_{12} &= g^{2i}\Gamma_{i12} = g^{22}\Gamma_{212} = 0,\\
\Gamma^2_{22} &= g^{2i}\Gamma_{i22} = g^{22}\Gamma_{222} = \frac{1}{2}g^{22}g_{22,2} = -\frac{1}{y}.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\nabla_{\partial_1}\partial_1 &= \Gamma^i_{11}\partial_i = \frac{1}{y}\partial_2,\\
\nabla_{\partial_2}\partial_1 = \nabla_{\partial_1}\partial_2 &= \Gamma^i_{12}\partial_i = -\frac{1}{y}\partial_1,\\
\nabla_{\partial_2}\partial_2 &= \Gamma^i_{22}\partial_i = -\frac{1}{y}\partial_2.
\end{aligned}
$$

◀

Continuing the preceding example, we have

$$
\begin{aligned}
(\nabla_{\partial_1}\circ\nabla_{\partial_2} - \nabla_{\partial_2}\circ\nabla_{\partial_1})\partial_1 &= \nabla_{\partial_1}(-\frac{1}{y}\partial_1) - \nabla_{\partial_2}(\frac{1}{y}\partial_2)\\
&= -\frac{1}{y}\nabla_{\partial_1}\partial_1 + \frac{1}{y^2}\partial_2 - \frac{1}{y}\nabla_{\partial_2}\partial_2\\
&= -\frac{1}{y^2}\partial_2 + \frac{1}{y^2}\partial_2 + \frac{1}{y^2}\partial_2\\
&= \frac{1}{y^2}\partial_2.
\end{aligned}
$$

So, unlike the differentiation of real functions, for the covariant differentiation of vector fields, mixed partial differentiation may depend on the order.

Let us conclude this section with a remark. The essence of covariant differentiation is *connection*: it connects or identifies the tensor spaces of a given type at any two points connected by a piece-wise smooth path. For example, to identify the cotangent space T^*_pM with the cotangent space T^*_qM via a piece-wise smooth map $\alpha\colon [t_0, t_1] \to M$ with $\alpha(t_0) = p$ and $\alpha(t_1) = q$, we solve the homogeneous linear ODE $\nabla_{\dot\alpha}\tilde\alpha = 0$ for the lifting $\tilde\alpha$ in diagram

Locally, the homogeneous linear ODE $\nabla_{\dot\alpha}\omega = 0$ can be written as equation $\omega'_j(t) - (\alpha^i)'(t)\,\Gamma^k_{ij}(\alpha(t))\omega_k(t) = 0$.

$$
\begin{tikzcd}
 & T^*M \arrow[d, "\pi"] \\
{[t_0, t_1]} \arrow[r, "\alpha"] \arrow[ur, dashed, "\tilde\alpha"] & M
\end{tikzcd}
$$

Then, by the existence and uniqueness theorem for the initial value problem in ODE, we get the linear invertible map $\alpha_*\colon T^*_pM \to T^*_qM$ that sends $\tilde\alpha(t_0)$ to $\tilde\alpha(t_1)$, thus the identification of T^*_pM with T^*_qM via the piece-wise smooth path α. By the way, $\tilde\alpha$ is called the parallel transport of $\tilde\alpha(t_0)$ along the path α.

Riemann Curvature Tensor

Let us start with a covariant differentiation ∇ on the smooth manifold M as well as its local representation Γ^k_{ij} defined via the equation

$$\nabla_{\partial_i}\partial_j = \Gamma^k_{ij}\,\partial_k.$$

For any vector fields X and Y, on the linear space of tensor fields, while $\mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X - \mathcal{L}_{[X,Y]} = 0$, the operator

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

may not be zero. For example, $R(\partial_x, \partial_y)\partial_x = \frac{1}{y^2}\partial_y \neq 0$ in Example 19.

On the other hand, the map $(X,Y,Z) \to R(X,Y)Z$ is clearly $\mathbb{R}$-trilinear, moreover, $R(fX, gY)(hZ) = fgh\ R(X,Y)Z$, as one can check via a direct computation. Consequently this map is $C^\infty(M)$-trilinear as well, i.e., a tensor field R_∇. On a coordinate patch with local coordinates x^i, the tensor field R_∇ can be represented by the set of local functions

$$R^k_{ijl} := \langle \mathrm{d}x^k, R(\partial_i, \partial_j)\partial_l \rangle.$$

A simple computation shows that

$$R^k_{ijl} = \partial_i \Gamma^k_{jl} + \Gamma^k_{im}\Gamma^m_{jl} - < i \leftrightarrow j >$$

Here $< i \leftrightarrow j >$ means all terms on the left, but with i and j interchanged, i.e., $\partial_j \Gamma^k_{il} + \Gamma^k_{jm}\Gamma^m_{il}$.

The tensor field R_∇, referred to as the Riemann curvature tensor of the covariant differentiation ∇, is a tensor field of type $(1,3)$. Here is a list of its basic properties: for any four vector fields X, Y, Z, and W, we have

i) Skew symmetry: As operator on tensor fields,

$$R(X,Y) = -R(Y,X),$$

In case ∇ is torsion free, we have

ii) Algebraic Bianchi identity:

$$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0,$$

and

iii) Differential Bianchi identity: As operator on tensor fields,

$$(\nabla_X R)(Y,Z) + (\nabla_Y R)(Z,X) + (\nabla_Z R)(X,Y) = 0.$$

By definition $(\nabla_X R)(Y,Z)T$ is

$$\nabla_X(R(Y,Z)T) - R(Y,Z)\nabla_X T - R(\nabla_X Y, Z)T - R(Y, \nabla_X Z)T$$

for any tensor field T.

In case ∇ is the Levi-Civita connection of the pseudo-Riemannian manifold (M,g), denoting $g(X,Y)$ by $\langle X, Y\rangle$, we have two more properties for R:

iv) Interchange symmetry: $\langle R(Z,W)X,Y\rangle = \langle R(X,Y)Z,W\rangle$,

v) Skew symmetry: $\langle R(X,Y)Z,W\rangle = -\langle R(X,Y)W,Z\rangle$.

Exercise 84. *Prove the above five properties for the curvature tensor R.*

`Hints: i) and ii) are trivial, and i) + iv)` $\Longrightarrow$ `v). For iii) and iv), one may assume that` $X = \partial_i$, $Y = \partial_j$, $Z = \partial_k$, $W = \partial_l$.

Since $\langle R(X,Y)Z,W\rangle$ is multi-linear in (X,Y,Z,W), we have a tensor field of type $(0,4)$, called the **covariant Riemann curvature tensor** for the pseudo-Riemannian manifold (M,g). Locally covariant Riemann curvature tensor is represented by the set of functions

$$R_{ijkl} = \langle R(\partial_i,\partial_j)\partial_k,\partial_l\rangle$$

Properties i) + iv) + v) implies that the value of the covariant curvature tensor at point p is a symmetric bilinear form $\mathcal{K}_p$ on $\wedge^2 T_pM$. In case P is a 2-dim linear subspace of T_pM, $\wedge^2 P$ is a 1-dim linear subspace of $\wedge^2 T_pM$.

By definition,

$$\mathcal{K}_p(u\wedge v, z\wedge w) = \langle R(U,V)Z,W\rangle$$

provided that the values at p of the vector fields U, V, Z, and W are u, v, z, and w respectively.

In the rest of this chapter, let us assume that ∇ is the Levi-Civita connection of the Riemannian manifold (M,g). In this case P is a sub Euclidean vector space of T_pM, so we can choose an orthonormal basis (u,v) for P. It is clear that the real number

$$K(P) := -\mathcal{K}_p(u\wedge v, u\wedge v)$$

is in dependent of the choice of the orthonormal basis (u,v). This number K_P is called the **sectional curvature** for the section plane P.

Exercise 85. *For any basis (u,v) of P, show that*

$$K(P) = -\frac{\mathcal{K}_p(u\wedge v, u\wedge v)}{|u\wedge v|^2} = \frac{\langle R(U,V)V,U\rangle\big|_p}{|u\wedge v|^2}$$

where $|u\wedge v|^2 = |u|^2\,|v|^2 - \langle u,v\rangle^2$ is the square of the area of the parallelogram spanned by vectors u and v.

`Hint: If` $[u,v] = [z,w]A$ `for a real square matrix of order 2, then` $u\wedge v = \det A\, z\wedge w$.

In case M is a regular surface with g being the first fundamental form, the sectional curvature for T_pM can be shown to be the Gauss curvature of M at point p.

A regular surface is a hypersurface in $\mathbb{E}^3$.

Exercise 86. *Please prove this statement.*

Proof. Without loss of generality we may assume that $p = (0,0,0)$ and M is the graph of function

$$z = f(x,y) = \frac{1}{2}\kappa_1 x^2 + \frac{1}{2}\kappa_2 y^2 + \cdots$$

A local parametrization ϕ around point p can be chosen such that $\phi(u, v) = (u, v, f(u, v))$, then $\mathrm{I} = g_{ij}\mathrm{d}x^i\,\mathrm{d}x^j$ where

$$\begin{aligned} g_{11} &= 1 + (\kappa_1 x^1)^2 + \cdots \\ g_{12} &= \kappa_1\kappa_2 x^1 x^2 + \cdots \\ g_{22} &= 1 + (\kappa_2 x^2)^2 + \cdots \end{aligned}$$

So

$$g_{ij}|_p = \delta_{ij}, \quad \partial_k g_{ij}|_p = 0, \quad \Gamma^k_{ij}|_p = 0$$

because $x^i(p) = 0$. These identities will be used many times in the the rest of this proof.

Since $(\partial_1|_p, \partial_2|_p)$ is an orthonormal basis for T_pM, one can use it to compute the sectional curvature of T_pM and get

$$K_{T_pM} = R^1_{122}|_p.$$

Since $R^1_{122} = \langle \mathrm{d}x^1, R(\partial_1, \partial_2)\partial_2\rangle$ and

$$\begin{aligned} R(\partial_1, \partial_2)\partial_2|_p &= (\nabla_{\partial_1}\nabla_{\partial_2}\partial_2 - \nabla_{\partial_2}\nabla_{\partial_1}\partial_2)|_p \\ &= (\nabla_{\partial_1}(\Gamma^i_{22}\partial_i) - \nabla_{\partial_2}(\Gamma^i_{12}\partial_i))|_p \\ &= (\partial_1\Gamma^i_{22} - \partial_2\Gamma^i_{12})|_p\partial_i|_p, \end{aligned}$$

we have

$$\begin{aligned} K(T_pM) &= (\partial_1\Gamma^1_{22} - \partial_2\Gamma^1_{12})|_p \\ &= (\partial_1\Gamma_{122} - \partial_2\Gamma_{112})|_p \\ &= \frac{1}{2}(\partial_1(g_{12,2} - g_{22,1} + g_{21,2}) - \partial_2(g_{12,1} - g_{12,1} + g_{11,2}))|_p \\ &= \frac{1}{2}((2g_{12,21} - g_{22,11}) - g_{11,22})|_p \\ &= \kappa_1\kappa_2, \end{aligned}$$

i.e., the Gauss curvature of M at point p. □

Therefore we say that Riemann curvature tensor is a vast extension of the *Gauss curvature for regular surfaces* to manifolds of higher dimension. It is clear that Riemann curvature tensor, sectional curvature, and the Gauss curvature for regular surfaces are all intrinsic because they are all completely determined by the first fundamental form I. However, the Gauss curvature for hypersurfaces in $\mathbb{E}^N$ is NOT intrinsic unless $N = 3$.

Let us return to Example 19. Recall from there that

$$R(\partial_2, \partial_1)\partial_1 = -\frac{1}{y^2}\partial_2,$$

so the Gauss curvature function for the Poincaré half-plane H is

$$K = \frac{g(R(\partial_2, \partial_1)\partial_1, \partial_2)}{g(\partial_1, \partial_1)g(\partial_2, \partial_2) - (g(\partial_1, \partial_2))^2} = \frac{-\frac{1}{y^4}}{\frac{1}{y^4}} = -1.$$

Remark 39. For the Euclidean plane $\mathbb{E}^2$, the Riemannian metric is

$$g = \mathrm{d}x^2 + \mathrm{d}y^2.$$

So $g_{ij} = g^{ij} = \delta_{ij}$ — all are constant functions. Then Γ_{ij}^k are all zero, so the Gauss curvature is zero. We say that $\mathbb{E}^2$ has constant Gauss curvature zero.

It is useful to keep the following table in mind.

Geometry	Analytic Model	Constant Curvature
Spheric	S^2	1
Euclidean	$\mathbb{E}^2$	0
Hyperbolic	H	-1

Remark 40. Since $R(\partial_i, \partial_j) = [\nabla_{\partial_i}, \nabla_{\partial_j}]$, we know that Riemann curvature tensor is the obstruction to the commutativity of partial differentiation of vector fields along two different directions.

Remark 41 (Gauss-Bonnet Formula)**.** Let Σ be a compact regular surface, $K\colon \Sigma \to \mathbb{R}$ be the Gauss curvature function, $\chi(\Sigma)$ be the Euler number of Σ. Then we have

$$\int_\Sigma K \, \mathrm{d}A = 2\pi\chi(\Sigma).$$

For example, if Σ is a sphere with radius a, then $K = \frac{1}{a^2}$, so the above formula says the the Euler number of the sphere is 2, which is indeed the case.

Let us conclude this chapter with a final exercise.

Exercise 87 (computations with isothermal coordinates)**.** *Let $\psi\colon U \to \Sigma$ be a smooth local parametrization of the regular surface Σ and u, v be the corresponding local coordinate functions. Assume that*

$$\mathrm{I} = e^{2\phi}(\mathrm{d}u^2 + \mathrm{d}v^2) \tag{45}$$

for some smooth real function ϕ on the coordinate patch $\psi(U)$.

1) Show that

$$K = -e^{-2\phi}(\partial_u^2 + \partial_v^2)\phi$$

on $\psi(U)$.

2) In our ordinary calculus, the Laplace operator Δ on functions is defined to be $\nabla \cdot \nabla$, i.e., Δf is the divergence of the gradient vector field of f. We have extended the notion of gradient vector field already. Here, we would like to say that the notion of divergence of a vector field can also be extended. Here is a definition in terms of local coordinates. Let X be a smooth vector field on Σ, the divergence of X, denoted by $\operatorname{div} X$, is defined as follows: on $\psi(U)$, we write $X = X^i \partial_{x^i}$, then

$$\operatorname{div} X := \langle \mathrm{d}x^i, \nabla_{\partial_{x^i}} X \rangle = X^i{}_{;i} = \partial_i X^i + \Gamma^i_{ij} X^j.$$

Of course, for this definition to make sense, you have to check that if there is another coordinate system $\tilde{x}^i$ with coordinate patch $\tilde{\psi}(\tilde{U})$, then on the common intersection of $\psi(U)$ with $\tilde{\psi}(\tilde{U})$, we must have

$$\langle \mathrm{d}x^i, \nabla_{\partial_{x^i}} X \rangle = \langle \mathrm{d}\tilde{x}^i, \nabla_{\partial_{\tilde{x}^i}} X \rangle,$$

a fact that can be easily verified.

Since gradient and divergence operator are intrinsic, so is the Laplace operator. Now show that

$$\Delta f = e^{-2\phi}(\partial_u^2 + \partial_v^2) f.$$

Therefore, the result in part 1) says that

$$K = -\Delta \phi. \tag{46}$$

3) (optional) Show that, around any point of Σ, one can find local coordinate functions u, v around that point such that the first fundamental form is of the form in equation (45).

Such coordinates are called isothermal coordinates.

4) Returning to Example 14, if we choose $v = \varphi$ and $u = \ln \tan \frac{\vartheta}{2}$ and $\phi = \ln \sin \vartheta = -\ln \cosh u$, then

$$\mathrm{I} = e^{2\phi}(\mathrm{d}u^2 + \mathrm{d}v^2).$$

Use formula (46) to verify that $K = 1$.

5) Use formula (46) to recompute the Gauss curvature for the Poincaré half-plane. In this case $u = x$, $v = y$ and $\phi = -\ln y$.

Exercise 88. *Let (M, g) be a Riemannian manifold and f be a smooth function on M. Locally we have $g = g_{ij}\, \mathrm{d}x^i\, \mathrm{d}x^j$. Let us write $\det[g_{ij}]$ as g (Sorry, this is the standard notation. So symbol g has two meanings.) Show that*

$$\Delta f = \frac{1}{\sqrt{g}} \partial_i (\sqrt{g}\, g^{ij} \partial_j f).$$

Exercise 89. *Let (M, g) be a 3-dimensional oriented Riemannian manifold. A system of local coordinates x^1, x^2 and x^3 are chosen such that $(\partial_1, \partial_2, \partial_3)$ represents the orientation of M at each point of the coordinate patch. Then* *the (point-wise) cross product of any two vector fields on M is well-defined and locally we have*

$$\partial_i \times \partial_j = \sqrt{g}\epsilon_{ijk}\, g^{kl}\partial_l.$$

Here ϵ_{ijk} is zero unless i, j, k are distinct, and is 1 (-1 resp.) if (i, j, k) is an even (odd resp.) permutation of $(1, 2, 3)$. In other words, locally we have

$$X \times Y = X^i Y^j \sqrt{g}\epsilon_{ij}{}^l \partial_l. \tag{47}$$

where $\epsilon_{ij}{}^l = \epsilon_{ijk}\, g^{kl}$.

(1) Show that this definition for cross product is well-defined. Consequently, the curl of vector fields is well-defined and locally we have

$$\mathrm{cur}\, X = \sqrt{g}\epsilon^i{}_j{}^k\, X^j{}_{;i}\, \partial_k. \tag{48}$$

where $\epsilon^i{}_j{}^k = \epsilon_{ljm} g^{li} g^{mk}$.

(2) Show that $\mathrm{cur}\,(\mathrm{grad} f) = 0$ for any real smooth function f on M.

(3) Show that $\mathrm{div}\,(\mathrm{cur}\, X) = 0$ for any real smooth vector field X on M.

On the Euclidean space $\mathbb{E}^3$, in terms of spherical coordinates (r, ϑ, φ), the Riemannian metric can be written as

$$g = \mathrm{d}r^2 + r^2\, \mathrm{d}\vartheta^2 + r^2(\sin\vartheta)^2 \mathrm{d}\varphi^2.$$

In Calculus II, people introduce the orthonormal tangent frame

$$\vec{e}_r := \partial_r, \quad \vec{e}_\vartheta := \frac{1}{r}\partial_\vartheta, \quad \vec{e}_\varphi := \frac{1}{r\sin\vartheta}\partial_\varphi.$$

and a vector field $\vec{F}$ is expanded as

$$\vec{F} = F_r\, \vec{e}_r + F_\vartheta\, \vec{e}_\vartheta + F_\varphi\, \vec{e}_\varphi.$$

Exercise 90. *Let f be a smooth function on $\mathbb{E}^3$ and $\vec{F}$ be a vector field on $\mathbb{E}^3$. Show that*

$$\mathrm{grad}\, f = \partial_r f\, \vec{e}_r + \frac{1}{r}\partial_\vartheta f\, \vec{e}_\vartheta + \frac{1}{r\sin\vartheta}\partial_\varphi f\, \vec{e}_\varphi$$

and

$$\mathrm{div}\, \vec{F} = \frac{1}{r^2}\partial_r(r^2 F_r) + \frac{1}{r\sin\vartheta}\partial_\vartheta(\sin\vartheta F_\vartheta) + \frac{1}{r\sin\vartheta}\partial_\varphi F_\varphi.$$

What about $\mathrm{cur}\, \vec{F}$?

§12 Schouten-Nijenhuis Bracket

Just as the exterior derivative operator on functions can be uniquely extended to differential forms, the Lie bracket on vector fields can be uniquely extended to multi-vector fields. This generalized Lie bracket is referred to as the Schouten-Nijenhuis bracket.

Here is one way to see it. To start, let us fix a torsion free covariant differentiation ∇. From definition we know that the Lie bracket of vector field X with vector field Y satisfies identity

$$[X, Y] = \nabla_X Y - \nabla_Y X.$$

So it is tempting to declare that, for multi-vector fields P of degree p and Q of degree q,

$$[P, Q] := \nabla_P Q - (-1)^{(p-1)(q-1)} \nabla_Q P.$$

Of course, you have to find the correct generalized derivative operators ∇_P, ∇_Q out of ∇ such that $[P, Q]$ defined this way is independent of the choice of ∇.

Indeed, this can be done such that ∇_P is $C^\infty(M)$-linear in P, moreover, $\nabla_P = 0$ if P is a function. Then, we have

1) $[f, g] = 0$,

2) $[X, Q] = \mathcal{L}_X Q$ if Q is a function or a vector field,

3) $[P, Q] = -(-1)^{(p-1)(q-1)}[P, Q]$. (graded skew-symmetry)

One can show that Schouten-Nijenhuis bracket is the unique $\mathbb{R}$-bilinear map

$$A.(M) \times A.(M) \to A.(M)$$

that satisfy properties 1), 2), 3) above and property

4) $[P, Q \wedge R] = [P, Q] \wedge R + (-1)^{(p-1)q} Q \wedge [P, R]$.

I leave the details to you. For further exploration, try to find the graded Jacobian identity for the Schouten-Nijenhuis bracket. If you like the duality principle as I do, you will believe that the exterior derivative operator d on differential forms and Schouten-Nijenhuis bracket on multi-vector fields should determine each other. Indeed, this is the case. For more details, please consult this paper by Charles-Michel Marle.

A Poisson structure on a manifold M is a bi-vector field Λ such that $[\Lambda, \Lambda] = 0$. The associated Poisson bracket of real smooth function f with real smooth function g on M, denoted by $\{f, g\}$, is defined as follows:

$$\{f, g\} = \langle \mathrm{d}f \wedge \mathrm{d}g, \Lambda \rangle.$$

Locally, if we write $\Lambda = \frac{1}{2!}\Lambda^{ij}\,\partial_i \wedge \partial_j$, then $\{f, g\} = \Lambda^{ij}\partial_i f \partial_j g$. One can show that the Poisson bracket

$$\{\,,\,\}: \quad C^\infty(M) \times C^\infty(M) \to C^\infty(M)$$

turns $C^\infty(M)$ into a real Lie algebra such that, for any $f \in C^\infty(M)$, the map

$$\{f, \cdot\}: \quad C^\infty(M) \to C^\infty(M)$$

is a derivation on $C^\infty(M)$.

A manifold with a Poisson structure is called a **Poisson manifold**. As an example, we remark that a symplectic manifold is naturally a Poisson manifold. That is because, a symplectic structure $\omega = \frac{1}{2!}\omega_{ij}\mathrm{d}x^i \wedge \mathrm{d}x^j$ defines a Poisson structure $\Lambda = \frac{1}{2!}\Lambda^{ij}\,\partial_i \wedge \partial_j$ with $[\Lambda^{ij}]$ be the matrix inverse of $[\omega_{ij}]$. As another example, the dual of a real Lie algebra $\mathfrak{g}$, $\mathfrak{g}^*$, is a Poisson manifold: its underlying manifold is the real affine space $(\mathfrak{g}^*)_{aff}$ and the Poisson structure is as follows: the elements in $\mathfrak{g}$ are real smooth functions on $\mathfrak{g}^*$, moreover, a basis x^i on $\mathfrak{g}$ is a global coordinate system on $(\mathfrak{g}^*)_{aff}$; with that understood, the Poisson structure is globally represented by functions

$$\Lambda^{ij} := [x^i, x^j], \quad \text{the Lie bracket of } x^i \text{ with } x^j.$$

§13 Integration

We have seen that partial differentiation can be formulated in a coordinate-free way. The same can be done for integration and the Stokes' formula as well. The coordinate-free version of integration is the abstraction and generalization of integration of the second kind in the Calculus II, i.e., integration over oriented curves or oriented surfaces. [4] The coordinate-free version of Stokes' formula is the common generalization of Newton-Leibniz formula, Green's theorem, Stokes' theorem, and Gauss's theorem in Calculus I & II, and its proof can be reduced to the Newton-Leibniz formula.

[4] They can be associated with the work done by a force or the flux of a magnetic field respectively

Manifolds with boundary

While n-manifolds are modelled on $\mathbb{E}^n$, n-manifolds with boundary are modelled on the n-dimensional upper half space

$$\mathrm{H}^n := \{x \in \mathbb{E}^n \mid x_1 \geq 0\}.$$

To be more precisely, an n-manifold with boundary is a subset X of an affine space such that for any point p in X, we have either

$$(X, p) \cong_{loc} (\mathrm{H}^n, 0)$$

or

$$(X, p) \cong_{loc} (\mathrm{H}^n, 0 + \vec{e}_1) \cong_{loc} (\mathbb{E}^n, 0).$$

If the former case is true we say p is a boundary point and if the later case is true we say that p is an interior point. It is easy to see that $VX_p \cong \mathbb{R}^n$ or $\mathbb{R}^{n-1}$ according as p is an interior point or a boundary point, so the notions of interior point and boundary points are unambiguously defined. It is clear that the set of interior points of X, denoted by $\operatorname{Int} X$, is an n-manifold, and the set of boundary points of X, denoted by ∂X, is an $(n-1)$-manifold. For example, the n-dimensional ball

$$\mathrm{B}^n := \{x \in \mathbb{E}^n \mid \vec{x} \cdot \vec{x} = 1\}$$

and n-dimensional upper half space H^n are all n-manifolds with boundary, moreover,

$$\operatorname{Int}\mathrm{B}^n \cong \operatorname{Int}\mathrm{H}^n \cong \mathbb{E}^n, \quad \partial\mathrm{B}^n \cong \mathrm{S}^{n-1}, \quad \partial\mathrm{H}^n \cong \mathbb{E}^{n-1}.$$

It is also clear that the empty set and any n-manifold is an n-manifold with boundary, but an n-manifold with non-empty boundary is not an n-manifold. For example, neither B^n nor H^n is a manifold.

Let X be an n-manifold with boundary. Suppose that X is inside an affine space $\mathbb{A}$ and $p \in X$. The tangent space of X at point p, denoted by T_pX, is defined as follows: T_pX is the linear span of the set of all tangent vectors of $\mathbb{A}$ of the form

$$\dot{\alpha}(0) := \left(\alpha(0), \lim_{t\to 0^-} \frac{\alpha(t)-\alpha(0)}{t}\right)$$

where $\alpha\colon (-\epsilon, 0] \to X$ is a smooth map with $\alpha(0) = p$ and ϵ is a positive real number. It is obvious that $V_pX \subseteq T_pX$. We claim that T_pX is always a real n-dimensional linear space, moreover, $T_pX = V_pX$ iff p is an interior point.

Exercise 91. *Prove this claim.*

`Hint: You may assume that` $X = \mathrm{H}^n$.

One can check that, if p is an interior point, then $T_pX = T_p(\operatorname{Int} X) = V_pX$, and if p is a boundary point, then $T_p(\partial X) = V_pX$ — a codimension 1 linear subspace of T_pX.

Just as before, we can introduce local coordinates x^i around a point p such that $x^i(p) = 0$. In case p is a boundary point, we may assume that $x^1 \geq 0$ on the coordinate patch U and $x^1 = 0$ on and only on $U \cap \partial X$; moreover, the tangent vector $-\frac{\partial}{\partial x^1}\big|_p$ or any of its positive multiple is called an **outward normal vector** of X at the boundary point p. Note that, when restricted to $U \cap \partial X$, $(x^2, \ldots, x^n)$ becomes a local coordinate system on ∂X.

Just as before, an **orientation** on X is a continuous assignment of orientation to each tangent space of X, and X together with an orientation on X is called an **oriented manifold with boundary**. In case X is an oriented manifold with boundary, ∂X is oriented manifold with the induced orientation given by this convention: $(v_1, v_2, \ldots, v_n)$ represents the orientation of T_pX if $(v_2, \ldots, v_n)$ represents the orientation of $T_p(\partial X)$ and v_1 is an outward normal vector at the boundary point p. Recalling that there are 4 different ways of describing orientation on a real vector space on page 95, then, in terms of the local coordinate system x^i described above, if $\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$ represents the orientation of X on U, then $-\mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n$ represents the orientation of ∂X on $U \cap \partial X$.

Integration of differential forms

Let X be an oriented n-manifold with boundary, ω a compactly supported differential n-form on X. The main goal of this chapter is to make sense of $\int_X \omega$, the integration of ω on X. This task shall be broken into a few steps.

Step 1. Let U be an open set of in H^n, ω a compactly supported differential n-form on U. Assume the standard orientation on U, and write $\omega = \alpha\, dx^1 \wedge \cdots \wedge dx^n$ and define

$$\int_U \omega := \int_U \alpha(x)\, \mathrm{d}^n x, \tag{49}$$

where $\mathrm{d}^n x$ is the Lebesque measure on H^n and the integral on the right is just the Riemann-Lebesque integral. The following lemma is a reformulation of a theorem in calculus.

Lemma 1. *Let $f\colon U \to V$ be an orientation preserving diffeomorphism between open sets of H^n, ω a compactly supported differential n-form, then*

$$\int_U f^*\omega = \int_V \omega. \tag{50}$$

Step 2. Let X be an oriented n-manifold with boundary, ω a compactly supported differential n-form on X whose support is inside a coordinate neighborhood, i.e., there is an open set N of X, an open set U of H^n and an orientation preserving diffeomorphism $f\colon U \to N$, and $\operatorname{supp}\omega$ is inside N. We define

$$\int_X \omega := \int_U f^*\omega. \tag{51}$$

We need to show that this is well-defined. Suppose that $\tilde{f}, \tilde{U}, \tilde{N}$ are another such data. Let $W = N \cap \tilde{N}$, then $\operatorname{supp}\omega$ is inside W and

$$\begin{aligned}
& \int_U f^*\omega \\
= \; & \int_{f^{-1}(W)} f^*\omega \quad \text{because } \operatorname{supp} f^*\omega \text{ is inside } f^{-1}(W) \\
= \; & \int_{\tilde{f}^{-1}(W)} (f^{-1}\tilde{f})^*(f^*\omega) \quad \text{Lemma 1} \\
= \; & \int_{\tilde{f}^{-1}(W)} \tilde{f}^*\omega \\
= \; & \int_{\tilde{U}} \tilde{f}^*\omega \quad \text{because } \operatorname{supp} \tilde{f}^*\omega \text{ is inside } \tilde{f}^{-1}(W).
\end{aligned}$$

The following lemma is clear.

Lemma 2. *Suppose that α and β are two compactly supported differential n-forms on an oriented n-manifold with boundary X and are both supported in a coordinate neighborhood. Let c_1 and c_2 be constants. Then*

$$\int_X (c_1\alpha + c_2\beta) = c_1 \int_X \alpha + c_2 \int_X \beta. \tag{52}$$

Step 3: We are now ready to define the integration in the general case. Let X be an oriented n-manifold with boundary, ω be a compactly supported differential n-form on X. Since ω is compactly supported, we can find finitely many coordinate neighborhoods N_1, ..., N_k to cover $\operatorname{supp}\omega$. Let N_0 be the complement of $\operatorname{supp}\omega$ in X. Let $\{\rho_i\}_{0\le i\le k}$ be the partition of unity subordinate to the covering $\{N_i\}_{0\le i\le k}$. Define

$$\int_X \omega := \sum_{i=1}^{k} \int_X \rho_i \omega \tag{53}$$

where each $\int_X \rho_i\omega$ is defined in step two. We need to show that this is well-defined. Suppose that $\{\tilde{N}_i\}_{0\le i\le \tilde{k}}$, $\{\tilde{\rho}_i\}_{0\le i\le k}$ are another such data, then

$$\begin{aligned}
\sum_{i=1}^{k}\int_X \rho_i\omega &= \sum_{i=1}^{k}\int_X \sum_{j=1}^{\tilde{k}} \tilde{\rho}_j\rho_i\omega \\
&= \sum_{i=1}^{k}\sum_{j=1}^{\tilde{k}}\int_X \tilde{\rho}_j\rho_i\omega \quad \text{Lemma 2} \\
&= \sum_{j=1}^{\tilde{k}}\int_X \sum_{i=1}^{k} \tilde{\rho}_j\rho_i\omega \quad \text{Lemma 2} \\
&= \sum_{j=1}^{\tilde{k}}\int_X \tilde{\rho}_j\omega.
\end{aligned}$$

The following lemma is clear.

Lemma 3. *Suppose that α and β are two compactly supported differential n-forms on an oriented n-manifold X. Let c_1 and c_2 be constants. Then*

$$\int_X (c_1\alpha + c_2\beta) = c_1\int_X \alpha + c_2\int_X \beta. \tag{54}$$

By convention, if X is empty, $\int_X \omega := 0$. We often write $\int_X \omega$ for $\int_X f^*\omega$ if f is an inclusion map.

Stokes Formula

The goal here is to state and prove the Stokes formula, i.e., the following generalization of the fundamental theorem of calculus:

Theorem 7 (Stokes Theorem)**.** *Let X be an oriented compact n-manifold with boundary, ω a differential $(n-1)$-form on X. Then*

$$\int_{\partial X} \omega = \int_X \mathrm{d}\omega. \tag{55}$$

Since both sides of the Stokes formula is additive, we just need to prove the theorem in the special case where X is either $\mathbb{R}^n$ or H^n.

These special case can each be easily checked by using the fundamental theorem of calculus.

To be precise, let $\{N_i\}_{1\le i\le k}$ be an coordinate patch covering of $\operatorname{supp}\omega$, and $f_i\colon U_i \to X$ be the parametrization of N_i for each i. We may assume that U_i is either $\mathbb{R}^n$ or H^n. Let $\{\rho_i\}_{1\le i\le k}$ be a partition of unity subordinate to $\{N_i\}_{0\le i\le k}$ (N_0 is the complement of $\operatorname{supp}\omega$). Suppose that the stokes formula is true for $\rho_i\omega$ for each i, then

$$\begin{aligned}
\int_{\partial X}\omega &= \int_{\partial X}\sum_i \rho_i\omega \\
&= \sum_i \int_{\partial X}\rho_i\omega && \text{Lemma 2 above} \\
&= \sum_i \int_X \mathrm{d}(\rho_i\omega) && \text{Assume the stokes formula is valid for } \rho_i\omega \\
&= \sum_i \int_X \mathrm{d}\rho_i \wedge \omega + \sum_i \int_X \rho_i \mathrm{d}\omega \\
&= \int_X \mathrm{d}\left(\sum_i \rho_i\right)\wedge\omega + \int_X \sum_i \rho_i\mathrm{d}\omega && \text{Lemma 2 above} \\
&= \int_X \mathrm{d}\omega.
\end{aligned}$$

Therefore, we just need to show that

$$\int_{\partial X}\rho_i\omega = \int_X \mathrm{d}(\rho_i\omega)$$

for every i. In view of our definition of integration of forms and the fact that d is natural, that is to say, we need to show that

$$\int_{\partial U}\alpha = \int_U \mathrm{d}\alpha$$

for every compactly supported differential $(n-1)$-form α on U with U being either $\mathbb{R}^n$ or H^n, equipped with the standard orientation.

Let $\alpha = \sum_i (-1)^{i-1} f_i\, \mathrm{d}x^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge \mathrm{d}x^n$. Then $\mathrm{d}\alpha = \sum_i \partial_i f_i\, \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$.

<u>Case 1</u>: $U = \mathbb{R}^n$. Then

$$\begin{aligned}
\int_{\mathbb{R}^n}\mathrm{d}\alpha &= \sum_i \int_{\mathbb{R}^n} \partial_i f_i\, \mathrm{d}x^1\wedge\cdots\wedge \mathrm{d}x^n = \sum_i \int_{\mathbb{R}^n} \partial_i f_i\, \mathrm{d}^n x \\
&= \sum_i \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \partial_i f_i\, \mathrm{d}x^i\right) \mathrm{d}^{n-1}x = 0.
\end{aligned}$$

Also, $\int_{\partial\mathbb{R}^n}\alpha = 0$.

<u>Case 2</u>: $U = \mathrm{H}^n$. We may assume that

$$U = \{(x_1,\dots,x_n) \in \mathbb{R}^n \mid x_1 \ge 0\}.$$

Then

$$\int_U \mathrm{d}\alpha = \sum_i \int_U \partial_i f_i\, \mathrm{d}x^1\wedge\cdots\wedge\mathrm{d}x^n$$

$$
\begin{aligned}
&= \sum_i \int_U \partial_i f_i \, \mathrm{d}^n x \\
&= \sum_{i>1} \int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{\infty} \partial_i f_i \, \mathrm{d}x^i \right) \mathrm{d}^{n-1}x + \int_{\mathbb{R}^{n-1}} \left(\int_0^{\infty} \partial_1 f_1 \, \mathrm{d}x^1 \right) \mathrm{d}^{n-1}x \\
&= \int_{\mathbb{R}^{n-1}} -f_1(0, x_2, \ldots, x_n) \, \mathrm{d}^{n-1}x \\
&= \int_{\partial U} -f_1(0, x_2, ..., x_n) \, (-\mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n) \\
&= \int_{\partial U} \alpha.
\end{aligned}
$$

Exercise 92. *Let $X \subset \mathbb{E}^3$ be a compact 3-manifold with boundary, $\vec{F}$ be a smooth vector field on an open neighborhood of X. Prove the Divergence Theorem*

$$\int_X \nabla \cdot \vec{F} \, \mathrm{d}V = \int_{\partial X} \vec{F} \cdot \mathrm{d}\vec{S}$$

in Calculus II. Here, the lefthand side is an integration of the first kind, and the righthand side is an integration of the second kind, with the boundary surface oriented by outward normal vectors.

`Hint:` `Orienting` X `by` $\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3$`, then the righthand side becomes` $\int_{\partial X} \omega$ `with`

$$\omega = F_1 \, \mathrm{d}x^2 \wedge \mathrm{d}x^3 + F_2 \, \mathrm{d}x^3 \wedge \mathrm{d}x^1 + F_3 \, \mathrm{d}x^1 \wedge \mathrm{d}x^2.$$

`So` $\mathrm{d}\omega = \nabla \cdot \vec{F} \, \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3$.

§14 Functions of more general type and their derivatives

The purpose here is to introduce the concept of **vector bundle** and the related concept of **connection** on smooth vector bundles.

vector bundle

Let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$. Roughly speaking, a $\mathbb{F}$-vector bundle of rank r is a family of r-dimensional vector spaces over $\mathbb{F}$, *bundled together nicely*. Let us list a few examples.

(i) (**product bundle**). If X is a topological space, for each $x \in X$, via the natural identification with $\mathbb{F}^r$, $\{x\} \times \mathbb{F}^r$ becomes a r-dimensional vector spaces over $\mathbb{F}$, so we arrive at a family of r-dimensional vector spaces over $\mathbb{F}$: $\{\{x\} \times \mathbb{F}^r \mid x \in X\} = X \times \mathbb{F}^r$. Bundled together nicely means the natural map π: $X \times \mathbb{F}^r \to X$ is continuous. This vector bundle is called a **product bundle** over X.

(ii) (**tangent bundle**). Let M be an n-manifold. The family of tangent spaces of M, $\{T_pM \mid p \in M\}$, is a (real) vector bundle. Bundled together nicely means the natural map π: $TM \to M$ is continuous and is locally trivial: each point $p \in M$ has an open neighborhood M_{loc} such that $(TM)|_{M_{loc}} \to M_{loc}$ is isomorphic to the product bundle $M_{loc} \times \mathbb{R}^n \to M_{loc}$. Here $(TM)|_{M_{loc}} := \cup_{q \in M_{loc}} T_qM = TM_{loc}$ and isomorphic means that there is a topological equivalence ϕ which makes the triangle

$\because\ T_qM = T_qM_{loc}$ for each $x \in M_{loc}$

$$\begin{tikzcd} TM_{loc} \arrow[rr, "\phi", "\cong"'] \arrow[dr] & & M_{loc} \times \mathbb{R}^n \arrow[dl] \\ & M_{loc} & \end{tikzcd}$$

commutative, moreover, ϕ is a fiberwise linear equivalence. For example, we can take M_{loc} to be the coordinate patch for a local coordinate system x^i, then $\phi^{-1}(q, \vec{u}) = u^i\, \partial_{x^i}|_q$.

Note 1: In case M is an open set of $\mathbb{E}^n$, the tangent bundle $TM \to M$ is globally trivial.

Note 2: The tangent bundle of a manifold is a smooth real vector bundle because both π and ϕ above are smooth.

Note 3: The product bundles are the *local models* for vector bundles, just as Euclidean spaces are the local models for smooth manifolds.

(iii) (**universal bundle**). Let r and n be integers such that $0 \leq r \leq n$. The family of r-dimensional subspaces of $\mathbb{F}^n$ is a vector bundle of rank r. Bundled together nicely means the natural map $\pi \colon E \to \mathrm{Gr}_r(\mathbb{F}^n)$ is continuous, here E is the disjoint union of this family of vector spaces and $\mathrm{Gr}_r(\mathbb{F}^n)$ is the set of r-dimensional subspaces of $\mathbb{F}^n$. This is a smooth vector bundle.

Definition 5 (Vector Bundle)**.** *Let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$. A rank-r $\mathbb{F}$-vector bundle over a topological space X is a continuous map $\pi \colon E \to X$ such that*

*1) For each point $p \in M$, $E_p := \pi^{-1}(p)$ (called the **fiber** over p) is a $\mathbb{F}$-vector space of dimension r.*

2) $\pi \colon E \to X$ is locally trivial, *i.e., isomorphic to a product bundle.*

Two families of vector spaces over X, say $\pi_i \colon E_i \to X$ $(i = 1, 2)$, are isomorphic if there is a topological equivalence $\phi \colon E_1 \to E_2$ which makes the relevant triangle diagram commutative (i.e., $\pi_2 \phi = \pi_1$) and is fiberwise linear equivalent.

In case all maps involved are smooth, we say the vector bundle is a smooth vector bundle. We say the vector bundle is a real (complex) vector bundle if the field $\mathbb{F}$ is $\mathbb{R}$ ($\mathbb{C}$). Here E is called the **total space**, X is called the **base space**, and π is called the **projection map**.

Remark 42. Keep in mind that all constructions in linear algebra, done fiberwisely, get transported to vector bundles over a fixed base space. For example, for two vector bundles $E \to X$ and $F \to X$, their direct sum $E \oplus F \to X$ is the vector bundle with $(E \oplus F)_x := E_x \oplus F_x$ for each point $x \in X$, their tensor product $E \otimes F \to X$ is the vector bundle with $(E \otimes F)_x := E_x \otimes F_x$ for each point $x \in X$. The **dual bundle** of $E \to X$, written as $E^* \to X$, is the family of vector spaces $\{E_x^* \mid x \in X\}$, the **endomorphism bundle** of $E \to X$, written as $\mathrm{End}(E) \to X$, is the family of vector spaces $\{\mathrm{Hom}(E_x, E_x) \mid x \in X\}$ and is naturally isomorphic to vector bundle $E^* \otimes E \to X$.

sections are twisted vector-valued functions

In Riemannian geometry, all vector bundles involved are smooth real vector bundles.

We shall use $C^\infty(E)$ to denote the space of smooth **sections** on the smooth real vector bundle $\pi \colon E \to M$, i.e., the space of smooth right inverse of π. So $A_k(M) = C^\infty(\wedge^k TM)$ and $\Omega^k(M) = C^\infty(\wedge^k T^*M)$. Note that the Riemann metric is a section of $S^2 T^*M \to M$. In general, if $E \to M$ is a **tensor bundle** (a bundle built via tensor product some copies of $TM \to M$ and $T^*M \to M$), its sections are **tensor fields** on M.

Note that $C^\infty(E)$ is a module over $C^\infty(M)$, with the scalar multiplication done point-wisely: $(fs)(p) = f(p)s(p)$.

For the product bundle $M \times \mathbb{R}^r \to M$, a section s must be of the form $s(p) = (p, f(p))$ for a unique smooth map $f\colon M \to \mathbb{R}^r$. In other words, $C^\infty(M \times \mathbb{R}^r) \equiv C^\infty(M, \mathbb{R}^r)$. Because vector bundles are locally trivial, locally, sections of vector bundles can be represented by vector-valued functions. Globally, we say sections are *twisted* vector-valued functions. It is a nice fact that differential calculus extends to sections of vector bundles.

partial differentiations on sections are connections

Definition 6 (Kozul connection)**.** *A connection ∇ on a real vector bundle $E \to M$ is a $\mathbb{R}$-bilinear map*

$$\begin{array}{rccc} \nabla: & C^\infty(TM) \times C^\infty(E) & \to & C^\infty(E) \\ & (X, s) & \mapsto & \nabla_X s \end{array}$$

such that the operator ∇_X is a derivation on $C^\infty(E)$, and its dependence on X is $C^\infty(M)$-linear.

i.e., ∇_X is additive and satisfies the product rule: $\nabla_X(fs) = \nabla_X f s + f \nabla_X s$ for all real smooth functions f and all sections s. Note $\nabla_X f$ always means $\mathcal{L}_X f$.

With the identification $C^\infty(M \times \mathbb{R}) \equiv C^\infty(M)$ in mind, one can see that the map

$$\nabla: \quad C^\infty(TM) \times C^\infty(M) \to C^\infty(M)$$

that sends (X, f) to $\mathcal{L}_X f$ is a Kozul connection on the product bundle $M \times \mathbb{R} \to M$. So a Kozul connection is a generalized partial differentiation.

An **affine connection** on M is a Kozul connection on the tangent bundle of M, and the Levi-Civita connection on the Riemannian manifold M is the unique Kozul connection on the tangent bundle of M that is metric compatible and torsion free.

The $C^\infty(M)$-linear dependence of ∇_X on X implies the definition $\nabla_x s := \nabla_X s|_p$ makes sense. Here, $x \in T_p X$ and X is a vector field on M that extends x, i.e., $X(p) = x$.

Kozul connections on any vector bundle $E \to M$ exist in abundance, in fact, they form an affine space modelled on the infinite dimensional vector space $C^\infty(\mathrm{Hom}(TM \otimes E, E)) \equiv C^\infty(T^*M \otimes \mathrm{End}(E))$.

Curvature operator R

A small computation shows that the curvature map

$$\begin{array}{rccc} R: & C^\infty(TM) \times C^\infty(TM) \times C^\infty(E) & \to & C^\infty(E) \\ & (X, Y, s) & \mapsto & ([\nabla_X, \nabla_Y] - \nabla_{[X,Y]})s \end{array}$$

is $C^\infty(M)$-trilinear and antisymmetric in (X, Y), so R is a section of the vector bundle

$$\wedge^2 T^*M \otimes \mathrm{End}(E) \to M.$$

In summary, for any two vector fields X and Y on M, operator

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

on $C^\infty(E)$ is a $C^\infty(M)$-linear operator.

pullback connection

Let $E \to N$ be a vector bundle and $f\colon M \to N$ a continuous map. The family of vector spaces $\{E_{f(m)} \mid m \in M\}$ is a vector bundle over M, denoted by $f^*E \to M$, and is called the **pullback** (bundle) of $E \to N$ by f.

Assume everything is smooth, then a section $s \in C^\infty(f^*E)$ is nothing but a smooth lifting:

$$\begin{tikzcd} & E \arrow[d] \\ M \arrow[ur, "s"] \arrow[r, "f"] & N \end{tikzcd}$$

Moreover, using partition of unity, one can see that the pullback on sections $f^*\colon C^\infty(E) \to C^\infty(f^*E)$ is surjective.

$$\begin{tikzcd} & E \arrow[d, shift right] \\ M \arrow[ur, "f^*\tilde{s}"] \arrow[r, "f"] & N \arrow[u, bend right, "\tilde{s}"'] \end{tikzcd}$$

Any Kozul connection ∇ on $E \to N$ can be pulled back to a Kozul connection on the pullback bundle $f^*E \to M$. The pullback connection of ∇ by f, denoted by $f^*\nabla$, is defined as follows:

$$(f^*\nabla_X s)_m := \nabla_{T_m f(X(m))}\tilde{s}, \quad \text{for any } \tilde{s} \in C^\infty(E) \text{ with } f^*\tilde{s} = s$$

Since $\nabla_{T_m f(X(m))}\tilde{s}$ is independent of the choice of $\tilde{s}$ (assumed fact), we rewrite it as $\nabla_{T_m f(X(m))} s$.

(i) Let $c\colon I \to M$ be a parametrized smooth curve in M, then a smooth vector field V along c is nothing but a smooth section of the pullback bundle $c^*TM \to I$. If t is the standard global coordinate on the open interval I, then, since $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t_0} = (t_0, \vec{e}_1)$, we have

$$T_{t_0} c\left(\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t_0}\right) := (c(t_0), J\hat{c}_{t_0}\vec{e}_1) = (c(t_0), c'(t_0)) = \dot{c}(t_0).$$

Then, the definition $\frac{\mathrm{D}V}{\mathrm{d}t} := (c^*\nabla)_{\frac{\mathrm{d}}{\mathrm{d}t}} V$ becomes the definition $\frac{\mathrm{D}V}{\mathrm{d}t}(t_0) := \nabla_{\dot{c}(t_0)} V$. The local representation for $\dot{c}$ is $\frac{\mathrm{d}c^i}{\mathrm{d}t}$, in the sense that

$$\dot{c}(t) = \frac{\mathrm{d}c^i}{\mathrm{d}t}(t) \left.\frac{\partial}{\partial x^i}\right|_{c(t)}.$$

Here c^i represent c locally with respect to local coordinate system x^i, i.e., $c^i(t) = x^i(c(t))$. The local representation for $\frac{\mathrm{D}V}{\mathrm{d}t}$ is

$$\left(\frac{\mathrm{D}V}{\mathrm{d}t}\right)^i := \frac{\mathrm{d}V^i}{\mathrm{d}t} + \Gamma^i_{jk} \frac{\mathrm{d}c^j}{\mathrm{d}t} V^k.$$

Here Γ^i_{jk} means the function $\Gamma^i_{jk}(c(t))$, and V^i represent V locally with respect to local coordinate system x^i in the sense that

$$V(t) = V^i(t) \left.\frac{\partial}{\partial x^i}\right|_{c(t)}.$$

(ii) Let $f\colon I_1 \times I_2 \to M$ be a parametrized smooth surface in M, then a smooth vector field V along f is nothing but a smooth section of the pullback bundle $f^*TM \to I$. Let (s, t) be the standard global coordinates on the open interval $I_1 \times I_2$, and

$$\frac{\partial f}{\partial s}(s_0, t_0) := T_{(s,t)} f\left(\left.\frac{\partial}{\partial s}\right|_{(s_0,t_0)}\right), \quad \frac{\partial f}{\partial t}(s, t) := T_{(s_0,t_0)} f\left(\left.\frac{\partial}{\partial t}\right|_{(s_0,t_0)}\right)$$

and the local representation are respectively

$$\left(\frac{\partial f}{\partial s}\right)^i := \frac{\partial f^i}{\partial s}, \quad \left(\frac{\partial f}{\partial t}\right)^i := \frac{\partial f^i}{\partial t}.$$

Here f^i represent f locally with respect to local coordinate system x^i, i.e., $f^i(s, t) = x^i(f(s, t))$.

Next, we let

$$\frac{\mathrm{D}V}{\partial t} = (f^*\nabla)_{\frac{\partial}{\partial t}} V, \quad \text{i.e.,} \quad \frac{\mathrm{D}V}{\partial t}(s_0, t_0) = \nabla_{\frac{\partial f}{\partial t}(s_0,t_0)} V.$$

and the local representation is

$$\left(\frac{\mathrm{D}V}{\partial t}\right)^i := \frac{\partial V^i}{\partial t} + \Gamma^i_{jk} \frac{\partial f^j}{\partial t} V^k.$$

Here Γ^i_{jk} means the function $\Gamma^i_{jk}(f(s, t))$.

Similarly, we have

$$\frac{\mathrm{D}V}{\partial s} := (f^*\nabla)_{\frac{\partial}{\partial s}} V, \quad \text{i.e.,} \quad \frac{\mathrm{D}V}{\partial s}(s_0, t_0) = \nabla_{\frac{\partial f}{\partial s}(s_0,t_0)} V.$$

With the help of local representations, one can check that

$$\frac{\mathrm{D}}{\partial s}\frac{\partial f}{\partial t} = \frac{\mathrm{D}}{\partial t}\frac{\partial f}{\partial s}.$$

(iii) Suppose that M is a sub-manifold of manifold $\overline{M}$ and ι is the inclusion map of M into $\overline{M}$. The pullback bundle $\iota^*T\overline{M} \to M$ is the restriction bundle $T\overline{M}|_M \to M$, i.e., the family of vector spaces $\{T_p\overline{M} \mid p \in M\}$. The pullback of the connection $\overline{\nabla}$ on the tangent bundle $T\overline{M} \to \overline{M}$ by ι shall still be written as $\overline{\nabla}$, as practiced in many textbooks.

The induced connection on splitting factors

Let ∇ be a Kozul connection on the real vector bundle $E \to M$. Suppose that $E \to M$ splits into the direct sum of two sub-bundles $E' \to M$ and $E'' \to M$ (so $E_m = E'_m \oplus E''_m$ for each $m \in M$). This splitting induces a Kozul connection ∇' on $E' \to M$ and a Kozul connection ∇'' on $E'' \to M$. By definition,

$$\nabla'_X s' := (\nabla_X s')'.$$

Here $s' \in C^\infty(E')$ and $(\nabla_X s')'$ means the E'-component of $\nabla_X s'$. Similarly, we have

$$\nabla''_X s'' := (\nabla_X s'')''.$$

One may ask, are maps $(X, s') \mapsto (\nabla_X s')''$ and $(X, s'') \mapsto (\nabla_X s'')'$ interesting? The answer is yes, though they are not connections.

Example. Suppose that M is a Riemannian sub-manifold of a Riemannian manifold $\overline{M}$, i.e., a sub-manifold of $\overline{M}$ with the induced metric $\iota^* ds^2_{\overline{M}}$ — the first fundamental form of M (w.r.t. the imbedding ι).

Using the Riemann metric, one has the orthogonal splitting $T\overline{M}|_M = TM \oplus NM$, i.e., the splitting into tangential component and the normal component. This is the point-wise splitting $T_p\overline{M} = T_pM \oplus N_pM$ bundled together. The splitting factor $NM \to M$ is called the **normal bundle** of M.

The pullback of Levi-Civita connection $\overline{\nabla}$ on $\overline{M}$ by the inclusion map ι, denoted also by $\overline{\nabla}$, is a Kozul connection on the restriction bundle $T\overline{M}|_M \to M$. The above splitting into tangential component and the normal component induces a connection $\nabla^{\|}$ on the tangent bundle $TM \to M$ and a connection $\nabla^{\perp}$ on the normal bundle $NM \to M$.

Here are the main facts:

(i) The connection $\nabla^{\|}$ on the tangent bundle of M coincides with the Levi-Civita connection ∇ on the Riemannian manifold M.

(ii) The 2nd fundamental form on M is the map $B\colon C^\infty(TM) \times C^\infty(TM) \to C^\infty(NM)$ that sends (X, Y) to $\overline{\nabla}_X Y - \nabla_X Y$ (i.e. $(\overline{\nabla}_X Y)^{\perp}$) is symmetric and bilinear over the commutative ring $C^\infty(M)$. So $B \in C^\infty(S^2TM \otimes NM)$.

(iii) The Gauss equation, Ricci equation, Codazzi equation are relations of the curvature of $\overline{\nabla}$ on the restriction bundle to M with the curvature of ∇ on the tangent bundle of M, the curvature of $\nabla^{\perp}$ on the normal bundle of M, 2nd fundamental form B on M as well as its covariant derivative.

Where to go from here?

We have demonstrated a canonical way of doing partial differentiation (i.e., covariant differentiation) to tensor fields on pseudo-Riemannian manifolds. There are infinitely many ways of doing partial differentiation to tensor fields of a given type on a manifold M; moreover, partial differentiation can be done to sections of a vector bundle, something that is more general than tensor fields.

We have seen that, just as partial differentiation, integration can be formulated in a coordinate-free way as well.

We have briefly mentioned the Gauss-Bonnet formula. This is the first formula that links geometry with topology just as the Newton-Leibniz formula is the first formula that links integration with differentiation. It is not a surprise that this formula has a generalization to higher even dimensions. [5] The ultimate theorem in this spirit is the celebrated Atiyah-Singer formula. Just as Stokes' formula is the fundamental theorem of calculus, Atiyah-Singer formula is the fundamental theorem in differential geometry/topology.

Of course, there is much more ahead, but let us stop here.

[5] The first beautiful intrinsic proof was given by Prof. S. S. Chern.

Appendix 1: Pullback, Push-forward, and Broken Duality

Let $f\colon M \to N$ be a smooth map between manifolds.

Pull back

Let λ be a rank-k covariant tensor field on N, i.e., a smooth function whose value at any point $y \in N$ is a k-form on $T_y N$. Then we can pull λ back to M to get a rank-k covariant tensor field on M, denoted by $f^*\lambda$:

i.e., $\lambda_y := \lambda(y)\colon (T_y N)^k \to \mathbb{R}$ is multilinear

If $k = 0$, then $\lambda = g$ is real smooth function on N, so f^*g is just the composition gf, i.e., $(f^*g)_x := (f^*g)(x) = g(f(x)) := g_{f(x)}$.

More generally, we have

please draw schematic diagram to help understand formula

$$(f^*\lambda)_x = \lambda_{f(x)} \circ (T_x f)^k.$$

Note that λ does not have to be a differential form.

Facts: 1) f^* preserves addition and multiplication (tensor product of various kind), 2) $(fg)^* = g^* f^*$.

Application of f^*

It makes various formulae much clean, hence save us a lot of RAM (random access memory). Here are some examples:

1. *Isometry*. We say a smooth equivalence $f\colon M \to N$ between two Riemannian manifolds is an isometry if

   $\mathrm{d}s^2_M$ denotes the Riemannian metric on M and $f^*\mathrm{d}s^2_N$ means $f^*(\mathrm{d}s^2_N)$ if we want to be clear

   $$f^*\mathrm{d}s^2_N = \mathrm{d}s^2_M.$$

2. *Induced Metric*. An immersion f from smooth manifold M into a Riemannian manifold N induced a Riemannian metric on M, namely $f^*\mathrm{d}s^2_N$.

3. *Flow map*. Any vector field X on M generates an incomplete flow $\phi_t\colon M \to M$, which in turns induces an incomplete linear action on

the space of covariant tensor fields

$$t \cdot T = \varphi_t^* T.$$

By definition, the Lie derivative of T with respect to X, is

$$\mathcal{L}_X T := \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} (t \cdot T).$$

4. A covariant tensor field λ on a Lie group G is called left (right resp.) invariant if $L_x^* \lambda = \lambda$ ($R_x^* \lambda = \lambda$ resp.) for any $x \in G$. In case λ is left-invariant and has rank k, we have

$$\lambda_x = (L_{x^{-1}}^* \lambda)_x = \lambda_{L_{x^{-1}} x} \circ (T_x L_{x^{-1}})^k = \lambda_e \circ (T_x L_{x^{-1}})^k.$$

Note 1: this relation can be used to find a left-invariant rank k covariant tensor field.

Note 2: A left invariant (or right invariant) function is nothing but a constant function. So left invariant (or right invariant) is a concept broader than the concept of "constant".

5. The averaging trick can turn a left invariant metric $\langle\, ,\, \rangle$ into a bi-invariant metric

$$\langle\langle\, ,\, \rangle\rangle := \frac{\int_G R_\cdot^* \langle\, ,\, \rangle\, \omega}{\int_G \omega}$$

Here ω is any top degree left-invariant non-vanishing differential form on a compact connected Lie group G.

Broken dual version: push forward

If P is a rank-k contra-variant tensor field on M, i.e., a smooth function whose value at any point $x \in M$ is a rank-k contra-variant tensor on $T_x M$. Then we can push P forward to N to get a rank-k contra-variant tensor field on N, denoted by $f_? P$, *when and only when f is a diffeomorphism*:

so $P(x) \in (T_x M)^{\otimes k}$ — the tesnor product of k copies of $T_x M$

this extra condition breaks the duality symmetry between contra-covariant and covariant tensor fields.

If $k = 0$, then $P = g$ is real smooth function on N, so $f_? g$ is just the composition $g f^{-1}$, i.e., $(f_? g)(y) = g(f^{-1}(y))$.

In set theory, $f_* g$ means fg, that is why I use $f_?$ rather than f_* here.

If $k = 1$, then $P = X$ is a vector field on M, so $f_? X$ is the vector field on N with $(f_? X)(y) = T_{f^{-1}(y)} f\, (X(f^{-1}(y)))$.

please draw schematic diagram to help understand formula

More generally, we have

$$(f_? P)(y) = (T_{f^{-1}(y)} f)^{\otimes k} (P(f^{-1}(y))).$$

Facts: 1) $f_?$ preserves addition and multiplication, 2) $(fg)_? = f_? g_?$, 3) adjoint relation: Let $f\colon M \to N$ be a diffeomorphism, P be a k-vector fields on M and α be a differential k-form on N, then we have

$$\langle f^* \alpha, P \rangle = f^*(\langle \alpha, f_? P \rangle)$$

Application of $f_?$

1. Let ϕ: $U \to M$ be a local parametrization of M. If we write the resulting local coordinate system as x^i, the push out $\bar{\phi}_? E_i$ of the standard tangent frame frame E_i on U is the resulting local tangent frame $\frac{\partial}{\partial x^i}$. Note: In case ϕ is the inclusion map into $\mathbb{E}^n$, $\frac{\partial}{\partial x^i} = E_i := (1, c_{\vec{e}_i})$.

2. A contra-variant tensor field P on a Lie group G is called left invariant if $(L_x)_? P = P$ for any $x \in G$.

3. Any vector field X on M generates an incomplete flow ϕ_t: $M \to M$, which in turns induces an incomplete action of $\mathbb{R}$ on the space of contra-variant tensor fields

$$t \cdot T := ((\varphi_t)_?)^{-1} T = (\varphi_{-t})_? T.$$

   Note: ϕ_t pushes forward rather than pullback, that is why we need its inverse. Note also that for $f \in C^\infty(M)$, we have $((\varphi_t)_?)^{-1} f = (\varphi_{-t})_? f = f \varphi_t = \varphi_t^* f$.

Appendix 2: Poisson Manifold and Hamilton Equation

Let M be a Poisson manifold with Poisson structure π — a 2-vector field such that $[\pi, \pi] = 0$. Locally we represent π by local functions π^{ab} in the sense that $\pi = \frac{1}{2}\pi^{ab}\partial_a \wedge \partial_b$. The Poisson bracket is

$$\{f, g\} := \langle \pi, \mathrm{d}f \wedge \mathrm{d}g \rangle = \pi^{ab}\partial_a f \partial_b g.$$

In particular $\pi^{ab} = \{x^a, x^b\}$.

In our sign convention in Poisson geometry, for any smooth functions f and g on a Poisson manifold with Poisson bi-vector π, $X_f g = \{f, g\} = \langle \mathrm{d}f \wedge \mathrm{d}g, \pi \rangle$, so $[X_f, X_g] = X_{\{f,g\}}$. Here X_f is called the the Hamiltonian vector field of the function f. Locally we write $\pi = \frac{1}{2}\pi^{ij}\partial_{x^i} \wedge \partial_{x^j}$, then $\pi^{ij} = \{x^i, x^j\}$, $\{f, g\} = \pi^{ij}\partial_{x^i} f \ \partial_{x^j} g$, so $X_f = \pi^{ij}\partial_{x^i} f \ \partial_{x^j}$.

By definition, the **Hamilton equation** with hamiltonian function H is the flow equation of X_H, i.e., $\dot{c}(t) = X_H(c(t))$ (more often written in this clean but less legible form: $\dot{c} = X_H$). Here c is a parametrized smooth curve in M and is called a flow curve of the vector field X_H if it satisfies the flow equation $\dot{c} = X_H$. Locally $\dot{c} = \frac{\mathrm{d}x^i}{\mathrm{d}t}\partial_{x^i}$ (i.e., $\dot{c}(t) = \frac{\mathrm{d}(x^i \circ c)}{\mathrm{d}t}(t)\partial_{x^i}|_{c(t)}$) and $X_H = \{H, x^i\}\partial_{x^i}$, so the flow equation is $\frac{\mathrm{d}x^i}{\mathrm{d}t} = \{H, x^i\}$, more precisely, $\frac{\mathrm{d}(x^i \circ c)}{\mathrm{d}t}(t) = \{H, x^i\}|_{c(t)}$. Note that, the Hamilton equation with hamiltonian function H more often refers to the following equivalent equation

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \{H, f\}, \quad f \in C^\infty(M)$$

which is a clean but less legible form of the equation $\frac{\mathrm{d}(f \circ c)}{\mathrm{d}t}(t) = \{H, f\}|_{c(t)}$.

In case the Poisson manifold is a symplectic manifold with symplectic form ω, the sign convention is chosen such that $\{f, g\} = \langle \omega, X_f \wedge X_g \rangle$, i.e., $-X_f \lrcorner \omega = \mathrm{d}f$. Locally we write $\omega = \frac{1}{2}\omega_{ij}\mathrm{d}x^i \wedge \mathrm{d}x^j$, then $\omega_{ij} = \langle \omega, \partial_{x^i} \wedge \partial_{x^j} \rangle$, $[\pi^{ij}] = -[\omega_{ij}]^{-1}$. In case the symplectic manifold is T^*M, the canonical symplectic form is $\omega_M = \mathrm{d}\theta_M$, where θ_M is the tautological one-form. Locally $\theta_M = p_i \mathrm{d}x^i$, so $\omega_M = \mathrm{d}p_i \wedge \mathrm{d}x^i$, and then $\{p_i, x^j\} = \delta_i^j$.

Assume M is a Riemannian manifold with the Riemannian metric tensor g. Under the vector bundle identification $TM \equiv T^*M$ via the Riemannian metric g, TM becomes a symplectic manifold with symplectic structure $\omega_M = \mathrm{d}\theta_M$, where θ_M is the tautological one-form. The Hamiltonian vector field G_M of the kinetic energy function K is called the **geodesic vector field** G on TM. That is because, its flow, after projected to M, becomes the geodesic flow on M. Locally, $g = g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j$ with $g_{ij} = g(\partial_{x^i}, \partial_{x^j})$, $K = \frac{1}{2}g_{ij}\dot{x}^i\dot{x}^j$, $\theta_M = g_{ij}\dot{x}^j \wedge \mathrm{d}x^i$, $\omega_M = \mathrm{d}(g_{ij}\dot{x}^j) \wedge \mathrm{d}x^i$, so $G_M := X_K = \dot{x}^k\partial_{x^k} - \dot{x}^i\dot{x}^j\Gamma^k_{ij}\partial_{\dot{x}^k}$. Note that, the hyper surface $(TM)_r$ is a contact manifold with its contact form being the restriction to $(TM)_r$ of the tautological one-form θ_M, moreover, the restriction of G_M to $(TM)_r$ is a vector field on $(TM)_r$.

Example 1. Let $\mathfrak{g}$ be a real Lie algebra and $\mathfrak{g}^*$ be its dual vector space. The paring $\mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$ enable us view each element $X \in \mathfrak{g}$ as a function on $\mathfrak{g}^*$. Then ${\mathfrak{g}^*}_{aff}$ is a Poisson manifold with Poisson bracket

$$\{X, Y\} := \langle -, [X, Y]\rangle$$

Note: The symplectic leaves of ${\mathfrak{g}^*}_{aff}$ are the co-adjoint orbits and the induced symplectic form on them is called the KKS-form. Indeed, if G is the adjoint group of $\mathfrak{g}$, and $\mathcal{O}$ is a co-adjoint orbit, then the KKS form ω_{KKS} on the co-adjoint orbit $\mathcal{O}$ is the symplectic form whose value at $f \in \mathcal{O}$ is

$$\omega_{KKS}(X_\xi, X_\eta) = \langle -, [\xi, \eta]\rangle$$

where X_ξ denotes the action-induced vector field on $\mathcal{O}$ of $\xi \in \mathfrak{g}$:

$$X_\xi(f) = \dot{c}_\xi(0), \quad \text{where } c_\xi(t) = \exp t\xi \cdot f$$