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

Preface iii

Notation 1

1 Prerequisites from Linear Algebra 3

2 Tensor Calculus 7

2.1 Vector Spaces and Bases 7
2.2 Dual Vector Spaces and Dual Bases 8
2.3 The Kronecker Tensor 10
2.4 Inner Products 11
2.5 Reciprocal Bases 14
2.6 Bases, Dual Bases, Reciprocal Bases: Mutual Relations 16
2.7 Examples of Vectors and Covectors 17
2.8 Tensors 18

2.8.1 Tensors in all Generality 18
2.8.2 Tensors Subject to Symmetries 22
2.8.3 Symmetry and Antisymmetry Preserving Product Operators 24
2.8.4 Vector Spaces with an Oriented Volume 31
2.8.5 Tensors on an Inner Product Space 34
2.8.6 Tensor Transformations 36

2.8.6.1 “Absolute Tensors” 37
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

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: $\mathcal{A}$, $\mathcal{B}$, 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: $A_{j_1...j_q}^{i_1...i_p}$, $B_{j_1...j_q}^{i_1...i_p}$, 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 $\{e_i\}_{i=1}^n$, such that for each $x \in V$ there exists a unique $n$-tuple $(x^1, \ldots, x^n) \in \mathbb{R}^n$ such that $x = \sum_{i=1}^n x^i e_i$. 

Reminder

- A linear operator $\mathcal{A} : V \to W$, in which $V$ and $W$ are vector spaces over $\mathbb{K}$, is a mapping that satisfies the following property: $\mathcal{A}(\lambda v + \mu w) = \lambda \mathcal{A}(v) + \mu \mathcal{A}(w)$ for all $\lambda, \mu \in \mathbb{K}$ and $v, w \in V$.

- The set $\mathcal{L}(V, W)$ of all linear operators of this type is itself a vector space, with the following definitions of vector addition and scalar multiplication: 
  \[
  (\lambda \mathcal{A} + \mu \mathcal{B})(v) = \lambda \mathcal{A}(v) + \mu \mathcal{B}(v)
  \]
  for all $\lambda, \mu \in \mathbb{K}$, $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V, W)$ 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

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

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 $\pm 1$ 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 $3 \times 3$ 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:

<table>
<thead>
<tr>
<th>product with row labels ordered</th>
<th>sign of column label permutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{11}A_{22}A_{33}$</td>
<td>$[123] = +1$</td>
</tr>
<tr>
<td>$A_{12}A_{23}A_{31}$</td>
<td>$[231] = +1$</td>
</tr>
<tr>
<td>$A_{13}A_{21}A_{32}$</td>
<td>$[312] = +1$</td>
</tr>
<tr>
<td>$A_{13}A_{22}A_{31}$</td>
<td>$[321] = -1$</td>
</tr>
<tr>
<td>$A_{11}A_{23}A_{32}$</td>
<td>$[132] = -1$</td>
</tr>
<tr>
<td>$A_{12}A_{21}A_{33}$</td>
<td>$[213] = -1$</td>
</tr>
</tbody>
</table>

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 \( A \in M_n \) the following identities hold:

\[
\sum_{i_1, \ldots, i_n=1}^n \begin{bmatrix} i_1, \ldots, i_n \end{bmatrix} A_{i_1 j_1} \cdots A_{i_n j_n} = \begin{bmatrix} j_1, \ldots, j_n \end{bmatrix} \det A,
\]

\[
\sum_{j_1, \ldots, j_n=1}^n \begin{bmatrix} j_1, \ldots, j_n \end{bmatrix} A_{i_1 j_1} \cdots A_{i_n j_n} = \begin{bmatrix} i_1, \ldots, i_n \end{bmatrix} \det A,
\]

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

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

\[
\begin{bmatrix} i_1, \ldots, i_n \end{bmatrix} \begin{bmatrix} j_1, \ldots, j_n \end{bmatrix} = \det \begin{pmatrix} \delta_{i_1 j_1} & \cdots & \delta_{i_1 j_n} \\ \vdots & \ddots & \vdots \\ \delta_{i_n j_1} & \cdots & \delta_{i_n j_n} \end{pmatrix} \overset{\text{def}}{=} \delta_{j_1 \cdots j_n}.
\]

and, more generally,

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

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 \(A \in M_n\) is the matrix \(\tilde{A} \in M_n\) with entries \(\tilde{A}^{ij}\) given by

\[
\tilde{A}^{ij} = \frac{\partial \det A}{\partial A_{ij}}.
\]

The adjugate matrix of \(A \in M_n\) is a synonymous term for \(\tilde{A}^T \in M_n\).

Notation
Cofactor and adjugate matrix of \(A \in M_n\) are sometimes indicated by \(\text{cof} A\) and \(\text{adj} A\), respectively.
**Prerequisites from Linear Algebra**

**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:

$$X_i^i \overset{\text{def}}{=} \sum_{i=1}^{n} X_i^i.$$

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 $i, j = 1, \ldots, n$: $\frac{\partial \tilde{A}^{ij}}{\partial A_{ij}} = \frac{\partial^2 \det A}{\partial A_{ij}^2} = 0$.

- If $A : \mathbb{R} \to M_n : t \mapsto A(t)$ is a matrix-valued function, $\frac{d \det A}{dt} = \frac{\partial \det A}{\partial A_{ij}} \frac{dA_{ij}}{dt} = \tilde{A}^{ij} \frac{dA_{ij}}{dt}$ (chain rule).

- If $A$ is invertible, then $\tilde{A}^T = \tilde{A}^T A = \det A I$, or $A \adj A = \adj A A = \det A I$, i.e. $A_{jk} \tilde{A}^{ik} = \tilde{A}^{kj} A_{kj} = \det A \delta^i_j$.

- Each entry $\tilde{A}^{ij}$ of $\tilde{A} = \text{cof} A$ is a homogeneous polynomial of degree $n-1$ in terms of the entries $A_{ij}$ of $A$.

- Unlike $A^{-1}$, $\tilde{A} = \text{cof} A$ and $\tilde{A}^T = \adj A$ 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 $\{e_i\}$.

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

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 $X^i = X^j$.

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

Terminology
The real numbers $x^i$ are referred to as the contravariant components of the vector $x$ relative to the vector basis $\{e_i\}$.

Definition 2.1 The Kronecker symbol, or Kronecker-$\delta$, is defined as

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

Theorem 2.1 Let $x = x^i e_i \in V$ and $f_j = A^i_j e_i$ define a change of basis. If $x^i e_i = y^j f_j$, then

$$x^i = A^i_j y^j$$
or, equivalently,

$$y^j = B^j_i x^i,$$
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 $e_i$ 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 \( \{ e_i = \partial_i = \partial/\partial x^i \} \) at any (implicit) fiducial point defines a vector space basis. Upon a coordinate transformation \( x^i = x^i(\bar{x}^j) \) the chain rule relates the corresponding bases by

\[
\partial_j = A^i_j \partial_i,
\]

in which \( f_j = \partial_j = \partial/\partial x^j \), and

\[
A^i_j = \frac{\partial x^i}{\partial \bar{x}^j}.
\]

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. \( v = \nu^i \partial_i \). (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, \( \hat{f} : V \rightarrow \mathbb{R} \) is a linear functional if for all \( x, y \in V \), \( \lambda, \mu \in \mathbb{R} \),

\[
\hat{f}(\lambda x + \mu y) = \lambda \hat{f}(x) + \mu \hat{f}(y).
\]

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

\[
(\lambda \hat{f} + \mu \hat{g})(x) = \lambda \hat{f}(x) + \mu \hat{g}(x)
\]

for all \( x \in V \), recall chapter 1.

Definition 2.3 The dual space is the vector space of linear functionals on \( V \): \( V^* = \mathcal{L}(V, \mathbb{R}) \).
2.2 Dual Vector Spaces and Dual Bases

Ansatz
An \( n \)-dimensional vector space \( V \) over \( \mathbb{R} \) furnished with a basis \( \{ e_i \} \), and its dual space \( V^* \).

Remark
Each basis \( \{ e_i \} \) of \( V \) induces a basis \( \{ \hat{e}^j \} \) of \( V^* \) as follows. If \( x = x^i e_i \), then \( \hat{e}^i(x) = x^i \).

Result 2.1 The dual basis \( \{ \hat{e}^j \} \) of \( V^* \) is uniquely determined through its defining property
\[
\hat{e}^i(e_j) = \delta^j_k.
\]

Proof of Result 2.1. Let \( \hat{f} \in V^* \).

- For any \( x = x^i e_i \in V \) we have \( \hat{f}(x) = \hat{f}(x^i e_i) = \hat{f}(e_i)x^i = \hat{f}(e_i)\hat{e}^i(x) = (\hat{f}(e_i))\hat{e}^i(x) = (f_i\hat{e}^i)(x) \) with \( f_i = \hat{f}(e_i) \). This shows that the set \( \{ \hat{e}^j \} \) spans \( V^* \).
- Suppose \( \alpha_i \hat{e}^i = 0 \in V^* \), then \( \alpha_i = \alpha_j \delta^j_i = \alpha_j \hat{e}^j(e_i) = (\alpha_j \hat{e}^j)(e_i) = 0 \in \mathbb{R} \) for all \( i = 1, \ldots, n \). This shows that \( \{ \hat{e}^j \} \) 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 \( \{ e_i \} \)) implies the existence of its dual space \( V^* \) (with corresponding dual basis \( \{ \hat{e}^j \} \)).

Observation
The dual basis \( V^* \) has the same dimension as \( V \).

Terminology
A linear functional \( \hat{f} \in V^* \) 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 \( \{ \hat{e}^j \} \) of \( V^* \) and each \( \hat{x} \in V^* \) there exists a unique collection of real numbers \( \{ x_i \} \) such that \( \hat{x} = x^i \hat{e}^i \).

Terminology
The real numbers \( x_i \) are referred to as the covariant components of the covector \( \hat{x} \) relative to covector basis \( \{ \hat{e}^j \} \).

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

Theorem 2.2 The change of basis \( f_j = A^j_i e_i \) induces a change of dual vector basis
\[
\hat{f}^i = A^j_i \hat{e}^j,
\]
in which \( A^{jk}_k \delta^i_j = \delta^i_j \). Consequently, if \( \hat{x} = x^i \hat{e}^i \in V^* \), then
\[
x_i = B^j_i y_j \quad \text{or, equivalently,} \quad y_i = A^j_i x_j .
\]

Proof of Theorem 2.2.

- Given \( f_j = A^j_i e_i \) and \( \hat{f}^i = B^i_k \hat{e}^k \) we obtain \( \delta^i_j = \hat{f}^i(f_j) = B^i_k \hat{e}^k(A^j_i e_i) = B^i_k A^j_i \hat{e}^k(e_i) = B^i_k A^j_i \delta^k \ell = B^i_k A^j_k, \) from which it follows that \( B = A^{-1} \).
Thus if we define \( x_i \hat{e}^i = y_i \hat{f}^i \), then \( y_j \hat{f}^j = y_j B^j_i \hat{e}^i \), 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 \( x^i = x^i(\xi^j) \), the chain rule (at any fiducial point) states that

\[
d\xi^j = B^j_i dx^i \quad \text{in which} \quad B^j_i = \frac{\partial \xi^j}{\partial x^i}.
\]

Identifying \( \hat{e}^i = dx^i \) and \( \hat{f}^j = d\xi^j \), 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. \( \bar{v} = v_i dx^i \). (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 \( \bar{f} \in V^* \) one often employs the bracket formalism \( \langle \bar{f}, \cdot \rangle \in V^* \). Function evaluation is achieved by dropping a vector argument \( x \in V \) into the empty slot: \( \langle \bar{f}, x \rangle = \bar{f}(x) \in \mathbb{R} \). In particular we have

\[
\langle \hat{e}^i, e_j \rangle = \hat{e}^i(e_j) = \delta^i_j
\]

for dual bases, recall Result 2.1 on page 9.

**Definition 2.4** The bilinear map \( \langle \cdot, \cdot \rangle : V^* \times V \to \mathbb{R} : (\bar{f}, x) \mapsto \langle \bar{f}, x \rangle \) 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:

- \( \forall \hat{f}, \hat{g} \in V^*, x \in V, \lambda, \mu \in \mathbb{R} : \langle \lambda \hat{f} + \mu \hat{g}, x \rangle = \lambda \langle \hat{f}, x \rangle + \mu \langle \hat{g}, x \rangle, \)
- \( \forall \hat{f} \in V^*, x, y \in V, \lambda, \mu \in \mathbb{R} : \langle \hat{f}, \lambda x + \mu y \rangle = \lambda \langle \hat{f}, x \rangle + \mu \langle \hat{f}, y \rangle. \)
- Result 2.1 on page 9 shows that the components of the Kronecker tensor are given by the Kronecker symbol.
- If \( \hat{f} = f_i \hat{e}^i \in V^* \) and \( x = x^i e_i \in V \), then bilinearity of the Kronecker tensor implies \( \langle \hat{f}, x \rangle = f_i x^j \langle \hat{e}^i, e_j \rangle = f_i x^j \delta^i_j = f_i x^i \). 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 \( \hat{f} \) and \( \langle \hat{f}, \cdot \rangle \) 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 \( x \), say, or as \( \langle \cdot, x \rangle \in V^{**} \sim V \).

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 \( \langle \cdot, \cdot \rangle : V \times V \to \mathbb{R} \) which satisfies

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

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, \( \bar{z} = a - bi \in \mathbb{C} \) 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 hermitian sesquilinear mapping of the form $(\cdot | \cdot) : V \times V \to \mathbb{R}$ which satisfies

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

Caveat

- Note that for the complex case $(\lambda x + \mu y|z) = \overline{\lambda} (x|z) + \overline{\mu} (y|z)$, 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

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

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 $\{e_i\}$.

Definition 2.8 The Gram matrix $G$ is the matrix with components $g_{ij} = (e_i|e_j)$. 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 exist $x = x^i e_i \in V, x \neq 0$, and a $0$-eigenvector $y = y^i e_i \in V, y \neq 0$, such that $g_{ij} y^j = 0$, whence $0 = x^i g_{ij} y^j = x^i (e_i|e_j) y^j = (x^i e_i|y^j e_j) = (x|y)$. This contradicts the non-degeneracy axiom of Definition 2.7.
2.4 Inner Products

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 $G : V \to V^*$, with inverse $G^{-1} : V^* \to V$, such that

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

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

Since $x^i, y^j \in \mathbb{R}$ are arbitrary it follows that $G_{ij} = g_{ij}$, recall Definition 2.8.

- Consequently, the inverse $G^{-1} : V^* \to V$ has a matrix representation $G^{-1}$ corresponding to the inverse Gram matrix: $(G^{-1})^{ij} = g^{ij}$.

Remark

- Note the synonymous formulations: $G(x) \equiv \hat{x} \equiv (G(x), \cdot) \equiv (x|\cdot)$. 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 $\hat{x} \in V^*$ there exists one and only one $x \in V$ such that $(\hat{x}, y) = (x|y)$ for all $y \in V$, is an immediate consequence of Theorem 2.3, viz. $x = G^{-1}(\hat{x})$.

---

\(^1\)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 $G$, respectively $G^{-1}$, viz. if $v \in V$, $\hat{v} \in V^*$, then

$$\sharp v \overset{\text{def}}{=} G(v) \in V^*$$

respectively

$$\flat \hat{v} \overset{\text{def}}{=} G^{-1}(\hat{v}) \in V.$$

**Terminology**

$G$ is commonly known as the *metric*, and $G^{-1}$ 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 $v = v^i e_i \in V$ and $\hat{v} = v_i \hat{e}^i \in V^*$, then

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

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

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

**Notation**

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

**Caveat**

In the foregoing $G \in \mathcal{L}(V, V^*)$ and $G^{-1} \in \mathcal{L}(V^*, V)$, with prototypes $G : V \to V^*$, resp. $G^{-1} : V^* \to V$. However, one often interprets the metric and its dual as bilinear mappings of type $G : V \times V \to \mathbb{R}$, resp. $G^{-1} : V^* \times V^* \to \mathbb{R}$. 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 $\{e_i\}$.
Definition 2.10 The basis \( \{ e_i \} \) of \( V \) induces a new basis \( \{ \hat{e}^i \} \) of \( V \) given by \( \hat{e}^i = g^{ij} e_j \), the so-called reciprocal basis.

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

Caveat

- Instead of \( \hat{e}^i \) one sometimes writes \( e^i \) for the reciprocal of \( e_i \).
- Likewise, instead of \( \check{e}_i \) one sometimes writes \( \hat{e}_i \) for the reciprocal of \( \hat{e}^i \).
- 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 \( \hat{\ } \) and \( \check{\ } \) operator.

Remark

- The reciprocal basis is uniquely determined by the inner product.
- The Gram matrix of the reciprocal basis is \( G^{-1} \), with components \( (\hat{e}^i | \hat{e}^j) = g^{ik} g^{jl} (e_k | e_l) = g_{ik} g_{jl} = \delta^{ij} \).
- We have \( (\hat{e}^i | e_j) = g^{ik} (e_k | e_j) = g^{ik} g_{kj} = \delta^i_j \).

Terminology

The vectors \( \hat{e}^i \) and \( e_j \) are said to be perpendicular for \( i \neq j \).

Remark

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

Terminology

The real numbers \( x_i \) are referred to as the covariant components of the vector \( x \) relative to the basis \( \{ e_i \} \).

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 \( \hat{\ } \) and \( \check{\ } \) 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 $G$, 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 $\{e_i\}$.

Observation
Result 2.3 on page 13 states that each vector $x \in V$ induces a dual vector $\hat{x} = G(x) \in V^*$. Let $x = x^i e_i$ relative to basis $\{e_i\}$ of $V$. Then $\hat{x} = \hat{x}_j \hat{e}^j$ relative to the corresponding dual basis $\{\hat{e}^j\}$ of $V^*$. According to Definition 2.10 on page 15 we also have $x = x_i \hat{e}^i$ relative to the reciprocal basis $\{\hat{e}^i\}$ 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 $G : V \rightarrow V^*$ and linear map $\mathcal{G} : V \rightarrow V$ represented by the Gram matrix $G$, i.e. $e_i = \mathcal{G}(\hat{e}^i)$.

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 ($\{\hat{e}^i\}$, respectively $\{\hat{e}^i\}$), 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 $^\hat{\cdot}$ symbol, i.e. $\hat{x} = x_i \hat{e}^i \in V^*$.

\footnote{For the sake of argument we temporarily write $\hat{x}_i$ for dual, and $x_i$ for reciprocal vector components.}
2.7 Examples of Vectors and Covectors

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 $G$—the components of a vector $x \in V$ relative to the reciprocal basis are identical to the components of the dual vector $\hat{x} = G(x)$ relative to the dual basis.
- Thus, in an inner product space with fixed metric $G$, one may loosely identify a vector $x \in V$ and its dual vector $\hat{x} \in V^*$ to the extent that $\hat{x} = G(x)$ and $(\cdot | x) = \langle \hat{x}, \cdot \rangle$.
- The metric and its dual, $G$, respectively $G^{-1}$, provide the natural mechanism to “toggle” between elements of $V$ and $V^*$. Abbreviations: $G(x) = \sharp x$, $G^{-1}(\hat{x}) = \bm{\flat} \hat{x}$.
- The Gram matrix and its inverse, $G$, respectively $G^{-1}$, or rather the associated linear mappings $\mathcal{G} : V \to V$ and $\mathcal{G}^{-1} : V \to V$, 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 $v$ provides an example of a vector.
- Particle momentum $\hat{p}$ provides an example of a covector.
- The metric tensor allows us to relate these two: $\hat{p} = G(v) = \sharp v$, or equivalently, $v = G^{-1}(\hat{p}) = \bm{\flat} \hat{p}$.

Example

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

Caveat

- In practice one often refrains from a notational distinction between covectors and vectors, writing e.g. $p$ and $k$ instead of $\hat{p}$ and $\hat{k}$, as we do here.
- By the same token it is a common malpractice to speak of “the momentum vector $p$” and the “wave vector $k$” 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 \( \{ e_i \} \) of \( V \).

**Definition 2.12** A tensor is a multilinear mapping of type

\[
T : V^* \times \ldots \times V^* \times V \times \ldots \times V \rightarrow \mathbb{R}
\]

for some \( p, q \in \mathbb{N}_0 \).

**Notation**

Recall Definition 2.12. The following notation will be used to indicate tensor prototypes:

\[
T \in V \otimes \ldots \otimes V \otimes V^* \otimes \ldots \otimes V^* \overset{\text{def}}{=} T^{p}_{q}(V).
\]

**Caveat**

- The ordering of arguments matters!
- Thus the prototype may be specified with a different arrangement of linear spaces \( V \) and \( V^* \).

**Terminology**

- We refer to \( T \in T^{p}_{q}(V) \) sometimes as a \( \binom{p}{q} \)-tensor.
- \( T \in T^{p}_{q}(V) \) is referred to as a mixed tensor of contravariant rank \( p \) and covariant rank \( q \).
- If \( q \neq 0 \), \( T \in T^{0}_{q}(V) \) is called a covariant tensor of rank \( q \).
- If \( p \neq 0 \), \( T \in T^{p}_{0}(V) \) is called a contravariant tensor of rank \( p \).
- \( T \in T^{0}_{0}(V) \) is called a scalar. We identify \( T \) with its value in this case, i.e. \( T \in \mathbb{R} \).
- The words cotensor/contratensor are synonymous to covariant/contravariant tensor.

**Example**

Let \( T \) be a \( \binom{1}{2} \)-tensor, and \( x, x_1, x_2, y, y_1, y_2 \in V, \hat{z}, \hat{z}_1, \hat{z}_2 \in V^* \), and \( \lambda_1, \lambda_2 \in \mathbb{R} \) arbitrary. Then we have

- \( T(\lambda_1 \hat{z}_1 + \lambda_2 \hat{z}_2, x, y) = \lambda_1 T(\hat{z}_1, x, y) + \lambda_2 T(\hat{z}_2, x, y) \);
- \( T(\hat{z}, \lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1 T(\hat{z}, x_1, y) + \lambda_2 T(\hat{z}, x_2, y) \);
- \( T(\hat{z}, x, \lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 T(\hat{z}, x, y_1) + \lambda_2 T(\hat{z}, x, y_2) \).
Result 2.3 Special cases of low rank tensors:

- \( A \left( \begin{array} {c} 0 \\ 0 \end{array} \right) \)-tensor is a scalar, i.e. \( T^0_0(V) = \mathbb{R} \).
- \( A \left( \begin{array} {c} 0 \\ 1 \end{array} \right) \)-tensor is a covector, i.e. \( T^1_0(V) = V^* \).
- \( A \left( \begin{array} {c} 1 \\ 0 \end{array} \right) \)-tensor is a vector, i.e. \( T^1_0(V) = V \).

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 \( f \otimes g : X \times Y \rightarrow \mathbb{R} \) of two \( \mathbb{R} \)-valued functions \( f : X \rightarrow \mathbb{R} \) and \( g : Y \rightarrow \mathbb{R} \) is defined as follows:

\[
(f \otimes g)(x, y) = f(x)g(y) \quad \text{for all } x \in X, y \in Y.
\]

Since both vectors as well as covectors can be interpreted as \( \mathbb{R} \)-valued mappings, the following definition is a natural one.

Definition 2.14 For each \((p+q)\)-tuple of indices \((i_1, \ldots, i_p, j_1, \ldots, j_q)\) the mixed \((\begin{array}{c} p \\ q \end{array})\)-tensor \( e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \cdots \otimes \hat{e}^{j_q} \in T^p_q(V) \) is defined by

\[
\left( e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \cdots \otimes \hat{e}^{j_q} \right) (x_1, \ldots, x_p, y_1, \ldots, y_q) = \langle x_1, e_{i_1} \rangle \cdots \langle x_p, e_{i_p} \rangle \langle \hat{e}^{j_1}, y_1 \rangle \cdots \langle \hat{e}^{j_q}, y_q \rangle,
\]

for all \( x_1, \ldots, x_p \in V \) and \( y_1, \ldots, y_q \in V \).

Caveat

- The symbol \( \otimes \) used in Definition 2.14 is related to, but should not be confused with the one used in Definition 2.12. Mnemonic: \( e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \cdots \otimes \hat{e}^{j_q} \in V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^* \).

Result 2.4 If \( T \in T^p_q(V) \) there exists a set of numbers \( t^{i_1 \cdots i_p}_{j_1 \cdots j_q} \in \mathbb{R} \) such that

\[
T = t^{i_1 \cdots i_p}_{j_1 \cdots j_q} e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \cdots \otimes \hat{e}^{j_q},
\]

viz. \( t^{i_1 \cdots i_p}_{j_1 \cdots j_q} = T(e_{i_1}, \ldots, \hat{e}^{i_p}, e_{j_1}, \ldots, e_{j_q}) \).

Proof of Result 2.4. We have

\[
T \left( \hat{e}^{i_1}, \ldots, \hat{e}^{i_p}, e_{j_1}, \ldots, e_{j_q} \right) = t^{k_1 \cdots k_p}_{k_1 \cdots k_q} (e_{k_1} \otimes \cdots \otimes e_{k_p} \otimes \hat{e}^{j_1} \otimes \cdots \otimes \hat{e}^{j_q}) (e^{i_1}, \ldots, \hat{e}^{i_p}, e_{j_1}, \ldots, e_{j_q})
\]

\[
= t^{k_1 \cdots k_p}_{k_1 \cdots k_q} (e_{k_1} \otimes \cdots \otimes e_{k_p}) (e_{k_1} \otimes \cdots \otimes e_{k_p}) (e^{i_1}, e_{j_1}) \cdots (e^{i_p}, e_{j_q})
\]

\[
= t^{k_1 \cdots k_p}_{k_1 \cdots k_q} g^{i_1}_{k_1} \cdots g^{i_p}_{k_p} f^{j_1}_{k_1} \cdots f^{j_q}_{k_q} = t^{i_1 \cdots i_p}_{j_1 \cdots j_q}.
\]
So, for general covector and vector arguments $\hat{\omega}_1, \ldots, \hat{\omega}_p \in V^*$, $\mathbf{v}_1, \ldots, \mathbf{v}_q \in V$,

$$T(\hat{\omega}_1, \ldots, \hat{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) = t_{j_1 \ldots j_q}^{i_1 \ldots i_p} \hat{\omega}_{i_1} \ldots \hat{\omega}_{i_p} \mathbf{v}_{j_1} \ldots \mathbf{v}_{j_q}.$$ 

**Terminology**

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

**Caveat**

Laziness has led to the habit of saying that $T = t_{j_1 \ldots j_q}^{i_1 \ldots i_p} e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \ldots \otimes \hat{e}^{j_q}$ is expressed “relative to basis $\{e_i\}$”, where it would have been more accurate to refer to the actual basis $\{e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \ldots \otimes \hat{e}^{j_q}\}$.

**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.

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

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

$$\langle e_i \otimes \hat{e}^i \rangle (\mathbf{y}, \mathbf{x}) = \langle \mathbf{y}, e_i \rangle (\hat{e}^i, \mathbf{x}) = y_j x^k (\hat{e}^i, e_k) = y_j x^k \delta^i_k = y_i \mathbf{x}^j (\hat{e}^i, e_j) = \langle \mathbf{y}, \mathbf{x} \rangle.$$ 

Note that the Kronecker tensor is parameter free. For this reason it is sometimes called a **constant tensor**. Since its coefficients $\delta^i_j$ do not depend on the chosen basis (an untypical property for tensor coefficients), they are said to constitute an **invariant holor**.

**Observation**

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

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

in which “(co)vectors” generically denotes an appropriate sequence of arbitrary vector and/or covector arguments.

- We have dim $T^p_q(V) = n^{p+q}$.

- Lexicographical ordering of the components $t_{j_1 \ldots j_q}^{i_1 \ldots i_p}$ of tensors $T \in T^p_q(V)$ establishes an isomorphism with the “column vector space” $\mathbb{R}^{n^{p+q}}$. 

2.8 Tensors

Terminology

- The indexed collection of tensor coefficients \( t_{j_1 \ldots j_q}^{i_1 \ldots i_p} \) 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, \( T = t_{j_1 \ldots j_q}^{i_1 \ldots i_p} e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \cdots \otimes \hat{e}^{j_q} \).
- If \( \{ e_i \} \) is a basis of \( V \), a basis of \( T^p_q(V) \) is given by \( \{ e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \cdots \otimes \hat{e}^{j_q} \} \). 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 \( G \in \mathcal{L}(V, V^*) \) as a covariant tensor of rank 2. In order to understand this, let us clarify a previous remark on page 14. For given metric \( G \), define \( g \in V^* \otimes V^* = T^0_2(V) \) as follows:

\[
g : V \times V \rightarrow \mathbb{R} : (v, w) \mapsto g(v, w) = \langle G(v), w \rangle = \langle v|w \rangle = \langle \hat{v}, w \rangle,
\]

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 \( g \) and \( G \). We will invariably write \( G \), whether it is intended as a covariant 2-tensor (\( g \) in the example, implied when we write \( G(v, w) \)), or as a vector-to-covector converter (\( \sharp \)-operator, implied when we write \( G(v) \)).
- 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.

Terminology

Irrespective of the intended interpretation, \( G \) will be referred to as the metric tensor.

Theorem 2.4

If \( S \in T^p_q(V) \), \( T \in T^r_s(V) \), with

\[
S = s_{j_1 \ldots j_q}^{i_1 \ldots i_p} e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \cdots \otimes \hat{e}^{j_q} \quad \text{and} \quad T = t_{j_1 \ldots j_q}^{i_1 \ldots i_p} e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \cdots \otimes \hat{e}^{j_q},
\]

then \( S \otimes T \in T^{p+r}_{q+s}(V) \), with \( S \otimes T = s_{j_1 \ldots j_q}^{i_1 \ldots i_p} t_{j_{p+1} \ldots j_q}^{i_{p+1} \ldots i_r} e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \cdots \otimes \hat{e}^{j_{q+s}} \).

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 \( S \otimes T \) 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 $T \in T^p_q(V)$ has contravariant rank $p$ and covariant rank $q$.

• The holor $t^{i_1...i_p}_{j_1...j_q} \in \mathbb{R}$ of $T \in T^p_q(V)$ is labeled with $p$ upper and $q$ lower indices.

• Elements of the underlying basis $\{e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \ldots \otimes \hat{e}^{j_q}\}$ of $T^p_q(V)$, 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 $\{e_i\}$ of $V$.

• Tensor spaces of type $T^p_q(V)$.

Certain subspaces of $T^p_q(V)$ 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

• $\pi = \{\pi(\alpha_1), \ldots, \pi(\alpha_p)\}$ denotes a permutation of the $p$ ordered symbols $\{\alpha_1, \ldots, \alpha_p\}$.

• $\pi(\alpha_i), i = 1, \ldots, p$, is shorthand notation for the $i$-th entry of $\pi$.

Definition 2.15 Let $\pi = \{\pi(1), \ldots, \pi(p)\}$ be any permutation of labels $\{1, \ldots, p\}$, then $T \in T^0_p(V)$ is called a symmetric covariant tensor if for all $v_1, \ldots, v_p \in V$

\[
T(v_{\pi(1)}, \ldots, v_{\pi(p)}) = T(v_1, \ldots, v_p).
\]

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

\[
T(\hat{v}_{\pi(1)}, \ldots, \hat{v}_{\pi(p)}) = T(\hat{v}_1, \ldots, \hat{v}_p).
\]
Notation

- The vector space of symmetric covariant $p$-tensors is denoted $\bigvee_p(V) \subset T^0_p(V)$.
- The vector space of symmetric contravariant $p$-tensors is denoted $\bigvee^p(V) \subset T^0_p(V)$.
- Alternatively one sometimes writes $\bigvee_p(V) = V^* \otimes \ldots \otimes V^*$ and $\bigvee^p(V) = V \otimes \ldots \otimes V$.

Example

An example of a symmetric covariant $p$-tensor on $\mathbb{R}^n$ is given by the set of $p$-th order partial derivatives of a sufficiently smooth function $f \in C^p(\mathbb{R}^n, \mathbb{R})$ at a fiducial point:

$$D_p f(0) = \frac{\partial^p f(0)}{\partial x^{i_1} \ldots \partial x^{i_p}} \hat{e}_{i_1} \otimes \ldots \otimes \hat{e}_{i_p}.$$ 

Notation

- For simplicity we write $f_{i_1 \ldots i_p}(0)$ for the partial derivatives in the above expression.
- The fiducial point (origin in this example) is often suppressed in the notation.

Caveat

The derivative tensor field $D_p f : \mathbb{R}^n \to T^0_p(V) : x \mapsto D_p f(x)$—as opposed to the isolated tensor $D_p f(0) \in T^0_p(V)$—fails to be a tensor field unless one replaces partial derivatives by covariant derivatives. (We return to this later.)

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

Definition 2.16 Let $\text{sgn}(\pi)$ denote the sign of the permutation $\pi$, i.e. $\text{sgn}(\pi) = 1$ if it is even, $\text{sgn}(\pi) = -1$ if it is odd. $\mathbf{T} \in T^0_p(V)$ 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 $\{1, \ldots, p\}$,

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

Likewise, $\mathbf{T} \in T^0_p(V)$ is called an antisymmetric contravariant tensor if for all $\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_p \in V^*$

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

Notation

- The vector space of antisymmetric covariant $p$-tensors is denoted $\bigwedge_p(V) \subset T^0_p(V)$.
- The vector space of antisymmetric contravariant $p$-tensors is denoted $\bigwedge^p(V) \subset T^0_p(V)$.
- Alternatively one sometimes writes $\bigwedge_p(V) = V^* \otimes_A \ldots \otimes_A V^*$ and $\bigwedge^p(V) = V \otimes_A \ldots \otimes_A V$.

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

- \( V_0(V) = V^0(V) = \bigwedge_0(V) = \bigwedge^0(V) = \mathbb{R} \). (This is consistent with Definitions 2.15 and 2.16 if we set \( \pi = \text{id} \), \( \text{sgn} \pi = 1 \) on the empty index set \( \emptyset \).
- \( V_1(V) = \bigwedge_1(V) = V^* \). (This follows from Definitions 2.15 and 2.16, with \( \pi = \text{id} \), \( \text{sgn} \pi = 1 \) on \( \{1\} \).
- \( V^1(V) = \bigwedge^1(V) = V \). (This follows from Definitions 2.15 and 2.16, with \( \pi = \text{id} \), \( \text{sgn} \pi = 1 \) on \( \{1\} \).

2.8.3. Symmetry and Antisymmetry Preserving Product Operators

Ansatz

- An \( n \)-dimensional vector space \( V \) over \( \mathbb{R} \).
- A basis \( \{e_i\} \) of \( V \).
- Tensor spaces of type \( \bigwedge^p(V) \), \( \bigvee^p(V) \), \( \bigwedge^p(V) \), \( \bigvee^p(V) \). 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 \( \bigwedge^p(V) \), \( \bigvee^p(V) \), 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:

\[
(\hat{\mathbf{v}}^1 \wedge \ldots \wedge \hat{\mathbf{v}}^k)(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \det(\angle \hat{\mathbf{v}}^i, \mathbf{x}_j) \quad \text{for all } \mathbf{x}_j \in V.
\]

Example

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

\[
(\hat{\mathbf{v}} \wedge \hat{\mathbf{w}})(\mathbf{x}, \mathbf{y}) = \det \left( \begin{array}{cc} \langle \hat{\mathbf{v}}, \mathbf{x} \rangle & \langle \hat{\mathbf{v}}, \mathbf{y} \rangle \\ \langle \hat{\mathbf{w}}, \mathbf{x} \rangle & \langle \hat{\mathbf{w}}, \mathbf{y} \rangle \end{array} \right) = \langle \hat{\mathbf{v}}, \mathbf{x} \rangle \langle \hat{\mathbf{w}}, \mathbf{y} \rangle - \langle \hat{\mathbf{w}}, \mathbf{x} \rangle \langle \hat{\mathbf{v}}, \mathbf{y} \rangle.
\]

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 \textit{wedge product} or the \textit{antisymmetric product}.

Observation

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

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

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

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 \( \text{perm} \) assigns a number to a square matrix in a way similar to \( \det \), but with a + sign for each term, recall page 4:

\[
\text{perm} A = \sum_{j_1, \ldots, j_n=1}^{n} |[j_1, \ldots, j_n]| A_{1j_1} \cdots A_{nj_n} = \frac{1}{n!} \sum_{i_1, \ldots, i_n=1}^{n} |[i_1, \ldots, i_n]| [j_1, \ldots, j_n] A_{i_1j_1} \cdots A_{inj_n}.
\]

**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 \( \text{perm} \) is called the **permanent**.

**Example**
Analogous to Table 1.1 on page 4 we have for a \( 3 \times 3 \) matrix \( A \),

\[
\text{perm} A = A_{11} A_{22} A_{33} + A_{12} A_{23} A_{31} + A_{13} A_{21} A_{32} + A_{12} A_{21} A_{33} + A_{13} A_{23} A_{31} + A_{12} A_{23} A_{31}.
\]

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

\[
(\hat{v}^1 \vee \ldots \vee \hat{v}^k)(x_1, \ldots, x_k) = \text{perm} (\hat{v}^1, x_j).
\]

**Example**
Let \( \hat{v}, \hat{w} \in V^* \), \( x, y \in V \), then

\[
(\hat{v} \vee \hat{w})(x, y) = \text{perm} \left( \begin{array}{c} \langle \hat{v}, x \rangle \\ \langle \hat{w}, x \rangle \\ \langle \hat{v}, y \rangle \\ \langle \hat{w}, y \rangle \end{array} \right) = \langle \hat{v}, x \rangle \langle \hat{w}, y \rangle + \langle \hat{w}, x \rangle \langle \hat{v}, y \rangle.
\]

Alternatively,

\[
\hat{v} \vee \hat{w} = \hat{v} \otimes \hat{w} + \hat{w} \otimes \hat{v}.
\]

**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 $\hat{u}, \hat{v}, \hat{w} \in V^\ast$, $\lambda, \mu \in \mathbb{R}$, then

1. $(\hat{u} \vee \hat{v}) \vee \hat{w} = \hat{u} \vee (\hat{v} \vee \hat{w})$.
2. $(\lambda \hat{u} + \mu \hat{v}) \vee \hat{w} = \lambda \hat{u} \vee \hat{v} + \mu \hat{v} \vee \hat{w}$.
3. $\hat{u} \vee (\lambda \hat{v} + \mu \hat{w}) = \lambda \hat{u} \vee \hat{v} + \mu \hat{u} \vee \hat{w}$.
4. $\hat{u} \vee \hat{v} = \hat{v} \vee \hat{u}$.

Observation

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

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

Proof of Theorem 2.5.

1. As for the first identity, a strictly increasing ordering of labels in a product of the form $\hat{e}^{i_1} \wedge \ldots \wedge \hat{e}^{i_p}$ can be realized in exactly as many distinct ways as one can select $p$ items out of $n$ different objects.

2. In particular, if $p > n$, then $\dim \bigwedge_p(V) = 0$, i.e. $\bigwedge_p(V) = \{0\}$.

3. As for the second identity, the number of ways in which a non-decreasing ordering of labels in a product of the form $\hat{e}^{i_1} \vee \ldots \vee \hat{e}^{i_p}$ can be realized can be inferred from the following argument. 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 $n = 5, p = 7$:

$$\bullet \times \bullet \times \bullet \times \bullet \times \bullet \equiv \hat{e}^1 \vee \hat{e}^3 \vee \hat{e}^2 \vee \hat{e}^4 \vee \hat{e}^4 \vee \hat{e}^4 \vee \hat{e}^5$$

$$\bullet \times \bullet \bullet \bullet \times \bullet \times \times \equiv \hat{e}^1 \vee \hat{e}^3 \wedge \hat{e}^2 \wedge \hat{e}^2 \wedge \hat{e}^2 \wedge \hat{e}^3$$

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.

\[\text{3} \text{Courtesy of Freek Paans.}\]
Example
Let $B^p$ and $B_p$ denote bases for $\bigwedge^p(V)$, respectively $\bigwedge_p(V)$, in Euclidean 3-space $V = \mathbb{R}^3$.

<table>
<thead>
<tr>
<th>$p$</th>
<th>$B^p$</th>
<th>dim</th>
<th>$p$</th>
<th>$B_p$</th>
<th>dim</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>${1}$</td>
<td>1</td>
<td>0</td>
<td>${1}$</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>${e_1, e_2, e_3}$</td>
<td>3</td>
<td>1</td>
<td>${\hat{e}^1, \hat{e}^2, \hat{e}^3}$</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>${e_1 \wedge e_2, e_1 \wedge e_3, e_2 \wedge e_3}$</td>
<td>3</td>
<td>2</td>
<td>${e^1 \wedge e^2, e^1 \wedge e^3, e^2 \wedge e^3}$</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>${e_1 \wedge e_2 \wedge e_3}$</td>
<td>1</td>
<td>3</td>
<td>${\hat{e}^1 \wedge \hat{e}^2 \wedge \hat{e}^3}$</td>
<td>1</td>
</tr>
</tbody>
</table>

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 $\mathcal{S} : T^0_0(V) \to \bigwedge^p(V)$, given by

$$(\mathcal{S}(T))(v_1, \ldots, v_p) = \frac{1}{p!} \sum_{\pi} T(v_{\pi(1)}, \ldots, v_{\pi(p)}) ,$$

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

Observation
- If $T = t_{i_1 \ldots i_p} \hat{e}^{i_1} \otimes \cdots \otimes \hat{e}^{i_p} \in T^0_0(V)$, then $\mathcal{S}(T) = t_{(i_1 \ldots i_p)} \hat{e}^{i_1} \otimes \cdots \otimes \hat{e}^{i_p} \in \bigwedge^p(V)$, in which

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

- If $T = \bigwedge^p(V)$ is symmetric, then $T = \mathcal{S}(T)$.

- The symmetrisation map is a projection, i.e. it is idempotent: $\mathcal{S} \circ \mathcal{S} = \mathcal{S}$.

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

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

$$\mathcal{S}(\hat{e}^{i_1} \lor \cdots \lor \hat{e}^{i_p}) = \frac{1}{p!} \hat{e}^{i_1} \lor \cdots \lor \hat{e}^{i_p} .$$

- $T \lor U = U \lor T$.

- The case $p = 0, q > 0$ is covered by defining $\lambda \lor U = \lambda U$ for all $\lambda \in \mathbb{R}$, $U \in \bigwedge_q(V)$.

- Rule of thumb: The action of $p! \mathcal{S}$ on a $\otimes$-product amounts to a formal replacement $\otimes \to \otimes_S \equiv \lor$.

The following result generalizes the example given on page 25:

Result 2.5 Let $T \in \bigwedge^p(V)$ and $U \in \bigwedge^q(V)$, then $T \lor U = \frac{(p + q)!}{p!q!} \mathcal{S}(T \otimes U)$.
Definition 2.21  The linear operator $\mathcal{A} : T^0_p(V) \to \bigwedge^p(V)$, given by
\[
(\mathcal{A}(T))(v_1, \ldots, v_p) = \frac{1}{p!} \sum_{\pi} \text{sgn}(\pi) T(v_{\pi(1)}, \ldots, v_{\pi(p)}),
\]
is referred to as the antisymmetrisation map.

Observation

- If $T = t_{i_1 \ldots i_p} \hat{e}^{i_1} \otimes \ldots \otimes \hat{e}^{i_p} \in T^0_p(V)$, then $\mathcal{A}(T) = t_{[i_1 \ldots i_p]} \hat{e}^{i_1} \otimes \ldots \otimes \hat{e}^{i_p} \in \bigwedge^p(V)$, in which
\[
t_{[i_1 \ldots i_p]} \overset{\text{def}}{=} \frac{1}{p!} \sum_{\pi} \text{sgn}(\pi) t_{\pi(i_1) \ldots \pi(i_p)}.
\]
- If $T = \bigwedge^p(V)$ is antisymmetric, then $T = \mathcal{A}(T)$.
- The antisymmetrisation map is a projection, i.e. it is idempotent: $\mathcal{A} \circ \mathcal{A} = \mathcal{A}$.

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
\[
\mathcal{A}(\hat{e}^{i_1} \otimes \ldots \otimes \hat{e}^{i_p}) = \frac{1}{p!} \hat{e}^{i_1} \wedge \ldots \wedge \hat{e}^{i_p}.
\]
- The case $p = 0, q > 0$ is covered by defining $\lambda \wedge U = \lambda U$ for all $\lambda \in \mathbb{R}, U \in \bigwedge^q(V)$.
- Rule of thumb: The action of $p!\mathcal{A}$ on a $\otimes$-product amounts to a formal replacement $\otimes \to \otimes \lambda \equiv \wedge$.

The following result generalizes the example given on page 24.

Result 2.6  Let $T \in \bigwedge^p(V)$ and $U \in \bigwedge^q(V)$, then $T \wedge U = \frac{(p + q)!}{p!q!} \mathcal{A}(T \otimes U)$.

Notation

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

Observation

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

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

$$\mathbf{T} \wedge \mathbf{U} = t_{i_1...i_p}u_{i_{p+1}...i_{p+q}}\hat{e}^{i_1} \wedge \ldots \wedge \hat{e}^{i_p} \wedge \ldots \wedge \hat{e}^{i_{p+q}} = \frac{1}{p!q!} t_{i_1...i_p}u_{i_{p+1}...i_{p+q}}\hat{e}^{i_1} \wedge \ldots \wedge \hat{e}^{i_{p+q}}$$

$$= \frac{(p+q)!}{p!q!} t_{i_1...i_p}u_{i_{p+1}...i_{p+q}} \mathcal{A} (\hat{e}^{i_1} \otimes \ldots \otimes \hat{e}^{i_{p+q}}) = \frac{(p+q)!}{p!q!} \mathcal{A} (\mathbf{T} \otimes \mathbf{U}) .$$

Notice that all $p + q$ indices can be assumed to be distinct. In step $\ast$ we have used the fact that there are precisely $p!$ (respectively $q!$) orderings of $p$ (respectively $q$) distinct indices. In step $\ast$ we have used the previous observation on the relationship between wedge product and antisymmetrised outer product. The last step exploits linearity of $\mathcal{A}$.

**Result 2.7** Let $\mathbf{T} \in \bigwedge_p^p (V)$ and $\mathbf{U} \in \bigwedge_q^q (V)$, 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 $\hat{e}^{i_1}, \ldots, \hat{e}^{i_p}$ in the product $\hat{e}^{i_1} \wedge \ldots \wedge \hat{e}^{i_{p+q}}$ so as to obtain the reordered product $\hat{e}^{i_{p+1}} \wedge \ldots \wedge \hat{e}^{i_{p+q}} \wedge \hat{e}^{i_1} \wedge \ldots \wedge \hat{e}^{i_p}$ requires $p$ (one for each $\hat{e}^{i_k}, k = 1, \ldots, p$) times $q$ (one for each $\hat{e}^{i_\ell}, \ell = p + 1, \ldots, p + q$) commutations, each of which brings in a minus sign.

**Remark**

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

**Remark**

- Both $\mathcal{I}$ as well as $\mathcal{A}$ are idempotent: $\mathcal{I} \circ \mathcal{I} = \mathcal{I}$, respectively $\mathcal{A} \circ \mathcal{A} = \mathcal{A}$. Thus they are, by definition, projections from $\mathbf{T}_p^p (V)$ onto $\bigwedge_p (V)$, 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 $\mathbf{S} = \mathcal{I} (\mathbf{T})$ and $\mathbf{A} = \mathcal{A} (\mathbf{T})$, 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 $\{e_1, e_2, e_3\}$. One conventionally denotes
\[ e_1 = \frac{\partial}{\partial x}, \quad e_2 = \frac{\partial}{\partial y}, \quad e_3 = \frac{\partial}{\partial z}. \]

It is then customary to denote the corresponding dual basis $\{\hat{e}^1, \hat{e}^2, \hat{e}^3\}$ of $V^*$ by
\[ \hat{e}^1 = dx, \quad \hat{e}^2 = dy, \quad \hat{e}^3 = dz. \]

Let $u = u^i e_i, v = v^i e_i, w = w^i e_i \in V$ be arbitrary vectors. Then we have
\[ dx(v) = \langle dx, v \rangle = v^1, \]
\[ dy(v) = \langle dy, v \rangle = v^2, \]
\[ dz(v) = \langle dz, v \rangle = v^3. \]

Example
Notation as in the previous example.
\[ (dx \wedge dy)(v, w) = \det \begin{vmatrix} \langle dx, v \rangle & \langle dx, w \rangle \\ \langle dy, v \rangle & \langle dy, w \rangle \end{vmatrix} = \det \begin{pmatrix} v^1 & w^1 \\ v^2 & w^2 \end{pmatrix} = v^1 w^2 - v^2 w^1, \]
\[ (dx \wedge dz)(v, w) = \det \begin{vmatrix} \langle dx, v \rangle & \langle dx, w \rangle \\ \langle dz, v \rangle & \langle dz, w \rangle \end{vmatrix} = \det \begin{pmatrix} v^1 & w^1 \\ v^3 & w^3 \end{pmatrix} = v^1 w^3 - v^3 w^1, \]
\[ (dy \wedge dz)(v, w) = \det \begin{vmatrix} \langle dy, v \rangle & \langle dy, w \rangle \\ \langle dz, v \rangle & \langle dz, w \rangle \end{vmatrix} = \det \begin{pmatrix} v^2 & w^2 \\ v^3 & w^3 \end{pmatrix} = v^2 w^3 - v^3 w^2. \]

Example
Notation as in the previous example.
\[ (dx \wedge dy \wedge dz)(u, v, w) = \det \begin{vmatrix} \langle dx, u \rangle & \langle dx, v \rangle & \langle dx, w \rangle \\ \langle dy, u \rangle & \langle dy, v \rangle & \langle dy, w \rangle \\ \langle dz, u \rangle & \langle dz, v \rangle & \langle dz, w \rangle \end{vmatrix} = \det \begin{pmatrix} u^1 & v^1 & w^1 \\ u^2 & v^2 & w^2 \\ u^3 & v^3 & w^3 \end{pmatrix} = u^1 v^2 w^3 + u^2 v^3 w^1 - u^3 v^1 w^2 - u^1 v^3 w^2 - u^2 v^1 w^3. \]

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 $\mathcal{V}_p(V), \mathcal{A}_p(V), \mathcal{V}^p(V), \mathcal{A}^p(V)$. 
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 \( v_1, \ldots, v_k, x_1, \ldots, x_k \in V \), set

\[
\langle v_1 \wedge \ldots \wedge v_k | x_1 \wedge \ldots \wedge x_k \rangle = \det \langle \sharp v_i, x_j \rangle, \quad \text{respectively}
\]

\[
\langle v_1 \lor \ldots \lor v_k | x_1 \lor \ldots \lor x_k \rangle = \operatorname{perm} \langle \sharp v_i, x_j \rangle.
\]

Likewise for the spaces \( \bigwedge^p(V) \) and \( \bigvee^p(V) \):

\[
\langle \hat{v}_1 \wedge \ldots \wedge \hat{v}_k | \hat{x}_1 \wedge \ldots \wedge \hat{x}_k \rangle = \det \langle \hat{\sharp} v_i, \flat \hat{x}_j \rangle, \quad \text{respectively}
\]

\[
\langle \hat{v}_1 \lor \ldots \lor \hat{v}_k | \hat{x}_1 \lor \ldots \lor \hat{x}_k \rangle = \operatorname{perm} \langle \hat{\sharp} v_i, \flat \hat{x}_j \rangle,
\]

for any \( \hat{v}_1, \ldots, \hat{v}_k, \hat{x}_1, \ldots, \hat{x}_k \in V^* \).

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 fixed basis \( \{ e_i \} \) of \( V \).
- A nonzero \( n \)-form \( \mu \in \bigwedge^n(V) \).

Terminology

If \( \mu(e_1, \ldots, e_n) = 1 \) then \( \mu \) is called the unit volume form.

Observation

- We have \( \dim \bigwedge^n(V) = 1 \).
- Consequently, if \( \mu_1, \mu_2 \in \bigwedge^n(V) \), then \( \mu_1 \) and \( \mu_2 \) differ by a constant factor \( \lambda \in \mathbb{R} \).
- For the unit volume form we have, according to the observation on page 28,

\[
\mu = \hat{e}^1 \wedge \cdots \wedge \hat{e}^n = \frac{1}{n!} [i_1, \ldots, i_n] \hat{e}^{i_1} \wedge \cdots \wedge \hat{e}^{i_n} = [i_1, \ldots, i_n] \mathcal{A} (\hat{e}^{i_1} \otimes \cdots \otimes \hat{e}^{i_n}).
\]

The \( \mathcal{A} \)-operator on the right hand side may be dropped, since its effect is void when combined with antisymmetric coefficients. We conclude that the covariant components of \( \mu \in \bigwedge^n(V) \) are given by \( \mu_{i_1 \cdots i_n} = [i_1, \ldots, i_n] \).

Caveat

The identification \( \mu_{i_1 \cdots i_n} = [i_1, \ldots, i_n] \) holds only relative to the fixed basis used above to construct \( \mu \). We refer to section 2.8.6 for a basis independent evaluation.

Observation

Thus \( \mu = \mu_{[i_1 \cdots i_n]} \hat{e}^{i_1} \wedge \cdots \wedge \hat{e}^{i_n} = \mu_{i_1 \cdots i_n} \hat{e}^{i_1} \otimes \cdots \otimes \hat{e}^{i_n} \). This shows the convenience of the two variants of Einstein summation.
Definition 2.22 Let \((V, \mu)\) denote the vector space \(V\) endowed with the \(n\)-form \(\mu \in \bigwedge^n(V)\).

- \((V, \mu)\) is called a vector space with an oriented volume.

- \((V, \mu)\) is said to have a positive orientation if \(\mu(e_1, \ldots, e_n) > 0\).

- \((V, \mu)\) is said to have a negative orientation if \(\mu(e_1, \ldots, e_n) < 0\).

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

- given a fixed basis \(\{e_i\}\) of \(V\) one refers to \(\mu \in \bigwedge^n(V)\) as a positive volume form if \(\mu(e_1, \ldots, e_n) > 0\), and

- given a basic volume form \(\mu \in \bigwedge^n(V)\) one refers to \(\{e_i\}\) as a positively oriented basis if \(\mu(e_1, \ldots, e_n) > 0\), etc.

Notation
Instead of \((V, \mu)\) we shall simply write \(V\). The existence and nature of an oriented volume should be clear from the context.

Observation
Using Definition 2.17 on page 24 we may write, for \(x_1, \ldots, x_n \in V\),

\[
\mu(x_1, \ldots, x_n) = \det(\hat{e}^j, x_i) = \det \begin{pmatrix}
\langle \hat{e}^1, x_1 \rangle & \cdots & \langle \hat{e}^1, x_n \rangle \\
\vdots & & \vdots \\
\langle \hat{e}^n, x_1 \rangle & \cdots & \langle \hat{e}^n, x_n \rangle
\end{pmatrix},
\]

which clearly reveals the role of the Kronecker tensor (and thus of the dual vector space \(V^*\)) in the definition of \(\mu\). Consistent with a foregoing observation it follows that the covariant components of \(\mu = \mu_{i_1 \ldots i_n} \hat{e}^{i_1} \otimes \cdots \otimes \hat{e}^{i_n}\) are given by \(\mu_{i_1 \ldots i_n} = \mu(e_{i_1}, \ldots, e_{i_n}) = [i_1, \ldots, i_n] \mu(e_1, \ldots, e_n) = [i_1, \ldots, i_n] \det \langle \hat{e}^j, e_i \rangle = [i_1, \ldots, i_n] \det \delta^i_{i_1} = [i_1, \ldots, i_n]\).

Definition 2.23 Let \(a_1, \ldots, a_k \in V\) for some \(k \in \{0, \ldots, n\}\), then we define the \((n-k)\)-form \(\mu \lrcorner a_1 \lrcorner \cdots \lrcorner a_k \in \bigwedge^{n-k}(V)\) as follows:

\[
(\mu \lrcorner a_1 \lrcorner \cdots \lrcorner a_k)(x_{k+1}, \ldots, x_n) = \mu(a_1, \ldots, a_k, x_{k+1}, \ldots, x_n).
\]
Observation

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

\[
\begin{pmatrix}
\langle \hat{e}^1, a_1 \rangle & \cdots & \langle \hat{e}^1, a_k \rangle & \langle \hat{e}^1, x_{k+1} \rangle & \cdots & \langle \hat{e}^1, x_n \rangle \\
\vdots & & \vdots & \vdots & & \vdots \\
\langle \hat{e}^n, a_1 \rangle & \cdots & \langle \hat{e}^n, a_k \rangle & \langle \hat{e}^n, x_{k+1} \rangle & \cdots & \langle \hat{e}^n, x_n \rangle \\
\end{pmatrix}
\]

- The holor of \( \mu \cdot a_1 \ldots a_k \in \bigwedge_{n-k}(V) \) for given \( a_1, \ldots, a_k \in V \) is established as follows:

\[
(\mu \cdot a_1 \ldots a_k)(e_{k+1}, \ldots, e_{i_n}) = a_1^{i_1} \ldots a_k^{i_k} \mu(e_{i_1}, \ldots, e_{i_n}) = \mu_{i_1 \ldots i_n} a_1^{i_1} \ldots a_k^{i_k},
\]

in which we have used the decompositions \( a_p = a_p^{i_p} e_{i_p} \) for \( p = 1, \ldots, k \) together with multilinearity.

- By the same token, expanding also \( x_q = x_q^{i_q} e_{i_q} \) for \( q = k + 1, \ldots, n \), we obtain

\[
(\mu \cdot a_1 \ldots a_k)(x_{k+1}, \ldots, x_n) = \mu_{i_1 \ldots i_n} a_1^{i_1} \ldots a_k^{i_k} x_{k+1}^{i_{k+1}} \ldots x_n^{i_n}.
\]

Remark

- An oriented volume \( \mu \in \bigwedge_n(V) \) on a vector space \( V \) does not require an inner product.
- An oriented volume \( \mu \in \bigwedge_n(V) \) exists by virtue of the existence of the dual space \( V^* \).
- If \( \mu(e_1, \ldots, e_n) = 1 \), then \( \mu = \hat{e}^1 \wedge \ldots \wedge \hat{e}^n \).
- More generally, if \( \mu(e_1, \ldots, e_n) = 1 \), then \( \mu \cdot e_1 \ldots e_k = \hat{e}^{k+1} \wedge \ldots \wedge \hat{e}^n \).

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 \( \{ \mu = dx \wedge dy \wedge dz \} \), which, according to the previous example, is given by

\[
\mu(u, v, w) = u^1 v^2 w^3 + u^3 v^1 w^2 + u^2 v^3 w^1 - u^3 v^2 w^1 - u^1 v^3 w^2 - u^2 v^1 w^3.
\]

This is the (signed) volume of a parallelepiped spanned by the (column vectors corresponding to the) vectors \( u, v, w \). The result is positive (negative) if the ordered triple \( u, v, w \) has the same (respectively opposite) orientation as the basis \( \{ e_1, e_2, e_3 \} \).

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

$$\langle dx, v \rangle = v^1 \quad \langle dy, v \rangle = v^2 \quad \langle dz, v \rangle = v^3.$$  

One recognizes the (signed) 1-dimensional volumes (lengths) of the parallel projections of the vector $v$ 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

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

This is clearly the (signed) 2-dimensional volume (surface area) of the parallelogram spanned by $v$ and $w$ 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](image.png)

Figure 2.1: Illustration of $(dx \wedge dy) (v, w)$ as the area of the cast shadow of a solar panel onto the ground ($x$-$y$-plane) when lit from above (along $z$-direction). Note that $(v \wedge w) (dx, dy) = (dx \wedge dy) (v, w)$ 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 $G$, also allows us to establish one-to-one relationships between $T^p_q(V)$ and $T^r_s(V)$, 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 $G$ as an element of $T^0_2(V)$, as previously discussed. We may consider the following four equivalent alternatives:

- $G : V \times V \rightarrow \mathbb{R} : (v, w) \mapsto G(v, w)$;
- $H : V^* \times V^* \rightarrow \mathbb{R} : (\tilde{v}, \tilde{w}) \mapsto H(\tilde{v}, \tilde{w}) \overset{\text{def}}{=} G(\tilde{v}, \tilde{w})$;
- $I : V^* \times V \rightarrow \mathbb{R} : (\tilde{v}, v) \mapsto I(\tilde{v}, v) \overset{\text{def}}{=} G(\tilde{v}, v)$;
- $J : V \times V^* \rightarrow \mathbb{R} : (v, \tilde{w}) \mapsto J(v, \tilde{w}) \overset{\text{def}}{=} G(v, \tilde{w})$.

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

- $G = g_{ij} \hat{e}^i \otimes \hat{e}^j = \hat{e}^i \otimes \hat{e}^j$;
- $H = h^{ij} e_i \otimes e_j = \hat{e}^i \otimes e_i = e_i \otimes \hat{e}^i$;
- $I = \delta^i_j e_i \otimes \hat{e}^j = e_i \otimes \hat{e}^i$;
- $J = \delta^i_j e_i \otimes e_j = \hat{e}^i \otimes e_i = \hat{e}^i \otimes \hat{e}^i$.

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, $I \neq J$, of equal rank $\binom{3}{1}$. Note that $I(\tilde{v}, v) = J(v, \tilde{w})$.

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 $\sharp T \in T^p_0(V)$ of $T \in T^0_q(V)$ as follows:

$$\sharp T(v_1, \ldots, v_{p+q}) = T(\sharp v_1, \ldots, \sharp v_p, v_{p+1}, \ldots, v_{p+q}) \text{ for all } v_1, \ldots, v_{p+q} \in V.$$  

Likewise, the contravariant representation $\flat T \in T^p_0(V)$ of $T \in T^0_q(V)$ is defined as follows:

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

In general, a mixed representation of $T \in T^p_0(V)$ 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 $T \in T^1_2(V)$ is given relative to a basis $\{e_i\}$ of $V$ by $T = t^i_{jk} e_i \otimes \hat{e}^j \otimes \hat{e}^k$. Let $\hat{u} = u_i \hat{e}^i, \hat{v} = v_j \hat{e}^j, \tilde{w} = w_k \hat{e}^k, u = u^i e_i, v = v^j e_j, w = w^k e_k$. Then $\sharp T = t^i_{jk} \hat{e}^i \otimes \hat{e}^j \otimes \hat{e}^k$ and $\flat T = t^{ij}_{jk} e_i \otimes e_j \otimes e_k$. In terms of components:

$$\sharp T(u, v, w) = T(\sharp u, \sharp v, w) = t^i_{jk} u_i v^j w^k,$$

$$\flat T(\hat{u}, \hat{v}, \tilde{w}) = T(\hat{u}, \hat{v}, \tilde{w}) = t_{ijk} u^i v^j w^k. \quad (2.1)$$

The remaining five mixed representations are

$$T_1(\hat{u}, \hat{v}, w) = T(\hat{u}, \hat{v}, w) = t^i_{jk} u_i v_j w^k,$$

$$T_2(\hat{u}, v, \tilde{w}) = T(\hat{u}, v, \tilde{w}) = t^i_{jk} u_i v^j w^k,$$

$$T_3(\hat{u}, v, \tilde{w}) = T(\hat{u}, v, \tilde{w}) = t_{ijk} u^i v_j w^k,$$

$$T_4(u, \hat{v}, \tilde{w}) = T(u, \hat{v}, \tilde{w}) = t_{ijk} u^i v^j w^k,$$

$$T_5(u, \hat{v}, \tilde{w}) = T(u, \hat{v}, \tilde{w}) = t_{ijk} u^i v^j w^k. \quad (2.2)$$

making up the total of $2^3 = 8$ different representations of the rank-3 tensor $T$, 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, $T = T_{000}$, $\sharp T = T_{100}$, $\flat T = T_{011}$, $T_1 = T_{010}$, $T_2 = T_{001}$, $T_3 = T_{111}$, $T_4 = T_{110}$, $T_5 = T_{101}$.

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: $T = T_{100}$, $\sharp T = T_{000}$, $\flat T = T_{111}$, $T_1 = T_{110}$, $T_2 = T_{101}$, $T_3 = T_{011}$, $T_4 = T_{010}$, $T_5 = T_{001}$.

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

Notation

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

Observation

- The invocation of the $\sharp$-operator to a vector argument of the tensor is reflected in the holor as an index lowering.
- The invocation of the $\flat$-operator to a dual vector argument of the tensor is reflected in the holor as an index raising.

Caveat

- Recall that $\sharp e_i = G(e_i) = g_{ij}\hat{e}^j$, respectively $\flat\hat{e}^i = G^{-1}(\hat{e}^i) = g^{ij}e_j$. 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 \( \{ e_i \} \) of \( V \).
- A linear space of mixed tensors, \( T^p_q(V) \).

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 tensor transformation law governing their holors. Here we present it essentially as a result.

**Result 2.8** Let \( f_j = A^i_j e_i \) 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

\[
T = t^{i_1 \ldots i_p}_{j_1 \ldots j_q} e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \ldots \otimes \hat{e}^{j_q} = t^{i_1 \ldots i_p}_{j_1 \ldots j_q} f_{i_1} \otimes \ldots \otimes f_{i_p} \otimes \hat{f}^{j_1} \otimes \ldots \otimes \hat{f}^{j_q},
\]

iff the holor adheres to the tensor transformation law,

\[
t^{i_1 \ldots i_p}_{j_1 \ldots j_q} = A^i_{\ell_1} \ldots A^i_{\ell_p} B^k_{j_1} \ldots B^k_{j_q} t^{k_1 \ldots k_p}_{\ell_1 \ldots \ell_q}.
\]

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

\[
\delta^i_j \overset{\text{def}}{=} (\hat{f}^i, f_j) = (C^i_j \hat{e}^k, A^i_j e_k) \overset{\text{def}}{=} C^i_j A^i_j \delta^k_k = C^i_j A^i_j,
\]

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

\[
e_i = B^i_f f_f \quad \text{and} \quad \hat{e}^i = A^i_k \hat{f}^k.
\]

Substitution into the basis expansion of \( T \) 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 \( v = v^i e_i = \pi^i f_i \):

\[
\pi^i = B^i_j v^j,
\]

also referred to as \textit{vector transformation law}. Recall Theorem 2.1 on page 7.
Example
Let \( df = \partial_i f \hat{e}^i \in V^* \) 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 \( v = v^i e_i \in V \), then
\[
\langle df, v \rangle = \partial_i f v^i = (\nabla f | v),
\]
the directional derivative of \( f \) along \( v \). 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 \( df = \partial_i f \hat{e}^i \), with \( \hat{e}^i = B^i_j \hat{e}^j \), and
\[
\partial_i f = A^\ell_i \partial_\ell f.
\]
Notice the analogy with “infinitesimal calculus”. Indeed, a change of coordinates, \( x \mapsto \pi = \phi(x) \), induces a reparametrization \( df = \partial_i f dx^i = \bar{\partial}_i f d\pi^i \), in which
\[
\bar{\partial}_i f = \frac{\partial f}{\partial \pi^i} = \frac{\partial f}{\partial x^\ell} \frac{\partial x^\ell}{\partial \pi^i} = A^\ell_i \partial_\ell f \quad \text{with} \quad A^\ell_i \overset{\text{def}}{=} \frac{\partial x^\ell}{\partial \pi^i},
\]
by the chain rule.

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

Example
Recall the observation on page 10, which confirms the tensor transformation law for the dual basis vectors \( dx^i \):
\[
dx^i = \frac{\partial \pi^i}{\partial x^j} dx^j = B^i_j dx^j,
\]
with \( B = A^{-1} \), cf. the previous example.

Caveat
Compare the latter example of the transformation law for a basis dual vector, \( dx^i = B^i_j dx^j \), to the one on page 37 for the components (holor) of a vector, \( v^i = B_i^j v^j \). The apparent formal correspondence has led practitioners of tensor calculus (mostly physicists and engineers) to interpret the dual basis vector \( dx^i \) 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.

2.8.6.2. “Relative Tensors”

Recall the unit volume form introduced in section 2.8.4. On page 31 it was anticipated with a modest amount of foresight that the holor \( \mu_{i_1 \ldots i_n} = [i_1, \ldots, i_n] \) depends on the chosen basis. 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 basis vectors. Let us scrutinize the situation more carefully.
Ansatz

- An $n$-dimensional vector space $V$ over $\mathbb{R}$.
- A fixed basis $\{e_i\}$ of $V$.
- The unit volume form $\mu = \mu_{[i_1, \ldots, i_n]} \hat{e}^{i_1} \wedge \cdots \wedge \hat{e}^{i_n} = \mu_{i_1 \ldots i_n} \hat{e}^{i_1} \otimes \cdots \otimes \hat{e}^{i_n} \in \Lambda^n(V)$, with $\mu_{i_1 \ldots i_n} = [i_1, \ldots, i_n]$. The following lemma allows us to compute the holor of $\mu$ relative to an arbitrary basis, given its canonical form $[i_1, \ldots, i_n]$ relative to a fiducial basis, via Result 2.8, page 37.

**Lemma 2.1** Let $\mu = \mu_{i_1 \ldots i_n} \hat{e}^{i_1} \otimes \cdots \otimes \hat{e}^{i_n} \in \Lambda^n(V)$, with $\mu_{i_1 \ldots i_n} = [i_1, \ldots, i_n]$ relative to basis $\{e_i\}$ of $V$, and $f_j = A^j_i e_i$, so that $\hat{e}^i = A^i_j \hat{f}_j$. Setting $\mu = \bar{\mu}_{i_1 \ldots i_n} \hat{f}^{i_1} \otimes \cdots \otimes \hat{f}^{i_n}$, we have

$$\bar{\mu}_{j_1 \ldots j_n} = A_{j_1}^{i_1} \cdots A_{j_n}^{i_n} \mu_{i_1 \ldots i_n} = \det A \mu_{j_1 \ldots j_n}.$$  

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

$$\mu = \mu_{i_1 \ldots i_n} \hat{e}^{i_1} \otimes \cdots \otimes \hat{e}^{i_n} = \mu_{i_1 \ldots i_n} A_{j_1}^{i_1} \cdots A_{j_n}^{i_n} \hat{f}^{j_1} \otimes \cdots \otimes \hat{f}^{j_n}.$$  

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

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

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 $\mu (e_1, \ldots, e_n) = 1$, then $\mu (f_1, \ldots, f_n) = \det A$.
- The holor $\mu_{i_1 \ldots i_n} = [i_1, \ldots, i_n]$ is not invariant under the absolute tensor transformation law, Result 2.8, in the sense that $\bar{\mu}_{i_1 \ldots i_n} = \det A \mu_{i_1 \ldots i_n} \neq [i_1, \ldots, i_n] = \mu_{i_1 \ldots i_n}$.

Heuristics

Clearly, the constant antisymmetric symbol fails to define an invariant holor of a 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 $\{e_i\}$ has a preferred status among all possible bases. The question 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. There are two ways to remedy the problem:

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

$$\mu = \mu_{[i_1, \ldots, i_n]} \hat{e}^{i_1} \wedge \cdots \wedge \hat{e}^{i_n} = \bar{\mu}_{[i_1, \ldots, i_n]} \hat{f}^{i_1} \wedge \cdots \wedge \hat{f}^{i_n},$$  

and, by definition,

$$\bar{\mu}_{i_1 \ldots i_n} \overset{\text{def}}{=} \left(\det A\right)^{-1} A_{j_1}^{i_1} \cdots A_{j_n}^{i_n} \mu_{j_1 \ldots j_n},$$  

then $\mu_{i_1 \ldots i_n} = \bar{\mu}_{i_1 \ldots i_n} = [i_1, \ldots, i_n]$ 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, $\mu$ 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

$$
\overline{\mu}_{i_1...i_n} \overset{\text{def}}{=} (\det A)^{-1} A^{j_1}_{i_1} ... A^{j_n}_{i_n} \mu_{j_1...j_n} = (\det A)^{-1} A^{j_1}_{i_1} ... A^{j_n}_{i_n} [j_1, ..., j_n] \\
= (\det A)^{-1} [i_1, ..., i_n] \det A = [i_1, ..., i_n] = \mu_{i_1...i_n}.
$$

Remark

- In Lemma 2.1 on page 39 the holor $\mu_{i_1...i_n} = [i_1, ..., i_n]$ is transformed according to the absolute tensor transformation law, recall Result 2.8 on page 37, which implies that $\overline{\mu}_{i_1...i_n} = \det A \mu_{i_1...i_n} \neq \mu_{i_1...i_n}$.
- In Result 2.9 the holor $\mu_{i_1...i_n} = [i_1, ..., i_n]$ is transformed according to the relative tensor transformation law, rendering the holor $\mu_{i_1...i_n} = [i_1, ..., i_n]$ invariant in the sense that $\overline{\mu}_{i_1...i_n} = [i_1, ..., i_n]$ regardless of the basis.
- Adopting the relative tensor transformation law allows us to interpret $\mu_{i_1...i_n}$ as a 'covariant alias' of the permutation symbol $[i_1, ..., i_n]$, 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

$$
\mu = \mu_{i_1...i_n} \hat{e}^{i_1} \otimes ... \otimes \hat{e}^{i_n} = \overline{\mu}_{i_1...i_n} \hat{f}^{i_1} \otimes ... \otimes \hat{f}^{i_n},
$$

or, equivalently,

$$
\mu = \mu_{[i_1...i_n]} \hat{e}^{i_1} \wedge ... \wedge \hat{e}^{i_n} = \overline{\mu}_{[i_1...i_n]} \hat{f}^{i_1} \wedge ... \wedge \hat{f}^{i_n}.
$$

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 $\overline{\mu}_{i_1...i_n}$.

- Henceforth we will adopt $\mu_{i_1...i_n}$ and $\mu^{i_1...i_n}$ as covariant and contravariant aliases for the permutation symbol $[i_1, ..., i_n]$, i.e.

$$
\mu_{i_1...i_n} \equiv \mu^{i_1...i_n} \equiv [i_1, ..., i_n].
$$

- Consequently, the holor $\mu_{i_1...i_n}$ 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 $\{e_i\}$ of $V$.
- The unit volume form $\mu = \mu_{[i_1...i_n]} \hat{e}^{i_1} \wedge ... \wedge \hat{e}^{i_n} = \mu_{i_1...i_n} \hat{e}^{i_1} \otimes ... \otimes \hat{e}^{i_n} \in \Lambda_n(V)$, with invariant holor definition $\mu_{i_1...i_n} = [i_1, ..., i_n]$. 

Notation
For brevity one denotes the determinant of the Gram matrix, \( g_{ij} = (e_i | e_j) \), 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:

\[ \epsilon = \sqrt{g} \; \mu = \sqrt{g} \; e^1 \wedge \ldots \wedge e^n = \epsilon_{[i_1 \ldots i_n]} e^{i_1} \wedge \ldots \wedge e^{i_n}. \]

Observation

- \( \epsilon_{i_1 \ldots i_n} = \sqrt{g} [i_1, \ldots, i_n] = \sqrt{g} \mu_{i_1 \ldots i_n} \).
- \( \epsilon = \epsilon_{i_1 \ldots i_n} e^{i_1} \wedge \ldots \wedge e^{i_n} = \epsilon_{[i_1 \ldots i_n]} e^{i_1} \wedge \ldots \wedge e^{i_n} \) in terms of the appropriate summation conventions.

**Theorem 2.6** The contravariant representation of the Levi-Civita tensor, \( \epsilon \in \bigwedge^n(V) \), recall Definition 2.24 on page 35, is given by \( \epsilon = \epsilon_{[i_1 \ldots i_n]} e_{i_1} \wedge \ldots \wedge e_{i_n} \), with holor representation

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

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

\[ e^{i_1 \ldots i_n} = g^{i_1 j_1} \ldots g^{i_n j_n} \epsilon_{j_1 \ldots j_n} = \epsilon_{[i_1 \ldots i_n]} e^{i_1} \wedge \ldots \wedge e^{i_n} \]

In the last step we have made use of the fact that \( g^{i_1 j_1} \ldots g^{i_n j_n} [j_1, \ldots, j_n] = \det g^{ij} [i_1, \ldots, i_n] = \det g^{ij} 1/ \det g_{ij} \).

Observation

- \( e_{i_1 \ldots i_n} = \frac{1}{\sqrt{g}} \, [i_1, \ldots, i_n] \) relative to the fiducial basis \( \{ e_i \} \).
- \( \epsilon = \epsilon_{i_1 \ldots i_n} e_{i_1} \wedge \ldots \wedge e_{i_n} = \epsilon_{[i_1 \ldots i_n]} e_{i_1} \wedge \ldots \wedge e_{i_n} \).

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

\[ \epsilon = \epsilon_{[i_1 \ldots i_n]} e^{i_1} \wedge \ldots \wedge e^{i_n} \]

relative to basis \( \{ e_i \} \), with \( \epsilon_{i_2 \ldots i_n} = \sqrt{g} [i_1, \ldots, i_n] \). Moreover, let

\[ \epsilon = \epsilon_{[i_1 \ldots i_n]} e^{i_1} \wedge \ldots \wedge e^{i_n} \]

relative to \( \{ f_i \} \) with \( f_i = A^j_i e_j \). Then the pseudotensor transformation law,

\[ \tau_{i_1 \ldots i_n} \overset{\text{def}}{=} \text{sgn}(\det A) A^{j_1}_{i_1} \ldots A^{j_n}_{i_n} \epsilon_{j_1 \ldots j_n}, \]

renders the holor form invariant, in the sense that \( \tau_{i_1 \ldots i_n} = \sqrt{g} [i_1, \ldots, i_n] \).

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

\[ g^*_{ij} = A^k_i A^l_j g_{kl}, \]
whence
\[ \sqrt{g} = |\det A| \sqrt{g}. \]

By definition of the pseudotensor transformation law we have, using an observation on page 5,
\[ \tau_{i_1 \ldots i_n} \overset{\text{def}}{=} \text{sgn} (\det A) A_{i_1}^{j_1} \ldots A_{i_n}^{j_n} \epsilon_{j_1 \ldots j_n} = \text{sgn} (\det A) A_{i_1}^{j_1} \ldots A_{i_n}^{j_n} \sqrt{g} \mu_{j_1 \ldots j_n} = \text{sgn} (\det A) \det A \sqrt{g} \mu_{i_1 \ldots i_n} = \sqrt{g} \mu_{i_1 \ldots i_n}. \]

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 \( \sqrt{g} \neq \sqrt{|g|} \), whence \( \epsilon_{i_1 \ldots i_n} \neq \epsilon_{i_1 \ldots i_n} \), unlike the formal identity \( \mu_{i_1 \ldots i_n} = \mu_{i_1 \ldots i_n} \equiv [i_1, \ldots, i_n] \).
- 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*
\[
\delta_{i_1 \ldots i_k}^{j_1 \ldots j_k} = \det \begin{pmatrix} \delta_{j_1}^{i_1} & \ldots & \delta_{j_k}^{i_1} \\ \vdots & \ddots & \vdots \\ \delta_{j_1}^{i_k} & \ldots & \delta_{j_k}^{i_k} \end{pmatrix} = \begin{cases} +1 & \text{if } (i_1, \ldots, i_k) \text{ even permutation of } (j_1, \ldots, j_k), \\ -1 & \text{if } (i_1, \ldots, i_k) \text{ odd permutation of } (j_1, \ldots, j_k), \\ 0 & \text{otherwise}. \end{cases}
\]

**Observation**

- \( \epsilon_{j_1 \ldots j_n} \epsilon^{i_1 \ldots i_n} = \mu_{j_1 \ldots j_n} \mu_{i_1 \ldots i_n} = \delta_{j_1 \ldots j_n}^{i_1 \ldots i_n}. \)
- \( \epsilon_{i_1 \ldots i_k} \epsilon_{j_1 \ldots j_{n-k}} \epsilon_{k_1 \ldots k_{n-k}} = \mu_{i_1 \ldots i_k} \mu_{j_1 \ldots j_{n-k}} \mu_{k_1 \ldots k_{n-k}} = k! \delta_{j_1 \ldots j_{n-k}}^{i_1 \ldots i_{n-k}}. \)
- For the matrix \( A \) of a mixed tensor \( A \in T_1^1(V) \) we have \( \det A^T_j = \frac{1}{n!} \delta_{i_1 \ldots i_n}^{j_1 \ldots j_n} A_{i_1}^{j_1} \ldots A_{i_n}^{j_n}. \)

### 2.8.7. Contractions

**Ansatz**

- An \( n \)-dimensional vector space \( V \) over \( \mathbb{R} \).
- A basis \( \{e_i\} \) of \( V \).
- A linear space of mixed tensors, \( T^p_q(V) \).
Definition 2.27 Let $T \in T^1_1(V)$. The trace, $\text{tr} \, T \in \mathbb{R}$, of $T$ is defined as follows:

$$\text{tr} \, T = T(\hat{e}^i, e_i).$$

Remark

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

Caveat

- In general, the notation $\text{tr} \, T$ is ambiguous for $T \in T^p_q(V)$, unless $T$ 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 $\{e_i\}$.
- The Levi-Civita tensor $\epsilon = \sqrt{g} \, \hat{e}^1 \wedge \ldots \wedge \hat{e}^n = \epsilon_{i_1 \ldots i_n} \hat{e}^{i_1} \wedge \ldots \wedge \hat{e}^{i_n} \in \wedge^n(V)$.

Definition 2.28 Let $a_1, \ldots, a_k \in V$. The Hodge star operator $* : \wedge^k(V) \to \wedge^{n-k}(V)$ is defined as follows:

$$\begin{cases}
*1 &= \epsilon \\
*(\sharp a_1 \wedge \ldots \wedge \sharp a_k) &= \epsilon \cdot a_1 \wedge \ldots \wedge a_k \\
&\quad \text{for } k = 1, \ldots, n,
\end{cases}$$

followed by linear extension.

This definition could be rephrased by replacing $\sharp a_1, \ldots, \sharp a_k \in V^*$ by $\hat{a}_1, \ldots, \hat{a}_k \in V^*$ on the l.h.s. and $a_1, \ldots, a_k \in V$ by $\flat a_1, \ldots, \flat a_k \in V$ on the r.h.s. A similar definition could be given for a contravariant counterpart $* : \wedge^k(V) \to \wedge^{n-k}(V)$, in which case $\sharp$-conversion of the vector arguments $a_1, \ldots, a_k \in V$ is obsolete.

Remark

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

$$\dim \wedge^k(V) = \binom{n}{k} = \binom{n}{n-k} = \dim \wedge^{n-k}(V).$$
Observation

- The holor of the tensor \( \{a_1 \wedge \ldots \wedge a_k\} \) for fixed \( a_p = a_p^i e_i \), with \( p = 1, \ldots, k \), is obtained as follows:

\[
( \{a_1 \wedge \ldots \wedge a_k\} \{ e_{i_1+1}, \ldots, e_{i_n} \} ) = a_1^1 \ldots a_k^k ( \{ e_{i_1} \wedge \ldots \wedge e_{i_k} \} ) \{ e_{i_{k+1}}, \ldots, e_{i_n} \}
\]

By the same token, expanding also \( x_q = x_q^i e_i \) for \( q = k + 1, \ldots, n \), we obtain

\[
( \{a_1 \wedge \ldots \wedge a_k\} \{ x_{k+1}, \ldots, x_n \} = \epsilon_{i_1 \ldots i_n} a_1^{i_1} \ldots a_k^{i_k} x_{k+1}^{i_k+1} \ldots x_n^{i_n}.
\]

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

\[
\left\{ \begin{array}{ll}
*1 & = \epsilon \\
(\hat{a}^1 \wedge \ldots \wedge \hat{a}^k) & = \epsilon_p \hat{a}_1^p \ldots \hat{a}_k^p \quad \text{for } k = 1, \ldots, n.
\end{array} \right.
\]

In this case we find, setting \( \hat{a}^p = a_p^j \hat{e}^j \) for \( p = 1, \ldots, k \), and using \( \hat{e}^j = g^{ij} e_i \),

\[
( \{ \hat{a}^1 \wedge \ldots \wedge \hat{a}^k \} \{ e_{i_{k+1}}, \ldots, e_{i_n} \} ) = a_1^1 \ldots a_k^k ( \{ \hat{e}^j_1 \wedge \ldots \wedge \hat{e}^j_k \} ) \{ e_{i_{k+1}}, \ldots, e_{i_n} \}
\]

\[
\hspace{1cm} = \epsilon^{j_1 \ldots j_k} g^{i_1 j_1} \ldots g^{i_k j_k} a_1^{j_1} \ldots a_k^{j_k} \epsilon_{i_1 \ldots i_n}
\]

\[
\hspace{1cm} = \epsilon_{i_1 \ldots i_n} g^{i_1 j_1} \ldots g^{i_k j_k} a_1^{j_1} \ldots a_k^{j_k}
\]

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 \( A \in \bigwedge_k(V) \), then the holor of \( *A \in \bigwedge_{n-k}(V) \) is given by

\[
* A_{i_{k+1} \ldots i_n} \overset{\text{def}}{=} (A) (e_{i_{k+1}}, \ldots, e_{i_n}) = \frac{1}{k!} \epsilon_{i_1 \ldots i_n} g^{i_1 j_1} \ldots g^{i_k j_k} A_{j_1 \ldots j_k}.
\]

In other words,

\[
* A = * A_{i_{k+1} \ldots i_n} \hat{e}^{i_k+1} \otimes \ldots \otimes \hat{e}^{i_n} = * A_{|j_{k+1} \ldots j_n} \hat{e}^{i_k+1} \wedge \ldots \wedge \hat{e}^{i_n}.
\]

In fact it suffices to remember that \( * (\{ e_{i_1} \wedge \ldots \wedge e_{i_k} \} \{ e_{i_{k+1}}, \ldots, e_{i_n} \}) = \epsilon_{i_1 \ldots i_n} \), by exploiting multilinearity.

Remark

- Taking \( k = 0 \) in the above observation is consistent with \( \*1 = \epsilon \).

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

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 $y = x + v$.

Caveat
Both $E$ and $V$ are usually written as $\mathbb{R}^n$. It should be clear from the context whether a Euclidean space or a Euclidean vector space is meant.

A point in $x \in E \sim \mathbb{R}^n$ can be represented by its Cartesian coordinates $x^i \in \mathbb{R}$ relative to a (globally defined) Cartesian coordinate system or standard basis $\{e_i\}$, $i = 1, \ldots, n$, in the usual way: $x = x^i e_i$. Here $e_i \in \mathbb{R}^n$ 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. $\Omega_\alpha$ 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 $\phi_\alpha \circ \phi_\beta^{-1} \bigg|_{\phi_\beta(\Omega_\alpha \cap \Omega_\beta)}$ 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 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

$$
\begin{align*}
\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(\Omega_\alpha \cap \Omega_\beta) \subset \mathbb{R}^n &\to \phi_\beta(\Omega_\alpha \cap \Omega_\beta) \subset \mathbb{R}^n : x \mapsto y = (\phi_\beta \circ \phi_\alpha^{-1})(x) \quad (3.1) \\
\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(\Omega_\alpha \cap \Omega_\beta) \subset \mathbb{R}^n &\to \phi_\alpha(\Omega_\alpha \cap \Omega_\beta) \subset \mathbb{R}^n : y \mapsto x = (\phi_\alpha \circ \phi_\beta^{-1})(y) \quad (3.2)
\end{align*}
$$

are called coordinate transformations.

**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 $p = \phi_\alpha^{-1}(x) = \phi_\beta^{-1}(y) \in M$. 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

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

**Observation**

The matrices $S$ and $T$ are regular, with $S = T^m$ 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 $x \in M$ of the base manifold, with $\dim V = \dim M = n$. For reason of clarity we write $TM_x$ for a local copy of $V$ attached to $x$, or, by abuse of notation, $TM_x$ if $x = \phi_\alpha(x) \in \mathbb{R}^n$ 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 $\gamma(0) = p \in M$, and $f \in C^\infty(M)$ an arbitrary smooth scalar field. Then the tangent vector of $\gamma$ at $p$ is the linear map given by

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

Smoothness of $\gamma$ and $f$ means that the numerically valued functions $\tau = \phi_\alpha \circ \gamma$ and $\tilde{f} = f \circ \phi_\alpha^{-1}$ 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 $p \in U_\alpha$, in the sense of standard calculus, i.e. $\tau^i \in C^\infty((-\varepsilon, \varepsilon))$ for each component $i = 1, \ldots, n$, and $\tilde{f} \in C^\infty(\phi_\alpha(U_\alpha))$.

Now recall the observation on page 8 concerning the abstract notation of basis vectors as a 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 $x = \phi_\alpha(x) \in \phi_\alpha(U_\alpha) \subset \mathbb{R}^n$, we observe that

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

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

$$\tau : (-\varepsilon, \varepsilon) \subset \mathbb{R} \to \mathbb{R}^n : t \mapsto x(t) \overset{\text{def}}{=} \tau(t), \quad (3.4)$$

$$\tilde{f} : \phi_\alpha(U_\alpha) \subset \mathbb{R}^n \to \mathbb{R} : x \mapsto \tilde{f}(x) \overset{\text{def}}{=} (f \circ \phi_\alpha^{-1})(x), \quad (3.5)$$

and of the convenient definition ($*$)

$$\frac{\partial}{\partial x^i} \bigg|_{x(0)} \tilde{f} = \frac{\partial}{\partial x^i} \bigg|_{\gamma(0)} f. \quad (3.6)$$

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

$$v^i_p = \left. \frac{\partial}{\partial x^i} \right|_{\gamma(0)} f,$$

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

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

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

In particular, if we take $f^i : M \to \mathbb{R} : x \mapsto x^i$ and $v_j = \partial_j \in TM_x$ in an arbitrary coordinate chart $x = \phi_\alpha(x)$, then $v_j f^i = \partial_j x^i = \delta^i_j = dx^i(\partial_j) = df^i(v_j)$, consistent with the formal notation for dual basis vectors $dx^i$ introduced on page 8.

**Result 3.1** Recall the Kronecker tensor definition in Section 2.3. Relative to a coordinate basis we have, at any implicit point $p \in M$,

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

### 3.4. Tangent and Cotangent Bundle

**Ansatz**
An $n$-dimensional manifold $M$.

**Definition 3.5** The tangent space at $x \in M$, $TM_x$, is the vector space of tangent vectors $v_x$ at point $x \in M$.

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

**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 $e_i$ is also often written as $\partial_i$, recall the observation on page 8.

**Definition 3.6** The tangent bundle, $TM = \bigcup_{x \in M} TM_x$, is the union of tangent spaces over all $x \in M$.

**Observation**

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

**Notation**

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

### 3.5. Exterior Derivative

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

3.6 Affine Connection

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

\[
d \omega^k = d \omega_{i_1...i_k} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} \in \bigwedge_{k+1}(TM),
\]

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

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

**Observation**
One readily verifies the following properties:

- \( d \omega^n = 0 \) for any \( \omega^n \in \bigwedge_n(TM) \).
- \( d^2 \omega^k \overset{\text{def}}{=} d d \omega^k = 0 \) for any \( \omega^k \in \bigwedge_k(TM) \) 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, \( \alpha = \alpha_1 dx + \alpha_2 dy + \alpha_3 dz \in T^*M \), with \( \alpha_i \in C^\infty(M) \) and \( \text{dim} \, M = 3 \). 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:

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

Special cases arise if \( \alpha = df \) for some \( f \in C^\infty(M) \), 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, \( v = v^i \partial_i \). Also recall from infinitesimal calculus that for a scalar function \( f \in C^\infty(M) \) we may define the differential \( df = \partial_i f \, dx^i \in T^*M \). 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 $\nu = v^i \partial_i \in TM$. Then

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

So, when acting on a scalar field it is tacitly understood that we make the formal identification $\nu = \langle d \cdot, \nu \rangle$.

Definition 3.9 An affine connection on a manifold $M$ is a mapping $\nabla : TM \times TM \to TM : (\nu, \omega) \mapsto \nabla_\nu \omega$, also referred to as the covariant derivative of $\omega$ with respect to $\nu$, with the following properties. Let $f, f_1, f_2 \in C^\infty(M)$, $\nu, \nu_1, \nu_2, \omega, \omega_1, \omega_2 \in TM$, and $\lambda, \lambda_1, \lambda_2 \in \mathbb{R}$, then

1. $\nabla_{f_1 \nu_1 + f_2 \nu_2} \omega = f_1 \nabla_{\nu_1} \omega + f_2 \nabla_{\nu_2} \omega$ for all $f_1, f_2 \in C^\infty(M)$,

2. $\nabla_\nu (\lambda_1 \omega_1 + \lambda_2 \omega_2) = \lambda_1 \nabla_\nu \omega_1 + \lambda_2 \nabla_\nu \omega_2$, for all $\lambda_1, \lambda_2 \in \mathbb{R}$,

3. $\nabla_\nu (f \omega) = f \nabla_\nu \omega + \nabla_\nu \omega f$ for all $f \in C^\infty(M)$. Here we have defined $\nabla_\nu f \equiv v f$.

Remark

If we leave out the directional vector argument we call $\nabla \omega$ the covariant differential of $\omega$, 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 $\Gamma^k_{ij}$ are defined by $\nabla_{e_j} e_i = \Gamma^k_{ij} e_k$.

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:

$$e^k = S^k_i e^i \quad \text{resp.} \quad e_j = T^m_j e_m .$$

The matrices $T$ and $S$ are the mutually inverse Jacobian matrices:

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

Set

$$\Gamma^k_{ij} = \langle e^k, \nabla_{e_j} e_i \rangle$$

Substitute the transformations, apply first and third rule of Definition 3.9, $\nabla_{e_j} = T^j_i \nabla_{e_i}$, respectively $\nabla_{e_i} (T^j_i e_k) = \partial_i T^j_k e_k + T^j_i \nabla_{e_i} e_k$, use the definition of the Christoffel symbols, $\nabla_{e_i} e_k = \Gamma^l_{ik} e_m$, and finally the chain rule, $\partial_j = T^j_i \partial_i$, to obtain the following result:

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

$$\Gamma^k_{ij} = S^k_i (T^m_j T^l_i \Gamma^l_{nm} + \bar{\partial}_j T^l_i) .$$
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 \( \Gamma^k_{ij} = 0 \) 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 \( v = v^i e_i, \ w = w^i e_i \in \text{TM} \) be two vector fields. Then

\[
\nabla_v w = v^i D_i w^j e_j.
\]

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 \( j^{th} \) component of \( w \), 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, \( T \in \text{T}^p_q(\text{TM}) \) and \( S \in \text{T}^r_s(\text{TM}) \), say, we require

\[
\nabla_v (S \otimes T) = \nabla_v S \otimes T + S \otimes \nabla_v T.
\]

**Observation**

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

\[
\nabla_v (\hat{z}, w) = \nabla_v (\text{tr} (\hat{z} \otimes v)) = \text{tr} (\nabla_v (\hat{z} \otimes v)) = \text{tr} (\nabla_v \hat{z} \otimes v) + \text{tr} (\hat{z} \otimes \nabla_v v) = \langle \nabla_v \hat{z}, w \rangle + \langle \hat{z}, \nabla_v w \rangle,
\]

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 \( T = T^{i_1 \ldots i_p}_{j_1 \ldots j_q} e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \ldots \otimes \hat{e}^{j_q} \) and \( v = v^i e_i \) be a tensor, respectively vector field on \( M \), then

\[
\nabla_v T = v^i D_i T^{i_1 \ldots i_p}_{j_1 \ldots j_q} e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes \hat{e}^{j_1} \otimes \ldots \otimes \hat{e}^{j_q}.
\]

Here we have defined the components of the covariant derivative

\[
D_i T^{i_1 \ldots i_p}_{j_1 \ldots j_q} \equiv \partial_i T^{i_1 \ldots i_p}_{j_1 \ldots j_q} + \Gamma^k_{ki} T^{i_1 \ldots i_p \ldots i_k}_{j_1 \ldots j_q} + \ldots + \Gamma^k_{ki} T^{i_1 \ldots i_p \ldots i_{k-1} k}_{j_1 \ldots j_q} - \Gamma^k_{kj} T^{i_1 \ldots i_p \ldots i_k}_{j_1 \ldots j_q} - \ldots - \Gamma^k_{kj} T^{i_1 \ldots i_p \ldots j_k}_{j_1 \ldots j_q}.
\]

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 ‘\( \Gamma \)-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 TM$^x$ is an inner product space for every $x \in M$, local bases are chosen in such a way that the coefficients of the Gram matrix, $g_{ij}(x) = (e_i(e_j)_x)$, vary smoothly across the manifold: $g_{ij} \in C^\infty(M)$. (This is guaranteed if the basis is holonomic.)

- Similar assumptions will henceforth be implicit in the case of tensor fields in general.

Notation

The 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 $w \in V$ with respect to $v \in V$ is defined as

$$\mathcal{L}_v w = [v, w],$$

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

$$[v, w] f = v(w f) - w(v f) = (v^i \partial_i w^j - w^i \partial_i v^j) \partial_j f.$$

for any smooth function $f$.

Observation

- $\mathcal{L}_v w$ is itself a vector, with components $[v, w]^i = v^j \partial_j w^i - w^j \partial_j v^i$.

- 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: $[u, [v, w]] + [w, [u, v]] + [v, [w, u]] = 0$. 
3.8. Torsion

Ansatz

- The tangent bundle $\mathcal{T}M$ over an $n$-dimensional manifold $\mathcal{M}$.
- An affine connection.

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

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

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 $\nabla_v w = v^i (\partial_i w^j + w^k \Gamma^i_{kj}) \partial_j$, from which you derive the components of the torsion tensor relative to a coordinate basis:

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

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

3.9. Levi-Civita Connection

Ansatz

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

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

Definition 3.14 Recall Definition 3.12, page 52 and Definition 3.13, page 53. A metric connection or Levi-Civita connection on a Riemannian manifold $(\mathcal{M}, g)$ satisfies the following properties:

- $g(\nabla_z (v \cdot w)) = g(\nabla_z v \cdot w + v \cdot \nabla_z w)$ for all $v, w, z \in \mathcal{T}M$.
- $\mathbf{T}(v, w) = 0$ for all $v, w \in \mathcal{T}M$.

Observation
The first requirement is tantamount to the product rule: $\nabla_z (v \cdot w) = \nabla_z v \cdot w + v \cdot \nabla_z w$ for all $v, w, z \in \mathcal{T}M$. 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 condensed written as $\nabla_z g = 0$ for all $z \in TM$.
- 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

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

Proof of Theorem 3.1. The proof is based on (i) the general properties of an affine connection, recall Definition 3.9 on page 50, and (ii) the defining properties of the Levi-Civita connection, Definition 3.14, page 53.

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

$$\partial_m g_{ij} - \partial_i g_{jm} - \partial_j g_{mi} = g_{ij} (\Gamma^m_{jm} - \Gamma^m_{mj}) + g_{ij} (\Gamma^m_{im} - \Gamma^m_{mi}) - g_{im} (\Gamma^m_{ij} + \Gamma^m_{ji}) .$$

Subsequently, according to Definition 3.13 on page 53 and the subsequent observation on page 53, the second condition of Definition 3.14 is equivalent to

$$\Gamma^k_{ij} = \Gamma^k_{ji} .$$

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 $\nabla v \in T^1_1(TM)$ of a vector field $v = v^i \partial_i \in TM$ is defined as

$$\nabla v = \nabla_{\partial_i} v dx^k .$$

**Observation**

If we define $\nabla v = Dv^j \partial_j$, 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 $T \in T^p_q(TM)$:

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

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 \( v(t) = \dot{x}^i(t) \partial_i \) for \( t \in I = (t_0, t_1) \), then

\[
\nabla v(t) = 0 \quad \text{for all } t \in I.
\]

**Remark**

- It is tacitly understood that \( \partial_i \in T_{x(t)}M \) indicates the coordinate basis vector at point \( x(t) \).
- One often identifies \( v(x(t)) \) with \( v(t) \). 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 \( v, w \in TM \) the (antisymmetric) curvature operator, \( R(v, w) : TM \to TM \): \( u \mapsto R(v, w)u \), is defined as

\[
R(v, w) = [\nabla v, \nabla w] - \nabla_{[v,w]}.
\]

**Observation**

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

**Definition 3.18** For \( u, v, w \in TM, \tilde{z} \in T^*M \), the Riemann curvature tensor \( \text{Riemann} \in T^*M \otimes TM \otimes TM \otimes TM \) is defined as

\[
\text{Riemann}(\tilde{z}, u, v, w) = \langle \tilde{z}, R(v, w)u \rangle.
\]
Observation

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

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

- 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 $\mathcal{R}(\partial_\mu, \partial_\nu) = [\nabla_{\partial_\nu}, \nabla_{\partial_\mu}] - \nabla_{[\partial_\nu, \partial_\mu]} = [\nabla_{\partial_\nu}, \nabla_{\partial_\mu}]$, 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

$$\text{Ric}(v, w) = R_{\mu
u} v^\mu w^\nu,$$

$$\text{R} = g^{\mu\nu} R_{\mu\nu},$$

with $v, w \in T^\ast M$, in which the Ricci holor is defined as $R_{\mu\nu} = R^\rho_{\mu\nu}$.

### 3.12. Push-Forward and Pull-Back

**Ansatz**

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

**Definition 3.20**

- The pull-back $\phi^* f \in C^\infty(M)$ of $f \in C^\infty(N)$ under $\phi$ is defined such that $\phi^* f = f \circ \phi$.
- The push-forward $\phi_* v \in T^\ast N$ of $v \in T^\ast M$ under $\phi$ is defined such that $\phi_* v(f) = v(\phi^* f)$.
- The pull-back $\phi^* \nu \in T^\ast M$ of $\nu \in T^\ast N$ under $\phi$ is defined such that $(\phi^* \nu, v)_M = (\nu, \phi_* v)_N$.

**Terminology**

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

Remark

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

Observation

• By decomposing \( \phi_* v = v^\mu S_\mu^a \frac{\partial}{\partial y^a} \in TN \) it follows that the components of the push-forward are given by the Jacobian matrix of the mapping: \( S_\mu^a = \frac{\partial y^a}{\partial x^\mu} \). (Note: \( \mu = 1, \ldots, m = \dim M, a = 1, \ldots, n = \dim N \).)

• By the same token, decomposing \( \phi^* \nu = \nu_a S_\mu^a dx^\mu \in T^*M \), it follows that the components of the pull-back are given by the same Jacobian.

\[
\begin{array}{ccc}
TM & \xrightarrow{\phi_*} & TN \\
\pi \downarrow & & \downarrow \pi \\
M & \xrightarrow{\phi} & N \\
\pi \uparrow & & \uparrow \pi \\
T^*M & \xleftarrow{\phi^*} & T^*N
\end{array}
\]
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),

$$e_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

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

$$e_i = \partial_i = \frac{\partial x^j}{\partial x^i} \partial_j = \frac{\partial x^j}{\partial x^i} e_j.$$

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

$$T^\ell_i = \frac{\partial x^\ell}{\partial x^i} \quad \text{resp.} \quad S^k_\ell = \frac{\partial r^k}{\partial x^\ell},$$

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

$$T^\ell_i = \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2 + y^2}} & -y \\ \frac{y}{\sqrt{x^2 + y^2}} & x \end{pmatrix},$$

$$S^k_\ell = \begin{pmatrix} \cos \phi & \sin \phi \\ -\frac{\sin \phi}{r} & \frac{\cos \phi}{r} \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \\ -\frac{y}{\sqrt{x^2 + y^2}} & \frac{x}{\sqrt{x^2 + y^2}} \end{pmatrix}.$$

Consequently, writing “$r$” and “$\phi$” for index values 1 and 2, respectively,

$$\mathbf{e}_r = \begin{pmatrix} \cos \phi & \sin \phi \end{pmatrix} \begin{pmatrix} e_x \\ e_y \end{pmatrix} = \cos \phi e_x + \sin \phi e_y = \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix},$$

$$\mathbf{e}_\phi = \begin{pmatrix} -r \sin \phi & r \cos \phi \end{pmatrix} \begin{pmatrix} e_x \\ e_y \end{pmatrix} = -r \sin \phi e_x + r \cos \phi e_y = \begin{pmatrix} -r \sin \phi \\ r \cos \phi \end{pmatrix}.$$

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 \( T_E \) relative to which all \( \Gamma_{nm}^\ell \) vanish identically \( (\ell, m, n = 1, \ldots, 2) \). In terms of polar coordinates, \( \mathbf{x} = (r, \phi) \), we find, according to Result 3.2 on page 50:

\[
\Gamma_{ij}^k = S_{ij}^k S_{ji}^k.
\]

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

\[
\Gamma_{ir}^k = \begin{pmatrix}
\cos \phi & \sin \phi \\
-r \sin \phi & r \cos \phi
\end{pmatrix}
\begin{pmatrix}
\cos \phi & -r \sin \phi \\
\sin \phi & r \cos \phi
\end{pmatrix}
= \begin{pmatrix}
0 & 0 \\
0 & -r
\end{pmatrix},
\]

\[
\Gamma_{i\phi}^k = \begin{pmatrix}
\cos \phi & -r \sin \phi \\
\sin \phi & r \cos \phi
\end{pmatrix}
\begin{pmatrix}
\cos \phi & -r \sin \phi \\
\sin \phi & r \cos \phi
\end{pmatrix}
= \begin{pmatrix}
0 & 0 \\
1 & r
\end{pmatrix}.
\]

In other words,

\[
\Gamma_{\phi r} = \frac{1}{r}, \quad \Gamma_{r \phi} = \frac{1}{r}, \quad \Gamma_{\phi \phi} = -r, \quad \text{all other } \Gamma_{ij}^k = 0.
\]

Consequently,

\[
\nabla_r \mathbf{e}_\phi = \frac{1}{r} \mathbf{e}_\phi, \quad \nabla_\phi \mathbf{e}_r = \frac{1}{r} \mathbf{e}_\phi, \quad \nabla_\phi \mathbf{e}_\phi = -r \mathbf{e}_r, \quad \text{all other } \nabla_{\mathbf{e}_j} \mathbf{e}_i = 0.
\]

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, \( x = (x, y, \phi) \), and \( \mathbf{x} = (\xi, \eta, \theta) \), related as follows:\(^1\)

\[
\begin{align*}
\begin{cases}
x(\xi, \eta, \theta) = \xi \cos \phi + \eta \sin \phi \\
y(\xi, \eta, \theta) = -\xi \sin \theta + \eta \cos \theta \\
\phi(\xi, \eta, \theta) = \theta,
\end{cases}
\end{align*}
\]

The Jacobian matrices are given by
\[
T_i^\ell = \frac{\partial x_\ell}{\partial x^i} = \begin{pmatrix}
\cos \theta & \sin \theta & -\xi \sin \theta + \eta \cos \theta \\
-\sin \theta & \cos \theta & -\xi \cos \theta - \eta \sin \theta \\
0 & 0 & 1
\end{pmatrix},
\]
respectively
\[
S_k^\ell = \frac{\partial x_k}{\partial x^\ell} = \begin{pmatrix}
\cos \phi & -\sin \phi & -x \sin \phi - y \cos \phi \\
\sin \phi & \cos \phi & x \cos \phi - y \sin \phi \\
0 & 0 & 1
\end{pmatrix}.
\]

Again, recalling Result 3.2 on page 50: \( \Gamma^k_{ij} = S_k^e \left( T^m_j T^n_i \Gamma^e_{nm} + \Gamma^e_{ji} T^l_i \right) \). There are different ways to proceed, depending on our assumptions about the \( \Gamma^e_{nm} \)-symbols in \( (x, y, \phi) \)-space. Here we pursue the simplest option, ignoring the remark in footnote 1, and assume \( \Gamma^e_{nm} = 0 \), i.e. \( (x, y, \phi) \) 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:

\[
\begin{align*}
\Gamma^k_{i\xi} &= \begin{pmatrix}
\cos \phi & -\sin \phi & -x \sin \phi - y \cos \phi \\
\sin \phi & \cos \phi & x \cos \phi - y \sin \phi \\
0 & 0 & 1
\end{pmatrix} \frac{\partial}{\partial \xi} \begin{pmatrix}
\cos \theta & \sin \theta & -\xi \sin \theta + \eta \cos \theta \\
-\sin \theta & \cos \theta & -\xi \cos \theta - \eta \sin \theta \\
0 & 0 & 1
\end{pmatrix}, \\
\Gamma^k_{i\eta} &= \begin{pmatrix}
\cos \phi & -\sin \phi & -x \sin \phi - y \cos \phi \\
\sin \phi & \cos \phi & x \cos \phi - y \sin \phi \\
0 & 0 & 1
\end{pmatrix} \frac{\partial}{\partial \eta} \begin{pmatrix}
\cos \theta & \sin \theta & -\xi \sin \theta + \eta \cos \theta \\
-\sin \theta & \cos \theta & -\xi \cos \theta - \eta \sin \theta \\
0 & 0 & 1
\end{pmatrix}, \\
\Gamma^k_{i\theta} &= \begin{pmatrix}
\cos \phi & -\sin \phi & -x \sin \phi - y \cos \phi \\
\sin \phi & \cos \phi & x \cos \phi - y \sin \phi \\
0 & 0 & 1
\end{pmatrix} \frac{\partial}{\partial \theta} \begin{pmatrix}
\cos \theta & \sin \theta & -\xi \sin \theta + \eta \cos \theta \\
-\sin \theta & \cos \theta & -\xi \cos \theta - \eta \sin \theta \\
0 & 0 & 1
\end{pmatrix}.
\end{align*}
\]

A tedious but straightforward computation then reveals that
\[
\begin{align*}
\Gamma^k_{i\xi} &= \begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & -1 \\
0 & 0 & 0
\end{pmatrix},
\Gamma^k_{i\eta} &= \begin{pmatrix}
0 & 0 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix},
\Gamma^k_{i\theta} &= \begin{pmatrix}
0 & 1 & -\xi \\
-1 & 0 & -\eta \\
0 & 0 & 0
\end{pmatrix},
\end{align*}
\]

i.e. all \( \Gamma^k_{ij} = 0 \), except \( \Gamma^\eta_{\xi\phi} = -1, \Gamma^\xi_{\eta\phi} = 1, \Gamma^\eta_{\eta\phi} = -1, \Gamma^\xi_{\xi\phi} = 1, \Gamma^\eta_{\xi\theta} = -\xi, \Gamma^\xi_{\eta\theta} = -\eta \). Note that \( \Gamma^k_{ij} - \Gamma^k_{ji} = 0 \), 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 \( S = \{ \phi \in [0, 2\pi) \ mod \ 2\pi \} \) as an independent dimension. The parametrization can then be seen as a "shift-twist" transformation, whereby the Cartesian basisvectors \( e_x \) and \( e_y \) of the underlying Euclidean 2-space \( E \) (at any fiducial point) are rotated in accordance to the periodic translation in \( e_{\phi} \)-direction (generating \( S \)).
Example
In the spirit of the previous example we consider a 3-space $M$, with coordinates $(x, y, \phi)$. However, we now introduce the following Riemannian metric $g \in T^*M \otimes T^*M$:

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

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

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

The induced Levi-Civita connection in $(x, y, \phi)$-coordinates (and therefore also in $(\xi, \eta, \theta)$-coordinates) follows from Theorem 3.1, page 54. In matrix notation we have for the metric and its dual, respectively,

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

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

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

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 $2 \times 2$ block in the upper left part:

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

Finally, the (dual) metric coefficients do not depend on $\phi$. 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 $\mathbf{r} = (r, \psi, \theta)$ instead of $\mathbf{x} = (x, y, \phi)$:

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

To this end we once again invoke Result 3.2, page 50:

$$
\Gamma^k_{ij} = S^k_{\ell} (T^m_j T^n_i \Gamma^\ell_{mn} + \partial_j T^\ell_i).
$$

The Jacobian matrices are given by

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

respectively

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


