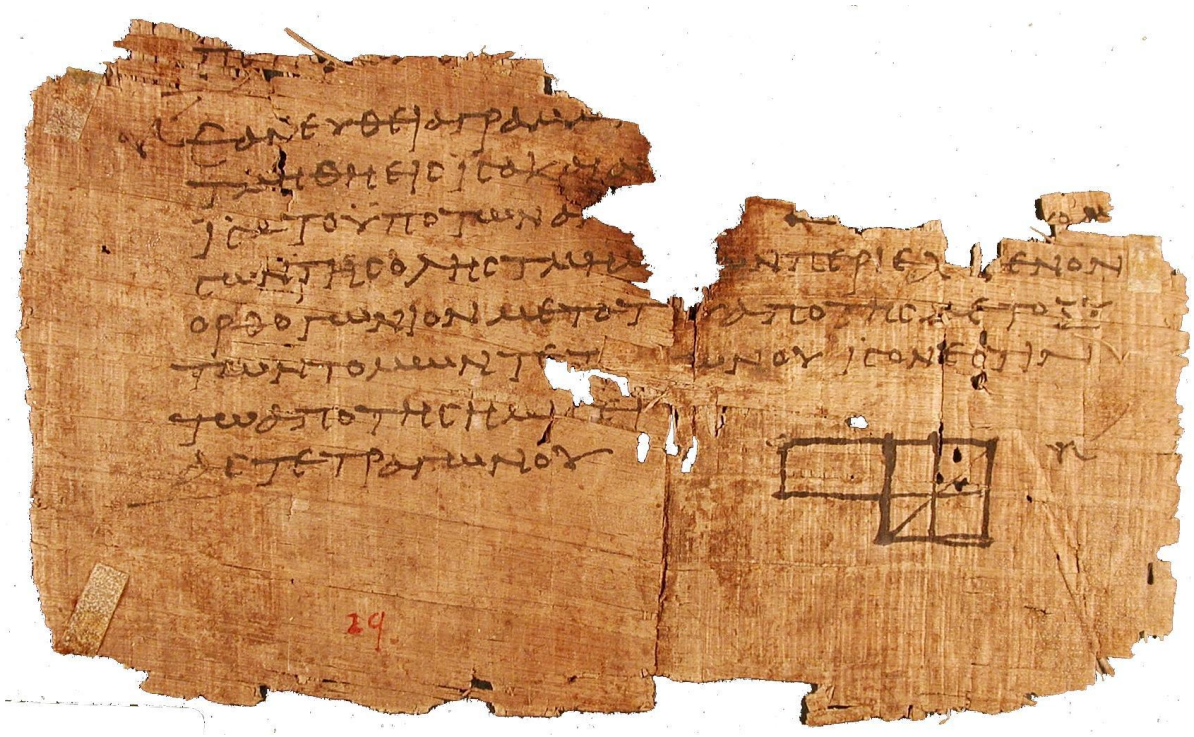
Course Notes

Tensor Calculus and Differential Geometry

2WAH0



Luc Florack

March 10, 2021

Cover illustration: papyrus fragment from Euclid's Elements of Geometry, Book II [8].

Contents

Preface

These course notes are intended for students of all TU/e departments that wish to learn the basics of tensor calculus and differential geometry. Prerequisites are linear algebra and vector calculus at an introductory level. The treatment is condensed, and serves as a complementary source next to more comprehensive accounts that can be found in the (abundant) literature. As a companion for classroom adoption it does provide a reasonably self-contained introduction to the subject that should prepare the student for further self-study.

These course notes are based on course notes written in Dutch by Jan de Graaf. Large parts are straightforward translations. I am therefore indebted to Jan de Graaf for many of the good things. I have, however, taken the liberty to skip, rephrase, and add material, and will continue to update these course notes (the date on the cover reflects the version). I am of course solely to blame for the errors that might have been introduced in this undertaking.

Rigor is difficult to reconcile with simplicity of terminology and notation. I have adopted Jan de Graaf's style of presentation in this respect, by letting the merit of simplicity prevail. The necessary sacrifice of rigor is compensated by a great number of interspersed "caveats", notational and terminological remarks, all meant to train the reader in coming to grips with the parlance of tensor calculus and differential geometry.

Luc Florack

Eindhoven, March 10, 2021.

Notation

Instead of rigorous notational declarations, a non-exhaustive list of examples is provided illustrating the notation for the most important object types used in these course notes:

- Linear mappings: $\mathscr{A}$, $\mathscr{B}$, etc.
- Matrices: A, B, etc.
- Tensors: $\mathbf{A}$, $\mathbf{B}$, etc.
- Vectors: $\mathbf{a}$, $\mathbf{b}$, etc.
- Covectors: $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$, etc.
- Basis vectors: $\mathbf{e}_i$, $\mathbf{f}_i$, etc. (for holonomic basis also $\partial/\partial x^i$, $\partial/\partial y^i$, etc.)
- Basis covectors: $\hat{\mathbf{e}}^i$, $\hat{\mathbf{f}}^i$, etc. (for holonomic basis also dx^i, dy^i, etc.)
- Vector components: a^i, b^i, etc.
- Covector components: a_i, b_i, etc.
- Tensor components: $A^{i_1 \ldots i_p}_{j_1 \ldots j_q}$, $B^{i_1 \ldots i_p}_{j_1 \ldots j_q}$, etc.

1. Prerequisites from Linear Algebra

Linear algebra forms the skeleton of tensor calculus and differential geometry. We recall a few basic definitions from linear algebra, which will play a pivotal role throughout this course.

Reminder

A *vector space* V over the field $\mathbb{K}$ ($\mathbb{R}$ or $\mathbb{C}$) is a set of objects that can be added and multiplied by scalars, such that the sum of two elements of V as well as the product of an element in V with a scalar are again elements of V (closure property). Moreover, the following properties are satisfied for any $u, v, w \in V$ and $\lambda, \mu \in \mathbb{K}$:

- $(u + v) + w = u + (v + w)$,
- there exists an element $o \in V$ such that $o + u = u$,
- there exists an element $-u \in V$ such that $u + (-u) = o$,
- $u + v = v + u$,
- $\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$,
- $(\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot u$,
- $(\lambda\mu) \cdot u = \lambda \cdot (\mu \cdot u)$,
- $1 \cdot u = u$.

The infix operator $+$ denotes *vector addition*, the infix operator $\cdot$ denotes *scalar multiplication*.

Terminology

Another term for vector space is *linear space*.

Remark

- We will almost exclusively consider real vector spaces, i.e. with scalar field $\mathbb{K} = \mathbb{R}$.
- We will invariably consider finite-dimensional vector spaces.

Reminder

A *basis* of an n-dimensional vector space V is a set $\{\mathbf{e}_i\}_{i=1}^n$, such that for each $x \in V$ there exists a unique n-tuple $(x^1, \ldots, x^n) \in \mathbb{R}^n$ such that $x = \sum_{i=1}^n x^i \mathbf{e}_i$.

Reminder

- A *linear operator* $\mathscr{A} : V \to W$, in which V and W are vector spaces over $\mathbb{K}$, is a mapping that satisfies the following property: $\mathscr{A}(\lambda\, v + \mu\, w) = \lambda\, \mathscr{A}(v) + \mu\, \mathscr{A}(w)$ for all $\lambda, \mu \in \mathbb{K}$ and $v, w \in V$.
- The set $\mathscr{L}(V, W)$ of all linear operators of this type is itself a vector space, with the following definitions of vector addition and scalar multiplication: $(\lambda\mathscr{A} + \mu\mathscr{B})\,(v) = \lambda\mathscr{A}(v) + \mu\mathscr{B}(v)$ for all $\lambda, \mu \in \mathbb{K}$, $\mathscr{A}, \mathscr{B} \in \mathscr{L}(V, W)$ and $v \in V$.

Notation

The set of square $n \times n$ matrices will be denoted by $\mathbb{M}_n$.

Reminder

The *determinant* of a square matrix $A \in \mathbb{M}_n$ with entries A_{ij} $(i, j = 1, \ldots, n)$ is given by

$$\det A = \sum_{j_1,\ldots,j_n=1}^{n} [j_1, \ldots, j_n]\, A_{1j_1} \ldots A_{nj_n} = \frac{1}{n!} \sum_{\substack{i_1,\ldots,i_n\,=\,1\\ j_1,\ldots,j_n\,=\,1}}^{n} [i_1, \ldots, i_n]\, [j_1, \ldots, j_n]\, A_{i_1 j_1} \ldots A_{i_n j_n}\,.$$

The *completely antisymmetric symbol* in n dimensions is defined as follows:

$$[i_1, \ldots, i_n] = \begin{cases} +1 & \text{if } (i_1, \ldots, i_n) \text{ is an even permutation of } (1, \ldots, n)\,, \\ -1 & \text{if } (i_1, \ldots, i_n) \text{ is an odd permutation of } (1, \ldots, n)\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Observation

- Exactly $n!$ of the n^n index realizations in $[i_1, \ldots, i_n]$ yield nontrivial results.
- To determine $\det A$, compose all $n!$ products of n matrix entries such that the row labels as well as the column labels of the factors in each product are all distinct. Rearrange factors such that the row labels are incrementally ordered. Assign a sign ± 1 to each product corresponding to the permutation (even/odd) of the column labels observed in this ordered form, and add up everything. Cf. the following example.

Example

Let $A \in \mathbb{M}_3$ be a 3×3 matrix:

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}\,.$$

Carrying out the recipe yields Table 1.1:

product with row labels ordered	sign of column label permutation
$A_{11}A_{22}A_{33}$	$[123] = +1$
$A_{12}A_{23}A_{31}$	$[231] = +1$
$A_{13}A_{21}A_{32}$	$[312] = +1$
$A_{13}A_{22}A_{31}$	$[321] = -1$
$A_{11}A_{23}A_{32}$	$[132] = -1$
$A_{12}A_{21}A_{33}$	$[213] = -1$

Table 1.1: $\det A = A_{11}A_{22}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{21}A_{32} - A_{13}A_{22}A_{31} - A_{11}A_{23}A_{32} - A_{12}A_{21}A_{33}$.

Observation

For a general square matrix $A \in \mathbb{M}_n$ the following identities hold:

$$\begin{aligned}
\sum_{i_1,\ldots,i_n=1}^{n} [i_1,\ldots,i_n]\, A_{i_1 j_1} \ldots A_{i_n j_n} &= [j_1,\ldots,j_n] \det A\,, \\
\sum_{j_1,\ldots,j_n=1}^{n} [j_1,\ldots,j_n]\, A_{i_1 j_1} \ldots A_{i_n j_n} &= [i_1,\ldots,i_n] \det A\,, \\
\sum_{i_1,\ldots,i_n=1}^{n} \sum_{j_1,\ldots,j_n=1}^{n} [i_1,\ldots,i_n]\, [j_1,\ldots,j_n]\, A_{i_1 j_1} \ldots A_{i_n j_n} &= n! \det A\,.
\end{aligned}$$

Products of completely antisymmetric symbols can be expressed in terms of determinants:

$$[i_1,\ldots,i_n]\,[j_1,\ldots,j_n] = \det \begin{pmatrix} \delta_{i_1 j_1} & \ldots & \delta_{i_1 j_n} \\ \vdots & & \vdots \\ \delta_{i_n j_1} & \ldots & \delta_{i_n j_n} \end{pmatrix} \stackrel{\text{def}}{=} \delta^{i_1 \ldots i_n}_{j_1 \ldots j_n}\,,$$

and, more generally,

$$\sum_{i_1,\ldots,i_k=1}^{n} [i_1,\ldots,i_k,i_{k+1},\ldots,i_n]\,[i_1,\ldots,i_k,j_{k+1},\ldots,j_n] = k! \det \begin{pmatrix} \delta_{i_{k+1} j_{k+1}} & \ldots & \delta_{i_{k+1} j_n} \\ \vdots & & \vdots \\ \delta_{i_n j_{k+1}} & \ldots & \delta_{i_n j_n} \end{pmatrix}.$$

Example

- $[i,j,k]\,[\ell,m,n] = \delta_{i\ell}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{k\ell} + \delta_{in}\delta_{j\ell}\delta_{km} - \delta_{im}\delta_{j\ell}\delta_{kn} - \delta_{i\ell}\delta_{jn}\delta_{km} - \delta_{in}\delta_{jm}\delta_{k\ell}$.
- $\sum_{i=1}^{3} [i,j,k]\,[i,m,n] = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}$.
- $\sum_{i,j=1}^{3} [i,j,k]\,[i,j,n] = 2\delta_{kn}$.
- $\sum_{i,j,k=1}^{3} [i,j,k]\,[i,j,k] = 6$.

Caveat

One often writes $\det A_{ij}$ instead of $\det A$. The labels i and j are *not* free indices in this case!

Reminder

The *cofactor matrix* of $A \in \mathbb{M}_n$ is the matrix $\tilde{A} \in \mathbb{M}_n$ with entries $\tilde{A}^{ij}$ given by

$$\tilde{A}^{ij} = \frac{\partial \det A}{\partial A_{ij}}\,.$$

The *adjugate matrix* of $A \in \mathbb{M}_n$ is a synonymous term for $\tilde{A}^{\mathrm{T}} \in \mathbb{M}_n$.

Notation

Cofactor and adjugate matrix of $A \in \mathbb{M}_n$ are sometimes indicated by $\mathrm{cof}A$ and $\mathrm{adj}A$, respectively.

Notation

In any expression of the form X^i_i, with a matching pair of sub- and superscript $i = 1, \ldots, n$, in which $n = \dim V$ is the dimension of an underlying vector space V, the Einstein summation convention applies, meaning that it is to be understood as an i-index-free sum:

$$X^i_i \stackrel{\text{def}}{=} \sum_{i=1}^{n} X^i_i \,.$$

This convention applies to each pair of matching sub- and superscript within the same index range. Indices to which the convention applies are referred to as spatial (or, depending on context, spatiotemporal) indices. Note that matching upper and lower indices are dummies that can be arbitrarily relabelled provided this does not lead to conflicting notation.

Observation

- Each term in the polynomial $\det A$ contains at most one factor A_{ij} for each fixed $i, j = 1, \ldots, n$.
- As a result, $\tilde{A}^{ij}$ does not depend on A_{ij} for fixed $i, j = 1, \ldots, n$: $\dfrac{\partial \tilde{A}^{ij}}{\partial A_{ij}} = \dfrac{\partial^2 \det A}{\partial A^2_{ij}} = 0$.
- If $A : \mathbb{R} \to \mathbb{M}_n : t \mapsto A(t)$ is a matrix-valued function, $\dfrac{d \det A}{dt} = \dfrac{\partial \det A}{\partial A_{ij}} \dfrac{dA_{ij}}{dt} = \tilde{A}^{ij} \dfrac{dA_{ij}}{dt}$ (chain rule).
- $A\tilde{A}^{\mathrm{T}} = \tilde{A}^{\mathrm{T}} A = \det A \, I$, or $A \,\mathrm{adj} A = \mathrm{adj} A \, A = \det A \, I$, i.e. $A_{jk}\tilde{A}^{ik} = \tilde{A}^{ki} A_{kj} = \det A \, \delta^i_j$.
- If A is invertible, then $\tilde{A}^{\mathrm{T}} = \mathrm{adj} A = \det A \, A^{-1}$.
- Each entry $\tilde{A}^{ij}$ of $\tilde{A} = \mathrm{cof} A$ is a homogeneous polynomial of degree $n-1$ in terms of the entries A_{ij} of A.
- Unlike A^{-1}, $\tilde{A} = \mathrm{cof} A$ and $\tilde{A}^{\mathrm{T}} = \mathrm{adj} A$ are always well-defined, even if A is singular.

2. Tensor Calculus

2.1. Vector Spaces and Bases

Ansatz

An n-dimensional vector space V over $\mathbb{R}$ furnished with a basis $\{\mathbf{e}_i\}$.

Notation

Unless stated otherwise the *Einstein summation convention* applies to identical pairs of upper and lower indices. Thus if $\{X^i_j\}$ is any collection of numbers, then $X^i_i \stackrel{\text{def}}{=} \sum_{i=1}^n X^i_i$.

Terminology

The act of constraining one upper and one lower index to be equal—and thus tacitly assuming summation, v.s.—is called an *index contraction*, or a *trace*.

Remark

Contracted indices are dummies that can be arbitrarily relabeled, as opposed to free indices. Thus for example $X^i_i = X^j_j$.

Remark

For each basis $\{\mathbf{e}_i\}$ of V and each $\mathbf{x} \in V$ there exists a unique collection of real numbers $\{x^i\}$ such that $\mathbf{x} = x^i\mathbf{e}_i$, recall chapter 1.

Terminology

The real numbers x^i are referred to as the *contravariant components* of the vector $\mathbf{x}$ relative to the vector basis $\{\mathbf{e}_i\}$.

Definition 2.1 *The* Kronecker symbol*, or* Kronecker-δ*, is defined as*

$$\delta^i_j = \begin{cases} 1 & \textit{if} \quad i = j \\ 0 & \textit{if} \quad i \neq j\,. \end{cases}$$

Theorem 2.1 *Let* $\mathbf{x} = x^i\mathbf{e}_i \in V$ *and* $\mathbf{f}_j = A^i_j\mathbf{e}_i$ *define a change of basis. If* $x^i\mathbf{e}_i = y^i\mathbf{f}_i$*, then*

$$x^i = A^i_j y^j \quad \textit{or, equivalently,} \quad y^i = B^i_j x^j\,,$$

in which, by definition, $A^i_k B^k_j = \delta^i_j$.

Caveat

Pay attention to the occurrence of the components A^i_j in the transformation of basis vectors $\mathbf{e}_i$ and contravariant components x^i, respectively.

Observation

Given a *smooth manifold*—to be defined later—with a coordinate system $\{x^i\}$, the collection of linear partial derivative operators $\{\mathbf{e}_i \equiv \partial_i \equiv \partial/\partial x^i\}$ at any (implicit) fiducial point defines a vector space basis. Upon a coordinate transformation $x^i = x^i(\overline{x}^j)$ the chain rule relates the corresponding bases by

$$\overline{\partial}_j = A^i_j \partial_i \,,$$

in which $\mathbf{f}_j \equiv \overline{\partial}_j \equiv \partial/\partial \overline{x}^j$, and

$$A^i_j = \frac{\partial x^i}{\partial \overline{x}^j} \,.$$

This is formally identical to the basis transformation in Theorem 2.1 on page 7.

Notation

On a smooth manifold one often encounters vector decomposition relative to a coordinate basis, i.e. $\mathbf{v} = v^i \partial_i$. (The fiducial point is implicit here.)

Terminology

The set $\{\partial_i\}$ is called a *coordinate (vector) basis*, or a *holonomic (vector) basis*. By definition it spans the *tangent space* of the manifold at the fiducial point.

Reminder

- Basis vectors are labeled with lower indices.
- Contravariant vector components are labeled with upper indices.

Remark

The distinction into upper and lower indices is relevant only if the indices are intrinsically constrained to take values in the range $\{1, \ldots, n\}$, where $n = \dim V$.

2.2. Dual Vector Spaces and Dual Bases

Ansatz

An n-dimensional vector space V over $\mathbb{R}$.

Definition 2.2 *A* linear functional *on V is a linear mapping from V into $\mathbb{R}$. In other words, $\hat{\mathbf{f}} : V \to \mathbb{R}$ is a linear functional if for all $\mathbf{x}, \mathbf{y} \in V$, $\lambda, \mu \in \mathbb{R}$,*

$$\hat{\mathbf{f}}(\lambda \mathbf{x} + \mu \mathbf{y}) = \lambda \hat{\mathbf{f}}(\mathbf{x}) + \mu \hat{\mathbf{f}}(\mathbf{y}) \,.$$

Remark

The set V^* consisting of all linear functionals on V can be turned into a vector space in the usual way, viz. let $\hat{\mathbf{f}}, \hat{\mathbf{g}} \in V^*$ and $\lambda, \mu \in \mathbb{R}$, define the linear combination $\lambda \hat{\mathbf{f}} + \mu \hat{\mathbf{g}} \in V^*$ by virtue of the identification

$$(\lambda \hat{\mathbf{f}} + \mu \hat{\mathbf{g}})(\mathbf{x}) = \lambda \hat{\mathbf{f}}(\mathbf{x}) + \mu \hat{\mathbf{g}}(\mathbf{x})$$

for all $\mathbf{x} \in V$, recall chapter 1.

Definition 2.3 *The* dual space *is the vector space of linear functionals on V: $V^* = \mathscr{L}(V, \mathbb{R})$.*

Ansatz

An n-dimensional vector space V over $\mathbb{R}$ furnished with a basis $\{\mathbf{e}_i\}$, and its dual space V^*.

Remark

Each basis $\{\mathbf{e}_i\}$ of V induces a basis $\left\{\hat{\mathbf{e}}^i\right\}$ of V^* as follows. If $\mathbf{x} = x^i\mathbf{e}_i$, then $\hat{\mathbf{e}}^i(\mathbf{x}) = x^i$.

Result 2.1 *The* dual basis $\left\{\hat{\mathbf{e}}^i\right\}$ *of* V^* *is uniquely determined through its defining property*

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

Proof of Result 2.1. Let $\hat{\mathbf{f}} \in V^*$.

- For any $\mathbf{x} = x^i\mathbf{e}_i \in V$ we have $\hat{\mathbf{f}}(\mathbf{x}) = \hat{\mathbf{f}}(x^i\mathbf{e}_i) = \hat{\mathbf{f}}(\mathbf{e}_i)x^i = \hat{\mathbf{f}}(\mathbf{e}_i)\hat{\mathbf{e}}^i(\mathbf{x}) = (\hat{\mathbf{f}}(\mathbf{e}_i)\hat{\mathbf{e}}^i)(\mathbf{x}) = (f_i\hat{\mathbf{e}}^i)(\mathbf{x})$ with $f_i = \hat{\mathbf{f}}(\mathbf{e}_i)$. This shows that the set $\{\hat{\mathbf{e}}^i\}$ spans V^*.
- Suppose $\alpha_i\hat{\mathbf{e}}^i = 0 \in V^*$, then $\alpha_i = \alpha_j\delta^j_i = \alpha_j\hat{\mathbf{e}}^j(\mathbf{e}_i) = (\alpha_j\hat{\mathbf{e}}^j)(\mathbf{e}_i) = 0 \in \mathbb{R}$ for all $i = 1, \ldots, n$. This shows that $\{\hat{\mathbf{e}}^i\}$ is a linearly independent set.

Assumption

In view of this result, it will henceforth be tacitly understood that a given vector space V (with basis $\{\mathbf{e}_i\}$) implies the existence of its dual space V^* (with corresponding dual basis $\left\{\hat{\mathbf{e}}^i\right\}$).

Observation

The dual basis V^* has the same dimension as V.

Terminology

A linear functional $\hat{\mathbf{f}} \in V^*$ is alternatively referred to as a *covector* or 1*-form*. The dual space V^* of a vector space V is also known as the *covector space* or the *space of* 1*-forms*.

Remark

For each covector basis $\left\{\hat{\mathbf{e}}^i\right\}$ of V^* and each $\hat{\mathbf{x}} \in V^*$ there exists a unique collection of real numbers $\{x_i\}$ such that $\hat{\mathbf{x}} = x_i\hat{\mathbf{e}}^i$.

Terminology

The real numbers x_i are referred to as the *covariant components* of the covector $\hat{\mathbf{x}}$ relative to covector basis $\left\{\hat{\mathbf{e}}^i\right\}$.

We end this section with the dual analogue of Theorem 2.1 on page 7:

Theorem 2.2 *The change of basis* $\mathbf{f}_j = A^i_j\mathbf{e}_i$ *induces a change of dual vector basis*

$$\hat{\mathbf{f}}^j = B^j_i\hat{\mathbf{e}}^i\,,$$

in which $A^k_jB^i_k = \delta^i_j$. *Consequently, if* $\hat{\mathbf{x}} = x_i\hat{\mathbf{e}}^i = y_i\hat{\mathbf{f}}^i \in V^*$, *then*

$$x_i = B^j_iy_j \quad \textit{or, equivalently,} \quad y_i = A^j_ix_j\,.$$

Proof of Theorem 2.2.

- Given $\mathbf{f}_j = A^\ell_j\mathbf{e}_\ell$ and $\hat{\mathbf{f}}^i = B^i_k\hat{\mathbf{e}}^k$ we obtain $\delta^i_j = \hat{\mathbf{f}}^i(\mathbf{f}_j) = B^i_k\hat{\mathbf{e}}^k(A^\ell_j\mathbf{e}_\ell) = B^i_kA^\ell_j\hat{\mathbf{e}}^k(\mathbf{e}_\ell) = B^i_kA^\ell_j\delta^k_\ell = B^i_kA^k_j$, from which it follows that $B = A^{-1}$.

- Thus if we define $x_i \hat{\mathbf{e}}^i = y_i \hat{\mathbf{f}}^i$, then $y_j \hat{\mathbf{f}}^j = y_j B^j_i \hat{\mathbf{e}}^i$, whence $x_i = B^j_i y_j$.
- Inversion yields the last identity.

Terminology

- Theorem 2.1, page 7, defines the so-called *vector transformation law*.
- Theorem 2.2, page 9, defines the so-called *covector transformation law*.

Observation

Recall the observation on page 8. Given a smooth manifold with a coordinate system $\{x^i\}$, and a coordinate transformation $x^i = x^i(\overline{x}^j)$, the chain rule (at any fiducial point) states that

$$d\overline{x}^j = B^j_i dx^i \quad \text{in which} \quad B^j_i = \frac{\partial \overline{x}^j}{\partial x^i} \,.$$

Identifying $\hat{\mathbf{e}}^i \equiv dx^i$ and $\hat{\mathbf{f}}^j \equiv d\overline{x}^j$, this is formally identical to the dual vector basis transformation in Theorem 2.2. In particular we have, recall Result 2.1,

$$dx^i(\partial_j) = \delta^i_j \,.$$

Notation

On a smooth manifold one often encounters covector decomposition relative to a coordinate basis, i.e. $\hat{\mathbf{v}} = v_i dx^i$. (The fiducial point is implicit here.)

Terminology

The set $\{dx^i\}$ is called a *coordinate (dual vector) basis*, or *holonomic (dual vector) basis*. By definition it spans the *cotangent space* of the manifold at the fiducial point.

Reminder

- Basis covectors are labeled with upper indices.
- Covariant components are labeled with lower indices.

2.3. The Kronecker Tensor

Ansatz

An n-dimensional vector space V over $\mathbb{R}$.

Notation

Instead of a standard function notation for $\hat{\mathbf{f}} \in V^*$ one often employs the *bracket formalism* $\langle \hat{\mathbf{f}}, \cdot \rangle \in V^*$. Function evaluation is achieved by dropping a vector argument $\mathbf{x} \in V$ into the empty slot: $\langle \hat{\mathbf{f}}, \mathbf{x} \rangle = \hat{\mathbf{f}}(\mathbf{x}) \in \mathbb{R}$. In particular we have

$$\langle \hat{\mathbf{e}}^i, \mathbf{e}_j \rangle = \hat{\mathbf{e}}^i(\mathbf{e}_j) = \delta^i_j$$

for dual bases, recall Result 2.1 on page 9.

Definition 2.4 *The bilinear map* $\langle \cdot, \cdot \rangle : V^* \times V \to \mathbb{R} : (\hat{\mathbf{f}}, \mathbf{x}) \mapsto \langle \hat{\mathbf{f}}, \mathbf{x} \rangle$ *is called the* Kronecker tensor.

Remark

The general concept of a *tensor* will be explained in section 2.8.

Caveat

Do not confuse the Kronecker tensor with an inner product. Unlike the Kronecker tensor an inner product requires two vector arguments from the *same* vector space.

Observation

The Kronecker tensor is bilinear:

- $\forall\, \hat{\mathbf{f}}, \hat{\mathbf{g}} \in V^*, \mathbf{x} \in V, \lambda, \mu \in \mathbb{R}: \quad \langle \lambda\hat{\mathbf{f}} + \mu\hat{\mathbf{g}}, \mathbf{x}\rangle = \lambda\langle \hat{\mathbf{f}}, \mathbf{x}\rangle + \mu\langle \hat{\mathbf{g}}, \mathbf{x}\rangle,$
- $\forall\, \hat{\mathbf{f}} \in V^*, \mathbf{x}, \mathbf{y} \in V, \lambda, \mu \in \mathbb{R}: \quad \langle \hat{\mathbf{f}}, \lambda\mathbf{x} + \mu\mathbf{y}\rangle = \lambda\langle \hat{\mathbf{f}}, \mathbf{x}\rangle + \mu\langle \hat{\mathbf{f}}, \mathbf{y}\rangle.$
- Result 2.1 on page 9 shows that the *components of the Kronecker tensor* are given by the Kronecker symbol.
- If $\hat{\mathbf{f}} = f_i\hat{\mathbf{e}}^i \in V^*$ and $\mathbf{x} = x^i\mathbf{e}_i \in V$, then bilinearity of the Kronecker tensor implies $\langle \hat{\mathbf{f}}, \mathbf{x}\rangle = f_i x^j \langle \hat{\mathbf{e}}^i, \mathbf{e}_j\rangle = f_i x^j \delta^i_j = f_i x^i$. Thus Kronecker tensor and index contraction are close-knit.

Remark

- The components of the Kronecker tensor are independent of the choice of basis. (Verify!)
- Given V, the assumption on page 9 implies that the Kronecker tensor exists.
- Since V^* is itself a vector space one may consider its dual $(V^*)^* = V^{**}$. Without proof we state that if V (and thus V^*) is finite-dimensional, then V^{**} is isomorphic to, i.e. may be identified with V. This does not generalize to infinite-dimensional vector spaces.

Notation

The use of the formally equivalent notations $\hat{\mathbf{f}}$ and $\langle \hat{\mathbf{f}}, \cdot\,\rangle$ for an element of V^* reflects our inclination to regard it either as an atomic entity, or as a linear mapping with unspecified argument. The previous remark justifies the use of similar alternative notations for an element of V, viz. either as $\mathbf{x}$, say, or as $\langle\, \cdot\,, \mathbf{x}\rangle \in V^{**} \sim V$.

2.4. Inner Products

Ansatz

An n-dimensional vector space V over $\mathbb{R}$.

Definition 2.5 *An* inner product *on a real vector space V is a nondegenerate positive symmetric bilinear mapping of the form $(\,\cdot\,|\,\cdot\,): V \times V \to \mathbb{R}$ which satisfies*

- $\forall\, \mathbf{x}, \mathbf{y} \in V: \quad (\mathbf{x}|\mathbf{y}) = (\mathbf{y}|\mathbf{x}),$
- $\forall\, \mathbf{x}, \mathbf{y}, \mathbf{z} \in V, \lambda, \mu \in \mathbb{R}: \quad (\mathbf{x}|\lambda\mathbf{y} + \mu\mathbf{z}) = \lambda\,(\mathbf{x}|\mathbf{y}) + \mu\,(\mathbf{x}|\mathbf{z}),$
- $\forall\, \mathbf{x} \in V: \quad (\mathbf{x}|\mathbf{x}) \geq 0 \quad$ *and* $\quad (\mathbf{x}|\mathbf{x}) = 0 \quad$ *iff* $\quad \mathbf{x} = \mathbf{0}.$

Definition 2.5 will be our default. In the context of smooth manifolds it is referred to as a *Riemannian inner product* or, especially in physics, as a 'Riemannian metric'. Occasionally we may encounter an inner product on a complex vector space. Below, $\overline{z} = a - bi \in \mathbb{C}$ denotes the complex conjugate of $z = a + bi \in \mathbb{C}$.

Definition 2.6 *An* inner product *on a complex vector space V is a nondegenerate positive hermitean sesquilinear mapping of the form $(\,\cdot\,|\,\cdot\,) : V \times V \to \mathbb{C}$ which satisfies*

- $\forall\, \mathbf{x}, \mathbf{y} \in V : \quad (\mathbf{x}|\mathbf{y}) = \overline{(\mathbf{y}|\mathbf{x})}$,
- $\forall\, \mathbf{x}, \mathbf{y}, \mathbf{z} \in V, \lambda, \mu \in \mathbb{C} : \quad (\mathbf{x}|\lambda\mathbf{y} + \mu\mathbf{z}) = \lambda\,(\mathbf{x}|\mathbf{y}) + \mu\,(\mathbf{x}|\mathbf{z})$,
- $\forall\, \mathbf{x} \in V : \quad (\mathbf{x}|\mathbf{x}) \geq 0$ *and* $(\mathbf{x}|\mathbf{x}) = 0$ *iff* $\mathbf{x} = \mathbf{0}$.

Caveat

- Note that for the complex case $(\lambda\mathbf{x} + \mu\mathbf{y}|\mathbf{z}) = \overline{\lambda}\,(\mathbf{x}|\mathbf{z}) + \overline{\mu}\,(\mathbf{y}|\mathbf{z})$, whence the attribute 'sesquilinear'.
- The term "inner product" is used in different contexts to denote bilinear or sesquilinear forms satisfying different sets of axioms. Examples include the *Lorentzian inner product* in the context of special relativity (v.i.) and the *symplectic inner product* in the Hamiltonian formalism of classical mechanics.
- Unlike with the Kronecker tensor, the existence of an inner product is never implied, but must always be stipulated explicitly.

Definition 2.7 *A* pseudo-Riemannian inner product *on a real vector space V is a nondegenerate symmetric bilinear mapping of the form $(\,\cdot\,|\,\cdot\,) : V \times V \to \mathbb{R}$ which satisfies*

- $\forall\, \mathbf{x}, \mathbf{y} \in V : \quad (\mathbf{x}|\mathbf{y}) = (\mathbf{y}|\mathbf{x})$,
- $\forall\, \mathbf{x}, \mathbf{y}, \mathbf{z} \in V, \lambda, \mu \in \mathbb{R} : \quad (\mathbf{x}|\lambda\mathbf{y} + \mu\mathbf{z}) = \lambda\,(\mathbf{x}|\mathbf{y}) + \mu\,(\mathbf{x}|\mathbf{z})$,
- $\forall\, \mathbf{x} \in V$ *with* $\mathbf{x} \neq \mathbf{0}\ \exists\, \mathbf{y} \in V$ *such that* $(\mathbf{x}|\mathbf{y}) \neq 0$.

A Lorentzian inner product *is a pseudo-Riemannian inner product with signature* $(+, -, \ldots, -)$ *or* $(-, +, \ldots, +)$.

Note that Definition 2.5 is a special case of Definition 2.7, viz. with signature $(+, \ldots, +)$. The signature of a Lorentzian inner product contains exactly one $+$ or one $-$ sign. The notion of signature is explained below in Definition 2.8.

Terminology

A vector space V endowed with an inner product $(\,\cdot\,|\,\cdot\,)$ is also called an *inner product space*.

Ansatz

An n-dimensional inner product space V over $\mathbb{R}$ furnished with a basis $\{\mathbf{e}_i\}$.

Definition 2.8 *The* Gram matrix *G is the matrix with components $g_{ij} = (\mathbf{e}_i|\mathbf{e}_j)$. Its* signature *$(s_1, \ldots, s_n)$ is defined as the (basis independent) n-tuple of signs $s_i = \pm$, $i = 1, \ldots, n$, of the (real) eigenvalues of the Gram matrix for a given ordering of its eigenvectors.*

Remark

Clearly, G is symmetric. Another consequence of the definition is that G is invertible in the pseudo-Riemannian case. For suppose G would be singular, then there exists $\mathbf{x} = x^i\mathbf{e}_i \in V$, $\mathbf{x} \neq \mathbf{0}$, such that $x^i g_{ij} = 0$. As a result we find that for all $\mathbf{y} = y^i\mathbf{e}_i \in V$, $0 = x^i g_{ij} y^j = x^i\,(\mathbf{e}_i|\mathbf{e}_j)\,y^j = \left(x^i\mathbf{e}_i|y^j\mathbf{e}_j\right) = (\mathbf{x}|\mathbf{y})$. This contradicts the non-degeneracy axiom of Definition 2.7.

Terminology

The components of the *inverse Gram matrix* G^{-1} are denoted by g^{ij}, i.e.

$$g^{ik} g_{kj} = \delta^i_j \,.$$

An inner product, if it exists, can be related to the Kronecker tensor, as we will see below.

Ansatz

An n-dimensional inner product space V over $\mathbb{R}$.

Theorem 2.3 *There exists a bijective linear map $\mathbf{G} : V \to V^*$, with inverse $\mathbf{G}^{-1} : V^* \to V$, such that*

- $\forall\, \mathbf{x}, \mathbf{y} \in V : \quad (\mathbf{x}|\mathbf{y}) = \langle \mathbf{G}(\mathbf{x}), \mathbf{y} \rangle$,
- $\forall\, \hat{\mathbf{x}} \in V^*, \mathbf{y} \in V : \quad \big(\mathbf{G}^{-1}(\hat{\mathbf{x}})|\mathbf{y}\big) = \langle \hat{\mathbf{x}}, \mathbf{y} \rangle$.

The matrix representations of $\mathbf{G}$ and $\mathbf{G}^{-1}$ are given by the Gram matrix G and its inverse G^{-1}, respectively, recall Definition 2.8 on page 12.

Proof of Theorem 2.3.

- For fixed $\mathbf{x} \in V$ define $\hat{\mathbf{x}} \in V^*$ as the linear map $\hat{\mathbf{x}} : V \to \mathbb{R} : \mathbf{y} \mapsto (\mathbf{x}|\mathbf{y})$. Subsequently define the linear map $\mathbf{G} : V \to V^* : \mathbf{x} \mapsto \mathbf{G}(\mathbf{x}) \equiv \hat{\mathbf{x}}$.
- For the second part we need to prove invertibility of $\mathbf{G}$. Suppose $\mathbf{x} \in V, \mathbf{x} \neq \mathbf{0}$, is such that $\mathbf{G}(\mathbf{x}) = \mathbf{0}$. Then $0 = \langle \mathbf{G}(\mathbf{x}), \mathbf{y} \rangle = (\mathbf{x}|\mathbf{y})$ for all $\mathbf{y} \in V$, contradicting the nondegeneracy property of the inner product. Hence $\mathbf{G}$ is one-to-one[1]. Since $\dim V = n < \infty$ it is also bijective, hence invertible.
- To compute the matrix representation G of $\mathbf{G}$, decompose $\mathbf{x} = x^i \mathbf{e}_i$, $\mathbf{y} = y^i \mathbf{e}_i$, and $\mathbf{G}(\mathbf{x}) = \mathbf{G}(x^i \mathbf{e}_i) = x^i G_{ij} \hat{\mathbf{e}}^j$, with matrix components defined via $\mathbf{G}(\mathbf{e}_i) = G_{ij} \hat{\mathbf{e}}^j$.
  - By construction (first item above) and by virtue of bilinearity of the inner product and the definition of the Gram matrix, we have $\langle \mathbf{G}(\mathbf{x}), \mathbf{y} \rangle = (\mathbf{x}|\mathbf{y}) = x^i y^j \, (\mathbf{e}_i|\mathbf{e}_j) = g_{ij} x^i y^j$.
  - By definition of duality, we have $\langle \mathbf{G}(\mathbf{x}), \mathbf{y} \rangle = G_{ij} x^i y^k \langle \hat{\mathbf{e}}^j, \mathbf{e}_k \rangle = G_{ij} x^i y^k \delta^j_k = G_{ij} x^i y^j$.

  Since $x^i, y^j \in \mathbb{R}$ are arbitrary it follows that $G_{ij} = g_{ij}$, recall Definition 2.8.
- Consequently, the inverse $\mathbf{G}^{-1} : V^* \to V$ has a matrix representation G^{-1} corresponding to the inverse Gram matrix: $(G^{-1})^{ij} = g^{ij}$.

Remark

- Note the synonymous formulations: $\mathbf{G}(\mathbf{x}) \equiv \hat{\mathbf{x}} \equiv \langle \mathbf{G}(\mathbf{x}), \cdot \rangle \equiv (\mathbf{x}| \cdot)$. Recall the comment on notation, page 11. Definition 2.9 below introduces yet another synonymous form.
- The *Riesz representation theorem*, which states that for each $\hat{\mathbf{x}} \in V^*$ there exists one and only one $\mathbf{x} \in V$ such that $\langle \hat{\mathbf{x}}, \mathbf{y} \rangle = (\mathbf{x}|\mathbf{y})$ for all $\mathbf{y} \in V$, is an immediate consequence of Theorem 2.3, viz. $\mathbf{x} = \mathbf{G}^{-1}(\hat{\mathbf{x}})$.

[1] The terms "one-to-one" and "injective" are synonymous. A function $f : D \to C$ is injective if it associates distinct arguments with distinct values. If the codomain C of the function is replaced by its actual range $f(D)$, then it is also surjective, and hence bijective.

Notation

- We reserve boldface symbols for intrinsic quantities such as vectors, covectors, and (other) linear mappings, in order to distinguish them from their respective components relative to a fiducial basis.

Definition 2.9 *The conversion operators* $\sharp : V \to V^*$ *and* $\flat : V^* \to V$ *are introduced as shorthand for* $\mathbf{G}$*, respectively* $\mathbf{G}^{-1}$*, viz. if* $\mathbf{v} \in V$*,* $\hat{\mathbf{v}} \in V^*$*, then*

$$\sharp\mathbf{v} \stackrel{\text{def}}{=} \mathbf{G}(\mathbf{v}) \in V^* \quad \textit{respectively} \quad \flat\hat{\mathbf{v}} \stackrel{\text{def}}{=} \mathbf{G}^{-1}(\hat{\mathbf{v}}) \in V\,.$$

Terminology

$\mathbf{G}$ is commonly known as the *metric*, and $\mathbf{G}^{-1}$ as the *dual metric*. The operators $\sharp : V \to V^*$ and $\flat : V^* \to V$ are pronounced as *sharp* and *flat*, respectively.

Caveat

In the literature, the roles of sharp and flat are often reversed! Make sure to adhere to an unambiguous convention.

Result 2.2 *Let* $\mathbf{v} = v^i\mathbf{e}_i \in V$ *and* $\hat{\mathbf{v}} = v_i\hat{\mathbf{e}}^i \in V^*$*, then*

- $\sharp\mathbf{v} = g_{ij}v^i\hat{\mathbf{e}}^j = v_j\hat{\mathbf{e}}^j = \hat{\mathbf{v}} \in V^*$*, with* $v_j = g_{ij}v^i$*;*
- $\flat\hat{\mathbf{v}} = g^{ij}v_i\mathbf{e}_j = v^j\mathbf{e}_j = \mathbf{v} \in V$*, with* $v^j = g^{ij}v_i$*.*

Proof of Result 2.2. This follows directly from the definition of $\sharp$ and $\flat$ using

- $\sharp\mathbf{e}_i = \mathbf{G}(\mathbf{e}_i) = g_{ij}\hat{\mathbf{e}}^j \in V^*$, respectively
- $\flat\hat{\mathbf{e}}^i = \mathbf{G}^{-1}(\hat{\mathbf{e}}^i) = g^{ij}\mathbf{e}_j \in V$.

Notation

Unless stated otherwise we will adhere to the following convention. If $v^i \in \mathbb{R}$, then $v_j \stackrel{\text{def}}{=} g_{ij}v^i$. reversely, if $v_i \in \mathbb{R}$, then $v^j \stackrel{\text{def}}{=} g^{ij}v_i$.

Caveat

In the foregoing $\mathbf{G} \in \mathscr{L}(V, V^*)$ and $\mathbf{G}^{-1} \in \mathscr{L}(V^*, V)$, with prototypes $\mathbf{G} : V \to V^*$, resp. $\mathbf{G}^{-1} : V^* \to V$. However, one often interprets the metric and its dual as bilinear mappings of type $\mathbf{G} : V \times V \to \mathbb{R}$, resp. $\mathbf{G}^{-1} : V^* \times V^* \to \mathbb{R}$. It should be clear from the context which prototype is implied. We return to this ambiguity in section 2.8.

Reminder

In an inner product space vectors can be seen as "covectors in disguise", vice versa.

2.5. Reciprocal Bases

Ansatz

An n-dimensional inner product space V over $\mathbb{R}$ furnished with a basis $\{\mathbf{e}_i\}$.

Definition 2.10 *The basis* $\{\mathbf{e}_i\}$ *of* V *induces a new basis* $\left\{\flat\hat{\mathbf{e}}^i\right\}$ *of* V *given by* $\flat\hat{\mathbf{e}}^i = g^{ij}\mathbf{e}_j$*, the so-called* reciprocal basis.

Definition 2.11 *The basis* $\left\{\hat{\mathbf{e}}^i\right\}$ *of* V^* *induces a new basis* $\{\sharp\mathbf{e}_i\}$ *of* V^* *given by* $\sharp\mathbf{e}_i = g_{ij}\hat{\mathbf{e}}^j$*, the so-called* reciprocal dual basis.

Caveat

- Instead of $\flat\hat{\mathbf{e}}^i$ one sometimes writes $\mathbf{e}^i$ for the reciprocal of $\mathbf{e}_i$.
- Likewise, instead of $\sharp\mathbf{e}_i$ one sometimes writes $\hat{\mathbf{e}}_i$ for the reciprocal of $\hat{\mathbf{e}}^i$.
- Since we conventionally label elements of V by *lower indices*, and elements of V^* by *upper indices*, we will adhere to the notation employing the $\flat$ and $\sharp$ operator.

Remark

- The reciprocal basis is uniquely determined by the inner product.
- The *Gram matrix of the reciprocal basis* is G^{-1}, with components $\left(\flat\hat{\mathbf{e}}^i|\flat\hat{\mathbf{e}}^j\right) = g^{ik}g^{j\ell}\left(\mathbf{e}_k|\mathbf{e}_\ell\right) = g^{ik}g^{j\ell}g_{k\ell} = g^{ij}$.
- We have $\left(\flat\hat{\mathbf{e}}^i|\mathbf{e}_j\right) = g^{ik}\left(\mathbf{e}_k|\mathbf{e}_j\right) = g^{ik}g_{kj} = \delta^i_j$.

Terminology

The vectors $\flat\hat{\mathbf{e}}^i$ and $\mathbf{e}_j$ are said to be *perpendicular* for $i \neq j$.

Remark

For each reciprocal basis $\left\{\flat\hat{\mathbf{e}}^i\right\}$ of V and each $\mathbf{x} \in V$ there exists a unique collection of real numbers $\{x_i\}$ such that $\mathbf{x} = x_i\flat\hat{\mathbf{e}}^i$.

Terminology

The real numbers x_i are referred to as the *covariant components* of the vector $\mathbf{x}$ relative to the basis $\{\mathbf{e}_i\}$.

Reminder

- Reciprocal basis vectors are labeled with upper indices.
- Covariant components are labeled with lower indices.
- No confusion is likely to arise if we adhere to the notation using the $\flat$ and $\sharp$ converters.

2.6. Bases, Dual Bases, Reciprocal Bases: Mutual Relations

Heuristics

Dual and reciprocal bases are close-knit. The exact relationship is clarified in this section.

Sofar, a vector space V and its associated dual space V^* have been introduced as a priori independent entities. V and V^* are mutually connected by virtue of the "interaction" between their respective elements, recall Result 2.1 on page 9. One nevertheless retains a clear distinction between a vector (element of V) and a dual vector (element of V^*) throughout. That is, there is no mechanism to convert a vector into a dual vector, vice versa.

An inner product provides us with an explicit mechanism to construct a one-to-one linear mapping (i.e. dual vector) associated with each vector by virtue of the metric $\mathbf{G}$, cf. Theorem 2.3 on page 13. An inner product also enables us to construct a reciprocal basis that allows us to mimic duality *within V proper*. "Interchangeability" of vectors and reciprocal vectors as well as that of vectors and dual vectors is governed by the Gram matrix G and its inverse.

Ansatz

An n-dimensional inner product space V over $\mathbb{R}$ furnished with a basis $\{\mathbf{e}_i\}$.

Observation

Result 2.3 on page 13 states that each vector $\mathbf{x} \in V$ induces a dual vector $\hat{\mathbf{x}} = \mathbf{G}(\mathbf{x}) \in V^*$. Let $\mathbf{x} = x^i\mathbf{e}_i$ relative to basis $\{\mathbf{e}_i\}$ of V. Then[2] $\hat{\mathbf{x}} = \hat{x}_i\hat{\mathbf{e}}^i$ relative to the corresponding dual basis $\{\hat{\mathbf{e}}^i\}$ of V^*. According to Definition 2.10 on page 15 we also have $\mathbf{x} = x_i\flat\hat{\mathbf{e}}^i$ relative to the reciprocal basis $\{\flat\hat{\mathbf{e}}^i\}$ of V. Given $x^i \in \mathbb{R}$ we have $x_i = \hat{x}_i = g_{ij}x^j$.

Caveat

Notice the distinct prototypes of metric $\mathbf{G} : V \to V^*$ and linear map $\mathscr{G} : V \to V$ represented by the Gram matrix G, i.e. $\mathbf{e}_i = \mathscr{G}(\flat\hat{\mathbf{e}}^i)$.

Notation

No confusion is likely to arise if we use identical symbols, $\hat{x}_i \equiv x_i$, to denote both dual as well as reciprocal vector components. The underlying basis ($\{\hat{\mathbf{e}}^i\}$, respectively $\{\flat\hat{\mathbf{e}}^i\}$), and thus the correct geometric interpretation, should be inferred from the context. By default we shall interpret x_i as dual vector components, henceforth omitting the $\hat{}$ symbol, i.e. $\hat{\mathbf{x}} = x_i\hat{\mathbf{e}}^i \in V^*$.

[2] For the sake of argument we temporarily write $\hat{x}_i$ for dual, and x_i for reciprocal vector components.

Reminder

- A dual basis does not require an inner product, a reciprocal basis does.
- A dual basis spans V^*, a reciprocal basis spans V.
- In an inner product space—with metric $\mathbf{G}$—the components of a vector $\mathbf{x} \in V$ relative to the reciprocal basis are identical to the components of the dual vector $\hat{\mathbf{x}} = \mathbf{G}(\mathbf{x})$ relative to the dual basis.
- Thus, in an inner product space with *fixed* metric $\mathbf{G}$, one may loosely identify a vector $\mathbf{x} \in V$ and its dual vector $\hat{\mathbf{x}} \in V^*$ to the extent that $\hat{\mathbf{x}} = \mathbf{G}(\mathbf{x})$ and $(\,\cdot\,|\mathbf{x}) = \langle \hat{\mathbf{x}}, \cdot \rangle$.
- The metric and its dual, $\mathbf{G}$, respectively $\mathbf{G}^{-1}$, provide the natural mechanism to "toggle" between elements of V and V^*. Abbreviations: $\mathbf{G}(\mathbf{x}) = \sharp\mathbf{x}$, $\mathbf{G}^{-1}(\hat{\mathbf{x}}) = \flat\hat{\mathbf{x}}$.
- The Gram matrix and its inverse, G, respectively G^{-1}, or rather the associated linear mappings $\mathscr{G} : V \to V$ and $\mathscr{G}^{-1} : V \to V$, provide natural mechanisms to "toggle" between elements of V and their reciprocals in V.

2.7. Examples of Vectors and Covectors

It is important to learn to recognize covectors and vectors in practice, and to maintain a clear distinction throughout. Some important examples from mechanics may illustrate this.

Example

- Particle velocity $\mathbf{v}$ provides an example of a vector.
- Particle momentum $\hat{\mathbf{p}}$ provides an example of a covector.
- The metric tensor allows us to relate these two: $\hat{\mathbf{p}} = \mathbf{G}(\mathbf{v}) = \sharp\mathbf{v}$, or equivalently, $\mathbf{v} = \mathbf{G}^{-1}(\hat{\mathbf{p}}) = \flat\hat{\mathbf{p}}$.

Example

- Displacement $\mathbf{x}$ relative to a fiducial point in a continuous medium (or in vacuo) provides an example of a vector.
- Spatial frequency $\hat{\mathbf{k}}$ provides an example of a covector.
- The Kronecker tensor allows us to construct a scalar from these two: $\langle \hat{\mathbf{k}}, \mathbf{x} \rangle \in \mathbb{R}$ ("phase").

Caveat

- In practice one often refrains from a notational distinction between covectors and vectors, writing e.g. $\mathbf{p}$ and $\mathbf{k}$ instead of $\hat{\mathbf{p}}$ and $\hat{\mathbf{k}}$, as we do here.
- By the same token it is a common malpractice to speak of "the momentum vector $\mathbf{p}$" and the "wave vector $\mathbf{k}$" even if covectors are intended. (Of course, technically speaking covectors *are* vectors.)

2.8. Tensors

2.8.1. Tensors in all Generality

Ansatz

- An n-dimensional vector space V over $\mathbb{R}$.
- A basis $\{\mathbf{e}_i\}$ of V.

Definition 2.12 *A* tensor *is a multilinear mapping of type* $\mathbf{T} : \underbrace{V^* \times \ldots \times V^*}_{p} \times \underbrace{V \times \ldots \times V}_{q} \to \mathbb{R}$, *with* $p, q \in \mathbb{N}_0$.

We collectively denote tensors of this type as follows: $\mathbf{T} \in \underbrace{V \otimes \ldots \otimes V}_{p} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{q} \stackrel{\text{def}}{=} \mathbf{T}^p_q(V)$.

Caveat

- The ordering of arguments matters! Our default ordering will be that of Definition 2.12.
- Thus the prototype may be specified with a different arrangement of linear spaces V and V^*.
- Without default ordering, the explicit $\otimes$-notation is the more informative one.

Observation

- $\mathbf{T}^p_q(V)$ constitutes a linear space subject to the usual rules for linear superposition of mappings, i.e. for all $\mathbf{S}, \mathbf{T} \in \mathbf{T}^p_q(V)$ and all $\lambda, \mu \in \mathbb{R}$ we declare

$$(\lambda\mathbf{S} + \mu\mathbf{T})\,((\text{co})\text{vectors}) = \lambda\mathbf{S}((\text{co})\text{vectors}) + \mu\mathbf{T}((\text{co})\text{vectors})\,,$$

  in which "(co)vectors" generically denotes an appropriate sequence of vector and/or covector arguments.
- Recall that the linear structure of $\mathbf{T}^p_q(V)$ has nothing to do with that of its domain of definition, but is induced by the linear structure on its codomain, in this case the trivial vector space $\mathbb{R}$ (cf. the r.h.s. above).

Terminology

- We refer to $\mathbf{T} \in \mathbf{T}^p_q(V)$ sometimes as a $\binom{p}{q}$-tensor.
- $\mathbf{T} \in \mathbf{T}^p_q(V)$ is referred to as a *mixed tensor* of *contravariant rank* p and *covariant rank* q.
- If $q \neq 0$, $\mathbf{T} \in \mathbf{T}^0_q(V)$ is called a *covariant tensor* of *rank* q. Covariant tensors take contravariant arguments, i.e. vectors.
- If $p \neq 0$, $\mathbf{T} \in \mathbf{T}^p_0(V)$ is called a *contravariant tensor* of *rank* p. Contravariant tensors take covariant arguments, i.e. covectors.
- $\mathbf{T} \in \mathbf{T}^0_0(V)$ is called a *scalar*. We identify $\mathbf{T}$ with its value in this case, i.e. $\mathbf{T} \in \mathbb{R}$.
- The words *cotensor*/*contratensor* are synonymous to covariant/contravariant tensor.

Example

Let $\mathbf{T}$ be a $\binom{1}{2}$-tensor, and $\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}, \mathbf{y}_1, \mathbf{y}_2 \in V$, $\hat{\mathbf{z}}, \hat{\mathbf{z}}_1, \hat{\mathbf{z}}_2 \in V^*$, and $\lambda_1, \lambda_2 \in \mathbb{R}$ arbitrary. Then we have

- $\mathbf{T}(\lambda_1\hat{\mathbf{z}}_1 + \lambda_2\hat{\mathbf{z}}_2, \mathbf{x}, \mathbf{y}) = \lambda_1\mathbf{T}(\hat{\mathbf{z}}_1, \mathbf{x}, \mathbf{y}) + \lambda_2\mathbf{T}(\hat{\mathbf{z}}_2, \mathbf{x}, \mathbf{y})$;
- $\mathbf{T}(\hat{\mathbf{z}}, \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2, \mathbf{y}) = \lambda_1\mathbf{T}(\hat{\mathbf{z}}, \mathbf{x}_1, \mathbf{y}) + \lambda_2\mathbf{T}(\hat{\mathbf{z}}, \mathbf{x}_2, \mathbf{y})$;
- $\mathbf{T}(\hat{\mathbf{z}}, \mathbf{x}, \lambda_1\mathbf{y}_1 + \lambda_2\mathbf{y}_2) = \lambda_1\mathbf{T}(\hat{\mathbf{z}}, \mathbf{x}, \mathbf{y}_1) + \lambda_2\mathbf{T}(\hat{\mathbf{z}}, \mathbf{x}, \mathbf{y}_2)$.

Result 2.3 *Special cases of low rank tensors:*

- *A* $\binom{0}{0}$*-tensor is a scalar, i.e.* $\mathbf{T}^0_0(V) = \mathbb{R}$.
- *A* $\binom{0}{1}$*-tensor is a covector, i.e.* $\mathbf{T}^0_1(V) = V^*$.
- *A* $\binom{1}{0}$*-tensor is a vector, i.e.* $\mathbf{T}^1_0(V) = V$.

Proof of Result 2.3. The first case is a convention previously explained. The second case is true by Definition 2.2, page 8, recall the terminology introduced on page 9. The third case is true by virtue of the isomorphism $V^{**} = V$, recall the remark on page 11.

Definition 2.13 *The outer product* $f \otimes g : X \times Y \to \mathbb{R}$ *of two* $\mathbb{R}$*-valued functions* $f : X \to \mathbb{R}$ *and* $g : Y \to \mathbb{R}$ *is defined as follows:*

$$(f \otimes g)(x, y) = f(x)g(y) \quad \textit{for all } x \in X,\ y \in Y.$$

Since vectors and covectors are $\mathbb{R}$-valued mappings, the following definition is a natural one.

Definition 2.14 *The mixed* $\binom{p}{q}$*-tensor* $\mathbf{e}_{i_1} \otimes \ldots \otimes \mathbf{e}_{i_p} \otimes \hat{\mathbf{e}}^{j_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{j_q} \in \mathbf{T}^p_q(V)$ *is defined, for each* $(p+q)$*-tuple of indices* $(i_1, \ldots, i_p, j_1, \ldots, j_q)$*, by*

$$\left(\mathbf{e}_{i_1} \otimes \ldots \otimes \mathbf{e}_{i_p} \otimes \hat{\mathbf{e}}^{j_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{j_q}\right)(\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_p, \mathbf{y}_1, \ldots, \mathbf{y}_q) = \langle \hat{\mathbf{x}}_1, \mathbf{e}_{i_1}\rangle \ldots \langle \hat{\mathbf{x}}_p, \mathbf{e}_{i_p}\rangle \langle \hat{\mathbf{e}}^{j_1}, \mathbf{y}_1\rangle \ldots \langle \hat{\mathbf{e}}^{j_q}, \mathbf{y}_q\rangle\,,$$

for all $\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_p \in V^*$ *and* $\mathbf{y}_1, \ldots, \mathbf{y}_q \in V$.

Caveat

- The symbol $\otimes$ used in Definitions 2.13 and 2.14 should not be confused with the one used in Definition 2.12.
- Mnemonic: $\underbrace{\mathbf{e}_{i_1} \otimes \ldots \otimes \mathbf{e}_{i_p}}_{p \text{ vectors}} \otimes \underbrace{\hat{\mathbf{e}}^{j_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{j_q}}_{q \text{ covectors}} \in \underbrace{V \otimes \ldots \otimes V}_{p \text{ copies of } V} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{q \text{ copies of } V^*}$.

Result 2.4 *If* $\mathbf{T} \in \mathbf{T}^p_q(V)$*, then there exist numbers* $t^{i_1 \ldots i_p}_{j_1 \ldots j_q} \in \mathbb{R}$ *such that* $\mathbf{T} = t^{i_1 \ldots i_p}_{j_1 \ldots j_q}\mathbf{e}_{i_1} \otimes \ldots \otimes \mathbf{e}_{i_p} \otimes \hat{\mathbf{e}}^{j_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{j_q}$*, viz.* $t^{i_1 \ldots i_p}_{j_1 \ldots j_q} = \mathbf{T}(\hat{\mathbf{e}}^{i_1}, \ldots, \hat{\mathbf{e}}^{i_p}, \mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_q})$.

Proof of Result 2.4. We have

$$
\begin{aligned}
&\mathbf{T}\left(\hat{\mathbf{e}}^{i_1},\ldots,\hat{\mathbf{e}}^{i_p},\mathbf{e}_{j_1},\ldots,\mathbf{e}_{j_q}\right)\\
&\quad= t^{k_1\ldots k_p}_{\ell_1\ldots\ell_q}\left(\mathbf{e}_{k_1}\otimes\ldots\otimes\mathbf{e}_{k_p}\otimes\hat{\mathbf{e}}^{\ell_1}\otimes\ldots\otimes\hat{\mathbf{e}}^{\ell_q}\right)\left(\hat{\mathbf{e}}^{i_1},\ldots,\hat{\mathbf{e}}^{i_p},\mathbf{e}_{j_1},\ldots,\mathbf{e}_{j_q}\right)\\
&\quad= t^{k_1\ldots k_p}_{\ell_1\ldots\ell_q}\langle\hat{\mathbf{e}}^{i_1},\mathbf{e}_{k_1}\rangle\ldots\langle\hat{\mathbf{e}}^{i_p},\mathbf{e}_{k_p}\rangle\langle\hat{\mathbf{e}}^{\ell_1},\mathbf{e}_{j_1}\rangle\ldots\langle\hat{\mathbf{e}}^{\ell_q},\mathbf{e}_{j_q}\rangle\\
&\quad= t^{k_1\ldots k_p}_{\ell_1\ldots\ell_q}\delta^{i_1}_{k_1}\ldots\delta^{i_p}_{k_p}\delta^{\ell_1}_{j_1}\ldots\delta^{\ell_q}_{j_q}=t^{i_1\ldots i_p}_{j_1\ldots j_q}\,.
\end{aligned}
$$

By multilinearity we have, for general covector and vector arguments $\hat{\boldsymbol{\omega}}_1,\ldots,\hat{\boldsymbol{\omega}}_p\in V^*$, $\mathbf{v}_1,\ldots,\mathbf{v}_q\in V$,

$$
\mathbf{T}\left(\hat{\boldsymbol{\omega}}_1,\ldots,\hat{\boldsymbol{\omega}}_p,\mathbf{v}_1,\ldots,\mathbf{v}_q\right)=t^{i_1\ldots i_p}_{j_1\ldots j_q}\omega_{1,i_1}\ldots\omega_{p,i_p}v^{j_1}_1\ldots v^{j_q}_q\,.
$$

Terminology

- If $\mathbf{T}\in\mathbf{T}^0_1(V)=V^*$ there exist numbers $t_i\in\mathbb{R}$ such that $\mathbf{T}=t_i\hat{\mathbf{e}}^i$, viz. $t_i=\mathbf{T}(\mathbf{e}_i)$. These are the *covariant components of* $\mathbf{T}$ *relative to basis* $\{\mathbf{e}_i\}$.
- If $\mathbf{T}\in\mathbf{T}^1_0(V)=V$ there exist numbers $t^i\in\mathbb{R}$ such that $\mathbf{T}=t^i\mathbf{e}_i$, viz. $t^i=\mathbf{T}(\hat{\mathbf{e}}^i)$. These are the *contravariant components of* $\mathbf{T}$ *relative to basis* $\{\mathbf{e}_i\}$.
- If $\mathbf{T}\in\mathbf{T}^p_q(V)$, the numbers $t^{i_1\ldots i_p}_{j_1\ldots j_q}\in\mathbb{R}$ are coined the *mixed components of* $\mathbf{T}$ *relative to basis* $\{\mathbf{e}_i\}$.

Caveat

Laziness has led to the habit of saying that $\mathbf{T}=t^{i_1\ldots i_p}_{j_1\ldots j_q}\mathbf{e}_{i_1}\otimes\ldots\otimes\mathbf{e}_{i_p}\otimes\hat{\mathbf{e}}^{j_1}\otimes\ldots\otimes\hat{\mathbf{e}}^{j_q}$ is expressed "relative to basis $\{\mathbf{e}_i\}$", where it would have been more accurate to refer to the actual basis $\{\mathbf{e}_{i_1}\otimes\ldots\otimes\mathbf{e}_{i_p}\otimes\hat{\mathbf{e}}^{j_1}\otimes\ldots\otimes\hat{\mathbf{e}}^{j_q}\}$.

Example

Recall Definition 2.4 on page 10. We have yet another notation for the Kronecker tensor $\langle\,\cdot\,,\,\cdot\,\rangle\in V\otimes V^*$, viz.

$$
\langle\,\cdot\,,\,\cdot\,\rangle=\mathbf{e}_i\otimes\hat{\mathbf{e}}^i=\delta^i_j\mathbf{e}_i\otimes\hat{\mathbf{e}}^j\,,
$$

in which we recognize the Kronecker symbol as the mixed components of the Kronecker tensor relative to basis $\{\mathbf{e}_i\}$. To see this, insert an arbitrary covector $\hat{\mathbf{y}}=y_i\hat{\mathbf{e}}^i\in V^*$ and vector $\mathbf{x}=x^i\mathbf{e}_i\in V$ on the r.h.s. The result is

$$
\left(\mathbf{e}_i\otimes\hat{\mathbf{e}}^i\right)(\hat{\mathbf{y}},\mathbf{x})=\langle\hat{\mathbf{y}},\mathbf{e}_i\rangle\langle\hat{\mathbf{e}}^i,\mathbf{x}\rangle=y_jx^k\langle\hat{\mathbf{e}}^j,\mathbf{e}_i\rangle\langle\hat{\mathbf{e}}^i,\mathbf{e}_k\rangle=y_jx^k\delta^j_i\delta^i_k=y_ix^j\delta^i_j=y_ix^j\langle\hat{\mathbf{e}}^i,\mathbf{e}_j\rangle=\langle\hat{\mathbf{y}},\mathbf{x}\rangle\,.
$$

Note that the Kronecker tensor is parameter free. For this reason it is sometimes called a *constant tensor*. Since its coefficients δ^i_j do not depend on the chosen basis (an untypical property for tensor coefficients, as we will see in Section 2.8.6), they are said to constitute an *invariant holor*.

Observation

- We have $\dim\mathbf{T}^p_q(V)=n^{p+q}$.
- Lexicographical ordering of the components $t^{i_1\ldots i_p}_{j_1\ldots j_q}$ of tensors $\mathbf{T}\in\mathbf{T}^p_q(V)$ establishes an isomorphism with the "column vector space" $\mathbb{R}^{n^{p+q}}$.

Terminology

- The indexed collection of tensor coefficients $t^{i_1\ldots i_p}_{j_1\ldots j_q}$ is also referred to as a *holor*.
- Physicists and engineers are wont to refer to holors as "tensors". In a mathematician's vocabulary the (basis dependent) holor is strictly to be distinguished from the associated (basis independent) tensor, $\mathbf{T} = t^{i_1\ldots i_p}_{j_1\ldots j_q}\mathbf{e}_{i_1}\otimes\ldots\otimes\mathbf{e}_{i_p}\otimes\hat{\mathbf{e}}^{j_1}\otimes\ldots\otimes\hat{\mathbf{e}}^{j_q}$.
- If $\{\mathbf{e}_i\}$ is a basis of V, a basis of $\mathbf{T}^p_q(V)$ is given by $\{\mathbf{e}_{i_1}\otimes\ldots\otimes\mathbf{e}_{i_p}\otimes\hat{\mathbf{e}}^{j_1}\otimes\ldots\otimes\hat{\mathbf{e}}^{j_q}\}$. But do recall the caveat on page 20 concerning terminology.

Caveat

Parlance and notation in tensor calculus tend to be somewhat sloppy. This is the price one must pay for simplicity. It is up to the reader to resolve ambiguities, if any, and provide correct mathematical interpretations.

Example

It is customary to think of the metric $\mathbf{G}\in\mathscr{L}(V,V^*)$ as a covariant tensor of rank 2. In order to understand this, let us clarify a previous remark on page 14. For given metric $\mathbf{G}$, define $\mathbf{g}\in V^*\otimes V^* = \mathbf{T}^0_2(V)$ as follows:

$$\mathbf{g}: V\times V\to\mathbb{R}:(\mathbf{v},\mathbf{w})\mapsto\mathbf{g}(\mathbf{v},\mathbf{w})\equiv(\mathbf{G}(\mathbf{v}))(\mathbf{w})\equiv(\mathbf{v}|\mathbf{w})\equiv\langle\sharp\mathbf{v},\mathbf{w}\rangle\,,$$

recall the sharp operator, introduced on page 14.

Notation

- Except when this might lead to confusion, no notational distinction will henceforth be made between such distinct prototypes as $\mathbf{g}$ and $\mathbf{G}$. We will invariably write $\mathbf{G}$, whether it is intended as a covariant 2-tensor ($\mathbf{g}$ in the example, implied when we write $\mathbf{G}(\mathbf{v},\mathbf{w})$), or as a vector-to-covector converter ($\sharp$-operator, implied when we write $\mathbf{G}(\mathbf{v})$).
- The $\sharp$- and $\flat$-operators provide convenient shorthands for the (dual) metric if one wishes to stress its action as a vector-to-covector (respectively covector-to-vector) converter.

Observation

- On an inner product space the $\sharp$- and $\flat$-operators may be invoked implicitly to convert arguments to the prototype that is compatible with each tensor slot. We return to this in Section 2.8.5.

Terminology

Irrespective of the intended interpretation, $\mathbf{G}$ (or $\mathbf{g}$) will be referred to as the *metric tensor*.

Theorem 2.4 *If* $\mathbf{S}\in\mathbf{T}^p_q(V)$, $\mathbf{T}\in\mathbf{T}^r_s(V)$, *with*

$$\mathbf{S} = s^{i_1\ldots i_p}_{j_1\ldots j_q}\mathbf{e}_{i_1}\otimes\ldots\otimes\mathbf{e}_{i_p}\otimes\hat{\mathbf{e}}^{j_1}\otimes\ldots\otimes\hat{\mathbf{e}}^{j_q}\quad\textit{and}\quad\mathbf{T} = t^{i_1\ldots i_r}_{j_1\ldots j_s}\mathbf{e}_{i_1}\otimes\ldots\otimes\mathbf{e}_{i_r}\otimes\hat{\mathbf{e}}^{j_1}\otimes\ldots\otimes\hat{\mathbf{e}}^{j_s}\,,$$

then $\mathbf{S}\otimes\mathbf{T}\in\mathbf{T}^{p+r}_{q+s}(V)$, *with* $\mathbf{S}\otimes\mathbf{T} = s^{i_1\ldots i_p}_{j_1\ldots j_q}t^{i_{p+1}\ldots i_{p+r}}_{j_{q+1}\ldots j_{q+s}}\mathbf{e}_{i_1}\otimes\ldots\otimes\mathbf{e}_{i_{p+r}}\otimes\hat{\mathbf{e}}^{j_1}\otimes\ldots\otimes\hat{\mathbf{e}}^{j_{q+s}}$.

Terminology

For $\mathbf{S}\in\mathbf{T}^p_q(V)$, $\mathbf{T}\in\mathbf{T}^r_s(V)$, the tensor $\mathbf{S}\otimes\mathbf{T}\in\mathbf{T}^{p+r}_{q+s}(V)$ is also referred to as the *outer product* or *tensor product* of $\mathbf{S}$ and $\mathbf{T}$.

Caveat

Notice that we have chosen to adhere to a default ordering of basis and dual basis vectors that span the tensor product space, recall Definition 2.12, page 18. This boils down to a rearrangement such that the first $p + r$ arguments of the product tensor $\mathbf{S} \otimes \mathbf{T}$ are covectors, and the remaining $q + s$ arguments are vectors.

Proof of Theorem 2.4. Apply Definition 2.13, page 19, accounting for the occurrence of multiplicative coefficients: $((\lambda f) \otimes (\mu g))(x, y) = ((\lambda f)(x))((\mu g)(y)) = \lambda\mu f(x)g(y)$. Further details are left to the reader (taking into account the previous caveat).

Caveat

The tensor product is associative, distributive with respect to addition and scalar multiplication, but not commutative. (Work out the exact meaning of these statements in formulae.)

Reminder

- A mixed tensor $\mathbf{T} \in \mathbf{T}^p_q(V)$ has contravariant rank p and covariant rank q.
- The holor $t^{i_1 \dots i_p}_{j_1 \dots j_q} \in \mathbb{R}$ of $\mathbf{T} \in \mathbf{T}^p_q(V)$ is labeled with p upper and q lower indices.
- Elements of the underlying basis $\{\mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_p} \otimes \hat{\mathbf{e}}^{j_1} \otimes \dots \otimes \hat{\mathbf{e}}^{j_q}\}$ of $\mathbf{T}^p_q(V)$, on the other hand, are labeled with p lower and q upper indices.
- A holor depends on the chosen basis, the tensor itself (by definition!) does not. (We return to this aspect in Section 2.8.6.)
- The outer product of two tensors yields another tensor. Ranks add up.

2.8.2. Tensors Subject to Symmetries

Ansatz

- An n-dimensional vector space V over $\mathbb{R}$.
- A basis $\{\mathbf{e}_i\}$ of V.
- Tensor spaces of type $\mathbf{T}^p_q(V)$.

Certain subspaces of $\mathbf{T}^p_q(V)$ are of special interest. For instance, it may be the case that the ordering of (co)vector arguments in a tensor evaluation does not affect the outcome.

Remark

Recall a previous remark on page 8 concerning the convention for index positioning. In particular, in Definitions 2.15 and 2.16 below, indices are used to label an a priori unconstrained number of arguments, whence the upper/lower index convention is not applicable.

Notation

- $\pi = \{\pi(\alpha_1), \dots, \pi(\alpha_p)\}$ denotes a *permutation* of the p ordered symbols $\{\alpha_1, \dots, \alpha_p\}$.
- $\pi(\alpha_i)$, $i = 1, \dots, p$, is shorthand notation for the i-th entry of π.

Definition 2.15 *Let $\pi = \{\pi(1), \ldots, \pi(p)\}$ be any permutation of labels $\{1, \ldots, p\}$, then $\mathbf{T} \in \mathbf{T}^0_p(V)$ is called a* symmetric covariant tensor *if for all $\mathbf{v}_1, \ldots, \mathbf{v}_p \in V$*

$$\mathbf{T}(\mathbf{v}_{\pi(1)}, \ldots, \mathbf{v}_{\pi(p)}) = \mathbf{T}(\mathbf{v}_1, \ldots, \mathbf{v}_p)\,.$$

Likewise, $\mathbf{T} \in \mathbf{T}^p_0(V)$ is called a symmetric contravariant tensor *if for all $\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_p \in V^*$*

$$\mathbf{T}(\hat{\mathbf{v}}_{\pi(1)}, \ldots, \hat{\mathbf{v}}_{\pi(p)}) = \mathbf{T}(\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_p)\,.$$

Notation

- The vector space of symmetric covariant p-tensors is denoted $\bigvee_p(V) \subset \mathbf{T}^0_p(V)$.
- The vector space of symmetric contravariant p-tensors is denoted $\bigvee^p(V) \subset \mathbf{T}^p_0(V)$.
- Alternatively one sometimes writes $\bigvee_p(V) = \underbrace{V^* \otimes_{\mathrm{S}} \ldots \otimes_{\mathrm{S}} V^*}_{p}$ and $\bigvee^p(V) = \underbrace{V \otimes_{\mathrm{S}} \ldots \otimes_{\mathrm{S}} V}_{p}$.

It may also happen that an interchange of two arguments causes the outcome to change sign.

Definition 2.16 *Let $\operatorname{sgn}(\pi)$ denote the sign of the permutation π, i.e. $\operatorname{sgn}(\pi) = 1$ if it is even, $\operatorname{sgn}(\pi) = -1$ if it is odd. $\mathbf{T} \in \mathbf{T}^0_p(V)$ is called an* antisymmetric covariant tensor *if for all $\mathbf{v}_1, \ldots, \mathbf{v}_p \in V$ and any permutation $\pi = \{\pi(1), \ldots, \pi(p)\}$ of labels $\{1, \ldots, p\}$,*

$$\mathbf{T}(\mathbf{v}_{\pi(1)}, \ldots, \mathbf{v}_{\pi(p)}) = \operatorname{sgn}(\pi)\mathbf{T}(\mathbf{v}_1, \ldots, \mathbf{v}_p)\,.$$

Likewise, $\mathbf{T} \in \mathbf{T}^p_0(V)$ is called an antisymmetric contravariant tensor *if for all $\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_p \in V^*$*

$$\mathbf{T}(\hat{\mathbf{v}}_{\pi(1)}, \ldots, \hat{\mathbf{v}}_{\pi(p)}) = \operatorname{sgn}(\pi)\mathbf{T}(\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_p)\,.$$

Notation

- The vector space of antisymmetric covariant p-tensors is denoted $\bigwedge_p(V) \subset \mathbf{T}^0_p(V)$.
- The vector space of antisymmetric contravariant p-tensors is denoted $\bigwedge^p(V) \subset \mathbf{T}^p_0(V)$.
- Alternatively one sometimes writes $\bigwedge_p(V) = \underbrace{V^* \otimes_{\mathrm{A}} \ldots \otimes_{\mathrm{A}} V^*}_{q}$ and $\bigwedge^p(V) = \underbrace{V \otimes_{\mathrm{A}} \ldots \otimes_{\mathrm{A}} V}_{p}$.

It is convenient to admit tensor ranks $p < 2$ in this notation by a natural extension.

Assumption

- $\bigvee_0(V) = \bigvee^0(V) = \bigwedge_0(V) = \bigwedge^0(V) = \mathbb{R}$. (This is consistent with Definitions 2.15 and 2.16 if we set $\pi = \mathrm{id}$, $\operatorname{sgn}\pi = 1$ on the empty index set $\emptyset$).
- $\bigvee_1(V) = \bigwedge_1(V) = V^*$. (This follows from Definitions 2.15 and 2.16, with $\pi = \mathrm{id}$, $\operatorname{sgn}\pi = 1$ on $\{1\}$).
- $\bigvee^1(V) = \bigwedge^1(V) = V$. (This follows from Definitions 2.15 and 2.16, with $\pi = \mathrm{id}$, $\operatorname{sgn}\pi = 1$ on $\{1\}$).

2.8.3. Symmetry and Antisymmetry Preserving Product Operators

Ansatz

- An n-dimensional vector space V over $\mathbb{R}$.
- A basis $\{\mathbf{e}_i\}$ of V.
- Tensor spaces of type $\bigvee_p(V)$, $\bigwedge_p(V)$, $\bigvee^p(V)$, $\bigwedge^p(V)$.

The outer product, recall Definition 2.13 on page 19 and in particular Theorem 2.4 on page 21, does not preserve (anti)symmetry. For this reason alternative product operators are introduced specifically designed for $\bigvee_p(V)$, $\bigwedge_p(V)$, and their contravariant counterparts.

Definition 2.17 *For every $\hat{\mathbf{v}}^1, \ldots, \hat{\mathbf{v}}^k \in V^*$ we define the antisymmetric covariant k-tensor $\hat{\mathbf{v}}^1 \wedge \ldots \wedge \hat{\mathbf{v}}^k \in \bigwedge_k(V)$ as follows:*

$$\left(\hat{\mathbf{v}}^1 \wedge \ldots \wedge \hat{\mathbf{v}}^k\right)(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \det\langle \hat{\mathbf{v}}^i, \mathbf{x}_j\rangle \quad \textit{for all } \mathbf{x}_j \in V.$$

Example

Let $\hat{\mathbf{v}}, \hat{\mathbf{w}} \in V^*$, $\mathbf{x}, \mathbf{y} \in V$, then

$$(\hat{\mathbf{v}} \wedge \hat{\mathbf{w}})(\mathbf{x}, \mathbf{y}) = \det\begin{pmatrix} \langle \hat{\mathbf{v}}, \mathbf{x}\rangle & \langle \hat{\mathbf{v}}, \mathbf{y}\rangle \\ \langle \hat{\mathbf{w}}, \mathbf{x}\rangle & \langle \hat{\mathbf{w}}, \mathbf{y}\rangle \end{pmatrix} = \langle \hat{\mathbf{v}}, \mathbf{x}\rangle\langle \hat{\mathbf{w}}, \mathbf{y}\rangle - \langle \hat{\mathbf{w}}, \mathbf{x}\rangle\langle \hat{\mathbf{v}}, \mathbf{y}\rangle\,.$$

Alternatively,

$$\hat{\mathbf{v}} \wedge \hat{\mathbf{w}} = \hat{\mathbf{v}} \otimes \hat{\mathbf{w}} - \hat{\mathbf{w}} \otimes \hat{\mathbf{v}}\,.$$

A more general result will be given on page 28.

Terminology

The infix operator $\wedge$ is referred to as the *wedge product* or the *antisymmetric product*.

Observation

- $\left(\hat{\mathbf{e}}^1 \wedge \ldots \wedge \hat{\mathbf{e}}^k\right)(\mathbf{e}_1, \ldots, \mathbf{e}_k) = \det\langle \hat{\mathbf{e}}^i, \mathbf{e}_j\rangle = 1$.
- $\left(\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_k}\right)(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_k}) = \delta^{i_1 \ldots i_k}_{j_1 \ldots j_k}$, recall the completely antisymmetric symbols on page 5, cf.
- $\left(\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_k}\right)(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_k}) = \delta^{i_1}_{j_1} \ldots \delta^{i_k}_{j_k}$.

Observation

Furthermore, the wedge product is associative, bilinear and antisymmetric (whence nilpotent, vice versa), i.e. if $\hat{\mathbf{u}}, \hat{\mathbf{v}}, \hat{\mathbf{w}} \in V^*$, $\lambda, \mu \in \mathbb{R}$, then

- $(\hat{\mathbf{u}} \wedge \hat{\mathbf{v}}) \wedge \hat{\mathbf{w}} = \hat{\mathbf{u}} \wedge (\hat{\mathbf{v}} \wedge \hat{\mathbf{w}})$.
- $(\lambda\hat{\mathbf{u}} + \mu\hat{\mathbf{v}}) \wedge \hat{\mathbf{w}} = \lambda\hat{\mathbf{u}} \wedge \hat{\mathbf{w}} + \mu\hat{\mathbf{v}} \wedge \hat{\mathbf{w}}$.
- $\hat{\mathbf{u}} \wedge (\lambda\hat{\mathbf{v}} + \mu\hat{\mathbf{w}}) = \lambda\hat{\mathbf{u}} \wedge \hat{\mathbf{v}} + \mu\hat{\mathbf{u}} \wedge \hat{\mathbf{w}}$.
- $\hat{\mathbf{u}} \wedge \hat{\mathbf{u}} = 0$, or, equivalently, $\hat{\mathbf{u}} \wedge \hat{\mathbf{v}} = -\hat{\mathbf{v}} \wedge \hat{\mathbf{u}}$.

A similar symmetry preserving product exists for symmetric tensors. To this end we need to introduce the symmetric counterpart of a determinant.

Definition 2.18 *The operator* perm *assigns a number to a square matrix in a way similar to* det*, but with a* $+$ *sign for each term, recall page 4:*

$$\operatorname{perm} A = \sum_{j_1,\ldots,j_n=1}^{n} |[j_1,\ldots,j_n]|\, A_{1j_1}\ldots A_{nj_n} = \frac{1}{n!}\sum_{\substack{i_1,\ldots,i_n\,=\,1\\ j_1,\ldots,j_n\,=\,1}}^{n} |[i_1,\ldots,i_n]\,[j_1,\ldots,j_n]|\, A_{i_1j_1}\ldots A_{i_nj_n}\,.$$

Remark

The formal role of the completely antisymmetric symbols in the above expressions is to restrict summation to relevant terms. The absolute value discards minus signs for odd parity terms, recall the last line in the observation made on page 4.

Terminology

The operator perm is called the *permanent*.

Example

Analogous to Table 1.1 on page 4 we have for a 3×3 matrix A,

$$\operatorname{perm} A = A_{11}A_{22}A_{33}+A_{12}A_{23}A_{31}+A_{13}A_{21}A_{32}+A_{13}A_{22}A_{31}+A_{11}A_{23}A_{32}+A_{12}A_{21}A_{33}\,.$$

Definition 2.19 *For every k-tuple $\hat{\mathbf{v}}^1,\ldots,\hat{\mathbf{v}}^k\in V^*$ we define the symmetric covariant k-tensor $\hat{\mathbf{v}}^1\vee\ldots\vee\hat{\mathbf{v}}^k\in\bigvee_k(V)$ as follows:*

$$\left(\hat{\mathbf{v}}^1\vee\ldots\vee\hat{\mathbf{v}}^k\right)(\mathbf{x}_1,\ldots,\mathbf{x}_k) = \operatorname{perm}\langle\hat{\mathbf{v}}^i,\mathbf{x}_j\rangle\,.$$

Example

Let $\hat{\mathbf{v}},\hat{\mathbf{w}}\in V^*$, $\mathbf{x},\mathbf{y}\in V$, then

$$(\hat{\mathbf{v}}\vee\hat{\mathbf{w}})(\mathbf{x},\mathbf{y}) = \operatorname{perm}\begin{pmatrix}\langle\hat{\mathbf{v}},\mathbf{x}\rangle & \langle\hat{\mathbf{v}},\mathbf{y}\rangle\\ \langle\hat{\mathbf{w}},\mathbf{x}\rangle & \langle\hat{\mathbf{w}},\mathbf{y}\rangle\end{pmatrix} = \langle\hat{\mathbf{v}},\mathbf{x}\rangle\langle\hat{\mathbf{w}},\mathbf{y}\rangle+\langle\hat{\mathbf{w}},\mathbf{x}\rangle\langle\hat{\mathbf{v}},\mathbf{y}\rangle\,.$$

Alternatively,

$$\hat{\mathbf{v}}\vee\hat{\mathbf{w}} = \hat{\mathbf{v}}\otimes\hat{\mathbf{w}}+\hat{\mathbf{w}}\otimes\hat{\mathbf{v}}\,.$$

Terminology

The infix operator $\vee$ is referred to as the *vee product* or the *symmetric product*.

Observation

The vee product is associative, bilinear and symmetric, i.e. if $\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{\mathbf{w}}\in V^*$, $\lambda,\mu\in\mathbb{R}$, then

- $(\hat{\mathbf{u}}\vee\hat{\mathbf{v}})\vee\hat{\mathbf{w}} = \hat{\mathbf{u}}\vee(\hat{\mathbf{v}}\vee\hat{\mathbf{w}})$.
- $(\lambda\hat{\mathbf{u}}+\mu\hat{\mathbf{v}})\vee\hat{\mathbf{w}} = \lambda\hat{\mathbf{u}}\vee\hat{\mathbf{w}}+\mu\hat{\mathbf{v}}\vee\hat{\mathbf{w}}$.
- $\hat{\mathbf{u}}\vee(\lambda\hat{\mathbf{v}}+\mu\hat{\mathbf{w}}) = \lambda\hat{\mathbf{u}}\vee\hat{\mathbf{v}}+\mu\hat{\mathbf{u}}\vee\hat{\mathbf{w}}$.
- $\hat{\mathbf{u}}\vee\hat{\mathbf{v}} = \hat{\mathbf{v}}\vee\hat{\mathbf{u}}$.

Observation

- $\dim \mathbf{T}^0_p(V) = n^p$.
- $\{\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_p} \,|\, 1 \leq i_1 < i_2 < \ldots < i_{p-1} < i_p \leq n\}$ is a basis of $\bigwedge_p(V)$.
- $\{\hat{\mathbf{e}}^{i_1} \vee \ldots \vee \hat{\mathbf{e}}^{i_p} \,|\, 1 \leq i_1 \leq i_2 \leq \ldots \leq i_{p-1} \leq i_p \leq n\}$ is a basis of $\bigvee_p(V)$.
- For $p \geq 2$, $\dim \bigwedge_p(V) + \dim \bigvee_p(V) \leq \dim \mathbf{T}^0_p(V)$, with strict equality iff $p = 2$.

Caveat

It is customary to suppress the infix operator $\vee$ in the notation.

Theorem 2.5 *Cf. the observations above. We have*

$$\dim \bigwedge_p(V) = \binom{n}{p} \quad \textit{and} \quad \dim \bigvee_p(V) = \binom{n+p-1}{p}.$$

Proof of Theorem 2.5.

- As for the first identity, a strictly increasing ordering of labels in a product of the form $\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_p}$ can be realized in exactly as many distinct ways as one can select p items out of n different objects.
- In particular, if $p > n$, then $\dim \bigwedge_p(V) = 0$, i.e. $\bigwedge_p(V) = \{0\}$.
- As for the second identity, the number of ways in which a non-decreasing ordering of labels in a product of the form $\hat{\mathbf{e}}^{i_1} \vee \ldots \vee \hat{\mathbf{e}}^{i_p}$ can be realized can be inferred from the following argument[3]. Any admissible ordering of labels can be represented symbolically by a sequence of p dots $\bullet$, split into subsequences of equal labels separated by $n-1$ crosses $\times$, each of which marks a unit label transition. Examples, taking $n = 5, p = 7$:

$$\begin{aligned} \bullet\bullet\times\bullet\times\times\bullet\bullet\bullet\times\bullet \quad &\equiv \quad \hat{\mathbf{e}}^1 \vee \hat{\mathbf{e}}^1 \vee \hat{\mathbf{e}}^2 \vee \hat{\mathbf{e}}^4 \vee \hat{\mathbf{e}}^4 \vee \hat{\mathbf{e}}^4 \vee \hat{\mathbf{e}}^5 \\ \bullet\bullet\times\bullet\bullet\bullet\bullet\times\bullet\times\times \quad &\equiv \quad \hat{\mathbf{e}}^1 \vee \hat{\mathbf{e}}^1 \vee \hat{\mathbf{e}}^2 \vee \hat{\mathbf{e}}^2 \vee \hat{\mathbf{e}}^2 \vee \hat{\mathbf{e}}^2 \vee \hat{\mathbf{e}}^3 \end{aligned}$$

  There are exactly $(n+p-1)!$ permutations of the $n+p-1$ symbols in such sequences, with exactly $p!$, respectively $(n-1)!$ indistinguishable permutations among the identical symbols $\bullet$, respectively $\times$, in any given realization.

Remark

Notice the qualitative differences between $\bigwedge_p(V)$ and $\bigvee_p(V)$ with respect to dimensionality scaling with p. In the former case we have $\dim \bigwedge_p(V) = 0$ as soon as $p > n = \dim V$, i.e. no nontrivial antisymmetric tensors with ranks p exceeding vector space dimension n exist. In contrast, nontrivial symmetric tensors do exist for any rank.

Example

Let $\mathscr{B}^p$ and $\mathscr{B}_p$ denote bases for $\bigwedge^p(V)$, respectively $\bigwedge_p(V)$, in Euclidean 3-space $V = \mathbb{R}^3$.

p	$\mathscr{B}^p$	dim
0	$\{1\}$	1
1	$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$	3
2	$\{\mathbf{e}_1 \wedge \mathbf{e}_2, \mathbf{e}_1 \wedge \mathbf{e}_3, \mathbf{e}_2 \wedge \mathbf{e}_3\}$	3
3	$\{\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3\}$	1

p	$\mathscr{B}_p$	dim
0	$\{1\}$	1
1	$\{\hat{\mathbf{e}}^1, \hat{\mathbf{e}}^2, \hat{\mathbf{e}}^3\}$	3
2	$\{\hat{\mathbf{e}}^1 \wedge \hat{\mathbf{e}}^2, \hat{\mathbf{e}}^1 \wedge \hat{\mathbf{e}}^3, \hat{\mathbf{e}}^2 \wedge \hat{\mathbf{e}}^3\}$	3
3	$\{\hat{\mathbf{e}}^1 \wedge \hat{\mathbf{e}}^2 \wedge \hat{\mathbf{e}}^3\}$	1

[3] Courtesy of Freek Paans.

Observation

Definitions 2.17, page 24, and 2.19, page 25, are straightforwardly complemented by similar ones applicable to (anti)symmetric contravariant tensors.

Definition 2.20 *The linear operator* $\mathscr{S} : \mathbf{T}^0_p(V) \to \bigvee_p(V)$*, given by*

$$(\mathscr{S}(\mathbf{T}))\,(\mathbf{v}_1, \ldots, \mathbf{v}_p) = \frac{1}{p!}\sum_\pi \mathbf{T}(\mathbf{v}_{\pi(1)}, \ldots, \mathbf{v}_{\pi(p)})\,,$$

in which summation runs over all permutations π *of the index set* $\{1, \ldots, p\}$*, is called the* symmetrisation map.

Observation

- If $\mathbf{T} = t_{i_1\ldots i_p}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_p} \in \mathbf{T}^0_p(V)$, then $\mathscr{S}(\mathbf{T}) = t_{(i_1\ldots i_p)}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_p} \in \bigvee_p(V)$, in which

$$t_{(i_1\ldots i_p)} \stackrel{\text{def}}{=} \frac{1}{p!}\sum_\pi t_{i_{\pi(1)}\ldots i_{\pi(p)}}\,.$$

- If $\mathbf{T} = \bigvee_p(V)$ is symmetric, then $\mathbf{T} = \mathscr{S}(\mathbf{T})$.
- The symmetrisation map is a projection, i.e. it is idempotent: $\mathscr{S} \circ \mathscr{S} = \mathscr{S}$.

These observations explain the normalization factor $\dfrac{1}{p!}$, since $\dfrac{1}{p!}\displaystyle\sum_\pi 1 = 1$.

Observation

- The $\vee$-product is particularly convenient in the context of the symmetrisation operator, for we have

$$\mathscr{S}(\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_p}) = \frac{1}{p!}\hat{\mathbf{e}}^{i_1} \vee \ldots \vee \hat{\mathbf{e}}^{i_p}\,.$$

- For $\mathbf{T} \in \bigvee_p(V)$ and $\mathbf{U} \in \bigvee_q(V)$, $\mathbf{T} \vee \mathbf{U} = \mathbf{U} \vee \mathbf{T}$.
- The case $p = 0$, $q > 0$ is covered by defining $\lambda \vee \mathbf{U} = \lambda\mathbf{U}$ for all $\lambda \in \mathbb{R}$, $\mathbf{U} \in \bigvee_q(V)$.
- Rule of thumb: The action of $p!\mathscr{S}$ on a $\otimes$-product amounts to a formal replacement $\otimes \to \otimes_{\mathrm{S}} \equiv \vee$.

The following result generalizes the example given on page 25:

Result 2.5 *Let* $\mathbf{T} \in \bigvee_p(V)$ *and* $\mathbf{U} \in \bigvee_q(V)$*, then* $\mathbf{T} \vee \mathbf{U} = \dfrac{(p+q)!}{p!q!}\,\mathscr{S}\,(\mathbf{T} \otimes \mathbf{U})$.

Proof of Result 2.5. Let $\mathbf{T} = t_{i_1\ldots i_p}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_p}$ and $\mathbf{U} = u_{i_{p+1}\ldots i_{p+q}}\hat{\mathbf{e}}^{i_{p+1}} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_{p+q}}$, with symmetric holors $t_{i_1\ldots i_p}$ and $u_{i_{p+1}\ldots i_{p+q}}$. Then

$$\begin{aligned}\mathbf{T} \vee \mathbf{U} &= t_{i_1\ldots i_p}u_{i_{p+1}\ldots i_{p+q}}(\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_p}) \vee (\hat{\mathbf{e}}^{i_{p+1}} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_{p+q}}) \stackrel{*}{=} \frac{1}{p!q!}t_{i_1\ldots i_p}u_{i_{p+1}\ldots i_{p+q}}\hat{\mathbf{e}}^{i_1} \vee \ldots \vee \hat{\mathbf{e}}^{i_{p+q}} \\ &\stackrel{*}{=} \frac{(p+q)!}{p!q!}t_{i_1\ldots i_p}u_{i_{p+1}\ldots i_{p+q}}\mathscr{S}\left(\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_{p+q}}\right) = \frac{(p+q)!}{p!q!}\mathscr{S}\,(\mathbf{T} \otimes \mathbf{U})\,.\end{aligned}$$

In step $*$ we have used one of the observations on page 27 concerning the relationship between vee product and symmetrised outer product. The last step exploits linearity of $\mathscr{S}$.

Definition 2.21 *The linear operator* $\mathscr{A} : \mathbf{T}^0_p(V) \to \bigwedge_p(V)$*, given by*

$$(\mathscr{A}(\mathbf{T}))\,(\mathbf{v}_1, \ldots, \mathbf{v}_p) = \frac{1}{p!}\sum_\pi \operatorname{sgn}(\pi)\mathbf{T}(\mathbf{v}_{\pi(1)}, \ldots, \mathbf{v}_{\pi(p)})\,,$$

is referred to as the antisymmetrisation map.

Observation

- If $\mathbf{T} = t_{i_1\ldots i_p}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_p} \in \mathbf{T}^0_p(V)$, then $\mathscr{A}(\mathbf{T}) = t_{[i_1\ldots i_p]}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_p} \in \bigwedge_p(V)$, in which

$$t_{[i_1\ldots i_p]} \stackrel{\text{def}}{=} \frac{1}{p!}\sum_\pi \operatorname{sgn}(\pi) t_{i_{\pi(1)}\ldots i_{\pi(p)}}\,.$$

- If $\mathbf{T} = \bigwedge_p(V)$ is antisymmetric, then $\mathbf{T} = \mathscr{A}(\mathbf{T})$.
- The antisymmetrisation map is a projection, i.e. it is idempotent: $\mathscr{A} \circ \mathscr{A} = \mathscr{A}$.

Recall a previous remark on the normalization factor (page 27).

Observation

- The $\wedge$-product is particularly convenient in the context of antisymmetrisation, for we have

$$\mathscr{A}(\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_p}) = \frac{1}{p!}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_p}\,.$$

- The case $p = 0$, $q > 0$ is covered by defining $\lambda \wedge \mathbf{U} = \lambda\mathbf{U}$ for all $\lambda \in \mathbb{R}$, $\mathbf{U} \in \bigwedge_q(V)$.
- Rule of thumb: The action of $p!\mathscr{A}$ on a $\otimes$-product amounts to a formal replacement $\otimes \to \otimes_{\mathrm{A}} \equiv \wedge$.

The following result generalizes the example given on page 24.

Result 2.6 *Let* $\mathbf{T} \in \bigwedge_p(V)$ *and* $\mathbf{U} \in \bigwedge_q(V)$*, then* $\mathbf{T} \wedge \mathbf{U} = \dfrac{(p+q)!}{p!q!}\,\mathscr{A}\,(\mathbf{T} \otimes \mathbf{U})$.

Notation

- The Einstein summation convention, recall page 7, is usually adapted in the context of antisymmetric tensors. In general, if $\mathbf{T} \in \bigwedge_p(V)$, then we may use either of the following equivalent decompositions:
  - $\mathbf{T} = t_{i_1\ldots i_p}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_p} = \displaystyle\sum_{i_1,\ldots,i_p} t_{i_1\ldots i_p}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_p}$,
  - $\mathbf{T} = t_{|i_1\ldots i_p|}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_p} = \displaystyle\sum_{i_1<\ldots<i_p} t_{i_1\ldots i_p}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_p}$.

Observation

Note that $t_{|i_1\ldots i_p|}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_p} = \frac{1}{p!}t_{i_1\ldots i_p}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_p}$, recall the previous observation.

Proof of Result 2.6. Let $\mathbf{T} = t_{|i_1\ldots i_p|}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_p}$ and $\mathbf{U} = u_{|i_{p+1}\ldots i_{p+q}|}\hat{\mathbf{e}}^{i_{p+1}} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_{p+q}}$, with antisymmetric

holors $t_{i_1\ldots i_p}$ and $u_{i_{p+1}\ldots i_{p+q}}$. Then

$$\begin{aligned}\mathbf{T}\wedge\mathbf{U} &= t_{|i_1\ldots i_p|}u_{|i_{p+1}\ldots i_{p+q}|}\hat{\mathbf{e}}^{i_1}\wedge\ldots\wedge\hat{\mathbf{e}}^{i_{p+q}} \stackrel{*}{=} \frac{1}{p!q!}t_{i_1\ldots i_p}u_{i_{p+1}\ldots i_{p+q}}\hat{\mathbf{e}}^{i_1}\wedge\ldots\wedge\hat{\mathbf{e}}^{i_{p+q}}\\ &\stackrel{\star}{=} \frac{(p+q)!}{p!q!}t_{i_1\ldots i_p}u_{i_{p+1}\ldots i_{p+q}}\mathscr{A}\left(\hat{\mathbf{e}}^{i_1}\otimes\ldots\otimes\hat{\mathbf{e}}^{i_{p+q}}\right) = \frac{(p+q)!}{p!q!}\mathscr{A}\left(\mathbf{T}\otimes\mathbf{U}\right)\,.\end{aligned}$$

In step $*$ we have used the fact that only distinct index values contribute nontrivially, together with the fact that there are precisely $p!$ ($q!$) orderings of p (q) distinct indices. In step $\star$ we have used the previous observation on the relationship between wedge product and antisymmetrised outer product. The last step exploits linearity of $\mathscr{A}$.

Result 2.7 *Let* $\mathbf{T}\in\bigwedge_p(V)$ *and* $\mathbf{U}\in\bigwedge_q(V)$*, then* $\mathbf{T}\wedge\mathbf{U}=(-1)^{pq}\mathbf{U}\wedge\mathbf{T}$.

Proof of Result 2.7. Cf. the proof of Result 2.6. To change the positions of the p basis covectors $\hat{\mathbf{e}}^{i_1},\ldots,\hat{\mathbf{e}}^{i_p}$ in the product $\hat{\mathbf{e}}^{i_1}\wedge\ldots\wedge\hat{\mathbf{e}}^{i_{p+q}}$ so as to obtain the reordered product $\hat{\mathbf{e}}^{i_{p+1}}\wedge\ldots\wedge\hat{\mathbf{e}}^{i_{p+q}}\wedge\hat{\mathbf{e}}^{i_1}\wedge\ldots\wedge\hat{\mathbf{e}}^{i_p}$ requires p (one for each $\hat{\mathbf{e}}^{i_k}$, $k=1,\ldots,p$) times q (one for each $\hat{\mathbf{e}}^{i_\ell}$, $\ell=p+1,\ldots,p+q$) commutations, each of which brings in a minus sign.

Remark

If $p+q>n=\dim V$, then $\mathbf{T}\wedge\mathbf{U}=0\in\bigwedge_{p+q}(V)$. Indeed, $\bigwedge_k(V)=\{0\}$ for all $k>n$.

Remark

- Both $\mathscr{S}$ as well as $\mathscr{A}$ are idempotent: $\mathscr{S}\circ\mathscr{S}=\mathscr{S}$, respectively $\mathscr{A}\circ\mathscr{A}=\mathscr{A}$. Thus they are, by definition, projections from $\mathbf{T}^0_p(V)$ onto $\bigvee_p(V)$, respectively onto $\bigwedge_p(V)$.
- Note that we have not defined the property of (anti)symmetry for a genuine mixed tensor.
- Most definitions also apply to the covariant and contravariant parts of a mixed tensor separately, i.e. treating the mixed tensor as a covariant tensor by freezing its covector arguments, respectively as a contravariant tensor by freezing its vector arguments.
- A (say covariant) 2-tensor can always be decomposed into a sum of a symmetric and antisymmetric 2-tensor, viz. $\mathbf{T}=\mathbf{S}+\mathbf{A}$, with $\mathbf{S}=\mathscr{S}(\mathbf{T})$ and $\mathbf{A}=\mathscr{A}(\mathbf{T})$, i.e. $\mathbf{S}(\mathbf{v},\mathbf{w})=\frac{1}{2}(\mathbf{T}(\mathbf{v},\mathbf{w})+\mathbf{T}(\mathbf{w},\mathbf{v}))$ and $\mathbf{A}(\mathbf{v},\mathbf{w})=\frac{1}{2}(\mathbf{T}(\mathbf{v},\mathbf{w})-\mathbf{T}(\mathbf{w},\mathbf{v}))$. This is consistent with the dimensionality equality for $p=2$ in the observation on page 26.
- A general tensor is typically neither symmetric nor antisymmetric, nor can it be split into a sum of a symmetric and antisymmetric tensor, recall the dimensionality inequality in the observation on page 26.
- The metric tensor $\mathbf{G}$ is symmetric, and so is its inverse, the dual metric tensor $\mathbf{G}^{-1}$.
- Both the $\wedge$-product as well as the $\vee$-product obey the "usual" rules for associativity, and distributivity with respect to tensor addition and scalar multiplication.

Example

Let $V = \mathbb{R}^3$, furnished with standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. One conventionally denotes

$$\mathbf{e}_1 = \frac{\partial}{\partial x} \qquad \mathbf{e}_2 = \frac{\partial}{\partial y} \qquad \mathbf{e}_3 = \frac{\partial}{\partial z}\,.$$

It is then customary to denote the corresponding dual basis $\{\hat{\mathbf{e}}^1, \hat{\mathbf{e}}^2, \hat{\mathbf{e}}^3\}$ of V^* by

$$\hat{\mathbf{e}}^1 = dx \qquad \hat{\mathbf{e}}^2 = dy \qquad \hat{\mathbf{e}}^3 = dz\,.$$

Let $\mathbf{u} = u^i\mathbf{e}_i, \mathbf{v} = v^i\mathbf{e}_i, \mathbf{w} = w^i\mathbf{e}_i \in V$ be arbitrary vectors. Then we have

$$\begin{aligned} dx(\mathbf{v}) = \langle dx, \mathbf{v}\rangle &= v^1\,, \\ dy(\mathbf{v}) = \langle dy, \mathbf{v}\rangle &= v^2\,, \\ dz(\mathbf{v}) = \langle dz, \mathbf{v}\rangle &= v^3\,. \end{aligned}$$

Example

Notation as in the previous example.

$$\begin{aligned} (dx \wedge dy)(\mathbf{v}, \mathbf{w}) = \det\begin{pmatrix} \langle dx, \mathbf{v}\rangle & \langle dx, \mathbf{w}\rangle \\ \langle dy, \mathbf{v}\rangle & \langle dy, \mathbf{w}\rangle \end{pmatrix} &= \det\begin{pmatrix} v^1 & w^1 \\ v^2 & w^2 \end{pmatrix} = v^1w^2 - v^2w^1\,, \\ (dx \wedge dz)(\mathbf{v}, \mathbf{w}) = \det\begin{pmatrix} \langle dx, \mathbf{v}\rangle & \langle dx, \mathbf{w}\rangle \\ \langle dz, \mathbf{v}\rangle & \langle dz, \mathbf{w}\rangle \end{pmatrix} &= \det\begin{pmatrix} v^1 & w^1 \\ v^3 & w^3 \end{pmatrix} = v^1w^3 - v^3w^1\,, \\ (dy \wedge dz)(\mathbf{v}, \mathbf{w}) = \det\begin{pmatrix} \langle dy, \mathbf{v}\rangle & \langle dy, \mathbf{w}\rangle \\ \langle dz, \mathbf{v}\rangle & \langle dz, \mathbf{w}\rangle \end{pmatrix} &= \det\begin{pmatrix} v^2 & w^2 \\ v^3 & w^3 \end{pmatrix} = v^2w^3 - v^3w^2\,. \end{aligned}$$

Example

Notation as in the previous example.

$$\begin{aligned} (dx \wedge dy \wedge dz)(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \det\begin{pmatrix} \langle dx, \mathbf{u}\rangle & \langle dx, \mathbf{v}\rangle & \langle dx, \mathbf{w}\rangle \\ \langle dy, \mathbf{u}\rangle & \langle dy, \mathbf{v}\rangle & \langle dy, \mathbf{w}\rangle \\ \langle dz, \mathbf{u}\rangle & \langle dz, \mathbf{v}\rangle & \langle dz, \mathbf{w}\rangle \end{pmatrix} = \det\begin{pmatrix} u^1 & v^1 & w^1 \\ u^2 & v^2 & w^2 \\ u^3 & v^3 & w^3 \end{pmatrix} \\ &= u^1v^2w^3 + u^3v^1w^2 + u^2v^3w^1 - u^3v^2w^1 - u^1v^3w^2 - u^2v^1w^3\,. \end{aligned}$$

Notice that sofar no inner product has been invoked in any of the definitions and results of this section. With the help of an inner product certain "straightforward" modifications can be made that generalize foregoing definitions and results. The underlying mechanism is that of the conversion tools provided by Definition 2.9 on page 14.

Ansatz

- An n-dimensional inner product space V over $\mathbb{R}$.
- Tensor spaces of type $\bigvee_p(V)$, $\bigwedge_p(V)$, $\bigvee^p(V)$, $\bigwedge^p(V)$.

Observation

Recall Theorem 2.3, page 13, Definition 2.9 page 14, and Result 2.2, page 14. An inner product on V induces an inner product on the spaces $\bigwedge^p(V)$ and $\bigvee^p(V)$, as follows. For any $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{x}_1, \ldots, \mathbf{x}_k \in V$, set

- $(\mathbf{v}_1 \wedge \ldots \wedge \mathbf{v}_k | \mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_k) = \det\langle \sharp\mathbf{v}_i, \mathbf{x}_j \rangle$, respectively
- $(\mathbf{v}_1 \vee \ldots \vee \mathbf{v}_k | \mathbf{x}_1 \vee \ldots \vee \mathbf{x}_k) = \text{perm}\,\langle \sharp\mathbf{v}_i, \mathbf{x}_j \rangle$.

Likewise for the spaces $\bigwedge_p(V)$ and $\bigvee_p(V)$:

- $\left(\hat{\mathbf{v}}^1 \wedge \ldots \wedge \hat{\mathbf{v}}^k | \hat{\mathbf{x}}^1 \wedge \ldots \wedge \hat{\mathbf{x}}^k\right) = \det\langle \hat{\mathbf{v}}^i, \flat\hat{\mathbf{x}}^j \rangle$, respectively
- $\left(\hat{\mathbf{v}}^1 \vee \ldots \vee \hat{\mathbf{v}}^k | \hat{\mathbf{x}}^1 \vee \ldots \vee \hat{\mathbf{x}}^k\right) = \text{perm}\,\langle \hat{\mathbf{v}}^i, \flat\hat{\mathbf{x}}^j \rangle$,

for any $\hat{\mathbf{v}}^1, \ldots, \hat{\mathbf{v}}^k, \hat{\mathbf{x}}^1, \ldots, \hat{\mathbf{x}}^k \in V^*$.

Section 2.8.5 covers tensors on an inner product space in more generality.

2.8.4. Vector Spaces with an Oriented Volume

Ansatz

- An n-dimensional vector space V over $\mathbb{R}$.
- A *preferred* basis $\{\mathbf{e}_i\}$ of V.
- A nonzero n-form $\boldsymbol{\mu} \in \bigwedge_n(V)$.

Terminology

If $\boldsymbol{\mu}(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$ then $\boldsymbol{\mu}$ is called the *unit volume form*.

Observation

- We have $\dim \bigwedge_n(V) = 1$.
- Consequently, if $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \bigwedge_n(V)$, then $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ differ by a constant factor $\lambda \in \mathbb{R}$.
- For the unit volume form we have, according to the observation on page 28 (no $\sum$-convention after $*$),

$$\boldsymbol{\mu} = \hat{\mathbf{e}}^1 \wedge \ldots \wedge \hat{\mathbf{e}}^n \overset{*}{=} [i_1, \ldots, i_n]\, \hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_n} = \sum_{i_1,\ldots,i_n}^{n} \frac{1}{n!}\, [i_1, \ldots, i_n]\, \hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_n}\,,$$

which, by definition, can be expressed in terms of the general tensor basis, viz.

$$\boldsymbol{\mu} = \sum_{i_1,\ldots,i_n}^{n} [i_1, \ldots, i_n]\, \mathscr{A}\left(\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n}\right) \overset{\star}{=} \sum_{i_1,\ldots,i_n}^{n} [i_1, \ldots, i_n]\, \hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n} \doteq \mu_{i_1 \ldots i_n}\, \hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n}\,.$$

Step $\star$ exploits automatic antisymmetrization: The $\mathscr{A}$-operator may be dropped, since its effect is void when combined with antisymmetric coefficients. Thus the covariant components of $\boldsymbol{\mu} \in \bigwedge_n(V)$ are given by $\mu_{i_1 \ldots i_n} = [i_1, \ldots, i_n]$.

Caveat

The normalisation of the unit volume form $\boldsymbol{\mu}$ requires a fiducial basis, viz. the preferred basis $\{\mathbf{e}_i\}$ chosen above. Consequently the identification $\mu_{i_1\ldots i_n} = [i_1, \ldots, i_n]$ holds only relative to that basis! We refer to section 2.8.6 for a basis independent evaluation.

Observation

Thus $\boldsymbol{\mu} = \mu_{|i_1\ldots i_n|}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_n} = \mu_{i_1\ldots i_n}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n}$. This shows, once again, the convenience of the two variants of the Einstein summation convention.

Definition 2.22 *Let $(V, \boldsymbol{\mu})$ denote the vector space V endowed with the n-form $\boldsymbol{\mu} \in \bigwedge_n(V)$, and let $\{\mathbf{e}_i\}$ be a given basis of V.*

- $(V, \boldsymbol{\mu})$ *is called a* vector space with an oriented volume.
- $(V, \boldsymbol{\mu})$ *is said to have a* positive orientation *if* $\boldsymbol{\mu}(\mathbf{e}_1, \ldots, \mathbf{e}_n) > 0$.
- $(V, \boldsymbol{\mu})$ *is said to have a* negative orientation *if* $\boldsymbol{\mu}(\mathbf{e}_1, \ldots, \mathbf{e}_n) < 0$.

Terminology

The use of attributes "positive" and "negative" in this context is usually self-explanatory, e.g.

- given $\{\mathbf{e}_i\}$ of V one refers to $\boldsymbol{\mu} \in \bigwedge_n(V)$ as a *positive volume form* if $\boldsymbol{\mu}(\mathbf{e}_1, \ldots, \mathbf{e}_n) > 0$, and
- given $\boldsymbol{\mu} \in \bigwedge_n(V)$ one refers to $\{\mathbf{e}_i\}$ as a *positively oriented basis* if $\boldsymbol{\mu}(\mathbf{e}_1, \ldots, \mathbf{e}_n) > 0$, etc.

Notation

Instead of $(V, \boldsymbol{\mu})$ we shall simply write V. The basis $\{\mathbf{e}_i\}$ relative to which $\boldsymbol{\mu}(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$ holds should be clear from the context.

Observation

Using Definition 2.17 on page 24 we may write, for $\mathbf{x}_1, \ldots, \mathbf{x}_n \in V$,

$$\boldsymbol{\mu}\left(\mathbf{x}_1, \ldots, \mathbf{x}_n\right) = \det\langle\hat{\mathbf{e}}^j, \mathbf{x}_i\rangle = \det\begin{pmatrix} \langle\hat{\mathbf{e}}^1, \mathbf{x}_1\rangle & \ldots & \langle\hat{\mathbf{e}}^1, \mathbf{x}_n\rangle \\ \vdots & & \vdots \\ \langle\hat{\mathbf{e}}^n, \mathbf{x}_1\rangle & \ldots & \langle\hat{\mathbf{e}}^n, \mathbf{x}_n\rangle \end{pmatrix},$$

which reveals the role of the Kronecker tensor (and thus of the dual space V^*) in the definition of $\boldsymbol{\mu}$. Consistent with a foregoing observation it follows that the covariant components of $\boldsymbol{\mu} = \mu_{i_1\ldots i_n}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n}$ are given by $\mu_{i_1\ldots i_n} = \boldsymbol{\mu}(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}) = [i_1, \ldots, i_n]\,\boldsymbol{\mu}(\mathbf{e}_1, \ldots, \mathbf{e}_n) = [i_1, \ldots, i_n]\det\langle\hat{\mathbf{e}}^j, \mathbf{e}_i\rangle = [i_1, \ldots, i_n]\det\delta^j_i = [i_1, \ldots, i_n]$ relative to (!) $\{\mathbf{e}_i\}$.

Definition 2.23 *Let $\mathbf{a}_1, \ldots, \mathbf{a}_k \in V$, $k \in \{0, \ldots, n\}$, then we define the $(n-k)$-form $\boldsymbol{\mu}\lrcorner\mathbf{a}_1\lrcorner\ldots\lrcorner\mathbf{a}_k \in \bigwedge_{n-k}(V)$ as follows:*

$$\left(\boldsymbol{\mu}\lrcorner\mathbf{a}_1\lrcorner\ldots\lrcorner\mathbf{a}_k\right)\left(\mathbf{x}_{k+1}, \ldots, \mathbf{x}_n\right) = \boldsymbol{\mu}\left(\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{x}_{k+1}, \ldots, \mathbf{x}_n\right).$$

Observation

- A trivial rewriting of the previous observation allows us to rewrite Definition 2.23 on page 32 as follows:

$$(\boldsymbol{\mu}\lrcorner\mathbf{a}_1\lrcorner\ldots\lrcorner\mathbf{a}_k)\,(\mathbf{x}_{k+1},\ldots,\mathbf{x}_n) = \det\begin{pmatrix} \langle\hat{\mathbf{e}}^1,\mathbf{a}_1\rangle & \ldots & \langle\hat{\mathbf{e}}^1,\mathbf{a}_k\rangle & \langle\hat{\mathbf{e}}^1,\mathbf{x}_{k+1}\rangle & \ldots & \langle\hat{\mathbf{e}}^1,\mathbf{x}_n\rangle \\ \vdots & & \vdots & \vdots & & \vdots \\ \langle\hat{\mathbf{e}}^n,\mathbf{a}_1\rangle & \ldots & \langle\hat{\mathbf{e}}^n,\mathbf{a}_k\rangle & \langle\hat{\mathbf{e}}^n,\mathbf{x}_{k+1}\rangle & \ldots & \langle\hat{\mathbf{e}}^n,\mathbf{x}_n\rangle \end{pmatrix}.$$

- The holor of $\boldsymbol{\mu}\lrcorner\mathbf{a}_1\lrcorner\ldots\lrcorner\mathbf{a}_k \in \bigwedge_{n-k}(V)$ for given $\mathbf{a}_1,\ldots,\mathbf{a}_k \in V$ is established as follows:

$$\begin{aligned}(\boldsymbol{\mu}\lrcorner\mathbf{a}_1\lrcorner\ldots\lrcorner\mathbf{a}_k)\left(\mathbf{e}_{i_{k+1}},\ldots,\mathbf{e}_{i_n}\right) &= a_1^{i_1}\ldots a_k^{i_k}\,(\boldsymbol{\mu}\lrcorner\mathbf{e}_{i_1}\lrcorner\ldots\lrcorner\mathbf{e}_{i_k})\left(\mathbf{e}_{i_{k+1}},\ldots,\mathbf{e}_{i_n}\right) \\ &= a_1^{i_1}\ldots a_k^{i_k}\,\boldsymbol{\mu}\left(\mathbf{e}_{i_1},\ldots,\mathbf{e}_{i_n}\right) = \mu_{i_1\ldots i_n}a_1^{i_1}\ldots a_k^{i_k}\,,\end{aligned}$$

  in which we have used the decompositions $\mathbf{a}_p = a_p^{i_p}\mathbf{e}_{i_p}$ for $p = 1,\ldots,k$ together with multilinearity.

- By the same token, expanding also $\mathbf{x}_q = x_q^{i_q}\mathbf{e}_{i_q}$ for $q = k+1,\ldots,n$, we obtain

$$(\boldsymbol{\mu}\lrcorner\mathbf{a}_1\lrcorner\ldots\lrcorner\mathbf{a}_k)\,(\mathbf{x}_{k+1},\ldots,\mathbf{x}_n) = \mu_{i_1\ldots i_n}a_1^{i_1}\ldots a_k^{i_k}x_{k+1}^{i_{k+1}}\ldots x_n^{i_n}\,.$$

Remark

- An oriented volume $\boldsymbol{\mu} \in \bigwedge_n(V)$ on a vector space V does not require an inner product.
- An oriented volume $\boldsymbol{\mu} \in \bigwedge_n(V)$ exists by virtue of the existence of the dual space V^*.
- If $\boldsymbol{\mu}(\mathbf{e}_1,\ldots,\mathbf{e}_n) = 1$, then $\boldsymbol{\mu} = \hat{\mathbf{e}}^1 \wedge \ldots \wedge \hat{\mathbf{e}}^n$ (unlike $\boldsymbol{\mu}$ itself, the factors are not unique).
- More generally, if $\boldsymbol{\mu}(\mathbf{e}_1,\ldots,\mathbf{e}_n) = 1$, then $\boldsymbol{\mu}\lrcorner\mathbf{e}_1\lrcorner\ldots\lrcorner\mathbf{e}_k = \hat{\mathbf{e}}^{k+1} \wedge \ldots \wedge \hat{\mathbf{e}}^n$.

The following example explains the terminology of a "volume form".

Example

Let the notation be as in the examples on page 30 and following. The space $\bigwedge_3(V)$ is spanned by $\{\boldsymbol{\mu} = dx \wedge dy \wedge dz\}$, which, according to the previous example, is given by

$$\boldsymbol{\mu}(\mathbf{u},\mathbf{v},\mathbf{w}) = u^1v^2w^3 + u^3v^1w^2 + u^2v^3w^1 - u^3v^2w^1 - u^1v^3w^2 - u^2v^1w^3\,.$$

This is the (signed) volume of a parallelepiped spanned by the (column vectors corresponding to the) vectors $\mathbf{u},\mathbf{v},\mathbf{w}$. The result is positive (negative) if the ordered triple $\mathbf{u},\mathbf{v},\mathbf{w}$ has the same (respectively opposite) orientation as the basis $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$.

The example suggests that similar interpretations hold for forms of lower degrees. This is indeed the case, as the next example shows.

Example

The spaces $\bigwedge_1(\mathbb{R}^3)$ and $\bigwedge_2(\mathbb{R}^3)$ are spanned by $\{dx, dy, dz\}$ and $\{dx \wedge dy, dx \wedge dz, dy \wedge dz\}$, respectively.

- According to the example on page 30 we have

$$\langle dx, \mathbf{v}\rangle = v^1 \qquad \langle dy, \mathbf{v}\rangle = v^2 \qquad \langle dz, \mathbf{v}\rangle = v^3 \,.$$

  One recognizes the (signed) 1-dimensional volumes (lengths) of the parallel projections of the vector $\mathbf{v}$ onto the x, y, respectively z axis. Recall the examples on page 17 for a physical interpretation in terms of particle kinematics or wave phenomena.

- According to the example on page 30 we have

$$(dx \wedge dy)\,(\mathbf{v}, \mathbf{w}) = \det \begin{pmatrix} \langle dx, \mathbf{v}\rangle & \langle dx, \mathbf{w}\rangle \\ \langle dy, \mathbf{v}\rangle & \langle dy, \mathbf{w}\rangle \end{pmatrix} = \det \begin{pmatrix} v^1 & w^1 \\ v^2 & w^2 \end{pmatrix} = v^1 w^2 - v^2 w^1 \,.$$

  This is clearly the (signed) 2-dimensional volume (surface area) of the parallelogram spanned by $\mathbf{v}$ and $\mathbf{w}$ as it exposes itself from a perspective along the z direction, i.e. the area of the parallelogram after projection along the z axis onto the x-y plane, cf. Fig. 2.1.

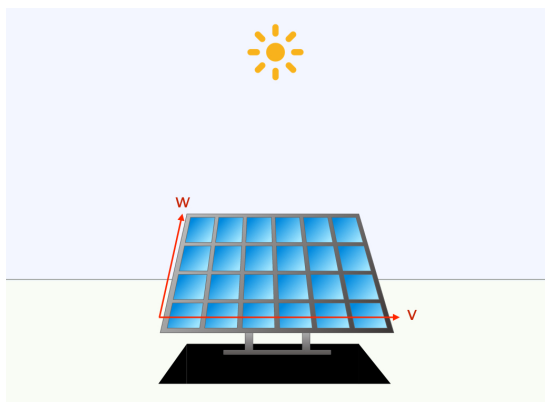


Figure 2.1: Interpretation of $(dx \wedge dy)\,(\mathbf{v}, \mathbf{w})$ as the area of the cast shadow of a solar panel onto the ground (xy-plane) when lit from above (along z-direction). Note that $(\mathbf{v} \wedge \mathbf{w})\,(dx, dy) = (dx \wedge dy)\,(\mathbf{v}, \mathbf{w})$ provides an equivalent interpretation of the same configuration.

2.8.5. Tensors on an Inner Product Space

Ansatz

An n-dimensional inner product space V over $\mathbb{R}$.

Heuristics

The same mechanism that allows us to toggle between vectors and dual vectors in an inner product space, viz. via the $\sharp$- and $\flat$-operator associated with the metric $\mathbf{G}$, also allows us to establish one-to-one relationships between $\mathbf{T}^p_q(V)$ and $\mathbf{T}^r_s(V)$, provided *total rank* is the same: $p + q = r + s$. This implies that—in an inner product space—tensor types can be ordered hierarchically according to their total ranks, and that, given fixed total rank, a tensor can be represented in as many equivalent ways as there are ways to choose its covariant and contravariant rank consistent with its total rank. The basic principle is most easily understood by an example.

Example

Let us interpret the metric $\mathbf{G}$ as an element of $\mathbf{T}^0_2(V)$, as previously discussed. We may consider the following four equivalent alternatives:

- $\mathbf{G} : V \times V \to \mathbb{R} : (\mathbf{v}, \mathbf{w}) \mapsto \mathbf{G}(\mathbf{v}, \mathbf{w})$;
- $\mathbf{H} : V^* \times V^* \to \mathbb{R} : (\hat{\mathbf{v}}, \hat{\mathbf{w}}) \mapsto \mathbf{H}(\hat{\mathbf{v}}, \hat{\mathbf{w}}) \stackrel{\text{def}}{=} \mathbf{G}(\flat\hat{\mathbf{v}}, \flat\hat{\mathbf{w}})$;
- $\mathbf{I} : V^* \times V \to \mathbb{R} : (\hat{\mathbf{v}}, \mathbf{w}) \mapsto \mathbf{I}(\hat{\mathbf{v}}, \mathbf{w}) \stackrel{\text{def}}{=} \mathbf{G}(\flat\hat{\mathbf{v}}, \mathbf{w})$;
- $\mathbf{J} : V \times V^* \to \mathbb{R} : (\mathbf{v}, \hat{\mathbf{w}}) \mapsto \mathbf{J}(\mathbf{v}, \hat{\mathbf{w}}) \stackrel{\text{def}}{=} \mathbf{G}(\mathbf{v}, \flat\hat{\mathbf{w}})$.

Note that $\mathbf{G} \in \mathbf{T}^0_2(V)$, $\mathbf{H} \in \mathbf{T}^2_0(V)$, $\mathbf{I} \in \mathbf{T}^1_1(V)$, $\mathbf{J} \in \mathbf{T}^1_1(V)$. If $\{\mathbf{e}_i\}$ is a basis of V, then

- $\mathbf{G} = g_{ij}\hat{\mathbf{e}}^i \otimes \hat{\mathbf{e}}^j = \sharp\mathbf{e}_i \otimes \hat{\mathbf{e}}^i = \hat{\mathbf{e}}^i \otimes \sharp\mathbf{e}_i$;
- $\mathbf{H} = g^{ij}\mathbf{e}_i \otimes \mathbf{e}_j = \flat\hat{\mathbf{e}}^i \otimes \mathbf{e}_i = \mathbf{e}_i \otimes \flat\hat{\mathbf{e}}^i$;
- $\mathbf{I} = \delta^i_j\mathbf{e}_i \otimes \hat{\mathbf{e}}^j = \mathbf{e}_i \otimes \hat{\mathbf{e}}^i = \flat\hat{\mathbf{e}}^i \otimes \sharp\mathbf{e}_i$;
- $\mathbf{J} = \delta^j_i\hat{\mathbf{e}}^i \otimes \mathbf{e}_j = \hat{\mathbf{e}}^i \otimes \mathbf{e}_i = \sharp\mathbf{e}_i \otimes \flat\hat{\mathbf{e}}^i$.

Caveat

The above example declines from the strict ordering of the spaces V and V^* in the domain of definition of a tensor, as introduced in Definition 2.12 on page 18. This explains the existence of two distinct tensors, $\mathbf{I} \neq \mathbf{J}$, of equal rank $\binom{1}{1}$. Note that $\mathbf{I}(\hat{\mathbf{v}}, \mathbf{w}) = \mathbf{J}(\mathbf{w}, \hat{\mathbf{v}})$.

Definition 2.24 *Recall Result 2.2 on page 14 and Definition 2.12 on page 18. For any $p, q \in \mathbb{N}_0$ we define the* covariant representation $\sharp\mathbf{T} \in \mathbf{T}^0_{p+q}(V)$ *of $\mathbf{T} \in \mathbf{T}^p_q(V)$ as follows:*

$$\sharp\mathbf{T}(\mathbf{v}_1, \ldots, \mathbf{v}_{p+q}) = \mathbf{T}(\sharp\mathbf{v}_1, \ldots, \sharp\mathbf{v}_p, \mathbf{v}_{p+1}, \ldots, \mathbf{v}_{p+q}) \quad \textit{for all } \mathbf{v}_1, \ldots, \mathbf{v}_{p+q} \in V.$$

Likewise, the contravariant representation $\flat\mathbf{T} \in \mathbf{T}^{p+q}_0(V)$ *of $\mathbf{T} \in \mathbf{T}^p_q(V)$ is defined as follows:*

$$\flat\mathbf{T}(\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_{p+q}) = \mathbf{T}(\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_p, \flat\hat{\mathbf{v}}_{p+1}, \ldots, \flat\hat{\mathbf{v}}_{p+q}) \quad \textit{for all } \hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_{p+q} \in V^*.$$

In general, a mixed representation *of $\mathbf{T} \in \mathbf{T}^p_q(V)$ is obtained by toggling an arbitrary subset of (dual) vector spaces V and V^* in the domain of definition of the tensor by means of the metric induced $\{\sharp, \flat\}$-mechanism in a similar fashion.*

Example

Suppose $\mathbf{T} \in \mathbf{T}^1_2(V)$ is given relative to a basis $\{\mathbf{e}_i\}$ of V by $\mathbf{T} = t^i{}_{jk}\mathbf{e}_i \otimes \hat{\mathbf{e}}^j \otimes \hat{\mathbf{e}}^k$. Let $\hat{\mathbf{u}} = u_i\hat{\mathbf{e}}^i$, $\hat{\mathbf{v}} = v_j\hat{\mathbf{e}}^j$, $\hat{\mathbf{w}} = w_k\hat{\mathbf{e}}^k$, $\mathbf{u} = u^i\mathbf{e}_i$, $\mathbf{v} = v^j\mathbf{e}_j$, $\mathbf{w} = w^k\mathbf{e}_k$. Then $\sharp\mathbf{T} = t_{ijk}\hat{\mathbf{e}}^i \otimes \hat{\mathbf{e}}^j \otimes \hat{\mathbf{e}}^k$ and $\flat\mathbf{T} = t^{ijk}\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k$. In terms of components:

$$\begin{array}{rclcl} & & \mathbf{T}(\hat{\mathbf{u}}, \mathbf{v}, \mathbf{w}) & = & t^i{}_{jk}u_iv^jw^k\,, \\ \sharp\mathbf{T}(\mathbf{u}, \mathbf{v}, \mathbf{w}) & = & \mathbf{T}(\sharp\mathbf{u}, \mathbf{v}, \mathbf{w}) & = & t_{ijk}u^iv^jw^k\,, \\ \flat\mathbf{T}(\hat{\mathbf{u}}, \hat{\mathbf{v}}, \hat{\mathbf{w}}) & = & \mathbf{T}(\hat{\mathbf{u}}, \flat\hat{\mathbf{v}}, \flat\hat{\mathbf{w}}) & = & t^{ijk}u_iv_jw_k\,. \end{array} \tag{2.1}$$

The remaining five mixed representations are

$$\begin{array}{rclcl} \mathbf{T}_1(\hat{\mathbf{u}}, \hat{\mathbf{v}}, \mathbf{w}) & = & \mathbf{T}(\hat{\mathbf{u}}, \flat\hat{\mathbf{v}}, \mathbf{w}) & = & t^{ij}{}_ku_iv_jw^k\,, \\ \mathbf{T}_2(\hat{\mathbf{u}}, \mathbf{v}, \hat{\mathbf{w}}) & = & \mathbf{T}(\hat{\mathbf{u}}, \mathbf{v}, \flat\hat{\mathbf{w}}) & = & t^i{}_j{}^ku_iv^jw_k\,, \\ \mathbf{T}_3(\mathbf{u}, \hat{\mathbf{v}}, \hat{\mathbf{w}}) & = & \mathbf{T}(\sharp\mathbf{u}, \flat\hat{\mathbf{v}}, \flat\hat{\mathbf{w}}) & = & t_i{}^{jk}u_iv^jw_k\,, \\ \mathbf{T}_4(\mathbf{u}, \hat{\mathbf{v}}, \mathbf{w}) & = & \mathbf{T}(\sharp\mathbf{u}, \flat\hat{\mathbf{v}}, \mathbf{w}) & = & t_i{}^j{}_ku^iv_jw^k\,, \\ \mathbf{T}_5(\mathbf{u}, \mathbf{v}, \hat{\mathbf{w}}) & = & \mathbf{T}(\sharp\mathbf{u}, \mathbf{v}, \flat\hat{\mathbf{w}}) & = & t_{ij}{}^ku^iv^jw_k\,, \end{array} \tag{2.2}$$

making up the total of $2^3 = 8$ different representations of the rank-3 tensor $\mathbf{T}$, corresponding to the number of possibilities to toggle the arguments of the tensor's 3 slots.

Caveat

The example shows that one quickly runs out of symbols if one insists on a unique naming convention for all distinct tensors that can be constructed in this way, especially for high ranks. One could opt for a systematic labeling, e.g. using a binary subscript to denote which argument requires a toggle (1) and which does not (0). In that case we would have, in the above example, $\mathbf{T} = \mathbf{T}_{000}$, $\sharp\mathbf{T} = \mathbf{T}_{100}$, $\flat\mathbf{T} = \mathbf{T}_{011}$, $\mathbf{T}_1 = \mathbf{T}_{010}$, $\mathbf{T}_2 = \mathbf{T}_{001}$, $\mathbf{T}_3 = \mathbf{T}_{111}$, $\mathbf{T}_4 = \mathbf{T}_{110}$, $\mathbf{T}_5 = \mathbf{T}_{101}$.

Other labeling policies are possible, e.g. using a similar 3-digit binary code to denote argument type, 0 for a vector and 1 for a covector, say. This would lead to the following relabeling: $\mathbf{T} = \mathbf{T}_{100}$, $\sharp\mathbf{T} = \mathbf{T}_{000}$, $\flat\mathbf{T} = \mathbf{T}_{111}$, $\mathbf{T}_1 = \mathbf{T}_{110}$, $\mathbf{T}_2 = \mathbf{T}_{101}$, $\mathbf{T}_3 = \mathbf{T}_{011}$, $\mathbf{T}_4 = \mathbf{T}_{010}$, $\mathbf{T}_5 = \mathbf{T}_{001}$.

Conventions like these are, despite the mathematical clarity they provide, rarely used. One often simply writes $\mathbf{T}$ for *all* cases, relying on the particular tensor prototype for disambiguation ("function overloading").

Notation

- By abuse of notation one invariably uses the same symbol, $\mathbf{T}$ say, to denote any of the (covariant, contravariant, or mixed) prototypes discussed above.
- By the same token the space $\mathbf{T}_k(V) = \bigcup_{p+q=k} \mathbf{T}^p_q(V)$ collectively refers to all tensors of total rank k.
- If a particular prototype $\mathbf{T} \in \mathbf{T}^p_q(V)$ is intended it should be clear from the context.
- Instead of the potentially confusing notation $t^{i_1\ldots i_r}_{j_1\ldots j_s}$ for a mixed holor one often writes $t^{i_1\ldots i_r}{}_{j_1\ldots j_s}$, i.e. one with a *manifest ordering of indices*.

Observation

- The invocation of the $\sharp$-operator to a vector argument of the tensor affects the holor by an *index lowering*.
- The invocation of the $\flat$-operator to a covector argument of the tensor affects the holor by an *index raising*.

Caveat

- Recall that $\sharp\mathbf{e}_i = \mathbf{G}(\mathbf{e}_i) = g_{ij}\hat{\mathbf{e}}^j$, respectively $\flat\hat{\mathbf{e}}^i = \mathbf{G}^{-1}(\hat{\mathbf{e}}^i) = g^{ij}\mathbf{e}_j$. Thus, *at the level of (dual) basis vectors*, $\sharp$ raises, and $\flat$ lowers indices!
- Notice the way indices are raised or lowered in a mixed holor, viz. such as to respect the index ordering, cf. a previous remark on mixed holor notation.

Reminder

- In an inner product space, tensor (holor) indices can be raised or lowered with the help of the (dual) metric tensor.
- As a rule of thumb, consistency of free index positioning determines whether and how to use metric and dual metric components, g_{ij} or g^{ij}, in an expression involving $\sharp$ or $\flat$.

2.8.6. Tensor Transformations

2.8.6.1. "Absolute Tensors"

Ansatz

- An n-dimensional vector space V over $\mathbb{R}$.
- A basis $\{\mathbf{e}_i\}$ of V.
- A linear space of mixed tensors, $\mathbf{T}^p_q(V)$.

Recall that a holor depends on the chosen basis, but the corresponding tensor itself does not. This implies that holors transform in a particular way under a change of basis, which is characteristic for tensors. In fact, in the literature tensors are often defined by virtue of the so-called *tensor transformation law* governing their *holors*. (Notice the misnomer.) Here we present it as a logical result of a tensor's inherently basis independent nature.

Result 2.8 *Let $\mathbf{f}_j = A^i_j \mathbf{e}_i$ define a change of basis, with transformation matrix A. Define $B = A^{-1}$, i.e. $AB = I$, or in terms of components, $A^i_k B^k_j = \delta^i_j$. Then*

$$\mathbf{T} = t^{i_1\ldots i_p}_{j_1\ldots j_q} \mathbf{e}_{i_1} \otimes \ldots \otimes \mathbf{e}_{i_p} \otimes \hat{\mathbf{e}}^{j_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{j_q} \stackrel{\text{def}}{=} \overline{t}^{i_1\ldots i_p}_{j_1\ldots j_q} \mathbf{f}_{i_1} \otimes \ldots \otimes \mathbf{f}_{i_p} \otimes \hat{\mathbf{f}}^{j_1} \otimes \ldots \otimes \hat{\mathbf{f}}^{j_q} ,$$

iff the holor adheres to the tensor transformation law,

$$\overline{t}^{i_1\ldots i_p}_{j_1\ldots j_q} = A^{\ell_1}_{j_1} \ldots A^{\ell_q}_{j_q} B^{i_1}_{k_1} \ldots B^{i_p}_{k_p} t^{k_1\ldots k_p}_{\ell_1\ldots \ell_q} .$$

Proof of Result 2.8. Suppose $\hat{\mathbf{f}}^i = C^i_j \hat{\mathbf{e}}^j$. Then, by definition of duality,

$$\delta^i_j \stackrel{\text{def}}{=} \langle \hat{\mathbf{f}}^i, \mathbf{f}_j \rangle = \langle C^i_k \hat{\mathbf{e}}^k, A^\ell_j \mathbf{e}_\ell \rangle \stackrel{*}{=} C^i_k A^\ell_j \langle \hat{\mathbf{e}}^k, \mathbf{e}_\ell \rangle \stackrel{\text{def}}{=} C^i_k A^\ell_j \delta^k_\ell = C^i_k A^k_j ,$$

whence $C = A^{-1} = B$. Bilinearity of the Kronecker tensor has been used in $*$. Consequently,

$$\mathbf{e}_i = B^\ell_i \mathbf{f}_\ell \quad \text{and} \quad \hat{\mathbf{e}}^j = A^j_k \hat{\mathbf{f}}^k .$$

Substitution into the basis expansion of $\mathbf{T}$ yields the stated result for the respective holors.

Remark

The homogeneous nature of the tensor transformation law implies that a holor equation of the form $t^{i_1\ldots i_p}_{j_1\ldots j_q} = 0$ holds relative to *any* basis if it holds relative to a particular one.

Example

With the same notation as used in Result 2.8 we have for a vector $\mathbf{v} = v^i \mathbf{e}_i = \overline{v}^i \mathbf{f}_i$:

$$\overline{v}^i = B^i_j v^j ,$$

also referred to as *vector transformation law*. Recall Theorem 2.1 on page 7.

Example

Let $df = \partial_i f \hat{\mathbf{e}}^i \in V^*$ denote the gradient of a scalar function $f \in C^1(\mathbb{R}^n)$, interpreted as a dual (!) vector and evaluated at a fixed point. The interpretation of this is as follows. If $\mathbf{v} = v^i \mathbf{e}_i \in V$, then

$$\langle df, \mathbf{v} \rangle = \partial_i f v^i \stackrel{*}{=} (\nabla f | \mathbf{v}) \ ,$$

the directional derivative of f along $\mathbf{v}$. The equality marked by $*$ holds only if V is an inner product space, in which case we identify $\nabla f = \flat df \in V$.
According to the tensor transformation law we have $df = \overline{\partial}_i f \hat{\mathbf{f}}^i$, with $\hat{\mathbf{f}}^i = B^i_j \hat{\mathbf{e}}^j$, and

$$\overline{\partial}_i f = A^\ell_i \partial_\ell f \, .$$

Notice the analogy with "infinitesimal calculus". Indeed, a change of coordinates, $x \mapsto \overline{x} = \phi(x)$, induces a reparametrization $df = \partial_i f dx^i = \overline{\partial}_i f d\overline{x}^i$, in which

$$\overline{\partial}_i f = \frac{\partial f}{\partial \overline{x}^i} = \frac{\partial f}{\partial x^\ell} \frac{\partial x^\ell}{\partial \overline{x}^i} = A^\ell_i \partial_\ell f \quad \text{with} \quad A^\ell_i \stackrel{\text{def}}{=} \frac{\partial x^\ell}{\partial \overline{x}^i} \, ,$$

by the chain rule.

Caveat

One should distinguish between the new variable $\overline{x}$ and the coordinate transformation function ϕ that relates it to the original variable x by $\overline{x} = \phi(x)$. In practice one often treats $\overline{x}$ either as a variable or as a function (i.e. $\overline{x} \equiv \phi$), depending on context. Likewise for x and ϕ^{-1}.

Example

Recall the observation on page 10, which confirms the tensor transformation law for the dual basis vectors dx^i:

$$d\overline{x}^i = \frac{\partial \overline{x}^i}{\partial x^j} dx^j = B^i_j dx^j \, ,$$

with $B = A^{-1}$, cf. the previous example.

Caveat

Compare the latter example of the transformation law for a *basis covector*, $d\overline{x}^i = B^i_j dx^j$, to the one on page 37 for the *components (holor) of a vector*, $\overline{v}^i = B^i_j v^j$. The apparent *formal* correspondence has led practitioners of tensor calculus (mostly physicists and engineers) to interpret the basis covectors dx^i as the "components of an infinitesimal displacement vector". This misinterpretation is, although potentially confusing, harmless and indeed justified to the extent that it is consistent with the stipulated tensor transformation laws for basis dual vectors, respectively regular vector components. Confusion may be avoided altogether if we denote the components of an "infinitesimal vector" by δx^i (i.e. a number, replacing the notation v^i above so as to express its small nature), reserving dx^i for the i-th dual basis covector (i.e. an element of V^*).

2.8.6.2. "Relative Tensors"

Recall the unit volume form $\boldsymbol{\mu} = \mu_{i_1 \ldots i_n} \hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n}$ introduced in section 2.8.4. Recall from page 32 that the identification $\mu_{i_1 \ldots i_n} = [i_1, \ldots, i_n]$ depends on the preferred basis used for normalisation. Heuristically this is clear from the geometric interpretation provided by the example on page 33 in terms of the volume of the parallelepiped spanned by the fiducial basis vectors. Let us scrutinise the situation more carefully.

Ansatz

- An n-dimensional vector space V over $\mathbb{R}$.
- A *fixed* basis $\{\mathbf{e}_i\}$ of V.
- The volume form $\boldsymbol{\mu} = \mu_{|i_1\ldots i_n|}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_n} = \mu_{i_1\ldots i_n}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n} \in \bigwedge_n(V)$.
- The normalisation $\mu_{i_1\ldots i_n} = [i_1, \ldots, i_n]$, declaring $\boldsymbol{\mu}$ as the unit volume form relative to $\{\mathbf{e}_i\}$.

The following lemma allows us to compute the holor of $\boldsymbol{\mu}$ relative to an arbitrary basis, given its canonical form $[i_1, \ldots, i_n]$ relative to our fiducial basis, via Result 2.8, page 37.

Lemma 2.1 *Let $\boldsymbol{\mu} = \mu_{i_1\ldots i_n}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n} \in \bigwedge_n(V)$, with $\mu_{i_1\ldots i_n} = [i_1, \ldots, i_n]$ relative to basis $\{\mathbf{e}_i\}$ of V, and $\mathbf{f}_j = A^i_j\mathbf{e}_i$, so that $\hat{\mathbf{e}}^i = A^i_j\hat{\mathbf{f}}^j$. Setting $\boldsymbol{\mu} = \overline{\mu}_{i_1\ldots i_n}\hat{\mathbf{f}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{f}}^{i_n}$, we have*

$$\overline{\mu}_{j_1\ldots j_n} = A^{i_1}_{j_1} \ldots A^{i_n}_{j_n}\mu_{i_1\ldots i_n} = \det A\, \mu_{j_1\ldots j_n}\,.$$

Proof of Lemma 2.1. In accordance with Result 2.8, page 37, we have

$$\boldsymbol{\mu} = \mu_{i_1\ldots i_n}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n} = \mu_{i_1\ldots i_n}A^{i_1}_{j_1} \ldots A^{i_n}_{j_n}\hat{\mathbf{f}}^{j_1} \otimes \ldots \otimes \hat{\mathbf{f}}^{j_n}\,.$$

This proves the first identity. Using $\mu_{i_1\ldots i_n} = [i_1, \ldots, i_n]$, together with an observation made on page 5 in the context of the determinant, this may be rewritten as

$$\boldsymbol{\mu} = \det A\, [j_1, \ldots, j_n]\, \hat{\mathbf{f}}^{j_1} \otimes \ldots \otimes \hat{\mathbf{f}}^{j_n}\,.$$

This establishes the proof of the last identity.

Observation

- Note that $\det A$ is the volume scaling factor of the transformation with matrix A.
- If $\boldsymbol{\mu}(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$, then $\boldsymbol{\mu}(\mathbf{f}_1, \ldots, \mathbf{f}_n) = \det A$.
- Thus $\boldsymbol{\mu}$ is, in general, *not* the unit volume form relative to $\{\mathbf{f}_i\}$, unless $\det A = 1$.
- The holor $\mu_{i_1\ldots i_n} = [i_1, \ldots, i_n]$ is, consequently, *not* invariant under the *absolute tensor transformation law*, Result 2.8, in the sense that $\overline{\mu}_{i_1\ldots i_n} = \det A\, \mu_{i_1\ldots i_n} \neq [i_1, \ldots, i_n] \doteq \mu_{i_1\ldots i_n}$.

Heuristics

Clearly, the constant antisymmetric symbol fails to define an invariant holor of a tensor (as opposed to the holor δ^i_j of the Kronecker tensor). The definition of the holor in terms of the antisymmetric symbol requires us to single out a preferred basis relative to which it holds, i.e. the fiducial basis $\{\mathbf{e}_i\}$ has a preferred status among all possible bases. Lack of invariance is not surprising, since a given holor is, by its very definition, defined relative to a particular basis. The question nevertheless arises whether it is possible to give an invariant expression for the holor, i.e. one that holds in *any* basis, so that all bases are treated on equal foot and no a priori knowledge of a preferred one is needed. This is indeed possible by a formal "trick", viz. through modification of the "tensor transformation law" for the (invariant counterpart of the) holor. There are two ways to remedy the lack of invariance in that case:

1. without the help of an inner product (Result 2.9), respectively
2. with the help of an inner product (Definition 2.25 and Result 2.10).

Result 2.9 *Recall Result 2.8, page 37, for notation. If*

$$\boldsymbol{\mu} = \mu_{|i_1\ldots i_n|}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_n} = \overline{\mu}_{|i_1\ldots i_n|}\hat{\mathbf{f}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{f}}^{i_n} ,$$

and, by definition,

$$\overline{\mu}_{i_1\ldots i_n} \stackrel{\text{def}}{=} (\det A)^{-1} A^{j_1}_{i_1} \ldots A^{j_n}_{i_n} \mu_{j_1\ldots j_n} ,$$

then $\mu_{i_1\ldots i_n} \equiv \overline{\mu}_{i_1\ldots i_n} \equiv [i_1, \ldots, i_n]$ *is an invariant holor.*

Terminology

- This adapted transformation law is also known as the *relative tensor transformation law*. The attribute "relative" refers to the occurrence of (a power of) the Jacobian determinant.
- By virtue of the relative tensor transformation law governing its holor, $\boldsymbol{\mu}$ is referred to as a *relative tensor*. Strictly speaking this terminology refers to its holor definition!

Proof of Result 2.9. Using an observation on page 5 we find

$$\begin{aligned}\overline{\mu}_{i_1\ldots i_n} &\stackrel{\text{def}}{=} (\det A)^{-1} A^{j_1}_{i_1} \ldots A^{j_n}_{i_n} \mu_{j_1\ldots j_n} = (\det A)^{-1} A^{j_1}_{i_1} \ldots A^{j_n}_{i_n} [j_1, \ldots, j_n] \\ &= (\det A)^{-1} [i_1, \ldots, i_n] \det A = [i_1, \ldots, i_n] = \mu_{i_1\ldots i_n} .\end{aligned}$$

Remark

- In Lemma 2.1 on page 39 the holor $\mu_{i_1\ldots i_n} = [i_1, \ldots, i_n]$ is transformed according to the absolute tensor transformation law, recall Result 2.8 on page 37, which implies that $\overline{\mu}_{i_1\ldots i_n} = \det A\, \mu_{i_1\ldots i_n} \neq \mu_{i_1\ldots i_n}$.
- In Result 2.9 the holor $\mu_{i_1\ldots i_n} = [i_1, \ldots, i_n]$ is transformed according to the relative tensor transformation law, rendering the holor $\mu_{i_1\ldots i_n} = [i_1, \ldots, i_n]$ *invariant* in the sense that $\overline{\mu}_{i_1\ldots i_n} = \mu_{i_1\ldots i_n} = [i_1, \ldots, i_n]$ regardless of the basis.
- Adopting the relative tensor transformation law allows us to interpret $\mu_{i_1\ldots i_n}$ as a 'covariant alias' of the permutation symbol $[i_1, \ldots, i_n]$, which by definition does not depend on a basis.
- There is no inconsistency other than abuse of notation (to be resolved below). In both interpretations we have
$$\boldsymbol{\mu} = \mu_{i_1\ldots i_n}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n} = \overline{\mu}_{i_1\ldots i_n}\hat{\mathbf{f}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{f}}^{i_n} ,$$
or, equivalently,
$$\boldsymbol{\mu} = \mu_{|i_1\ldots i_n|}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_n} = \overline{\mu}_{|i_1\ldots i_n|}\hat{\mathbf{f}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{f}}^{i_n} .$$

Notation

The following convention is used henceforth to support consistent use of the Einstein summation convention and to resolve ambiguity w.r.t. the meaning of $\overline{\mu}_{i_1\ldots i_n}$:

- Henceforth we will adopt $\mu_{i_1\ldots i_n}$ and $\mu^{i_1\ldots i_n}$ as covariant and contravariant aliases for the permutation symbol, i.e. $\mu_{i_1\ldots i_n} \equiv \mu^{i_1\ldots i_n} \equiv [i_1, \ldots, i_n]$.
- Consequently, the holor $\mu_{i_1\ldots i_n}$ is subject to the *relative tensor transformation law*, Result 2.9.

2.8.6.3. "Pseudo Tensors"

Ansatz

- An n-dimensional inner product space V over $\mathbb{R}$, based on our default Definition 2.5.
- A *fixed* basis $\{\mathbf{e}_i\}$ of V.
- The unit volume form $\boldsymbol{\mu} = \mu_{|i_1\ldots i_n|}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_n} = \mu_{i_1\ldots i_n}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n} \in \bigwedge_n(V)$, with invariant holor definition $\mu_{i_1\ldots i_n} = [i_1, \ldots, i_n]$.

Notation

For brevity one denotes the determinant of the Gram matrix, $g_{ij} = (\mathbf{e}_i|\mathbf{e}_j)$, by $g = \det g_{ij}$, recall page 5 concerning a remark on the use of indices, and Definition 2.8 on page 12.

Definition 2.25 *The* Levi-Civita tensor *is the unique unit n-form of positive orientation:*

$$\boldsymbol{\epsilon} = \sqrt{g}\,\boldsymbol{\mu} = \sqrt{g}\,\hat{\mathbf{e}}^1 \wedge \ldots \wedge \hat{\mathbf{e}}^n = \epsilon_{|i_1\ldots i_n|}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_n}\,.$$

Observation

- $\epsilon_{i_1\ldots i_n} = \sqrt{g}\,[i_1, \ldots, i_n] = \sqrt{g}\,\mu_{i_1\ldots i_n}$.
- $\boldsymbol{\epsilon} = \epsilon_{i_1\ldots i_n}\hat{\mathbf{e}}^{i_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n} = \epsilon_{|i_1\ldots i_n|}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_n}$ in terms of the appropriate summation conventions.

Theorem 2.6 *The contravariant representation of the Levi-Civita tensor, $\flat\boldsymbol{\epsilon} \in \bigwedge^n(V)$, recall Definition 2.24 on page 35, is given by $\flat\boldsymbol{\epsilon} = \epsilon^{|i_1\ldots i_n|}\mathbf{e}_{i_1} \wedge \ldots \wedge \mathbf{e}_{i_n}$, with holor representation*

$$\epsilon^{i_1\ldots i_n} = \frac{1}{\sqrt{g}}\,[i_1, \ldots, i_n]\,.$$

Proof of Theorem 2.6. It suffices to prove the validity of the holor representation by index raising:

$$\epsilon^{i_1\ldots i_n} = g^{i_1 j_1}\ldots g^{i_n j_n}\epsilon_{j_1\ldots j_n} = g^{i_1 j_1}\ldots g^{i_n j_n}\sqrt{g}\,[j_1, \ldots, j_n] = \frac{1}{\sqrt{g}}\,[i_1, \ldots, i_n]\,.$$

In the last step we have made use of the fact that $g^{i_1 j_1}\ldots g^{i_n j_n}\,[j_1, \ldots, j_n] = \det g^{ij}\,[i_1, \ldots, i_n]$, an observation made on page 5, together with the fact that $\det g^{ij} = 1/\det g_{ij}$.

Observation

- $\epsilon^{i_1\ldots i_n} = \dfrac{1}{\sqrt{g}}\,[i_1, \ldots, i_n] = \dfrac{1}{\sqrt{g}}\mu^{i_1\ldots i_n}$ relative to the fiducial basis $\{\mathbf{e}_i\}$.
- $\flat\boldsymbol{\epsilon} = \epsilon^{i_1\ldots i_n}\mathbf{e}_{i_1} \otimes \ldots \otimes \mathbf{e}_{i_n} = \epsilon^{|i_1\ldots i_n|}\mathbf{e}_{i_1} \wedge \ldots \wedge \mathbf{e}_{i_n}$.

Result 2.10 *Recall Result 2.8, page 37, for notation. Let*

$$\boldsymbol{\epsilon} = \epsilon_{|i_1\ldots i_n|}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_n}$$

be the basis n-form relative to basis $\{\mathbf{e}_i\}$, with $\epsilon_{i_1\ldots i_n} = \sqrt{g}\,[i_1,\ldots,i_n]$. Moreover, let

$$\overline{\boldsymbol{\epsilon}} = \overline{\epsilon}_{|i_1\ldots i_n|}\hat{\mathbf{f}}^{i_1}\wedge\ldots\wedge\hat{\mathbf{f}}^{i_n}$$

be the basis n-form relative to basis $\{\mathbf{f}_i\}$, with $\mathbf{f}_i = A_i^j\mathbf{e}_j$. Then the pseudotensor transformation law*,*

$$\overline{\epsilon}_{i_1\ldots i_n} \stackrel{\text{def}}{=} \operatorname{sgn}(\det A)A_{i_1}^{j_1}\ldots A_{i_n}^{j_n}\epsilon_{j_1\ldots j_n}\,,$$

renders the holor form invariant, in the sense that $\overline{\epsilon}_{i_1\ldots i_n} = \sqrt{\overline{g}}\,[i_1,\ldots,i_n]$. Note that $\overline{\boldsymbol{\epsilon}} = \operatorname{sgn}(\det A)\,\boldsymbol{\epsilon}$.

Proof of Theorem 2.10. By virtue of the transformation properties of g_{ij}, recall Result 2.8 on page 37, one obtains

$$\overline{g}_{ij} = A_i^k A_j^\ell g_{k\ell}\,,$$

whence

$$\sqrt{\overline{g}} = |\det A|\,\sqrt{g}\,.$$

By definition of the pseudotensor transformation law we have, using an observation on page 5,

$$\begin{aligned}\overline{\epsilon}_{i_1\ldots i_n} &\stackrel{\text{def}}{=} \operatorname{sgn}(\det A)A_{i_1}^{j_1}\ldots A_{i_n}^{j_n}\epsilon_{j_1\ldots j_n} = \operatorname{sgn}(\det A)\,A_{i_1}^{j_1}\ldots A_{i_n}^{j_n}\sqrt{g}\,\mu_{j_1\ldots j_n}\\ &= \operatorname{sgn}(\det A)\,\det A\sqrt{g}\,\mu_{i_1\ldots i_n} = |\det A|\,\sqrt{g}\,\mu_{i_1\ldots i_n} = \sqrt{\overline{g}}\,\overline{\mu}_{i_1\ldots i_n}\,.\end{aligned}$$

Recall that $\mu_{i_1\ldots i_n}$ and $\overline{\mu}_{i_1\ldots i_n}$ have been used here as aliases for the same (basis independent) permutation symbol $[i_1,\ldots,i_n]$.

Remark

- The Levi-Civita tensor is referred to as a *pseudotensor*, because its *form invariant holor* (!) transforms according to the *pseudotensor transformation law* of Result 2.10, *not* the absolute tensor transformation law of Result 2.8.
- Form invariance, achieved by virtue of the factor $\sqrt{g}$, should not be confused with numerical invariance. Clearly $\sqrt{\overline{g}} \neq \sqrt{g}$, whence $\overline{\epsilon}_{i_1\ldots i_n} \neq \epsilon_{i_1\ldots i_n}$, unlike the formal identity $\overline{\mu}_{i_1\ldots i_n} = \mu_{i_1\ldots i_n} \equiv [i_1,\ldots,i_n]$.
- In the case of a non-positive inner product, recall Definition 2.7, slight adaptations may be needed, a typical one being the replacement of the $\sqrt{g}$ factor by $\sqrt{|g|}$ if g is negative.

The following definition is based on an earlier observation, recall page 5.

Definition 2.26 *The* generalised Kronecker*, or* permutation symbol*, is defined as*

$$\delta_{j_1\ldots j_k}^{i_1\ldots i_k} = \det\begin{pmatrix}\delta_{j_1}^{i_1} & \ldots & \delta_{j_k}^{i_1}\\ \vdots & & \vdots\\ \delta_{j_1}^{i_k} & \ldots & \delta_{j_k}^{i_k}\end{pmatrix} = \begin{cases}+1 & \textit{if } (i_1,\ldots,i_k) \textit{ even permutation of } (j_1,\ldots,j_k)\,,\\ -1 & \textit{if } (i_1,\ldots,i_k) \textit{ odd permutation of } (j_1,\ldots,j_k)\,,\\ 0 & \textit{otherwise}\,.\end{cases}$$

Observation

- $\epsilon_{j_1\ldots j_n}\epsilon^{i_1\ldots i_n} = \mu_{j_1\ldots j_n}\mu^{i_1\ldots i_n} = \delta_{j_1\ldots j_n}^{i_1\ldots i_n}$.
- $\epsilon_{\ell_1\ldots\ell_k j_1\ldots j_{n-k}}\epsilon^{\ell_1\ldots\ell_k i_1\ldots i_{n-k}} = \mu_{\ell_1\ldots\ell_k j_1\ldots j_{n-k}}\mu^{\ell_1\ldots\ell_k i_1\ldots i_{n-k}} = k!\delta_{j_1\ldots j_{n-k}}^{i_1\ldots i_{n-k}}$.
- For the matrix A of a mixed tensor $\mathbf{A}\in\mathbf{T}_1^1(V)$ we have $\det A_j^i = \dfrac{1}{n!}\delta_{j_1\ldots j_n}^{i_1\ldots i_n}A_{i_1}^{j_1}\ldots A_{i_n}^{j_n}$.

2.8.7. Contractions

Ansatz

- An n-dimensional vector space V over $\mathbb{R}$.
- A basis $\{\mathbf{e}_i\}$ of V.
- A linear space of mixed tensors, $\mathbf{T}^p_q(V)$.

Definition 2.27 *Let* $\mathbf{T} \in \mathbf{T}^1_1(V)$. *The* trace, $\operatorname{tr}\mathbf{T} \in \mathbb{R}$, *of* $\mathbf{T}$ *is defined as follows:*

$$\operatorname{tr}\mathbf{T} = \mathbf{T}(\hat{\mathbf{e}}^i, \mathbf{e}_i)\,.$$

Remark

- The tr -operator is applicable to any mixed tensor $\mathbf{T} \in \mathbf{T}^p_q(V)$ provided $p, q \geq 1$, viz. by arbitrarily fixing $p-1$ covector and $q-1$ vector arguments.
- Without "frozen" arguments one may interpret $\operatorname{tr}\mathbf{T}$ as a mixed tensor $\operatorname{tr}\mathbf{T} \in \mathbf{T}^{p-1}_{q-1}(V)$.
- The tr -operator lowers both covariant and contravariant ranks by one.

Caveat

- In general, the notation $\operatorname{tr}\mathbf{T}$ is ambiguous for $\mathbf{T} \in \mathbf{T}^p_q(V)$, unless $\mathbf{T}$ is symmetric with respect to both covector as well as vector arguments. (This is manifest if $p = q = 1$.)
- One could decorate the tr -operator with disambiguating subscripts and superscripts so as to indicate the covector/vector argument pair it applies to. This is hardly ever done in the literature. Instead, ambiguities should be resolved from the context.

2.9. The Hodge Star Operator

Ansatz

- An n-dimensional inner product space V over $\mathbb{R}$ furnished with a basis $\{\mathbf{e}_i\}$.
- The Levi-Civita tensor $\boldsymbol{\epsilon} = \sqrt{g}\,\hat{\mathbf{e}}^1 \wedge \ldots \wedge \hat{\mathbf{e}}^n = \epsilon_{|i_1 \ldots i_n|}\hat{\mathbf{e}}^{i_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_n} \in \bigwedge_n(V)$.

Definition 2.28 *Let* $\mathbf{a}_1, \ldots, \mathbf{a}_k \in V$. *The* Hodge star operator $* : \bigwedge_k(V) \to \bigwedge_{n-k}(V)$ *is defined as follows:*

$$\begin{cases} *1 & = & \boldsymbol{\epsilon} & \textit{for } k = 0, \\ *(\sharp\mathbf{a}_1 \wedge \ldots \wedge \sharp\mathbf{a}_k) & = & \boldsymbol{\epsilon} \lrcorner \mathbf{a}_1 \lrcorner \ldots \lrcorner \mathbf{a}_k & \textit{for } k = 1, \ldots, n, \end{cases}$$

followed by linear extension.

This definition could be rephrased by replacing $\sharp\mathbf{a}_1, \ldots, \sharp\mathbf{a}_k \in V^*$ by $\hat{\mathbf{a}}_1, \ldots, \hat{\mathbf{a}}_k \in V^*$ on the l.h.s. and $\mathbf{a}_1, \ldots, \mathbf{a}_k \in V$ by $\flat\hat{\mathbf{a}}_1, \ldots, \flat\hat{\mathbf{a}}_k \in V$ on the r.h.s. A similar definition could be given for a contravariant

counterpart $* : \bigwedge^k(V) \to \bigwedge^{n-k}(V)$, in which case $\sharp$-conversion of the vector arguments $\mathbf{a}_1, \ldots, \mathbf{a}_k \in V$ is obsolete.

Remark

The Hodge star operator establishes an isomorphism between $\bigwedge_k(V)$ and $\bigwedge_{n-k}(V)$, recall the dimensionality argument on page 26:

$$\dim \bigwedge_k(V) = \binom{n}{k} = \binom{n}{n-k} = \dim \bigwedge_{n-k}(V)\,.$$

Observation

- The holor of the tensor $*\left(\sharp\mathbf{a}_1 \wedge \ldots \wedge \sharp\mathbf{a}_k\right)$ for fixed $\mathbf{a}_p = a_p^{i_p}\mathbf{e}_{i_p}$, with $p = 1, \ldots, k$, is obtained as follows:

$$\begin{aligned}
\left(*\left(\sharp\mathbf{a}_1 \wedge \ldots \wedge \sharp\mathbf{a}_k\right)\right)\left(\mathbf{e}_{i_{k+1}}, \ldots, \mathbf{e}_{i_n}\right) &= a_1^{i_1} \ldots a_k^{i_k}\left(*\left(\sharp\mathbf{e}_{i_1} \wedge \ldots \wedge \sharp\mathbf{e}_{i_k}\right)\right)\left(\mathbf{e}_{i_{k+1}}, \ldots, \mathbf{e}_{i_n}\right) \\
&\overset{\text{def}}{=} a_1^{i_1} \ldots a_k^{i_k}\left(\boldsymbol{\epsilon} \lrcorner \mathbf{e}_{i_1} \lrcorner \ldots \lrcorner \mathbf{e}_{i_k}\right)\left(\mathbf{e}_{i_{k+1}}, \ldots, \mathbf{e}_{i_n}\right) \\
&= a_1^{i_1} \ldots a_k^{i_k}\boldsymbol{\epsilon}\left(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}\right) = \epsilon_{i_1 \ldots i_n} a_1^{i_1} \ldots a_k^{i_k}\,.
\end{aligned}$$

- By the same token, expanding also $\mathbf{x}_q = x_q^{i_q}\mathbf{e}_{i_q}$ for $q = k+1, \ldots, n$, we obtain

$$\left(*\left(\sharp\mathbf{a}_1 \wedge \ldots \wedge \sharp\mathbf{a}_k\right)\right)\left(\mathbf{x}_{k+1}, \ldots, \mathbf{x}_n\right) = \epsilon_{i_1 \ldots i_n} a_1^{i_1} \ldots a_k^{i_k} x_{k+1}^{i_{k+1}} \ldots x_n^{i_n}\,.$$

Result 2.11 *Alternatively, we may rewrite Definition 2.28 as follows:*

$$\begin{cases}
*1 &= \boldsymbol{\epsilon} & \textit{for } k = 0, \\
*\left(\hat{\mathbf{a}}^1 \wedge \ldots \wedge \hat{\mathbf{a}}^k\right) &= \boldsymbol{\epsilon} \lrcorner \flat\hat{\mathbf{a}}^1 \lrcorner \ldots \lrcorner \flat\hat{\mathbf{a}}^k & \textit{for } k = 1, \ldots, n.
\end{cases}$$

In this case we find, setting $\hat{\mathbf{a}}^p = a_{j_p}^p \hat{\mathbf{e}}^{j_p}$ *for* $p = 1, \ldots, k$*, and using* $\hat{\mathbf{e}}^j = g^{ij}\sharp\mathbf{e}_i$*,*

$$\begin{aligned}
\left(*\left(\hat{\mathbf{a}}^1 \wedge \ldots \wedge \hat{\mathbf{a}}^k\right)\right)\left(\mathbf{e}_{i_{k+1}}, \ldots, \mathbf{e}_{i_n}\right) &= a_{j_1}^1 \ldots a_{j_k}^k\left(*\left(\hat{\mathbf{e}}^{j_1} \wedge \ldots \wedge \hat{\mathbf{e}}^{j_k}\right)\right)\left(\mathbf{e}_{i_{k+1}}, \ldots, \mathbf{e}_{i_n}\right) \\
&= g^{i_1 j_1} \ldots g^{i_k j_k} a_{j_1}^1 \ldots a_{j_k}^k\left(*\left(\sharp\mathbf{e}_{i_1} \wedge \ldots \wedge \sharp\mathbf{e}_{i_k}\right)\right)\left(\mathbf{e}_{i_{k+1}}, \ldots, \mathbf{e}_{i_n}\right) \\
&\overset{\text{def}}{=} g^{i_1 j_1} \ldots g^{i_k j_k} a_{j_1}^1 \ldots a_{j_k}^k\left(\boldsymbol{\epsilon} \lrcorner \mathbf{e}_{i_1} \lrcorner \ldots \lrcorner \mathbf{e}_{i_k}\right)\left(\mathbf{e}_{i_{k+1}}, \ldots, \mathbf{e}_{i_n}\right) \\
&= g^{i_1 j_1} \ldots g^{i_k j_k} a_{j_1}^1 \ldots a_{j_k}^k \boldsymbol{\epsilon}\left(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}\right) \\
&= \epsilon_{i_1 \ldots i_n} g^{i_1 j_1} \ldots g^{i_k j_k} a_{j_1}^1 \ldots a_{j_k}^k\,.
\end{aligned}$$

Observation

As a rule of thumb, in order to compute the holor of the Hodge star of a *general* antisymmetric cotensor of rank k, the holor of which has free indices $i_1, \ldots, i_k$, say, raise all indices and contract with the antisymmetric $\epsilon_{i_1 \ldots i_n}$-symbol. That is, if $\mathbf{A} \in \bigwedge_k(V)$, then the holor of $*\mathbf{A} \in \bigwedge_{n-k}(V)$ is given by

$$*A_{i_{k+1} \ldots i_n} \overset{\text{def}}{=} (*\mathbf{A})\left(\mathbf{e}_{i_{k+1}}, \ldots, \mathbf{e}_{i_n}\right) = \frac{1}{k!}\epsilon_{i_1 \ldots i_n} g^{i_1 j_1} \ldots g^{i_k j_k} A_{j_1 \ldots j_k}\,.$$

In other words,

$$*\mathbf{A} = *A_{i_{k+1} \ldots i_n}\hat{\mathbf{e}}^{i_{k+1}} \otimes \ldots \otimes \hat{\mathbf{e}}^{i_n} = *A_{|i_{k+1} \ldots i_n|}\hat{\mathbf{e}}^{i_{k+1}} \wedge \ldots \wedge \hat{\mathbf{e}}^{i_n}\,.$$

In fact it suffices to remember that $*\left(\sharp\mathbf{e}_{i_1} \wedge \ldots \wedge \sharp\mathbf{e}_{i_k}\right)\left(\mathbf{e}_{i_{k+1}}, \ldots, \mathbf{e}_{i_n}\right) = \epsilon_{i_1 \ldots i_n}$, by exploiting multilinearity.

Remark

- Taking $k = 0$ in the above observation is consistent with $*1 = \boldsymbol{\epsilon}$.
- Taking $k = n$ shows that $*\boldsymbol{\epsilon} = ** 1 = 1$. This result relies essentially on Definition 2.5. In the general case of Definition 2.7 you may verify that $*\boldsymbol{\epsilon} = ** 1 = \operatorname{sgn} g$, and that $** \boldsymbol{A} = (-1)^{k(n-k)} \operatorname{sgn} g \, \boldsymbol{A}$ for a general k-form $\boldsymbol{A} \in \bigwedge_k(V)$.

3. Differential Geometry

3.1. Euclidean Space: Cartesian and Curvilinear Coordinates

In this chapter we consider *tensor fields*, i.e. tensors, in the widest sense as introduced in Chapter 2, attached to all points of an open subset $\Xi \subset \mathbb{R}^n$ and considered as "smoothly varying" functions of those points. The space $\mathbb{R}^n$ represents n-dimensional Euclidean space, i.e. "classical" space endowed with a flat, homogeneous and isotropic structure as we usually take for granted [8, 9, 10].

Terminology

Recall Definition 2.5, page 11. The terminology "*n-dimensional Euclidean vector space*" is synonymous to "n-dimensional (real, positive definite) inner product space".

Definition 3.1 *An* n-dimensional Euclidean space *E is a metric space, i.e. an affine space equipped with a distance function $d : E \times E \to \mathbb{R}$, furnished with a mapping $+ : E \times V \to E$, in which V is an n-dimensional Euclidean vector space, such that*

- *for all $x, y \in E$ there exists a unique $v \in V$ such that $y = x + v$, and $d(x, y) = \sqrt{(v|v)}$;*
- *for all $x, y \in E$ and all $v \in V$ $d(x + v, y + v) = d(x, y)$;*
- *for all $x \in E$ and all $u, v \in V$ $(x + u) + v = x + (u + v)$.*

Heuristics

Roughly speaking, a Euclidean space is a Euclidean vector space V that has "forgotten" its origin, or for which any point may be taken as origin. To make this more precise, let us introduce a scalar $t \in \mathbb{R}$, replace y by $x(t)$ and x by $x(0)$, and consider the parametrized (straight) path $x(t) = x(0) + vt$. Recall that $+ : \mathrm{E} \times V \to \mathrm{E}$ does *not* denote vector addition here. If we now *arbitrarily* designate an origin $\omega \in \mathrm{E}$, and reinterpret $x(t), x(0) \in \mathrm{E}$, by abuse of notation, as the (unique) *vectors* that connect them to this fiducial origin, then we may reinterpret each term as a genuine *vector*, and the transport operator as *vector addition*, $+ : V \times V \to V$. It is thus legitimate to read $x(t) = x(0) + vt$ as a linear combination in V, with the identification $V = \mathrm{E}$ as a set, regardless of the choice of origin. Note that $\ddot{x}(t) = \dot{v}(t) = 0$—with $v(t) \doteq \dot{x}(t)$—is the defining second order ordinary differential equation for the family of all such (straight) lines. These are the so-called geodesics of Euclidean geometry, to be generalized for curved manifolds later on.

Observation

The distance function naturally arises from the inner product: $d(x, y) = \sqrt{(y - x|y - x)}$, in which $y - x$ denotes the vector $v \in V$ such that $y = x + v$. According to the explanation above we may reinterpret $x, y \in \mathrm{E}$ as vectors $x, y \in V$ after the choice of a global origin, which boils down to replacing the distance function prototype $d : \mathrm{E} \times \mathrm{E} \to \mathbb{R}$ by one defined on V, viz. $d : V \times V \to \mathbb{R}$ for the difference vector (with the $- : \mathrm{E} \times \mathrm{E} \to V$ operator reinterpreted as $- : V \times V \to V$ in the usual way, referring to vector addition with the antivector of its second argument).

Observation

Another important observation is that if $x(t) = x(0) + vt$ and $y(t) = y(0) + wt$ are two Euclidean geodesics, then $z(t) \doteq y(t) - x(t)$ is generically linear in t (constant iff $v = w$), whence its second order derivative vanishes identically, $\ddot{z}(t) = 0$. This is the so-called geodesic deviation equation of Euclidean geometry. It relates to Euclid's fifth postulate about Euclidean geometry [8]: "If a line segment intersects two straight lines forming two interior angles on the same side that sum to less than two right angles, then the two lines, if extended indefinitely, meet on that side on which the angles sum to less than two right angles." (You may want to sketch the situation and relate Euclid's fifth postulate to the geodesic deviation vector $z(t)$.)

Caveat

Both E and V are usually written as $\mathbb{R}^n$, implying the choice of a (global) basis. It should be clear from the context whether a Euclidean space or a Euclidean vector space is meant.

A point in $\mathbf{x} \in \mathrm{E} \sim \mathbb{R}^n$ can be represented by its *Cartesian coordinates* $x^i \in \mathbb{R}$ relative to a (globally defined) *Cartesian coordinate system* or *standard basis* $\{\mathbf{e}_i\}$, $i = 1, \ldots, n$, in the usual way: $\mathbf{x} = x^i \mathbf{e}_i$. Here $\mathbf{e}_i \in \mathbb{R}^n$ is the coordinate vector with an entry 1 in i-th position and with remaining $n-1$ entries equal to 0.

It is sometimes convenient to consider *(curvilinear) coordinates* on (an open subset of) $\mathbb{R}^n$. The formal definition of coordinates, generally applicable to both Euclidean as well as curved spaces, is given in the next section.

3.2. Differentiable Manifolds

Differentiable manifolds, aka *differential manifolds*, are spaces that *locally* look Euclidean. The points of such manifolds can be labelled with the help of an *atlas*, i.e. a collection of charts that provide us with coordinates on open sets that cover the entire manifold.

Definition 3.2 *An n-dimensional smooth* differentiable manifold *is a Hausdorff topological space M furnished with a family of smooth diffeomorphisms $\phi_\alpha : \Omega_\alpha \subset M \to \phi_\alpha(\Omega_\alpha) \subset \mathbb{R}^n$, with the following properties:*

- *$\{\Omega_\alpha\}$ is an open covering of M, i.e. Ω_α is open and $\cup_\alpha \Omega_\alpha = M$.*
- *If $\Omega_\alpha \cap \Omega_\beta \neq \emptyset$, then $\phi_\alpha(\Omega_\alpha \cap \Omega_\beta), \phi_\beta(\Omega_\alpha \cap \Omega_\beta) \subset \mathbb{R}^n$ are open sets, and $\phi_\alpha \circ \phi_\beta^{-1}\big|_{\phi_\beta(\Omega_\alpha \cap \Omega_\beta)}$ and (thus) $\phi_\beta \circ \phi_\alpha^{-1}\big|_{\phi_\alpha(\Omega_\alpha \cap \Omega_\beta)}$ are diffeomorphisms.*
- *The atlas $\{(\Omega_\alpha, \phi_\alpha)\}$ is maximal w.r.t. the previous two axioms.*

The last axiom ensures that any chart we may come up with is tacitly assumed to be already contained in the atlas.

To keep notation simple one often writes $x \doteq \phi_\alpha(p)$ and $y \doteq \phi_\beta(p)$ for the coordinates of a fiducial point $p = \phi_\alpha^{-1}(x) = \phi_\beta^{-1}(y)$ relative to their corresponding charts. The maps

$$\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(\Omega_\alpha \cap \Omega_\beta) \subset \mathbb{R}^n \to \phi_\beta(\Omega_\alpha \cap \Omega_\beta) \subset \mathbb{R}^n : x \mapsto y = (\phi_\beta \circ \phi_\alpha^{-1})(x) \tag{3.1}$$

$$\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(\Omega_\alpha \cap \Omega_\beta) \subset \mathbb{R}^n \to \phi_\alpha(\Omega_\alpha \cap \Omega_\beta) \subset \mathbb{R}^n : y \mapsto x = (\phi_\alpha \circ \phi_\beta^{-1})(y) \tag{3.2}$$

are called *coordinate transformations*. You may read ϕ_α as "the coordinates of ...", and ϕ_α^{-1} as "the point on the manifold with coordinates ...", both relative to the coordinate chart labelled by α. Hence $y = (\phi_\beta \circ \phi_\alpha^{-1})(x)$ may be pronounced as "the coordinates y on 'page' β of the same physical point as the one referred to by coordinates x on 'page' α in the atlas". The intersection $\Omega_\alpha \cap \Omega_\beta \subset \mathrm{M}$ is, loosely speaking, that part of the physical manifold that is 'visible' on both 'pages' α and β of the atlas.

Notation

By abuse of notation Eqs. (3.1–3.2) are often simplified as $y = y(x)$, respectively $x = x(y)$, with proper domains tacitly suppressed.

Caveat

The thing to keep in mind is that coordinates $x, y \in \mathbb{R}^n$ are subjective in the sense that they can be rather arbitrarily chosen. Eqs. (3.1–3.2) ensure that, given their associated charts $(\phi_\alpha, \phi_\beta)$, both refer to the *same* physical point $p = \phi_\alpha^{-1}(x) = \phi_\beta^{-1}(y) \in \mathrm{M}$. The consequence of this observation is that, in all that follows, we must ensure that numerical statements expressed in terms of coordinates are invariant under coordinate transformations. Here lies the connection with Chapter 2.

Definition 3.3 *The* Jacobian *of the coordinate transformation $\phi_\beta \circ \phi_\alpha^{-1}$, respectively $\phi_\alpha \circ \phi_\beta^{-1}$, is the $n{\times}n$ matrix of first order partial derivatives*

$$S^i_j(x) = \frac{\partial(\phi_\beta \circ \phi_\alpha^{-1})^i(x)}{\partial x^j} \doteq \frac{\partial y^i}{\partial x^j} \quad \textit{resp.} \quad T^i_j(y) = \frac{\partial(\phi_\alpha \circ \phi_\beta^{-1})^i(y)}{\partial y^j} \doteq \frac{\partial x^i}{\partial y^j}\,.$$

Observation

The matrices S and T are regular, with $S = T^{\text{inv}}$ by virtue of the chain rule. In simplified notation we have

$$S^i_k T^k_j = \frac{\partial y^i}{\partial x^k}\frac{\partial x^k}{\partial y^j} = \delta^i_j\,.$$

3.3. Tangent Vectors

In differential geometry, the *tangent space* takes over the role of the archetypical vector space V that lies at the core of all tensor calculus concepts introduced in Chapter 2. You may see this as a local copy of V attached to each point $\mathbf{x} \in \mathrm{M}$ of the base manifold, with $\dim V = \dim \mathrm{M} = n$. For reason of clarity we write $\mathrm{TM}_{\mathbf{x}}$ for a local copy of V attached to $\mathbf{x}$, or, by abuse of notation, TM_x if $x = \phi_\alpha(\mathbf{x}) \in \mathbb{R}^n$ in some given coordinate chart.

The tangent space is the vector space consisting of all tangent vectors at a fiducial point.

Definition 3.4 *Let $\gamma : (-\varepsilon, \varepsilon) \to M : t \mapsto \gamma(t)$ be a smooth parametrized curve, with $\gamma(0) = \mathbf{p} \in M$, and $f \in C^\infty(M)$ an arbitrary smooth scalar field. Then the* tangent vector *of γ at $\mathbf{p}$ is the linear map given by*

$$\dot\gamma(0) : C^\infty(M) \to \mathbb{R} : f \mapsto \frac{d}{dt}\,(f \circ \gamma)\Big|_{t=0}\,.$$

Smoothness of γ and f means that the numerically valued functions $\overline{\gamma} = \phi_\alpha \circ \gamma$ and $\overline{f} = f \circ \phi_\alpha^{-1}$ are smooth on the open interval $(-\varepsilon, \varepsilon) \subset \mathbb{R}$, respectively on the open neighbourhood $\phi_\alpha(U_\alpha) \subset \mathbb{R}^n$ with $\mathbf{p} \in U_\alpha$, in the sense of standard calculus, i.e. $\overline{\gamma}^i \in C^\infty((-\varepsilon, \varepsilon))$ for each component $i = 1, \ldots, n$, and $\overline{f} \in C^\infty(\phi_\alpha(U_\alpha))$.

Now recall the observation on page 8 concerning the abstract notation of basis vectors as partial derivative operators, initially justified by the formal consistency of their 'coordinate transformation properties' (chain rule) with basis transformations. On a differentiable manifold we can make this more concrete, because we now actually have something to take derivatives of.

Evaluating the expression in Definition 3.4 on a dummy scalar field with the help of a coordinate chart such that

$x = \phi_\alpha(\mathbf{x}) \in \phi_\alpha(U_\alpha) \subset \mathbb{R}^n$, we observe that

$$\dot\gamma(0)f = \frac{d}{dt}\,(f\circ\gamma)\bigg|_{t=0} = \frac{d}{dt}\,\left(\overline{f}\circ\phi_\alpha\circ\phi_\alpha^{-1}\circ\overline{\gamma}\right)\bigg|_{t=0} = \frac{d}{dt}\,\left(\overline{f}\circ\overline{\gamma}\right)\bigg|_{t=0} = \dot x^i(0)\,\frac{\partial}{\partial x^i}\bigg|_{x(0)}\overline{f} \doteq \dot x^i(0)\,\frac{\partial}{\partial x^i}\bigg|_{\gamma(0)} f\,, \tag{3.3}$$

in which $\overline{\gamma}^i(t) \doteq x^i(t)$ and $\dot{\overline{\gamma}}^i(t) \doteq \dot x^i(t)$. Here we have made use of the aforementioned coordinate prototypes

$$\overline{\gamma} : (-\varepsilon,\varepsilon) \subset \mathbb{R} \to \mathbb{R}^n : t \mapsto x(t) \doteq \overline{\gamma}(t)\,, \tag{3.4}$$

$$\overline{f} : \phi_\alpha(U_\alpha) \subset \mathbb{R}^n \to \mathbb{R} : x \mapsto \overline{f}(x) \doteq (f\circ\phi_\alpha^{-1})(x)\,, \tag{3.5}$$

and of the convenient definition

$$\frac{\partial}{\partial x^i}\bigg|_{x(0)}\overline{f} \doteq \frac{\partial}{\partial x^i}\bigg|_{\gamma(0)} f\,. \tag{3.6}$$

The dummy nature of f on left and right hand sides of Eq. (3.3) justifies the notation introduced without derivative connotation on page 8.

Notation

Recall Definition 3.4 and Eqs. (3.3–3.6). Instead of $\dot\gamma(0)$ we write

$$\mathbf{v}|_p = v^i\,\frac{\partial}{\partial x^i}\bigg|_p \tag{3.7}$$

in which $\mathbf{v}|_p = \dot\gamma(0)$ and $v^i = \dot x^i(0)$. The subscript is reminiscent of the fact that the tangent vector is defined at $\mathbf{p} = \gamma(0)$. This construction can be repeated for any point $\mathbf{x} \in \mathrm{M}$ and any smooth curve $\gamma \in C^\infty((-\varepsilon,\varepsilon),\mathrm{M})$.

Observation

Omitting explicit reference to the fiducial point $\mathbf{p}$ for the sake of simplicity, we observe from Eq. (3.3) that

$$\mathbf{v}f = df(\mathbf{v}) = \langle df, \mathbf{v}\rangle\,.$$

In particular, if we take the coordinate scalar field $f^i : \mathrm{M} \to \mathbb{R} : \mathbf{x} \mapsto x^i$ and the coordinate basis vector $\mathbf{v}_j = \partial_j \in \mathrm{TM}_\mathbf{x}$ in an arbitrary coordinate chart $x = \phi_\alpha(\mathbf{x})$, then $\mathbf{v}_j f^i = \partial_j x^i = \delta^i_j = dx^i(\partial_j) = \langle dx^i, \partial_j\rangle$, which shows that, indeed consistent with the formal notation introduced on page 8, the 1-forms dx^i are just the dual basis covectors for the coordinate vectors ∂_j.

Result 3.1 *Recall Section 2.3. Relative to a coordinate basis we have, at any implicit point* $\mathbf{p} \in M$*,*

$$\langle dx^i, \partial_j\rangle = \delta^i_j\,.$$

Remark

The base point $\mathbf{p} \in \mathrm{M}$ of interest is important. It is ususally taken for granted that $\langle\_,\_\rangle \in \mathrm{T}^*\mathrm{M}_\mathbf{p} \otimes \mathrm{TM}_\mathbf{p}$, $dx^i \in \mathrm{T}^*\mathrm{M}_\mathbf{p}$, $\partial_j \in \mathrm{TM}_\mathbf{p}$, et cetera, without explicit reference to basepoint $\mathbf{p} \in \mathrm{M}$ for simplicity of notation. If one wishes to be explicit $\mathbf{p}$ may be attached as a label to each local geometric entity as a reminder.

3.4. Tangent and Cotangent Bundle

Ansatz

An n-dimensional manifold M.

Definition 3.5 *The* tangent space *at* $\mathbf{x} \in M$*,* $\mathrm{TM}_\mathbf{x}$*, is the vector space of tangent vectors* $\mathbf{v}_\mathbf{x}$ *at point* $\mathbf{x} \in M$*.*

Remark

Recall from the previous section that the definition can be stated in intrinsic terms by defining $\mathrm{TM}_{\mathbf{x}}$ as the span of n tangent vectors of a family of mutually transversal curves intersecting at $\mathbf{x} \in \mathrm{M}$. A special case is obtained by considering the n local coordinate curves induced by the coordinate system at $\mathbf{x} \in \mathrm{M}$.

Terminology

In the latter case the basis is referred to as a *coordinate basis* or *holonomic basis*, in which case the i-th coordinate basis vector $\mathbf{e}_i$ is also often written as ∂_i, recall the observation on page 8.

Definition 3.6 *The* tangent bundle, $\mathrm{TM} = \cup_{\mathbf{x}\in M}\mathrm{T}M_{\mathbf{x}}$, *is the union of tangent spaces over all* $\mathbf{x} \in M$.

Observation

- In order to identify base points explicitly, one often incorporates the *projection map*, $\pi : \mathrm{TM} \to \mathrm{M} : (\mathbf{x}, \mathbf{v}) \mapsto \mathbf{x}$, into the definition of the tangent bundle. In particular $\pi(\mathrm{TM}_{\mathbf{x}}) = \mathbf{x}$ and $\pi(\mathrm{TM}) = \mathrm{M}$.

Notation

- Recall Definition 3.5. The *cotangent space* at $\mathbf{x} \in \mathrm{M}$ is denoted by $\mathrm{T}^*\mathrm{M}_{\mathbf{x}}$.
- Recall Definition 3.6. The *cotangent bundle* is written as $\mathrm{T}^*\mathrm{M} = \cup_{\mathbf{x}\in\mathrm{M}}\mathrm{T}^*\mathrm{M}_{\mathbf{x}}$.

3.5. Exterior Derivative

Ansatz

The tangent bundle TM over an n-dimensional manifold M.

Definition 3.7 *Let* $\omega^k = \omega_{|i_1\ldots i_k|}\, dx^{i_1}\wedge\ldots\wedge dx^{i_k} = \frac{1}{k!}\omega_{i_1\ldots i_k}\, dx^{i_1}\wedge\ldots\wedge dx^{i_k} \in \bigwedge_k(\mathrm{TM})$ *be an antisymmetric covariant tensor field of rank* k*, or* k*-form for brevity, with* $x \in M$*. The* exterior derivative *of* ω *is the* $(k+1)$*-form given by*

$$d\omega^k = \frac{1}{k!}d\omega_{i_1\ldots i_k}\wedge dx^{i_1}\wedge\ldots\wedge dx^{i_k} \doteq \frac{1}{k!}\partial_i\omega_{i_1\ldots i_k}dx^i\wedge dx^{i_1}\wedge\ldots\wedge dx^{i_k} \in \textstyle\bigwedge_{k+1}(\mathrm{TM})\,,$$

with $df \stackrel{\text{def}}{=} \partial_i f\, dx^i$ *for any sufficiently smooth scalar field* $f : M \to \mathbb{R}$.

Remark

The factor $1/k!$ has been provided so that we may interpret $\omega_{i_1\ldots i_k}$ as the holor of ω^k, since

$$\omega^k = \omega_{|i_1\ldots i_k|}\, dx^{i_1}\wedge\ldots\wedge dx^{i_k} = \frac{1}{k!}\omega_{i_1\ldots i_k}\, dx^{i_1}\wedge\ldots\wedge dx^{i_k} = \omega_{i_1\ldots i_k}\,\mathscr{A}(dx^{i_1}\otimes\ldots\otimes dx^{i_k}) = \omega_{i_1\ldots i_k}\, dx^{i_1}\otimes\ldots\otimes dx^{i_k}\,,$$

in which the last step relies on automatic antisymmetrisation due to antisymmetry of the holor. Also note that

$$\frac{1}{k!}\partial_i\omega_{i_1\ldots i_k}dx^i\wedge dx^{i_1}\wedge\ldots\wedge dx^{i_k} = \frac{(k+1)!}{k!}\partial_{|i}\omega_{i_1\ldots i_{k|}}dx^i\wedge dx^{i_1}\wedge\ldots\wedge dx^{i_k}$$

Caveat

We have loosely identified $\mathrm{M} \sim \mathbb{R}^n$ via an implicit choice of coordinates. Strictly speaking this can be done for sufficiently small neighbourhoods of any fiducial point $x \in \mathrm{M}$ and a corresponding open set of $\mathbb{R}^n$.

Observation

One readily verifies the following properties:

- $d\omega^n = 0$ for any $\omega^n \in \bigwedge_n(\mathrm{TM})$.
- $d^2\omega^k \stackrel{\text{def}}{=} dd\omega^k = 0$ for any $\omega^k \in \bigwedge_k(\mathrm{TM})$ and $k = 0, \ldots, n$, in other words: $d^2 = 0$.
- One sometimes writes $d\wedge$ instead of d.

Exterior derivatives are often used in combination with the Hodge star to extract derivatives from a k-form field in some desired format. Note that d increases the rank of a k-form by 1, whereas $*$ converts a k-form into an $(n-k)$-form. The following example illustrates this up to first order derivatives.

Example

In the following examples, $\alpha = \alpha_1 dx + \alpha_2 dy + \alpha_3 dz \in \mathrm{T}^*\mathrm{M}$, with $\alpha_i \in C^\infty(\mathrm{M})$ and $\dim \mathrm{M} = 3$. For simplicity the manifold is assumed to be Euclidean, and identified with $\mathbb{R}^3$ in Cartesian coordinates (in which the inner product defining Gram matrix equals the identity matrix). Using $*1 = dx \wedge dy \wedge dz$, $*dx = dy \wedge dz$, $*dy = -dx \wedge dz$, $*dz = dx \wedge dy$, $*(dx \wedge dy) = dz$, $*(dx \wedge dz) = -dy$, $*(dy \wedge dz) = dx$, $*(dx \wedge dy \wedge dz) = 1$, we obtain the following results:

- $*\alpha = \alpha_3 dx \wedge dy - \alpha_2 dx \wedge dz + \alpha_1 dy \wedge dz$;
- $d\alpha = (\partial_x\alpha_2 - \partial_y\alpha_1)\, dx \wedge dy + (\partial_x\alpha_3 - \partial_z\alpha_1)\, dx \wedge dz + (\partial_y\alpha_3 - \partial_z\alpha_2)\, dy \wedge dz$;
- $*d\alpha = (\partial_x\alpha_2 - \partial_y\alpha_1)\, dz - (\partial_x\alpha_3 - \partial_z\alpha_1)\, dy + (\partial_y\alpha_3 - \partial_z\alpha_2)\, dx$;
- $d * \alpha = (\partial_x\alpha_1 + \partial_y\alpha_2 + \partial_z\alpha_3)\, dx \wedge dy \wedge dz$;
- $*d * \alpha = \partial_x\alpha_1 + \partial_y\alpha_2 + \partial_z\alpha_3$;
- et cetera.

Special cases arise if $\alpha = df$ for some $f \in C^\infty(\mathrm{M})$, a so-called "exact 1-form".

3.6. Affine Connection

Ansatz

The tangent bundle TM over an n-dimensional manifold M.

Recall the observation on page 8 concerning the formal notation for vector decomposition relative to a coordinate basis, $\mathbf{v} = v^i\partial_i$. Also recall from infinitesimal calculus that for a scalar function $f \in C^\infty(\mathrm{M})$ we may define the differential $df = \partial_i f dx^i \in \mathrm{T}^*\mathrm{M}$. This suggests that we interpret a vector as a *directional derivative operator*, as follows.

Definition 3.8 *Let $f \in C^\infty(M)$ be a scalar function, and $\mathbf{v} = v^i\partial_i \in \mathrm{TM}$. Then*

$$\mathbf{v}f \stackrel{\text{def}}{=} \langle df, \mathbf{v}\rangle \stackrel{\text{def}}{=} v^i\partial_i f\,.$$

So, when acting on a scalar field it is tacitly understood that we make the formal identication $\mathbf{v} = \langle d\cdot, \mathbf{v}\rangle$.

Definition 3.9 *An* affine connection *on a manifold M is a mapping $\nabla : \mathrm{TM} \times \mathrm{TM} \to \mathrm{TM} : (\mathbf{v}, \mathbf{w}) \mapsto \nabla_{\mathbf{v}}\mathbf{w}$, also referred to as the* covariant derivative *of $\mathbf{w}$ with respect to $\mathbf{v}$, with the following properties. Let $f, f_1, f_2 \in C^\infty(M)$, $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{w}, \mathbf{w}_1, \mathbf{w}_2 \in \mathrm{TM}$, and $\lambda, \lambda_1, \lambda_2 \in \mathbb{R}$, then*

1. $\nabla_{f_1\mathbf{v}_1+f_2\mathbf{v}_2}\mathbf{w} = f_1\nabla_{\mathbf{v}_1}\mathbf{w} + f_2\nabla_{\mathbf{v}_2}\mathbf{w}$ *for all* $f_1, f_2 \in C^\infty(M)$,
2. $\nabla_{\mathbf{v}}(\lambda_1\mathbf{w}_1 + \lambda_2\mathbf{w}_2) = \lambda_1\nabla_{\mathbf{v}}\mathbf{w}_1 + \lambda_2\nabla_{\mathbf{v}}\mathbf{w}_2$, *for all* $\lambda_1, \lambda_2 \in \mathbb{R}$,
3. $\nabla_{\mathbf{v}}(f\mathbf{w}) = f\nabla_{\mathbf{v}}\mathbf{w} + \nabla_{\mathbf{v}}f\,\mathbf{w}$ *for all* $f \in C^\infty(M)$. *Here we have defined* $\nabla_{\mathbf{v}}f \equiv \mathbf{v}f$.

Remark

If we leave out the directional vector argument we call $\nabla\mathbf{w}$ the covariant differential of $\mathbf{w}$, and identify $\nabla f = df$.

Thus the covariant derivative of a local frame vector (i.e. basis vector of the local tangent space) is a vector within the same tangent space. Consequently it can be written as a linear combination of frame vectors.

Definition 3.10 *The* Christoffel symbols Γ^k_{ij} *are defined by* $\nabla_{\mathbf{e}_j}\mathbf{e}_i = \Gamma^k_{ij}\mathbf{e}_k$.

Caveat

1. Conventions for the ordering of the lower indices of the Christoffel symbols differ among literature sources. As a rule of thumb I put the differentiation index in second position.
2. The Christoffel symbols are *not* the components of a tensor, v.i.

To see how the Christoffel symbols transform, consider a vector and dual covector basis transformation:

$$\overline{\mathbf{e}}^k = S^k_\ell\mathbf{e}^\ell \quad \text{resp.} \quad \overline{\mathbf{e}}_j = T^m_j\mathbf{e}_m\,.$$

The matrices $\mathbf{T}$ and $\mathbf{S}$ are the mutually inverse Jacobian matrices:

$$T^i_j = \frac{\partial x^i}{\partial \overline{x}^j} \quad \text{resp.} \quad S^i_j = \frac{\partial \overline{x}^i}{\partial x^j}\,.$$

Set

$$\overline{\Gamma}^k_{ij} = \left\langle \overline{\mathbf{e}}^k, \nabla_{\overline{\mathbf{e}}_j}\overline{\mathbf{e}}_i \right\rangle$$

Substitute the transformations, apply first and third rule of Definition 3.9, $\nabla_{\overline{\mathbf{e}}_j} = T^\ell_j\nabla_{\mathbf{e}_\ell}$, respectively $\nabla_{\mathbf{e}_\ell}(T^\ell_i\mathbf{e}_k) = \partial_\ell T^k_i\mathbf{e}_k + T^\ell_i\nabla_{\mathbf{e}_\ell}\mathbf{e}_k$, use the definition of the Christoffel symbols, $\nabla_{\mathbf{e}_\ell}\mathbf{e}_k = \Gamma^m_{k\ell}\mathbf{e}_m$, and finally the chain rule, $\overline{\partial}_j = T^\ell_j\partial_\ell$, to obtain the following result:

Result 3.2 *The components of the Christoffel symbols in coordinate system* $\overline{\mathbf{x}}$ *are given relative to those in coordinate system* $\mathbf{x}$ *by*

$$\overline{\Gamma}^k_{ij} = S^k_\ell\left(T^m_jT^n_i\Gamma^\ell_{nm} + \overline{\partial}_jT^\ell_i\right)\,.$$

Thus the tensor transformation law, recall Result 2.8 on page 37, fails to hold due to the inhomogeneous term, which arises after any non-affine coordinate transformation.

Remark

For any fixed point on the manifold you can solve the system of equations $\overline{\Gamma}^k_{ij} = 0$ by slick choice of Jacobian matrix, i.e. by transforming to a suitable local coordinate system. However, in general it is not possible to achieve this globally, unless the manifold happens to be flat (zero Riemann curvature, v.i.).

Using the above definitions one can easily prove the following.

Result 3.3 *Let* $\mathbf{v} = v^i \mathbf{e}_i$, $\mathbf{w} = w^i \mathbf{e}_i \in \mathrm{TM}$ *be two vector fields. Then*

$$\nabla_{\mathbf{v}} \mathbf{w} = v^i D_i w^j \mathbf{e}_j \,.$$

Here we have defined the components of the covariant derivative $D_i w^j \equiv \partial_i w^j + \Gamma^j_{ki} w^k$.

Observation

- The meaning of D_i applied to a holor is *rank-dependent*; the above definition applies to a holor with only one, contravariant index.
- Moreover, $D_i w^j$ does not only depend on the jth component of $\mathbf{w}$, but on *all* components w^k, $k = 1, \ldots, n$.

Definition 3.11 *See Definition 3.9. The affine connection can be extended to arbitrary tensor fields by generalizing the product rule, viz. for any pair of tensor fields,* $\mathbf{T} \in \mathbf{T}^p_q(\mathrm{TM})$ *and* $\mathbf{S} \in \mathbf{T}^r_s(\mathrm{TM})$*, say, we require*

$$\nabla_{\mathbf{v}} \left(\boldsymbol{S} \otimes \boldsymbol{T} \right) = \nabla_{\mathbf{v}} \boldsymbol{S} \otimes \boldsymbol{T} + \boldsymbol{S} \otimes \nabla_{\mathbf{v}} \boldsymbol{T} \,.$$

Observation

Recall the linear trace operator from Definition 2.27. For a covector field $\hat{\mathbf{z}} \in \mathrm{T}^*\mathrm{M}$ and a vector field $\mathbf{w} \in \mathrm{TM}$ we have $\langle \hat{\mathbf{z}}, \mathbf{w} \rangle = \mathrm{tr}\,(\hat{\mathbf{z}} \otimes \mathbf{v})$. Observe that

$$\nabla_{\mathbf{v}} \langle \hat{\mathbf{z}}, \mathbf{w} \rangle = \nabla_{\mathbf{v}} (\mathrm{tr}\,(\hat{\mathbf{z}} \otimes \mathbf{v})) = \mathrm{tr}\,(\nabla_{\mathbf{v}} (\hat{\mathbf{z}} \otimes \mathbf{v})) = \mathrm{tr}\,(\nabla_{\mathbf{v}} \hat{\mathbf{z}} \otimes \mathbf{v}) + \mathrm{tr}\,(\hat{\mathbf{z}} \otimes \nabla_{\mathbf{v}} \mathbf{v}) = \langle \nabla_{\mathbf{v}} \hat{\mathbf{z}}, \mathbf{w} \rangle + \langle \hat{\mathbf{z}}, \nabla_{\mathbf{v}} \mathbf{w} \rangle \,,$$

implying the product rule $D_i (z_j w^j) = D_i z_j \, w^j + z_j D_i \, w^j$.

Result 3.4 *See Result 3.3. Let* $\mathbf{T} = T^{i_1 \ldots i_p}_{j_1 \ldots j_q} \mathbf{e}_{i_1} \otimes \ldots \otimes \mathbf{e}_{i_p} \otimes \hat{\mathbf{e}}^{j_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{j_q}$ *and* $\mathbf{v} = v^i \mathbf{e}_i$ *be a tensor, respectively vector field on M, then*

$$\nabla_{\mathbf{v}} \mathbf{T} = v^i D_i T^{i_1 \ldots i_p}_{j_1 \ldots j_q} \mathbf{e}_{i_1} \otimes \ldots \otimes \mathbf{e}_{i_p} \otimes \hat{\mathbf{e}}^{j_1} \otimes \ldots \otimes \hat{\mathbf{e}}^{j_q} \,.$$

Here we have defined the components of the covariant derivative

$$D_i T^{i_1 \ldots i_p}_{j_1 \ldots j_q} \equiv \partial_i T^{i_1 \ldots i_p}_{j_1 \ldots j_q} + \Gamma^{i_1}_{ki} T^{k i_2 \ldots i_p}_{j_1 \ldots j_q} + \ldots + \Gamma^{i_p}_{ki} T^{i_1 \ldots i_{p-1} k}_{j_1 \ldots j_q} - \Gamma^k_{j_1 i} T^{i_1 \ldots i_p}_{k j_2 \ldots j_q} - \ldots - \Gamma^k_{j_q i} T^{i_1 \ldots i_p}_{j_1 \ldots j_{q-1} k} \,.$$

With the latter definition the covariant derivative operator D_i satisfies the product rule for general holor products. Note that D_i is holor-type dependent, involving a 'Γ-correction term' for each index.

Terminology

- With *affine geometry* one indicates the geometrical structure of a manifold endowed with an affine connection. Note that this does *not* require the existence of a metric. Without a Riemannian metric, M is referred to as an *affine space* or *affine manifold*.
- If a metric (induced by an inner product on each TM_x) does exist we are in the realm of *Riemannian geometry*. Endowed with a Riemannian metric, M is called a *Riemannian space* or *Riemannian manifold*. We refer to section 3.9 for further details.

Assumption

- Recall Definition 2.8, page 12. It is assumed that, if $\mathrm{TM}_{\mathbf{x}}$ is an inner product space for every $\mathbf{x} \in \mathrm{M}$, local bases are chosen in such a way that the coefficients of the Gram matrix, $g_{ij}(\mathbf{x}) = (\mathbf{e}_i | \mathbf{e}_j)_{\mathbf{x}}$, vary smoothly across the manifold: $g_{ij} \in C^\infty(\mathrm{M})$. (This is guaranteed if the basis is holonomic.)
- Similar assumptions will henceforth be implicit in the case of tensor fields in general.

Notation

The $\mathbf{x}$-dependency of tensor fields is tacitly understood and suppressed in the notation.

3.7. Lie Derivative

Ansatz

The tangent bundle TM over an n-dimensional manifold M.

Definition 3.12 *The* Lie derivative *of* $\mathbf{w} \in V$ *with respect to* $\mathbf{v} \in V$ *is defined as*

$$\mathscr{L}_{\mathbf{v}}\mathbf{w} = [\mathbf{v}, \mathbf{w}] \ ,$$

in which the so-called Lie bracket *is defined by the commutator*

$$[\mathbf{v}, \mathbf{w}]\, f = \mathbf{v}(\mathbf{w}f) - \mathbf{w}(\mathbf{v}f) = \left(v^i \partial_i w^j - w^i \partial_i v^j\right) \partial_j f \ .$$

for any smooth function f.

Observation

- $\mathscr{L}_{\mathbf{v}}\mathbf{w}$ is itself a vector, with components $[\mathbf{v}, \mathbf{w}]^i = v^j \partial_j w^i - w^j \partial_j v^i$.
- No connection is involved, since only derivations of $\mathbb{R}$-valued functions are considered.
- Each term separately in the subtraction does *not* constitute a vector (it fails to satisfy the covariance requirement), but the difference *does*.
- The Lie-bracket is antisymmetric.
- The Lie-bracket satisfies the *Jacobi identity*: $[\mathbf{u}, [\mathbf{v}, \mathbf{w}]] + [\mathbf{w}, [\mathbf{u}, \mathbf{v}]] + [\mathbf{v}, [\mathbf{w}, \mathbf{u}]] = 0$.

3.8. Torsion

Ansatz

- The tangent bundle TM over an n-dimensional manifold M.
- An affine connection.

Definition 3.13 *Notation as before. The* torsion tensor *is given by*

$$\mathbf{T}(\mathbf{v}, \mathbf{w}) = \nabla_{\mathbf{v}}\mathbf{w} - \nabla_{\mathbf{w}}\mathbf{v} - [\mathbf{v}, \mathbf{w}] \ .$$

Remark

The torsion tensor thus reveals the difference between a commutator of ordinary derivatives, as in the Lie derivative, and of covariant derivatives induced by the affine connection.

Observation

Recall that $\nabla_{\mathbf{v}}\mathbf{w} = v^i(\partial_i w^j + w^k\Gamma^j_{ki})\partial_j$, from which you derive the components of the torsion tensor relative to a coordinate basis:

$$\mathbf{T}(\mathbf{v},\mathbf{w}) = T^j_{ki}v^k w^i \partial_j \quad \text{with} \quad T^j_{ki} = \Gamma^j_{ik} - \Gamma^j_{ki}\,.$$

Observation

The torsion tensor is antisymmetric: $\mathbf{T}(\mathbf{v},\mathbf{w}) = -\mathbf{T}(\mathbf{w},\mathbf{v})$.

3.9. Levi-Civita Connection

Ansatz

- The tangent bundle TM over an n-dimensional manifold M.
- A Riemannian metric $\mathbf{g} \in \mathrm{T}^*\mathrm{M} \otimes \mathrm{T}^*\mathrm{M}$.

An important result is that a Riemannian geometry induces a unique torsion-free connection.

Definition 3.14 *Recall Definition 3.12, page 55 and Definition 3.13, page 55. A* metric connection *or* Levi-Civita connection *on a Riemannian manifold* $(M, \mathbf{g})$ *satisfies the following properties:*

- $\mathbf{z}\,(\mathbf{g}(\mathbf{v},\mathbf{w})) = \mathbf{g}(\nabla_{\mathbf{z}}\mathbf{v},\mathbf{w}) + \mathbf{g}(\mathbf{v},\nabla_{\mathbf{z}}\mathbf{w})$ *for all* $\mathbf{v},\mathbf{w},\mathbf{z} \in \mathrm{TM}$.
- $\mathbf{T}(\mathbf{v},\mathbf{w}) = 0$ *for all* $\mathbf{v},\mathbf{w} \in \mathrm{TM}$.

Observation

The first requirement is tantamount to the *product rule*: $\nabla_{\mathbf{z}}\,(\mathbf{v}\cdot\mathbf{w}) = \nabla_{\mathbf{z}}\mathbf{v}\cdot\mathbf{w} + \mathbf{v}\cdot\nabla_{\mathbf{z}}\mathbf{w}$ for all $\mathbf{v},\mathbf{w},\mathbf{z} \in \mathrm{TM}$. Alternatively, in terms of the $\sharp$-operator, $\nabla \circ \sharp = \sharp \circ \nabla$, recall Definition 2.9 on page 14.

Terminology

- The product rule is variously phrased as "covariant constancy of metric", "preservation of metric", "metric compatibility", et cetera, and can be condensely written as $\nabla_{\mathbf{z}}\mathbf{g} = 0$ for all $\mathbf{z} \in \mathrm{TM}$.
- The second requirement is referred to as the requirement of "zero torsion".

Theorem 3.1 *Recall Definition 3.14. The Levi-Civita connection is uniquely given by the Christoffel symbols*

$$\Gamma^k_{ij} = \frac{1}{2}g^{k\ell}\left(\partial_i g_{\ell j} + \partial_j g_{\ell i} - \partial_\ell g_{ij}\right)\,.$$

Proof of Theorem 3.1. The proof is based on (i) the general properties of an affine connection, recall Definition 3.9 on page 52, and (ii) the defining properties of the Levi-Civita connection, Definition 3.14, page 56.

Here are the details. Show that relative to a basis the first condition of Definition 3.14 is equivalent to $D_m g_{ij} = 0$, then cyclicly permute the three free indices; subtract the two resulting expressions from the given one. The result is

$$\partial_m g_{ij} - \partial_i g_{jm} - \partial_j g_{mi} = g_{i\ell}\left(\Gamma^\ell_{jm} - \Gamma^\ell_{mj}\right) + g_{\ell j}\left(\Gamma^\ell_{im} - \Gamma^\ell_{mi}\right) - g_{\ell m}\left(\Gamma^\ell_{ij} + \Gamma^\ell_{ji}\right)\,.$$

Subsequently, according to Definition 3.13 on page 55 and the subsequent observation on page 56, the second condition of Definition 3.14 is equivalent to

$$\Gamma^{\ell}_{ij} = \Gamma^{\ell}_{ji} \,.$$

Apply this to the previous equality. Contraction with $-\frac{1}{2}g^{km}$ on both sides, using the fact that $g^{ik}g_{kj} = \delta^i_j$, finally yields the stated result.

Observation

A metric connection is *symmetric*:

$$\Gamma^{k}_{ij} = \Gamma^{k}_{ji} \,.$$

Recall that this is not necessarily the case for an affine connection in general.

3.10. Geodesics

Definition 3.15 *The* covariant differential $\nabla\mathbf{v} \in \mathbf{T}^1_1(\mathrm{TM})$ *of a vector field* $\mathbf{v} = v^i\partial_i \in \mathrm{TM}$ *is defined as*

$$\nabla\mathbf{v} = \nabla_{\partial_k}\mathbf{v}dx^k \,.$$

Observation

If we define $\nabla\mathbf{v} = Dv^j\partial_j$, then $Dv^j = dv^j + \Gamma^j_{ki}v^k dx^i = D_i v^j dx^i$, with $D_i v^j = \partial_i v^j + \Gamma^j_{ki}v^k$.

Remark

The definition can be extended to arbitrary tensor fields $\mathbf{T} \in \mathbf{T}^p_q(\mathrm{TM})$:

$$\nabla\mathbf{T} = DT^{i_1\ldots i_p}_{j_1\ldots j_q} dx^{j_1} \otimes \ldots \otimes dx^{j_q} \otimes \partial_{i_1} \otimes \ldots \otimes \partial_{i_p} \,,$$

with $DT^{i_1\ldots i_p}_{j_1\ldots j_q} = D_i T^{i_1\ldots i_p}_{j_1\ldots j_q} dx^i$, recall Result 3.4.

Heuristics

The conceptual advantage of a (covariant) differential over a (covariant) derivative is that the former is applicable to vectors known only along a curve, whereas the latter requires us to specify a vector field on a full neighbourhood, as the following definition illustrates.

Definition 3.16 *A* geodesic *is a curve with "covariantly constant tangent", or a curve obtained by "parallel transport" of its tangent vector at any fiducial point. In other words, if* $\mathbf{v}(t) = \dot{x}^i(t)\partial_i$ *for* $t \in I = (t_0, t_1)$*, then*

$$\nabla\mathbf{v}(t) = 0 \quad \textit{for all } t \in I.$$

Remark

- It is tacitly understood that $\partial_i \in \mathrm{TM}_{\mathbf{x}(t)}$ indicates the coordinate basis vector *at point* $\mathbf{x}(t)$.
- One often identifies $\mathbf{v}(\mathbf{x}(t))$ with $\mathbf{v}(t)$. It should be clear from the context which of these prototypes is applicable.

Observation

In terms of components we have $Dv^k(t)/dt = 0$, or $\ddot{x}^k + \Gamma^k_{ij}\dot{x}^i\dot{x}^j = 0$.

Terminology

The latter second order ordinary differential equation is known as the *geodesic equation*.

Heuristics

- A geodesic generalizes the concept of a straight line, and can in general be seen as the "straightest line" between any two sufficiently close points on the geodesic.
- This should not be confused with the "shortest line", for this would require the additional notion of a "length" or "distance" on the manifold (i.e. a Riemannian metric).

3.11. Curvature

Definition 3.17 *For* $\mathbf{v}, \mathbf{w} \in \mathrm{TM}$ *the (antisymmetric)* curvature operator, $\mathscr{R}(\mathbf{v}, \mathbf{w}) : \mathrm{TM} \to \mathrm{TM} : \mathbf{u} \mapsto \mathscr{R}(\mathbf{v}, \mathbf{w})\mathbf{u}$, *is defined as*

$$\mathscr{R}(\mathbf{v}, \mathbf{w}) = [\nabla_{\mathbf{v}}, \nabla_{\mathbf{w}}] - \nabla_{[\mathbf{v},\mathbf{w}]} \, .$$

Observation

Alternatively, in order to obtain a scalar valued mapping, we would need a covector field $\widehat{\mathbf{z}} \in \mathrm{T}^*\mathrm{M}$ besides the three vector field arguments $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathrm{TM}$ involved in Definition 3.17. This naturally leads to the definition of a fourth order mixed tensor field.

Definition 3.18 *For* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathrm{TM}$, $\widehat{\mathbf{z}} \in \mathrm{T}^*\mathrm{M}$, *the* Riemann curvature tensor $\mathbf{Riemann} \in \mathrm{T}^*\mathrm{M} \otimes \mathrm{TM} \otimes \mathrm{TM} \otimes \mathrm{TM}$ *is defined as*

$$\mathbf{Riemann}(\widehat{\mathbf{z}}, \mathbf{u}, \mathbf{v}, \mathbf{w}) = \langle \widehat{\mathbf{z}} \, , \, \mathscr{R}(\mathbf{v}, \mathbf{w})\mathbf{u} \rangle \, .$$

Observation

- One should first verify that Definition 3.17 does *not* require any derivatives of $\mathbf{u}$!
- In order to find the components of the Riemann curvature tensor, write $\widehat{\mathbf{z}} = z_\rho dx^\rho$, $\mathbf{u} = u^\sigma \partial_\sigma$, $\mathbf{v} = v^\mu \partial_\mu$, $\mathbf{w} = w^\nu \partial_\nu$. Then, using previous result, you find $\mathbf{R}(\widehat{\mathbf{z}}, \mathbf{u}, \mathbf{v}, \mathbf{w}) = R^\rho_{\sigma\mu\nu} z_\rho u^\sigma v^\mu w^\nu$, in which

$$R^\rho_{\sigma\mu\nu} = \langle dx^\rho \, , \, \mathscr{R}(\partial_\mu, \partial_\nu)\partial_\sigma \rangle = \partial_\mu \Gamma^\rho_{\sigma\nu} - \partial_\nu \Gamma^\rho_{\sigma\mu} + \Gamma^\rho_{\lambda\mu}\Gamma^\lambda_{\sigma\nu} - \Gamma^\rho_{\lambda\nu}\Gamma^\lambda_{\sigma\mu} \, .$$

- Instead of $R^\rho_{\sigma\mu\nu}$ the holor is often written as $R^\rho{}_{\sigma\mu\nu}$ in order to stress that the first slot of the tensor is intended for the covector argument.
- Note that $\mathscr{R}(\partial_\mu, \partial_\nu) = [\nabla_{\partial_\mu}, \nabla_{\partial_\nu}] - \nabla_{[\partial_\mu, \partial_\nu]} = [\nabla_{\partial_\mu}, \nabla_{\partial_\nu}]$, i.e. the curvature operator and, a fortiori, the Riemann curvature tensor apparently capture the intrinsic non-commutativity of covariant derivatives.
- Thus the 'second order' nature of the curvature operator pertains to the geometry of the manifold, *not* the differential structure of the vector field it operates on.
- In the case of a Riemannian manifold (with the induced Levi-Civita connection) this is clear from the appearance of (up to) second order partial derivatives of the Gram matrix $g_{\alpha\beta}$ in $R^\rho_{\sigma\mu\nu}$.

Definition 3.19 *The* Ricci curvature tensor *and the* Ricci curvature scalar *are defined as*

$$\begin{aligned} \mathbf{Ric}(\mathbf{v},\mathbf{w}) &= R_{\mu\nu}v^\mu w^\nu\,, \\ \mathbf{R} &= g^{\mu\nu}R_{\mu\nu}\,, \end{aligned}$$

with $\mathbf{v},\mathbf{w}\in\mathrm{TM}$*, in which the Ricci holor is defined as* $R_{\mu\nu}=R^\rho_{\mu\rho\nu}$.

3.12. Push-Forward and Pull-Back

Ansatz

- A smooth mapping $\phi:\mathrm{M}\to\mathrm{N}:x\mapsto y=\phi(x)$ between manifolds.
- A smooth scalar field $f:\mathrm{N}\to\mathbb{R}:y\mapsto z=f(y)$.
- A smooth vector field $\mathbf{v}\in\mathrm{TM}$.
- A smooth covector field $\boldsymbol{\nu}\in\mathrm{T}^*\mathrm{N}$.

Definition 3.20

- *The pull-back* $\phi^*f\in C^\infty(M)$ *of* $f\in C^\infty(N)$ *under* ϕ *is defined such that* $\phi^*f=f\circ\phi$.
- *The push-forward* $\phi_*\mathbf{v}\in\mathrm{TN}$ *of* $\mathbf{v}\in\mathrm{TM}$ *under* ϕ *is defined such that* $\phi_*\mathbf{v}(f)=\mathbf{v}(\phi^*f)$.
- *The pull-back* $\phi^*\boldsymbol{\nu}\in\mathrm{T}^*\mathrm{M}$ *of* $\boldsymbol{\nu}\in\mathrm{T}^*\mathrm{N}$ *under* ϕ *is defined such that* $\langle\phi^*\boldsymbol{\nu},\mathbf{v}\rangle_\mathrm{M}=\langle\boldsymbol{\nu},\phi_*\mathbf{v}\rangle_\mathrm{N}$.

Terminology

The terminology reflects the "direction" of the mappings involved between (tangent or cotangent bundles over) the respective manifolds, cf. the diagram (3.8).

Remark

- Note that the dimensions of the manifolds M and N need not be the same.

Observation

- By decomposing $\phi_*\mathbf{v}=v^\mu S^a_\mu\frac{\partial}{\partial y_a}\in\mathrm{TN}$ it follows that the components of the push-forward are given by the Jacobian matrix of the mapping: $S^a_\mu=\dfrac{\partial y^a}{\partial x^\mu}$. (Note: $\mu=1,\ldots,m=\dim\mathrm{M}$, $a=1,\ldots,n=\dim\mathrm{N}$.)
- By the same token, decomposing $\phi^*\boldsymbol{\nu}=\nu_aS^a_\mu dx^\mu\in\mathrm{T}^*\mathrm{M}$, it follows that the components of the pull-back are given by the same Jacobian.

$$\begin{tikzcd}
\mathrm{TM} \arrow[r, "\phi_*"] \arrow[d, "\pi"'] & \mathrm{TN} \arrow[d, "\pi"] \\
\mathrm{M} \arrow[r, "\phi"'] & \mathrm{N} \\
\mathrm{T^*M} \arrow[u, "\pi"] & \mathrm{T^*N} \arrow[l, "\phi^*"'] \arrow[u, "\pi"']
\end{tikzcd} \tag{3.8}$$

3.13. Examples

3.13.1. Polar Coordinates in the Euclidean Plane

Example

Starting out from a Cartesian coordinate basis at an arbitrary point of the Euclidean plane E (i.e. Euclidean 2-space),

$$\mathbf{e}_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

we obtain the local polar coordinate basis by the following transformation:

$$\overline{\mathbf{e}}_i = \overline{\partial}_i = \frac{\partial x^j}{\partial \overline{x}^i} \partial_j = \frac{\partial x^j}{\partial \overline{x}^i} \mathbf{e}_j \,.$$

Here we identify $\mathbf{x} = (x, y)$, $\overline{\mathbf{x}} = (r, \phi)$, with

$$\begin{cases} x(r,\phi) &= r\cos\phi \\ y(r,\phi) &= r\sin\phi\,. \end{cases}$$

Recall that the Jacobian matrix and its inverse are given by

$$T_i^\ell = \frac{\partial x^\ell}{\partial \overline{x}^i} \quad \text{resp.} \quad S_\ell^k = \frac{\partial \overline{x}^k}{\partial x^\ell}\,,$$

or, in explicit matrix notation (upper/lower index = row/column index):

$$\begin{aligned} T_i^\ell &= \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix} = \begin{pmatrix} \dfrac{x}{\sqrt{x^2+y^2}} & -y \\ \dfrac{y}{\sqrt{x^2+y^2}} & x \end{pmatrix}, \\ S_\ell^k &= \begin{pmatrix} \cos\phi & \sin\phi \\ -\dfrac{\sin\phi}{r} & \dfrac{\cos\phi}{r} \end{pmatrix} = \begin{pmatrix} \dfrac{x}{\sqrt{x^2+y^2}} & \dfrac{y}{\sqrt{x^2+y^2}} \\ -\dfrac{y}{x^2+y^2} & \dfrac{x}{x^2+y^2} \end{pmatrix}. \end{aligned}$$

Consequently, writing "r" and "ϕ" for index values 1 and 2, respectively,

$$\begin{aligned} \overline{\mathbf{e}}_r &= \begin{pmatrix} \cos\phi & \sin\phi \end{pmatrix} \begin{pmatrix} \mathbf{e}_x \\ \mathbf{e}_y \end{pmatrix} = \cos\phi\,\mathbf{e}_x + \sin\phi\,\mathbf{e}_y = \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}, \\ \overline{\mathbf{e}}_\phi &= \begin{pmatrix} -r\sin\phi & r\cos\phi \end{pmatrix} \begin{pmatrix} \mathbf{e}_x \\ \mathbf{e}_y \end{pmatrix} = -r\sin\phi\,\mathbf{e}_x + r\cos\phi\,\mathbf{e}_y = \begin{pmatrix} -r\sin\phi \\ r\cos\phi \end{pmatrix}. \end{aligned}$$

Note that the polar basis depends on the base point of the tangent plane, whereas the Cartesian basis is identical in form at all base points.

Example

Cartesian coordinates $\mathbf{x} = (x, y)$ in the Euclidean plane E induce a local basis for each $\mathrm{TE}_{\mathbf{x}}$ relative to which all Γ^{ℓ}_{nm} vanish identically ($\ell, m, n = 1, \ldots, 2$). In terms of polar coordinates, $\overline{\mathbf{x}} = (r, \phi)$, we find, according to Result 3.2 on page 53:

$$\overline{\Gamma}^{k}_{ij} = S^{k}_{\ell}\overline{\partial}_{j}T^{\ell}_{i}\,.$$

From the previous example it follows that for any (row) index $k = 1, 2$ and (column) index $i = 1, 2$,

$$\begin{aligned}
\overline{\Gamma}^{k}_{ir} &= \begin{pmatrix} \cos\phi & \sin\phi \\ -\dfrac{\sin\phi}{r} & \dfrac{\cos\phi}{r} \end{pmatrix} \frac{\partial}{\partial r} \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \dfrac{1}{r} \end{pmatrix}, \\
\overline{\Gamma}^{k}_{i\phi} &= \begin{pmatrix} \cos\phi & \sin\phi \\ -\dfrac{\sin\phi}{r} & \dfrac{\cos\phi}{r} \end{pmatrix} \frac{\partial}{\partial \phi} \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix} = \begin{pmatrix} 0 & -r \\ \dfrac{1}{r} & 0 \end{pmatrix}.
\end{aligned}$$

In other words,

$$\overline{\Gamma}^{\phi}_{\phi r} = \frac{1}{r}\,, \quad \overline{\Gamma}^{\phi}_{r\phi} = \frac{1}{r}\,, \quad \overline{\Gamma}^{r}_{\phi\phi} = -r\,, \quad \text{all other } \overline{\Gamma}^{k}_{ij} = 0.$$

Consequently,

$$\nabla_{r}\overline{\mathbf{e}}_{\phi} = \frac{1}{r}\overline{\mathbf{e}}_{\phi}\,, \quad \nabla_{\phi}\overline{\mathbf{e}}_{r} = \frac{1}{r}\overline{\mathbf{e}}_{\phi}\,, \quad \nabla_{\phi}\overline{\mathbf{e}}_{\phi} = -r\overline{\mathbf{e}}_{r}\,, \quad \text{all other } \nabla_{\overline{\mathbf{e}}_{j}}\overline{\mathbf{e}}_{i} = 0.$$

This can be immediately confirmed by inspection of the (partial derivatives of the) coordinates of the polar basis vectors relative to the standard (Cartesian) basis, recall the previous example.

3.13.2. A Helicoidal Extension of the Euclidean Plane

Example

Consider two coordinatizations of 3-space, $\mathbf{x} = (x, y, \phi)$, and $\overline{\mathbf{x}} = (\xi, \eta, \theta)$, related as follows[1]:

$$\left\{\begin{array}{lcl} x(\xi,\eta,\theta) & = & \xi\cos\theta + \eta\sin\theta \\ y(\xi,\eta,\theta) & = & -\xi\sin\theta + \eta\cos\theta \\ \phi(\xi,\eta,\theta) & = & \theta\,, \end{array}\right. \quad \text{or} \quad \left\{\begin{array}{lcl} \xi(x,y,\phi) & = & x\cos\phi - y\sin\phi \\ \eta(x,y,\phi) & = & x\sin\phi + y\cos\phi \\ \theta(x,y,\phi) & = & \phi\,. \end{array}\right.$$

The Jacobian matrices are given by

$$T_i^\ell = \frac{\partial x^\ell}{\partial \overline{x}^i} = \begin{pmatrix} \cos\theta & \sin\theta & -\xi\sin\theta + \eta\cos\theta \\ -\sin\theta & \cos\theta & -\xi\cos\theta - \eta\sin\theta \\ 0 & 0 & 1 \end{pmatrix},$$

respectively

$$S_\ell^k = \frac{\partial \overline{x}^k}{\partial x^\ell} = \begin{pmatrix} \cos\phi & -\sin\phi & -x\sin\phi - y\cos\phi \\ \sin\phi & \cos\phi & x\cos\phi - y\sin\phi \\ 0 & 0 & 1 \end{pmatrix}.$$

Again, recalling Result 3.2 on page 53: $\overline{\Gamma}_{ij}^k = S_\ell^k \left(T_j^m T_i^n \Gamma_{nm}^\ell + \overline{\partial}_j T_i^\ell\right)$. There are different ways to proceed, depending on our assumptions about the Γ_{nm}^ℓ-symbols in (x, y, ϕ)-space. Here we pursue the simplest option, ignoring the remark in footnote 1, and assume $\Gamma_{nm}^\ell = 0$, i.e. (x, y, ϕ) is identified with a Cartesian coordinate triple in Euclidean 3-space. In this case we may evaluate the Christoffel symbols as in the previous example:

$$\begin{aligned}
\overline{\Gamma}_{i\xi}^k &= \begin{pmatrix} \cos\phi & -\sin\phi & -x\sin\phi - y\cos\phi \\ \sin\phi & \cos\phi & x\cos\phi - y\sin\phi \\ 0 & 0 & 1 \end{pmatrix} \frac{\partial}{\partial\xi} \begin{pmatrix} \cos\theta & \sin\theta & -\xi\sin\theta + \eta\cos\theta \\ -\sin\theta & \cos\theta & -\xi\cos\theta - \eta\sin\theta \\ 0 & 0 & 1 \end{pmatrix}, \\
\overline{\Gamma}_{i\eta}^k &= \begin{pmatrix} \cos\phi & -\sin\phi & -x\sin\phi - y\cos\phi \\ \sin\phi & \cos\phi & x\cos\phi - y\sin\phi \\ 0 & 0 & 1 \end{pmatrix} \frac{\partial}{\partial\eta} \begin{pmatrix} \cos\theta & \sin\theta & -\xi\sin\theta + \eta\cos\theta \\ -\sin\theta & \cos\theta & -\xi\cos\theta - \eta\sin\theta \\ 0 & 0 & 1 \end{pmatrix}, \\
\overline{\Gamma}_{i\theta}^k &= \begin{pmatrix} \cos\phi & -\sin\phi & -x\sin\phi - y\cos\phi \\ \sin\phi & \cos\phi & x\cos\phi - y\sin\phi \\ 0 & 0 & 1 \end{pmatrix} \frac{\partial}{\partial\theta} \begin{pmatrix} \cos\theta & \sin\theta & -\xi\sin\theta + \eta\cos\theta \\ -\sin\theta & \cos\theta & -\xi\cos\theta - \eta\sin\theta \\ 0 & 0 & 1 \end{pmatrix}.
\end{aligned}$$

A tedious but straightforward computation then reveals that

$$\overline{\Gamma}_{i\xi}^k = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \quad \overline{\Gamma}_{i\eta}^k = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \overline{\Gamma}_{i\theta}^k = \begin{pmatrix} 0 & 1 & -\xi \\ -1 & 0 & -\eta \\ 0 & 0 & 0 \end{pmatrix},$$

i.e. all $\overline{\Gamma}_{ij}^k = 0$, except $\overline{\Gamma}_{\theta\xi}^\eta = -1\,, \overline{\Gamma}_{\theta\eta}^\xi = 1\,, \overline{\Gamma}_{\xi\theta}^\eta = -1\,, \overline{\Gamma}_{\eta\theta}^\xi = 1\,, \overline{\Gamma}_{\theta\theta}^\xi = -\xi\,, \overline{\Gamma}_{\theta\theta}^\eta = -\eta$. Note that $\overline{\Gamma}_{ij}^k - \overline{\Gamma}_{ji}^k = 0$, consistent with the fact that torsion vanishes identically by construction.

[1] We interpret this space as an extension of Euclidean 2-space E by attaching to each point a copy of the periodic circle $\mathrm{S} = \{\phi \in [0, 2\pi) \bmod 2\pi\}$ as an independent dimension. The parametrization can then be seen as a "shift-twist" transformation, whereby the Cartesian basisvectors $\mathbf{e}_x$ and $\mathbf{e}_y$ of the underlying Euclidean 2-space E (at any fiducial point) are rotated in accordance to the periodic translation in $\mathbf{e}_\phi$-direction (generating S).

Example

In the spirit of the previous example we consider a 3-space M, with coordinates (x, y, ϕ). However, we now introduce the following Riemannian metric $\mathbf{g} \in \mathrm{T}^*\mathrm{M} \otimes \mathrm{T}^*\mathrm{M}$:

$$\mathbf{g}(x, y, \phi) = dx \otimes dx + dy \otimes dy + (x^2 + y^2) d\phi \otimes d\phi \,.$$

Note that the metric is form invariant under the reparametrization of the previous example:

$$\mathbf{g}(\xi, \eta, \theta) = d\xi \otimes d\xi + d\eta \otimes d\eta + (\xi^2 + \eta^2) d\theta \otimes d\theta \,.$$

The induced Levi-Civita connection in (x, y, ϕ)-coordinates (and therefore also in (ξ, η, θ)-coordinates) follows from Theorem 3.1, page 56. In matrix notation we have for the metric and its dual, respectively,

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & x^2 + y^2 \end{pmatrix} \quad \text{and} \quad \mathbf{G}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/(x^2 + y^2) \end{pmatrix} .$$

The fact that metric and its dual are diagonal greatly simplifies the computation, for we have that

$$\Gamma^k_{ij} = \frac{1}{2} g^{kk} \left(\partial_i g_{kj} + \partial_j g_{ki} - \partial_k g_{ij} \right) \quad \text{(no summation over } k\text{)}.$$

Moreover, by virtue of the fact that the coefficients g_{ab} are constants for $a, b = 1, 2$, the above matrix (with row index i and column index j, and k fixed) will have a null 2×2 block in the upper left part:

$$\Gamma^k_{ij} = \begin{pmatrix} 0 & 0 & * \\ 0 & 0 & * \\ * & * & * \end{pmatrix} \quad \text{(row index } i\text{, column index } j\text{, } k \text{ fixed.)}$$

Finally, the (dual) metric coefficients do not depend on ϕ. By direct computation, taking these simplifying remarks into account, we obtain

$$\Gamma^x_{ij} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -x \end{pmatrix} \quad \Gamma^y_{ij} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -y \end{pmatrix} \quad \Gamma^\phi_{ij} = \frac{1}{x^2 + y^2} \begin{pmatrix} 0 & 0 & x \\ 0 & 0 & y \\ x & y & 0 \end{pmatrix} .$$

Note that $\Gamma^k_{ij} - \Gamma^k_{ji} = 0$, consistent with the fact that torsion always vanishes in the case of the Levi-Civita connection.

Example

Rotational invariance in the x–y plane of the previous example suggests the use of cylinder coordinates $\overline{\mathbf{x}} = (r, \psi, \theta)$ instead of $\mathbf{x} = (x, y, \phi)$:

$$\begin{cases} x(r,\psi,\theta) &= r\cos\psi \\ y(r,\psi,\theta) &= r\sin\psi \\ \phi(r,\psi,\theta) &= \theta\,. \end{cases}$$

To this end we once again invoke Result 3.2, page 53:

$$\overline{\Gamma}^k_{ij} = S^k_\ell \left(T^m_j T^n_i \Gamma^\ell_{nm} + \overline{\partial}_j T^\ell_i \right) .$$

The Jacobian matrices are given by

$$T^\ell_i = \frac{\partial x^\ell}{\partial \overline{x}^i} = \begin{pmatrix} \cos\psi & -r\sin\psi & 0 \\ \sin\psi & r\cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} ,$$

respectively

$$S^k_\ell = \frac{\partial \overline{x}^k}{\partial x^\ell} = \begin{pmatrix} x/(x^2+y^2) & y/(x^2+y^2) & 0 \\ -y/(x^2+y^2) & x/(x^2+y^2) & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Bibliography

[1] R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*, volume 75 of *Applied Mathematical Sciences*. Springer-Verlag, New York, second edition, 1988.

[2] D. Bao, S.-S. Chern, and Z. Shen. *An Introduction to Riemann-Finsler Geometry*, volume 2000 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000.

[3] M. Berger. *A Panoramic View of Riemannian Geometry*. Springer-Verlag, Berlin, 2003.

[4] R. L. Bishop and S. I. Goldberg. *Tensor Analysis on Manifolds*. Dover Publications, Inc., New York, 1980. Originally published by The Macmillan Company in 1968.

[5] M. P. do Carmo. *Differential Geometry of Curves and Surfaces*. Mathematics: Theory & Applications. Prentice-Hall, Englewood Cliffs, New Jersey, 1976.

[6] M. P. do Carmo. *Riemannian Geometry*. Mathematics: Theory & Applications. Birkhäuser, Boston, second edition, 1993.

[7] E. Cartan. *Leçons sur la Géométrie des Espaces de Riemann*. Gauthiers-Villars, Paris, second edition, 1963.

[8] Sir Thomas L. Heath. *The Thirteen Books of Euclid's Elements*, volume 1 (Books I and II). Dover Publications, Inc., New York, second edition, 1956.

[9] Sir Thomas L. Heath. *The Thirteen Books of Euclid's Elements*, volume 2 (Books III–IX). Dover Publications, Inc., New York, second edition, 1956.

[10] Sir Thomas L. Heath. *The Thirteen Books of Euclid's Elements*, volume 3 (Books X–XIII). Dover Publications, Inc., New York, second edition, 1956.

[11] J. Jost. *Riemannian Geometry and Geometric Analysis*. Springer-Verlag, Berlin, fourth edition, 2005.

[12] J. J. Koenderink. *Solid Shape*. MIT Press, Cambridge, 1990.

[13] A. P. Lightman, W. H. Press, R. H. Price, and S. A. Teukolsky. *Problem Book in Relativity and Gravitation*. Princeton University Press, Princeton, 1975.

[14] D. Lovelock and H. Rund, editors. *Tensors, Differential Forms, and Variational Principles*. Dover Publications, Inc., Mineola, N.Y., 1988.

[15] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. Freeman, San Francisco, 1973.

[16] H. Rund. *The Hamilton-Jacobi Theory in the Calculus of Variations*. Robert E. Krieger Publishing Company, Huntington, N.Y., 1973.

[17] M. Spivak. *Calculus on Manifolds*. W. A. Benjamin, New York, 1965.

[18] M. Spivak. *Differential Geometry*, volume 1–5. Publish or Perish, Berkeley, 1975.