Multilinear Algebra and Applications

July 15, 2014.

Contents

Chapter 1. Introduction 1

Chapter 2. Review of Linear Algebra 5
2.1. Vector Spaces and Subspaces 5
2.2. Bases 7
2.3. The Einstein convention 10
2.3.1. Change of bases, revisited 12
2.3.2. The Kronecker delta symbol 13
2.4. Linear Transformations 14
2.4.1. Similar matrices 18
2.5. Eigenbases 19

Chapter 3. Multilinear Forms 23
3.1. Linear Forms 23
3.1.1. Definition, Examples, Dual and Dual Basis 23
3.1.2. Transformation of Linear Forms under a Change of Basis 26
3.2. Bilinear Forms 30
3.2.1. Definition, Examples and Basis 30
3.2.2. Tensor product of two linear forms on V 32
3.2.3. Transformation of Bilinear Forms under a Change of Basis 33
3.3. Multilinear forms 34
3.4. Examples 35
3.4.1. A Bilinear Form 35
3.4.2. A Trilinear Form 36
3.5. Basic Operation on Multilinear Forms 37

Chapter 4. Inner Products 39
4.1. Definitions and First Properties 39
4.1.1. Correspondence Between Inner Products and Symmetric Positive Definite Matrices 40
4.1.1.1. From Inner Products to Symmetric Positive Definite Matrices 42
4.1.1.2. From Symmetric Positive Definite Matrices to Inner Products 42
4.1.2. Orthonormal Basis 42
4.2. Reciprocal Basis 46
4.2.1. Properties of Reciprocal Bases 48
4.2.2. Change of basis from a basis $\mathcal{B}$ to its reciprocal basis $\mathcal{B}^g$ 50

4.2.3. Isomorphisms Between a Vector Space and its Dual 52
4.2.4. Geometric Interpretation 54

Chapter 5. Tensors 57
5.1. Generalities 57
5.1.1. Canonical isomorphism between V and $(V^*)^*$. 57
5.1.2. Towards general tensors 58
5.1.3. Tensor product of $(1,0)$-tensors on V^* 59
5.1.4. Components of a $(2,0)$-tensor and their contravariance 60
5.2. Tensors of type (p,q) 61
5.3. Tensor product 61

Chapter 6. Applications 65
6.1. Inertia tensor 65
6.1.1. Moment of inertia with respect to the axis determined by the angular velocity 65
6.1.2. Moment of inertia about any axis through the fixed point. 70
6.1.3. Moment of inertia with respect to an eigenbasis of the inertia tensor 72
6.1.4. Angular momentum 73
6.2. Stress tensor (Spannung) 75
6.2.1. Special forms of the stress tensor (written with respect to an orthonormal eigenbasis or another special basis) 80
6.2.2. Contravariance of the stress tensor 82
6.3. Strain tensor (Verzerrung) 83
The antisymmetric case 84
The symmetric case 85
6.3.1. Special forms of the strain tensor 87
6.4. Elasticity tensor 87
6.5. Conductivity tensor 88
6.5.1. Electrical conductivity 88
6.5.2. Heat conductivity 90

Chapter 7. Solutions 93

CHAPTER 1

Introduction

The main protagonists of this course are **tensors** and **multilinear maps**, just like the main protagonists of a Linear Algebra course are vectors and linear maps.

Tensors are geometric objects that describe linear relations among objects in space, and are represented by *multidimensional arrays of numbers*:

The indices can be upper or lower or, in tensor of order at least 2, some of them can be upper and some lower. The numbers in the arrays are called **components** of the tensor and give the representation of the tensor *with respect to a given basis*.

There are two natural questions that arise:

(1) Why do we need tensors?
(2) What are the important features of tensors?

(1) Scalars are no enough to describe directions, for which we need to resort to vectors. At the same time, vectors might not be enough, in that they lack the ability to "modify" vectors.

EXAMPLE 1.1. We denote by $\mathbf{B}$ the magnetic fluid density measured in Volt $\cdot$ sec/m^2 and by $\mathbf{H}$ the megnetizing intensity measured in Amp/m. They are related by the formula

$$\mathbf{B} = \mu \mathbf{H}\,,$$

where μ is the permeability of the medium in H/m. In free space, $\mu = \mu_0 = 4\pi \times 10^{-7}$ H/m is a scalar, so that the flux density and the magnetization are vectors differ only by their magnitude.

Other material however have properties that make these terms differ both in magnitude and direction. In such materials the scalar permeability is replaced by the tensor permeability μ and

$$\mathbf{B} = \boldsymbol{\mu} \cdot \mathbf{H}\,.$$

Being vectors, $\mathbf{B}$ and $\mathbf{H}$ are tensors of order 1, and $\boldsymbol{\mu}$ is a tensor of order 2. We will see that they are of different type, and in fact *the order of* $\mathbf{H}$ *"cancels out" with the order of* $\boldsymbol{\mu}$ *to give a tensor of order 1.* □

(2) Physical laws do not change with different coordinate systems, hence tensors describing them must satisfy some *invariance* properties. So tensors must have invariance properties with respect to changes of bases, but their coordinates will of course not stay invariant.

Here is an example of a familiar tensor:

EXAMPLE 1.2. We recall here the familiar transformation property that vectors enjoy according to which they are an example of a **contravariant tensor of first order**. We use here freely notions and properties that will be recalled in the next chapter.

Let $\mathcal{B} := \{b_1, b_2, b_3\}$ and $\widetilde{\mathcal{B}} := \{\tilde{b}_1, \tilde{b}_2, \tilde{b}_3\}$ be two basis of a vector space V. A vector $v \in V$ can be written as

$$v = v^1 b_1 + v^2 b_2 + v^3 b_3 ,$$

or

$$v = \tilde{v}^1 \tilde{b}_1 + \tilde{v}^2 \tilde{b}_2 + \tilde{v}^3 \tilde{b}_3 ,$$

where v^1, v^2, v^3 (resp. $\tilde{v}^1, \tilde{v}^2, \tilde{v}^3$) are the coordinate of v with respect to the basis $\mathcal{B}$ (resp. $\widetilde{\mathcal{B}}$).

Warning: *Please keep the lower indices as lower indices and the upper ones as upper ones. You will see later that there is a reason for it!*

We use the following notation:

$$[v]_{\mathcal{B}} = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} \qquad \text{and} \qquad [v]_{\widetilde{\mathcal{B}}} = \begin{pmatrix} \tilde{v}^1 \\ \tilde{v}^2 \\ \tilde{v}^3 \end{pmatrix} , \tag{1.1}$$

and we are interested in finding the relation between the coordinates of v in the two bases.

The vectors $\tilde{b}_j$, $j = 1, 2, 3$, in the basis $\widetilde{\mathcal{B}}$ can be written as a linear combination of vectors in $\mathcal{B}$ as follows:

$$\tilde{b}_j = L^1_j b_1 + L^2_j b_2 + L^3_j b_3 ,$$

for some $L^i_j \in \mathbb{R}$. We consider the matrix of the change of basis from $\mathcal{B}$ to $\widetilde{\mathcal{B}}$,

$$L := L_{\widetilde{\mathcal{B}}\mathcal{B}} = \begin{bmatrix} L^1_1 & L^1_2 & L^1_3 \\ L^2_1 & L^2_2 & L^2_3 \\ L^3_1 & L^3_2 & L^3_3 \end{bmatrix}$$

whose jth-column consists of the coordinates of the vectors $\tilde{b}_j$ with respect to the basis $\mathcal{B}$. The equalities

$$\begin{cases} \tilde{b}_1 = L^1_1 b_1 + L^2_1 b_2 + L^3_1 b_3 \\ \tilde{b}_2 = L^1_2 b_1 + L^2_2 b_2 + L^3_2 b_3 \\ \tilde{b}_3 = L^1_3 b_1 + L^2_3 b_2 + L^3_3 b_3 \end{cases}$$

can simply be written as

$$\begin{pmatrix} \tilde{b}_1 & \tilde{b}_2 & \tilde{b}_3 \end{pmatrix} = \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix} L . \tag{1.2}$$

(Check this symbolic equation using the rules of matrix multiplication.) Analogously, writing basis vectors in a row and vector coordinates in a column, we can write

$$v = v^1 b_1 + v^2 b_2 + v^3 b_3 = \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} \tag{1.3}$$

and

$$v = \tilde{v}^1 \tilde{b}_1 + \tilde{v}^2 \tilde{b}_2 + \tilde{v}^3 \tilde{b}_3 = \begin{pmatrix} \tilde{b}_1 & \tilde{b}_2 & \tilde{b}_3 \end{pmatrix} \begin{pmatrix} \tilde{v}^1 \\ \tilde{v}^2 \\ \tilde{v}^3 \end{pmatrix} = \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix} L \begin{pmatrix} \tilde{v}^1 \\ \tilde{v}^2 \\ \tilde{v}^3 \end{pmatrix}, \tag{1.4}$$

where we used (1.2) in the last equality. Comparing the expression of v in (1.3) and in (1.4), we conclude that

$$L \begin{pmatrix} \tilde{v}^1 \\ \tilde{v}^2 \\ \tilde{v}^3 \end{pmatrix} = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix}$$

or equivalently

$$\begin{pmatrix} \tilde{v}^1 \\ \tilde{v}^2 \\ \tilde{v}^3 \end{pmatrix} = L^{-1} \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix}$$

We say that the components of a vector v are **contravariant**[1] because they change by L^{-1} when the basis changes by L. A vector v is hence a **contravariant 1-tensor** or **tensor of order** $(1, 0)$. □

Example 1.3 (A numerical example). Let

$$\mathcal{B} = \{e_1, e_2, e_3\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \tag{1.5}$$

be the standard basis or $\mathbb{R}^3$ and let

$$\widetilde{\mathcal{B}} = \{\tilde{b}_1, \tilde{b}_2, \tilde{b}_3\} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \\ 0 \end{bmatrix} \right\}$$

[1] In Latin *contra* means "contrary', against".

be another basis of $\mathbb{R}^3$. The vector[2] $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ has coordinates

$$[v]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad [v]_{\widetilde{\mathcal{B}}} = \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ 0 \end{pmatrix} .$$

Since it is easy to check that

$$\begin{cases} \tilde{b}_1 = 1 \cdot e_1 + 4 \cdot e_2 + 7 \cdot e_3 \\ \tilde{b}_2 = 2 \cdot e_1 + 5 \cdot e_2 + 8 \cdot e_3 \\ \tilde{b}_3 = 3 \cdot e_1 + 6 \cdot e_2 \end{cases} ,$$

the matrix of the change of coordinates from $\mathcal{B}$ to $\widetilde{\mathcal{B}}$ is

$$L = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 0 \end{bmatrix} .$$

It is easy to check that

$$\begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ 0 \end{pmatrix} = L^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

or equivalently

$$L \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} .$$

□

[2]The vector v here is meant here as an element in $\mathbb{R}^3$. As such, it is identified by three real numbers that we write in column surrounded by square brackets. This should not be confused with the coordinates of v with respect to a basis $\mathcal{B}$, that are indicated by round parentheses as in (1.1), while $[\,\cdot\,]_{\mathcal{B}}$ indicates the "operation" of taking the vector v and looking at its coordinates in the basis $\mathcal{B}$. Of course with this convention there is the – slightly confusing – fact that if $\mathcal{B}$ is the basis in (1.5), then $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $[v]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

CHAPTER 2

Review of Linear Algebra

2.1. Vector Spaces and Subspaces

DEFINITION 2.1. A **vector space** V over $\mathbb{R}$ is a set V equipped with two operations:

(1) *Vector addition:* $V \times V \to V$, $(v, w) \mapsto v + w$, and
(2) *Multiplication by a scalar:* $\mathbb{R} \times V \to V$, $(\alpha, v) \mapsto \alpha v$,

satisfying the following properties:

(1) (associativity) $(u + v) + w = u + (v + w)$ for every $u, v, w \in V$;
(2) (commutativity) $u + v = v + u$ for every $u, v \in V$;
(3) (existence of the zero vector) there exists $0 \in V$ such that $v + 0 = v$ for every $v \in V$;
(4) (existence of additive inverse) For every $v \in V$, there exists $w_v \in V$ such that $v + w_v = 0$. The vector w_v is denoted by $-v$.
(5) $\alpha(\beta v) = (\alpha\beta)v$ for every $\alpha, \beta \in \mathbb{R}$ and every $v \in \mathbb{R}$;
(6) $1v = v$ for every $v \in V$;
(7) $\alpha(u + w) = \alpha u + \alpha v$ for all $\alpha \in \mathbb{R}$ and $u, v \in V$;
(8) $(\alpha + \beta)v = \alpha u + \beta v$ for all $\alpha, \beta \in \mathbb{R}$ and $v \in V$.

An element of the vector space is called a **vector**.

EXAMPLE 2.2 (Prototypical example). The Euclidean space $\mathbb{R}^n$, $n = 1, 2, 3, \dots,$ is a vector space with componentwise addition and multiplication by scalars. Vectors in $\mathbb{R}^n$ are denoted by $v = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, with $x_1, \dots, x_n \in \mathbb{R}$. □

EXAMPLES 2.3 (Other examples). (1) The set of real polynomials of degree $\leq n$ is a vector space, denoted by

$$V = \mathbb{R}[x]_n := \{a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n \,:\, a_j \in \mathbb{R}\}.$$

(2) The set of real matrices of size $m \times n$,

$$V = M_{m\times n}(\mathbb{R}) := \left\{ \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} \,:\, a_{ij} \in \mathbb{R} \right\}.$$

(3) The space of solutions of a homogeneous linear (ordinary or partial) differential equation.

(4) The space $\{f : W \to \mathbb{R}\}$, where W is a vector space.

□

EXERCISE 2.4. Are the following vector spaces?

(1) The set V of all vectors in $\mathbb{R}^3$ perpendicular to the vector $\begin{bmatrix}1\\2\\3\end{bmatrix}$.

(2) The set of invertible 2×2 matrices, that is

$$V := \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : ad - bc =\neq 0 \right\} .$$

(3) The set of polynomials of degree exactly n, that is

$$V := \{a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n : a_j \in \mathbb{R}, a_n \neq 0\} .$$

(4) The set V of 2×4 matrices with last column zero, that is

$$V := \left\{ \begin{bmatrix} a & b & c & 0 \\ d & e & f & 0 \end{bmatrix} : a, b, c, d, e, f \in \mathbb{R} \right\}$$

(5) The set of solutions $f : \mathbb{R} \to \mathbb{R}$ of the equation $f' = 5$, that is

$$V := \{f : \mathbb{R} \to \mathbb{R} : f(x) = 5x + C, C \in \mathbb{R}\} .$$

(6) The set of all linear transformations $T : \mathbb{R}^2 \to \mathbb{R}^3$.

Before we pass to the notion of subspace, recall that a **linear combination** of vectors $v_1, \ldots, v_n \in V$ is a vector of the form $\alpha_1v_1 + \cdots + \alpha_nv_n$ for $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$.

DEFINITION 2.5. A subset W of a vector space V that is itself a vector space is a **subspace**.

In other words, a subset $W \subseteq V$ is a subspace if the following conditions are verified:

(1) The 0 element is in V;
(2) W is *closed under addition*, that is $v + w \in W$ for every $v, w \in W$;
(3) W is *closed under multiplication by scalars*, that is $\alpha v \in W$ for every $\alpha \in \mathbb{R}$ and every $v \in W$.

Condition (1) in fact follows from (2) and (3), but it is often emphasized because it is an easy way to check that a subset is not a subspace. In any case the above three conditions are equivalent to the following ones:

(1)' W is nonempty;
(2)' W is *closed under linear combinations*, that is $\alpha v + \beta w \in W$ for all $\alpha, \beta \in \mathbb{R}$ and all $v, w \in W$.

2.2. Bases

The key yo study vector spaces is the concept of basis.

DEFINITION 2.6. The vectors $\{b_1, \dots, b_n\} \in V$ form a **basis** of V if:
(1) they are *linearly independent* and
(2) the *span* V.

Warning: *We consider only vector spaces that have bases consisting of a finite number of elements.*

We recall here the notions of liner dependence/independence and the notion of span.

DEFINITION 2.7. The vectors $\{b_1, \dots, b_n\} \in V$ are **linearly independent** if $\alpha_1 b_1 + \cdots + \alpha_n b_n = 0$ implies that $\alpha_1 = \cdots = \alpha_n = 0$. In other words if the only linear combination that represents zero is the trivial one.

EXAMPLE 2.8. The vectors

$$b_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \qquad b_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \qquad b_3 = \begin{bmatrix} 7 \\ 8 \\ 0 \end{bmatrix}$$

are linearly independent in $\mathbb{R}^3$. In fact,

$$\mu_1 b_1 + \mu_2 b_2 + \mu_3 b_3 = 0 \iff \begin{cases} \mu_1 + 4\mu_2 + 7\mu_3 = 0 \\ 2\mu_1 + 5\mu_2 + 8\mu_3 0 \\ 3\mu_1 + 6\mu_2 = 0 \end{cases} \iff \dots \iff \mu_1 = \mu_2 = \mu_3 = 0\,.$$

(If you are unsure how to fill in the dots look at Example 2.13.) □

EXAMPLE 2.9. The vectors

$$b_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \qquad b_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \qquad b_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

are linearly dependent in $\mathbb{R}^3$. In fact,

$$\mu_1 b_1 + \mu_2 b_2 + \mu_3 b_3 = 0 \iff \begin{cases} \mu_1 + 4\mu_2 + 7\mu_3 = 0 \\ 2\mu_1 + 5\mu_2 + 8\mu_3 0 \\ 3\mu_1 + 6\mu_2 + 9\mu_3 = 0 \end{cases} \iff \dots \iff \begin{cases} \mu_1 = \mu_2 \\ \mu_2 = -2\mu_3\,, \end{cases}$$

so

$$b_1 - 2b_2 + b_3 = 0$$

and b_1, b_2, b_3 are not linearly independent. For example $b_1 = 2b_2 - b_3$ is a non-trivial linear relation between the vectors b_1, b_2 and b_3. □

DEFINITION 2.10. The vectors $\{b_1, \dots, b_n\} \in V$ **span** V if every vector $v \in V$ can be written as a linear combination $v = \alpha_1 b_1 + \cdots + \alpha_n b_n$, for some $\alpha_1, \dots, \alpha_n \in \mathbb{R}$.

EXAMPLE 2.11. The vectors in Example 2.8 span $\mathbb{R}^3$, while the vectors in Example 2.9 do not span $\mathbb{R}^3$. To see this, we recall the following facts about bases. □

FACTS ABOUT BASES: Let V be a vector space:

(1) All bases of V have the same number of elements. This number is called the **dimension** of V and indicated with $\dim V$.
(2) If $\mathcal{B} := \{b_1, \dots, b_n\}$ form a basis of V, there is a unique way of writing v as a linear combination

$$v = v^1 b_1 + \dots v^n b_n$$

of elements in $\mathcal{B}$. We denote by

$$[v]_{\mathcal{B}} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

the coordinate vector of v with respect to $\mathcal{B}$.
(3) If we know that $\dim V = n$, then:
  (a) More than n vectors in V must be linearly dependent;
  (b) Fewer than n vectors in V cannot span V;
  (c) Any n linearly independent vectors span V;
  (d) Any n vectors that span V must be linearly independent;
  (e) If k vectors span V, then $k \geq n$ and some subset of those k vectors must be a basis of V;
  (f) If a set of m vectors is linearly independent, then $m \leq n$ and we can always complete the set to form a basis of V.

EXAMPLE 2.12. The vectors in Example 2.8 form a basis of $\mathbb{R}^3$ since they are linearly independent and they are exactly as many as the dimension of $\mathbb{R}^3$. □

EXAMPLE 2.13 (Gauss-Jordan elimination). We are going to compute here the coordinates of $v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ with respect to the basis $\mathcal{B} = \{b_1, b_2, b_3\}$ in Example 2.8. The seeked coordinates $[v]_{\mathcal{B}} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$ must satisfy the equation

$$v^1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + v^2 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + v^3 \begin{bmatrix} 7 \\ 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

so to find them we have to solve the following system of linear equations:

$$\begin{cases} v^1 + 4v^2 + 7v^3 = 1 \\ 2v^1 + 5v^2 + 8v^3 = 1 \\ 3v^1 + 6v^2 = 1 \end{cases}$$

or, equivalently, reduced the following augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 4 & 7 & 1 \\ 2 & 5 & 8 & 1 \\ 3 & 6 & 0 & 1 \end{array}\right]$$

in echelon form using the Gauss–Jordan elimination method. We are going to perform both calculations in parallel, which will also point out that they are indeed seemingly different incarnation of the same method.

By multiplying the first equation/row by 2 (reps. 3) and subtracting it from the second (reps. third) equation/row we obtain

$$\begin{cases} v^1 + 4v^2 + 7v^3 = 1 \\ \quad - 3v^2 - 6v^3 = -1 \\ \quad - 6v^2 - 21v^3 = -2 \end{cases} \quad \leftrightsquigarrow \quad \left[\begin{array}{ccc|c} 1 & 4 & 7 & 1 \\ 0 & -3 & -6 & -1 \\ 0 & -6 & -21 & -2 \end{array}\right].$$

By multiplying the second equation/row by $-1/3$ and by adding to the first (resp. third) equation/row the second equation/row multiplied by $-4/3$ (resp. 2) we obtain

$$\begin{cases} v^1 \qquad - v^3 = 1 \\ \qquad v^2 + 2v^3 = \dfrac{1}{3} \\ \qquad\quad - 9v^3 = 0 \end{cases} \quad \leftrightsquigarrow \quad \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & \frac{1}{3} \\ 0 & 0 & -9 & 0 \end{array}\right].$$

The last equation/row shows that $v^3 = 0$, hence the above becomes

$$\begin{cases} v^1 \qquad\qquad = 1 \\ \qquad v^2 \qquad = \dfrac{1}{3} \\ \qquad\qquad v^3 = 0 \end{cases} \quad \leftrightsquigarrow \quad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 0 \end{array}\right].$$

□

EXERCISE 2.14. Let V be the vector space consisting of all 2×2 matrices with trace zero, namely

$$V := \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbb{R} \text{ and } a + d = 0 \right\}.$$

(1) Show that

$$\mathcal{B} := \left\{ \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}}_{b_1}, \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{b_2}, \underbrace{\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}}_{b_3} \right\}$$

is a basis of V.

(2) Show that

$$\widetilde{\mathcal{B}} := \left\{ \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}}_{\tilde{b}_1}, \underbrace{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}_{\tilde{b}_2}, \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{\tilde{b}_3} \right\}$$

is another basis of V.

(3) Compute the coordinates of

$$v = \begin{bmatrix} 2 & 1 \\ 7 & -2 \end{bmatrix}$$

with respect to $\mathcal{B}$ and with respect to $\widetilde{\mathcal{B}}$.

2.3. The Einstein convention

We start by setting a notation that will turn out to be useful later on. Recall that if $\mathcal{B} = \{b_1, b_2, b_3\}$ is a basis of a vector space V, any vector $v \in V$ can be written as

$$v = v^1 b_1 + v^2 b_2 + v^3 b_3 \tag{2.1}$$

for appropriate $v^1, v^2, v^3 \in \mathbb{R}$.

Notation. From now on, expressions like the one in (2.1) will be written as

$$v = \cancel{v^1 b_1 + v^2 b_2 + v^3 b_3} = v^j b_j \,. \tag{2.2}$$

That is, from now on when an index appear *twice* (that is, *once as a subscript and once as a superscript*) in a term, we know that it implies that there is a summation over all possible values of that index. The summation symbol will not be displayed.

Analogously, indices that are not repeated in expressions like $a_{ij}x^k y^j$ are free indices not subject to summation.

Examples 2.15. For indices ranging over $\{1, 2, 3\}$, i.e. $n = 3$:

(1) The expression $a_{ij}x^i y^k$ means

$$a_{ij}x^i y^k = a_{1j}x^1 y^k + a_{2j}x^2 y^k + a_{3j}x^3 y^k \,,$$

and could be called R_j^k (meaning that R_j^k and $a_{ij}x^i y^k$ both depend on the indices j and k).

(2) Likewise,

$$a_{ij}x^k y^j = a_{i1}x^k y^1 + a_{i2}x^k y^2 + a_{i3}x^k y^3 =: Q_i^k \,.$$

(3) Further

$$\begin{aligned} a_{ij}x^iy^j &= a_{11}x^1y^1 + a_{12}x^1y^2 + a_{13}x^1y^3 \\ &+ a_{21}x^2y^1 + a_{22}x^2y^2 + a_{23}x^2y^3 \\ &+ a_{31}x^3y^1 + a_{32}x^3y^2 + a_{33}x^3y^3 =: P \end{aligned}$$

(4) An expression like

$$A^iB^j_{k\ell}C^\ell =: D^{ij}_k$$

makes sense. Here the indices i, j, k are free (i.e. free to range in $\{1, 2, \dots, n\}$) and ℓ is a summation index.

(5) On the other hand an expression like

$$E_{ij}F_\ell{}^{jk}G^\ell = H^{jk}_i$$

does not make sense because the expression on the left has only two free indices, i and k, while j and ℓ are summation indices and neither of them can appear on the right hand side.

Notation. Since v^jb_j denotes a sum, the generic term of a sum will be denoted with *capital letters*. For example we write v^Ib_I and the above expressions could have been written as

(1)

$$a_{ij}x^iy^k = \sum_{i=1}^{3} a_{IJ}x^Iy^K = a_{1j}x^1y^k + a_{2j}x^2y^k + a_{3j}x^3y^k\,,$$

(2)

$$a_{ij}x^ky^j = \sum_{j=1}^{3} a_{IJ}x^Ky^J = a_{i1}x^ky^1 + a_{i2}x^ky^2 + a_{i3}x^ky^3\,.$$

(3)

$$\begin{aligned} a_{ij}x^iy^j &= \sum_{j=1}^{3}\sum_{i=1}^{3} a_{IJ}x^Iy^J = \\ &= a_{11}x^1y^1 + a_{12}x^1y^2 + a_{13}x^1y^3 \\ &+ a_{21}x^2y^1 + a_{22}x^2y^2 + a_{23}x^2y^3 \\ &+ a_{31}x^3y^1 + a_{32}x^3y^2 + a_{33}x^3y^3\,. \end{aligned}$$

□

2.3.1. Change of bases, revisited. Let $\mathcal{B}$ and $\widetilde{\mathcal{B}}$ be two bases of a vector space V and let

$$L := L_{\widetilde{\mathcal{B}}\mathcal{B}} = \begin{bmatrix} L_1^1 & \dots & L_n^1 \\ \vdots & & \vdots \\ L_1^n & \dots & L_n^n \end{bmatrix} \tag{2.3}$$

be the matrix of the change of basis from $\mathcal{B}$ to $\widetilde{\mathcal{B}}$. [Recall that the entries of the j-th column of L are the coordinates of the $\tilde{b}_j$s with respect to the basis $\mathcal{B}$.] With the Einstein convention we can write

$$\boxed{\tilde{b}_j = L_j^i b_i}, . \tag{2.4}$$

or, equivalently

$$\begin{pmatrix} \tilde{b}_1 & \dots & \tilde{b}_n \end{pmatrix} = \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix} L \, .$$

If $\Lambda = L^{-1}$ denotes the matrix of the change of basis from $\widetilde{\mathcal{B}}$ to $\mathcal{B}$, then

$$\begin{pmatrix} b_1 & \dots & b_n \end{pmatrix} = \begin{pmatrix} \tilde{b}_1 & \dots & \tilde{b}_n \end{pmatrix} \Lambda \, .$$

Equivalently, this can be written in compact form using the Einstein notation as

$$b_j = \Lambda_j^i \tilde{b}_i \, .$$

Analogously, the corresponding relations for the vector coordinates are

$$\begin{pmatrix} v^1 \\ \vdots \\ v^i \\ \vdots \\ v^n \end{pmatrix} = \begin{pmatrix} L_1^1 & \dots & L_n^1 \\ \vdots & & \vdots \\ L_1^i & \dots & L_n^i \\ \vdots & & \vdots \\ L_1^n & \dots & L_n^n \end{pmatrix} \begin{pmatrix} \tilde{v}^1 \\ \vdots \\ \tilde{v}^n \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \tilde{v}^1 \\ \vdots \\ \tilde{v}^i \\ \vdots \\ \tilde{v}^n \end{pmatrix} = \begin{pmatrix} \Lambda_1^1 & \dots & \Lambda_n^1 \\ \vdots & & \vdots \\ \Lambda_1^i & \dots & \Lambda_n^i \\ \vdots & & \vdots \\ \Lambda_1^n & \dots & \Lambda_n^n \end{pmatrix} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

and these can be written with the Einstein convention respectively as

$$v^i = \cancel{L_1^i \tilde{v}^1 + \dots + L_n^i \tilde{v}^n} = L_j^i \tilde{v}^j \quad \text{and} \quad \tilde{v}^i = \cancel{\Lambda_1^i v^1 + \dots + \Lambda_n^i v^n} = \Lambda_j^i v^j \, , \tag{2.5}$$

or, in matrix notation,

$$[v]_{\mathcal{B}} = L_{\widetilde{\mathcal{B}}\mathcal{B}} [v]_{\widetilde{\mathcal{B}}} \quad \text{and} \quad [v]_{\widetilde{\mathcal{B}}} = (L_{\widetilde{\mathcal{B}}\mathcal{B}})^{-1} [v]_{\mathcal{B}} = L_{\mathcal{B}\widetilde{\mathcal{B}}} [v]_{\mathcal{B}} \, .$$

Example 2.16. We consider the following two bases of $\mathbb{R}^2$

$$\begin{aligned} \mathcal{B} &= \Bigg\{ \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{b_1}, \underbrace{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{b_2} \Bigg\} \\ \widetilde{\mathcal{B}} &= \Bigg\{ \underbrace{\begin{bmatrix} 3 \\ 1 \end{bmatrix}}_{\tilde{b}_1}, \underbrace{\begin{bmatrix} -1 \\ -1 \end{bmatrix}}_{\tilde{b}_2} \Bigg\} \end{aligned} \tag{2.6}$$

and we look for the matrix of the change of basis. Namely we look for a matrix L such that

$$\begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix} = \begin{pmatrix} \tilde{b}_1 & \tilde{b}_2 \end{pmatrix} = \begin{pmatrix} b_1 & b_2 \end{pmatrix} L = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} L\,.$$

There are two alternative ways of finding L:

(1) *With matrix inversion:* Recall that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{D}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \tag{2.7}$$

where $D = \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)$. Thus

$$L = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

(2) *With the Gauss-Jordan elimination:*

$$\left[\begin{array}{cc|cc} 1 & 2 & 3 & -1 \\ 0 & 1 & 1 & -1 \end{array}\right] \quad \leftrightsquigarrow \quad \left[\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \end{array}\right]$$

□

2.3.2. The Kronecker delta symbol.

NOTATION. The **Kronecker delta symbol** δ^i_j is defined as

$$\delta^i_j := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j\,. \end{cases} \tag{2.8}$$

EXAMPLES 2.17. If L is a matrix, the (i, j)-entry of L is the coefficient in the i-th row and j-th column, and is denoted by L^i_j.

(1) The $n \times n$ identity matrix

$$I = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix}$$

has (i, j)-entry equal to δ^i_j.

(2) Let L and Λ be two square matrices. The (i, j)-th entry of the product ΛL

$$\Lambda L = \begin{bmatrix} \Lambda^1_1 & \dots & \Lambda^1_n \\ \vdots & & \vdots \\ \Lambda^i_1 & \dots & \Lambda^i_n \\ \vdots & & \vdots \\ \Lambda^n_1 & \dots & \Lambda^n_n \end{bmatrix} \begin{bmatrix} L^1_1 & \dots & L^1_j & \dots & L^1_n \\ \vdots & & \vdots & & \vdots \\ L^n_1 & \dots & L^n_j & \dots & L^n_n \end{bmatrix}$$

equals the dot product of the i-th row and j-th column,

$$\begin{pmatrix}\Lambda^i_1 & \dots & \Lambda^i_n\end{pmatrix} \cdot \begin{pmatrix} L^1_j \\ \vdots \\ L^n_j \end{pmatrix} = \Lambda^i_1 L^1_j + \dots + \Lambda^i_n L^n_j \,,$$

or, using the Einstein convention,

$$\Lambda^i_k L^k_j$$

Notice that since in general $\Lambda L \neq L\Lambda$, it follows that

$$\Lambda^i_k L^k_j \neq L^i_k \Lambda^k_j = \Lambda^k_j L^i_k \,.$$

On the other hand, if $\Lambda = L^{-1}$, that is $\Lambda L = L\Lambda = I$, then we can write

$$\Lambda^i_k L^k_j = \delta^i_j = L^i_k \Lambda^k_j \,.$$

□

Remark 2.18. Using the Kronecker delta symbol we can check that the notations in (2.5) are all consistent. In fact, from (2.2) we should have

$$v^i b_i = v = \tilde{v}^i \tilde{b}_i \,, \tag{2.9}$$

and, in fact, using (2.5),

$$\tilde{v}^i \tilde{b}_i = \Lambda^i_j v^j L^k_i b_k = \delta^k_j v^j b_k = v^j b_j \,,$$

where we used that $\Lambda^i_j L^k_i = \delta^k_j$ since $\Lambda = L^{-1}$.

Two words of warning:

- The two expressions $v^j b_j$ and $v^k b_k$ are identical, as the indices j and k are dummy indices.
- When multiplying two different expressions in Einstein notation, you should be careful to distinguish by different letters different summation indices. For example, if $\tilde{v}^i = \Lambda^i_j v^j$ and $\tilde{b}_i = L^j_i b_j$, in order to perform the multiplication $\tilde{v}^i \tilde{b}_i$ we have to make sure to replace one of the dummy indices in the two expressions. So, for example, we can write $\tilde{b}_i = L^k_i b_k$, so that $\tilde{v}^i \tilde{b}_i = \Lambda^i_j v^j L^k_i b_k$.

2.4. Linear Transformations

Let $T : V \to V$ be a linear transformation, that is a transformation that satisfies the property

$$T(\alpha v + \beta w) = \alpha T(v) + \beta T(w) \,,$$

for all $\alpha, \beta \in \mathbb{R}$ and all $v, w \in V$. Once we choose a basis of V, the transformation T is represented by a matrix A with respect to that basis, and that matrix gives the effect of T on the coordinate vectors. In other words, if $T(v)$ is the effect of the transformation T on the vector v, with respect a basis $\mathcal{B}$ we have that

$$[v]_{\mathcal{B}} \longmapsto [T(v)]_{\mathcal{B}} = A[v]_{\mathcal{B}} \,. \tag{2.10}$$

If $\widetilde{\mathcal{B}}$ is another basis, we have also

$$[v]_{\widetilde{\mathcal{B}}} \longmapsto [T(v)]_{\widetilde{\mathcal{B}}} = \widetilde{A}[v]_{\widetilde{\mathcal{B}}} \,, \tag{2.11}$$

where now $\widetilde{A}$ is the matrix of the transformation T with respect to the basis $\widetilde{\mathcal{B}}$.

We want to find now the relation between A and $\widetilde{A}$. Let $L := L_{\widetilde{\mathcal{B}}\mathcal{B}}$ be the matrix of the change of basis from $\mathcal{B}$ to $\widetilde{\mathcal{B}}$. Then, for any $v \in V$,

$$[v]_{\widetilde{\mathcal{B}}} = L^{-1}[v]_{\mathcal{B}} \,. \tag{2.12}$$

In particular the above equation holds for the vector $T(v)$, that is

$$[T(v)]_{\widetilde{\mathcal{B}}} = L^{-1}[T(v)]_{\mathcal{B}} \,. \tag{2.13}$$

Using (2.12), (2.11), (2.13) and (2.10) in this order, we have

$$\widetilde{A}L^{-1}[v]_{\mathcal{B}} = \widetilde{A}[v]_{\widetilde{\mathcal{B}}} = [T(v)]_{\widetilde{\mathcal{B}}} = L^{-1}[T(v)]_{\mathcal{B}} = L^{-1}A[v]_{\mathcal{B}}$$

for every vector $v \in V$. If follows that $\widetilde{A}L^{-1} = L^{-1}A$ or equivalently

$$\widetilde{A} = L^{-1}AL \,, \tag{2.14}$$

which in Einstein notation reads

$$\widetilde{A}^i_j = \Lambda^i_k A^k_m L^m_j \,.$$

We say that the linear transformation T is a **tensor of type** $(1,1)$.

Example 2.19. Let $V = \mathbb{R}^2$ and let $\mathcal{B}$ and $\widetilde{\mathcal{B}}$ be the bases in Example 2.16. The matrices corresponding to the change of coordinates are

$$L := L_{\widetilde{\mathcal{B}}\mathcal{B}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad \text{and} \qquad L^{-1} = \frac{1}{-2}\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} ,$$

where in the last equality we used the formula for the inverse of a matrix in (2.7).

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation that in the basis $\mathcal{B}$ takes the form

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} .$$

Then according to (2.14) the matrix $\widetilde{A}$ of the linear transformation T with respect to the basis $\widetilde{\mathcal{B}}$ is

$$\widetilde{A} = L^{-1}AL = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 5 & -2 \\ -1 & 0 \end{bmatrix} .$$

□

EXAMPLE 2.20. We now look for the **standard matrix of** T, that is the matrix M that represents T with respect to the standard basis of $\mathbb{R}^2$, which we denote by

$$\mathcal{E} := \Big\{ \underbrace{\begin{bmatrix}1\\0\end{bmatrix}}_{e_1}, \underbrace{\begin{bmatrix}0\\1\end{bmatrix}}_{e_2} \Big\}.$$

We want to apply again the formula (2.14) and hence we first need to find the matrix $S := L_{\mathcal{B}\mathcal{E}}$ of the change of basis from $\mathcal{E}$ to $\mathcal{B}$. Recall that the columns of S are the coordinates of b_j with respect to the basis $\mathcal{E}$, that is

$$S = \begin{bmatrix}1 & 2\\0 & 1\end{bmatrix}.$$

According to (2.14),

$$A = S^{-1}MS\,,$$

from which, using again (2.7), we obtain

$$M = SAS^{-1} = \begin{bmatrix}1 & 2\\0 & 1\end{bmatrix}\begin{bmatrix}1 & 3\\2 & 4\end{bmatrix}\begin{bmatrix}1 & -2\\0 & 1\end{bmatrix} = \begin{bmatrix}1 & 2\\0 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\2 & 0\end{bmatrix} = \begin{bmatrix}5 & 1\\2 & 0\end{bmatrix}.$$

□

EXAMPLE 2.21. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the *orthogonal projection* onto the plane $\mathcal{P}$ of equation

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

This means that the transformation T is characterized by the fact that

- it does not change vectors in the plane $\mathcal{P}$, and
- it takes to zero vectors perpendicular to $\mathcal{P}$.

We want to find the standard matrix for T.

Idea: First compute the matrix of T with respect to a basis $\mathcal{B}$ of $\mathbb{R}^3$ well adapted to the problem, then use (2.14) after having found the matrix $L_{\mathcal{B}\mathcal{E}}$ of the change of basis.

To this purpose, we choose two linearly independent vectors in the plane $\mathcal{P}$ and a third vector perpendicular to $\mathcal{P}$. For instance, we set

$$\mathcal{B} := \Big\{ \underbrace{\begin{bmatrix}1\\0\\2\end{bmatrix}}_{b_1}, \underbrace{\begin{bmatrix}0\\1\\1\end{bmatrix}}_{b_2}, \underbrace{\begin{bmatrix}2\\1\\-1\end{bmatrix}}_{b_3} \Big\},$$

where the coordinates of b_1 and b_2 satisfy the equation of the plane, while the coordinates of b_3 are the coefficients of the equation describing $\mathcal{P}$. Let $\mathcal{E}$ be the standard basis of $\mathbb{R}^3$.

Since

$$T(b_1) = b_1, \qquad T(b_2) = b_2 \qquad \text{and } T(b_3) = 0\,,$$

the matrix of T with respect to $\mathcal{B}$ is

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{2.15}$$

where we recall that the j-th column are the coordinates $[T(b_j)]_{\mathcal{B}}$ of the vector $T(b_j)$ with respect to the basis $\mathcal{B}$.

The matrix of the change of basis from $\mathcal{E}$ to $\mathcal{B}$ is

$$L = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 2 & 1 & -1 \end{bmatrix},$$

hence, by Gauss–Jordan elimination,

$$L^{-1} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{5}{6} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & -\frac{1}{6} \end{bmatrix}.$$

Therefore

$$M = LAL^{-1} = \cdots = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{5}{6} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{5}{6} \end{bmatrix}.$$

□

Example 2.22. Let $V := \mathbb{R}[x]_2$ be the vector space of polynomials of degree ≤ 2, and let $T : \mathbb{R}[x]_2 \to \mathbb{R}[x]_2$ be the linear transformation given by differentiating a polynomial and then multiplying the derivative by x,

$$T(p(x)) := xp'(x)\,,$$

so that $T(a + bx + cx^2) = x(b + 2cx) = bx + 2cx^2$. Let

$$\mathcal{B} := \{1, x, x^2\} \qquad \text{and} \qquad \widetilde{\mathcal{B}} := \{x, x-1, x^2-1\}$$

be two bases of $\mathbb{R}[x]_2$. Since

$$\begin{aligned} T(1) &= 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ T(x) &= x = 0 \cdot 1 + 1 \cdot x + 0 \cdot x^2 \\ T(x^2) &= 2x^2 = 0 \cdot 1 + 0 \cdot x + 2 \cdot x^2 \end{aligned}$$

and

$$\begin{aligned} T(x) &= x = 1 \cdot x + 0 \cdot (x-1) + 0 \cdot (x^2-1) \\ T(x-1) &= x = 1 \cdot x + 0 \cdot (x-1) + 0 \cdot (x^2-1) \\ T(x^2-1) &= 2x^2 = 2 \cdot x - 2 \cdot (x-1) + 2 \cdot (x^2-1)\,, \end{aligned}$$

then

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad \text{and} \qquad \widetilde{A} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & -2 \\ 0 & 0 & 2 \end{bmatrix}.$$

One can check that $L = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and that indeed $AL = L\widetilde{A}$ or, equivalently $\widetilde{A} = L^{-1}AL$. □

2.4.1. Similar matrices. The above calculations can be summarized by the *commutativity* of the following diagram. Here the vertical arrows correspond to the operation of the change of basis from $\mathcal{B}$ to $\widetilde{\mathcal{B}}$ (recall that the coordinate vectors are contravariant tensors, that is they transform as $[v]_{\widetilde{\mathcal{B}}} = L^{-1}[v]_{\mathcal{B}}$) and the horizontal arrows to the operation of applying the transformation T in the two different basis

$$\begin{array}{ccc} [v]_{\mathcal{B}} & \xmapsto{A} & [T(v)]_{\mathcal{B}} \\ {\scriptstyle L^{-1}}\big\downarrow & & \big\downarrow{\scriptstyle L^{-1}} \\ [v]_{\widetilde{\mathcal{B}}} & \xmapsto[\widetilde{A}]{} & [T(v)]_{\widetilde{\mathcal{B}}}\,. \end{array}$$

Saying the the diagram is commutative is exactly the same thing as saying that if one starts from the upper left hand corner, reaching the lower right hand corner following either one of the two paths has exactly the same effect. In other words, changing coordinates first then applying the transformation T yields exactly the same affect as applying first the transformation T and then the change of coordinates, that is

$$AL^{-1} = L^{-1}M$$

or equivalently

$$A = L^{-1}ML\,.$$

In this case we say that A and $\widetilde{A}$ are **similar** matrices. This means that A and $\widetilde{A}$ represent the same transformation with respect to different bases.

Definition 2.23. We say that two matrices A and $\widetilde{A}$ are **similar** if there exists and invertible matrix L such that $\widetilde{A} = L^{-1}AL$.

Examples 2.24. (1) The matrices in Example 2.19 and Example 2.20

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \qquad M = \begin{bmatrix} 5 & 1 \\ 2 & 0 \end{bmatrix} \qquad \text{and} \qquad \widetilde{A} = \begin{bmatrix} 5 & -2 \\ -1 & 0 \end{bmatrix}$$

are similar.

(2) The matrices A in (1) and $A' = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ are not similar. In fact, A is invertible, as $\det A = -2 \neq 0$, while $\det A' = 0$, so that A' is not invertible.

□

We collect here few facts about similar matrices. Recall that the **eigenvalues** of a matrix A are the roots of the **characteristic polynomial**

$$p_A(\lambda) := \det(A - \lambda I)\,.$$

Moreover

(1) the **determinant** of a matrix is the product of its eigenvalues, and
(2) the **trace** of a matrix is the sum of its eigenvalues.

Let us assume that A and $\widetilde{A}$ are similar matrices, that is $\widetilde{A} = L^{-1}AL$ for some invertible matrix L. Then

$$\begin{aligned} p_{\widetilde{A}}(\lambda) &= \det(\widetilde{A} - \lambda I) = \det(L^{-1}AL - \lambda L^{-1}IL) \\ &= \det(L^{-1}(A - \lambda I)L) \\ &= (\det L^{-1}) \det(A - \lambda I)(\det L) = p_A(\lambda)\,, \end{aligned} \tag{2.16}$$

which means that any two similar matrices have the same characteristic polynomial.

Facts about similar matrices: From (2.16) if follows immediately that if the matrices A and $\widetilde{A}$ are similar, then:

- A and $\widetilde{A}$ have the same size;
- the eigenvalues of A (as well as their multiplicity) are the same as those of $\widetilde{A}$;
- $\det A = \det \widetilde{A}$;
- $\operatorname{tr} A = \operatorname{tr} \widetilde{A}$;
- A is invertible if and only if $\widetilde{A}$ is invertible.

2.5. Eigenbases

The possibility of choosing different bases is very important and often simplifies the calculations. Example 2.21 is such an example, where we choose an appropriate basis according to the specific problem. Other times a basis can be chosen according to the symmetries and, completely at the opposite side, sometime there is just not a basis that is a preferred one. One basis that is particularly important, when it exists, is an **eigenbasis** with respect to some linear transformation A of V.

Recall that an **eigenvector** of a linear transformation A corresponding to an eigenvalue λ is a non-zero vector $v \in E_\lambda := \ker(A - \lambda I)$. An **eigenbasis** of a vector space V is a basis consisting of eigenvectors of a linear transformation A of V. The point of having an eigenbasis is that, with respect to this eigenbasis, the linear transformation is as simple as possible, that is is as close as possible to be diagonal. This diagonal matrix similar to A is called the **Jordan canonical form** of A.

Given a linear transformation $T : V \to V$, in order to find an eigenbasis of T, we need to perform the following steps:

(1) Compute the eigenvalues
(2) Compute the eigenspaces

(3) Find a eigenbasis.

We will do this in the following example.

EXAMPLE 2.25. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation given by the matrix $A = \begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix}$ with respect to the standard basis of $\mathbb{R}^2$.

(1) The eigenvalues are the roots of the characteristic polynomial $p_\lambda(A)$. Since
$$\begin{aligned} p_A(\lambda) = \det(A - \lambda I) &= \det \begin{bmatrix} 3-\lambda & -4 \\ -4 & -3-\lambda \end{bmatrix} \\ &= (3-\lambda)(-3-\lambda) - 16 = \lambda^2 - 25 = (\lambda - 5)(\lambda + 5)\,, \end{aligned}$$
hence $\lambda = \pm 5$ are the eigenvalues of A.

(2) If λ is an eigenvalue of A, the eigenspace corresponding to λ is given by $E_\lambda = \ker(A - \lambda I)$. Note that
$$v \in E_\lambda \quad \Longleftrightarrow \quad Av = \lambda v\,.$$
With our choice of A and with the resulting eigenvalues, we have
$$\begin{aligned} E_5 &= \ker(A - 5I) = \ker \begin{bmatrix} -2 & -4 \\ -4 & -8 \end{bmatrix} = \operatorname{span} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \\ E_{-5} &= \ker(A + 5I) = \ker \begin{bmatrix} -8 & -4 \\ -4 & -2 \end{bmatrix} = \operatorname{span} \begin{bmatrix} 1 \\ 2 \end{bmatrix}\,. \end{aligned}$$

(3) The following is an eigenbasis of $\mathbb{R}^2$
$$\widetilde{\mathcal{B}} = \left\{ \tilde{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \tilde{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$$
and
$$\begin{aligned} T(\tilde{b}_1) &= 5\tilde{b}_1 = 5 \cdot \tilde{b}_1 + 0 \cdot \tilde{b}_2 \\ T(\tilde{b}_2) &= -5\tilde{b}_2 = 0 \cdot \tilde{b}_1 - 5 \cdot \tilde{b}_2\,, \end{aligned}$$
so that $A = \begin{bmatrix} 5 & 0 \\ 0 & -5 \end{bmatrix}$.

Notice that the eigenspace E_5 consists of vectors on the line $x + 2y = 0$ and these vectors get scaled by the transformation T by a factor of 5. On the other hand, the eigenspace E_{-5} consists of vectors perpendicular to the line $x + 2y = 0$ and these vectors get flipped by the transformation T and then also scaled by a factor of 5. Hence T is just the reflection across the line $x + 2y = 0$ followed by multiplication by 5.

□

Summarizing, in Examples 2.19 and 2.20 we looked at how the matrix of a transformation changes with respect to two different basis that we were given. In Example 2.21 we looked for a particular basis appropriate to the transformation at hand. In Example 2.25 we looked for an eigenbasis with respect to the given transformation. Example 2.21 in this respect fits into the same framework as Example 2.25, but the orthogonal projection has a zero eigenvalue (see (2.15)).

CHAPTER 3

Multilinear Forms

3.1. Linear Forms

3.1.1. Definition, Examples, Dual and Dual Basis.

DEFINITION 3.1. Let V be a vector space. A **linear form** on V is a map $\alpha : V \to \mathbb{R}$ such that for every $a, b \in \mathbb{R}$ and for every $v, w \in V$

$$\alpha(av + bw) = a\alpha(v) + b\alpha(w)\,.$$

Alternative terminologies for "linear form" are **tensor of type** $(0, 1)$, **1-form**, **linear functional** and **covector**.

EXERCISE 3.2. If $V = \mathbb{R}^3$, which of the following is a linear form?

(1) $\alpha(x, y, z) := xy + z$;
(2) $\alpha(x, y, z) := x + y + z + 1$;
(3) $\alpha(x, y, z) := \pi x - \frac{7}{2}z$.

EXERCISE 3.3. If V is the infinite dimensional vector space of continuous functions $f : \mathbb{R} \to \mathbb{R}$, which of the following is a linear form?

(1) $\alpha(f) := f(7) - f(0)$;
(2) $\alpha(f) := \int_0^{33} e^x f(x)dx$;
(3) $\alpha(f) := e^{f(x)}$.

EXAMPLE 3.4. [Coordinate forms] This is the most important example of linear form. Let $\mathcal{B} := \{b_1, \dots, b_n\}$ be a basis of V and let $v = v^i b_i \in V$ be a generic vector. Define $\beta^i : V \to \mathbb{R}$ by

$$\beta^i(v) := v^i\,, \tag{3.1}$$

that is β^i will extract the i-th coordinate of a vector with respect to the basis $\mathcal{B}$. The linear form β^i is called **coordinate form**. Notice that

$$\beta^i(b_j) = \delta^i_j\,, \tag{3.2}$$

since the i-th coordinate of the basis vector b_j with respect to the basis $\mathcal{B}$ is equal to 1 if $i = j$ and 0 otherwise. □

EXAMPLE 3.5. Let $V = \mathbb{R}^3$ and let $\mathcal{E}$ be its standard basis. The three coordinate forms are defined by

$$\beta^1 \begin{bmatrix} x \\ y \\ z \end{bmatrix} := x \,, \quad \beta^2 \begin{bmatrix} x \\ y \\ z \end{bmatrix} := y \,, \quad \beta^3 \begin{bmatrix} x \\ y \\ z \end{bmatrix} := z \,.$$

□

EXAMPLE 3.6. Let $V = \mathbb{R}^2$ and let $\mathcal{B} := \Big\{ \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{b_1}, \underbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix}}_{b_2} \Big\}$. We want to describe the elements of $\mathcal{B}^* := \{\beta^1, \beta^2\}$, in other words we want to find

$$\beta^1(v) \qquad \text{and} \qquad \beta^2(v)$$

for a generic vector $v \in V$.

To this purpose we need to find $[v]_{\mathcal{B}}$. Recall that if $\mathcal{E}$ denotes the standard basis of $\mathbb{R}^2$ and $L := L_{\mathcal{B}\mathcal{E}}$ the matrix of the change of coordinate from $\mathcal{E}$ to $\mathcal{B}$, then

$$[v]_{\mathcal{B}} = L^{-1}[v]_{\mathcal{E}} = L^{-1} \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} .$$

Since

$$L = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

and hence

$$L^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} ,$$

then

$$[v]_{\mathcal{B}} = \begin{pmatrix} \frac{1}{2}(v^1 + v^2) \\ \frac{1}{2}(v^1 - v^2) \end{pmatrix} .$$

Thus, according to (3.1), we deduce that

$$\beta^1(v) = \frac{1}{2}(v^1 + v^2) \quad \text{and} \quad \beta^2(v) = \frac{1}{2}(v^1 - v^2) \,.$$

□

Let us define

$$V^* := \{\text{all linear forms } \alpha : V \to \mathbb{R}\} \,.$$

EXERCISE 3.7. Check that V^* is a vector space whose null vector is the linear form identically equal to zero.

We called V^* the **dual** of V.

PROPOSITION 3.8. *Let $\mathcal{B} := \{b_1, \dots, b_n\}$ be a basis of V and $\beta^1, \dots, \beta^n$ are the corresponding coordinate forms. Then $\mathcal{B}^* := \{\beta^1, \dots, \beta^n\}$ is a basis of V^*. As a consequence*

$$\dim V = \dim V^* .$$

PROOF. According to Definition 2.6 , we need to check that the linear forms in $\mathcal{B}^*$

(1) span V and
(2) are linearly independent.

(1) To check that $\mathcal{B}^*$ spans V we need to verify that any $\alpha \in V^*$ is a linear combination of $\beta^1, \dots, \beta^n$, that is that

$$\alpha = \alpha_i \beta^i \tag{3.3}$$

for some $\alpha_i \in \mathbb{R}$. Because of (3.2), if we apply both sides of (3.3) to the j-th basis vector b_i, we obtain

$$\alpha(b_j) = \alpha_i \beta^i(b_j) = \alpha_i \delta^i_j = \alpha_j , \tag{3.4}$$

which identifies the coefficients in (3.3).

Now let $v = v^i b_i \in V$ be an arbitrary vector. Then

$$\alpha(v) = \alpha(v^i b_i) = v^i \alpha(b_i) = v^i \alpha_i ,$$

where the second equality follows form the definition of linear form and the third from (3.4).

On the other hand

$$\alpha_i \beta^i(v) = \alpha_i \beta^i(v^j b_j) = \alpha_i v^j \beta^i(b_j) = \alpha_i v^j \delta^i_j = \alpha_i v^i .$$

Thus (3.3) is verified.

(2) We need to check that the only linear combination of $\beta^1, \dots, \beta^n$ that gives the zero linear form is the trivial linear combination. Let $c_i \beta^i = 0$ be a linear combination of the β^i. Then for every basis vector b_j, with $j = 1, \dots, n$,

$$0 = (c_i \beta^i)(b_j) = c_i(\beta^i(b_j)) = c_i \delta^i_j = c_j ,$$

thus showing the linear independence. □

The basis $\mathcal{B}^*$ of V^* is called the **basis of V dual to $\mathcal{B}$**. We emphasize that the coordinates (or components) of a linear form α with respect to $\mathcal{B}^*$ are exactly the values of α on the elements of $\mathcal{B}$,

$$\alpha_i = \alpha(b_i) .$$

EXAMPLE 3.9. Let $V = \mathbb{R}[x]_2$ be the vector space of polynomials of degree ≤ 2, let $\alpha : V \to \mathbb{R}$ be the linear form given by

$$\alpha(p(x)) := p(2) - p'(2) \tag{3.5}$$

and let $\mathcal{B} := \{1, x, x^2\}$ be a basis of V. We want to:

(1) find the components of α with respect to $\mathcal{B}^*$;
(2) describe the basis $\mathcal{B}^* = \{\beta^1, \beta^2, \beta^3\}$;

(1) Since

$$\begin{aligned}\alpha_1 &= \alpha(b_1) = \alpha(1) = 1 - 0 = 1\\ \alpha_2 &= \alpha(b_2) = \alpha(x) = 2 - 1 = 1\\ \alpha_3 &= \alpha(b_3) = \alpha(x^2) = 4 - 4 = 0\,,\end{aligned}$$

then

$$[\alpha]_{\mathcal{B}^*} = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}\,. \tag{3.6}$$

(2) The generic element $p(x) \in \mathbb{R}[x]_2$ written as combination of basis elements $1, x$ and x^2 is

$$p(x) = a + bx + cx^2\,.$$

Hence $\mathcal{B}^* = \{\beta^1, \beta^2, \beta^3\}$, is given by

$$\begin{aligned}\beta^1(a + bx + cx^2) &= a\\ \beta^2(a + bx + cx^2) &= b\\ \beta^3(a + bx + cx^2) &= c\,.\end{aligned} \tag{3.7}$$

□

Remark 3.10. Note that it does not make sense to talk about a "dual basis" of V^*, as for every basis $\mathcal{B}$ of V there is going to be a basis $\mathcal{B}^*$ of V^* dual to the basis $\mathcal{B}$. In the next section we are going to see how the dual basis transform with a change of basis.

3.1.2. Transformation of Linear Forms under a Change of Basis. We want to study how a linear form $\alpha : V \to \mathbb{R}$ behaves with respect to a change a basis in V. To this purpose, let

$$\mathcal{B} := \{b_1, \dots, b_n\} \quad \text{and} \quad \widetilde{\mathcal{B}} := \{\tilde{b}_1, \dots, \tilde{b}_n\}$$

be two bases of V and let

$$\mathcal{B}^* := \{\beta^1, \dots, \beta^n\} \quad \text{and} \quad \widetilde{\mathcal{B}}^* := \{\tilde{\beta}^1, \dots, \tilde{\beta}^n\}$$

the corresponding dual bases. Let

$$[\alpha]_{\mathcal{B}^*} = \begin{pmatrix} \alpha_1 & \dots & \alpha_n \end{pmatrix} \quad \text{and} \quad [\alpha]_{\widetilde{\mathcal{B}}^*} = \begin{pmatrix} \widetilde{\alpha}_1 & \dots & \widetilde{\alpha}_n \end{pmatrix}$$

be the coordinate vectors of α with respect to $\mathcal{B}^*$ and $\widetilde{\mathcal{B}}^*$, that is

$$\alpha(b_i) = \alpha_i \quad \text{and} \quad \alpha(\tilde{b}_i) = \widetilde{\alpha}_i\,.$$

Let $L := L_{\widetilde{\mathcal{B}}\mathcal{B}}$ be the matrix of the change of basis in (2.3)

$$\tilde{b}_j = L^i_j b_i\,.$$

Then

$$\widetilde{\alpha}_j = \alpha(\tilde{b}_j) = \alpha(L^i_j b_i) = L^i_j \alpha(b_i) = L^i_j \alpha_i = \alpha_i L^i_j \,, \tag{3.8}$$

so that

$$\widetilde{\alpha}_j = \alpha_i L^i_j \,. \tag{3.9}$$

Exercise 3.11. Verify that (3.9) is equivalent to saying that

$$[\alpha]_{\widetilde{\mathcal{B}}^*} = [\alpha]_{\mathcal{B}^*} L \,. \tag{3.10}$$

Note that we have exchanged the order of α_i and L^i_j in the last equation in (3.8) to respect the order in which the matrix multiplication in (3.10) has to be performed. This was possible because both α_i and L^i_j are real numbers.

We say that the component of a linear form α are **covariant**[1] because they change by L when the basis changes by L. A linear form α is hence a **covariant tensor** or a **tensor of type** $(1,0)$.

Example 3.12. We continue with Example 3.9. We consider the bases as in Example 2.22, that is

$$\mathcal{B} := \{1, x, x^2\} \qquad \text{and} \qquad \widetilde{\mathcal{B}} := \{x, x-1, x^2-1\}$$

and the linear form $\alpha : V \to \mathbb{R}$ as in (3.5). We will:

(1) find the components of α with respect to $\mathcal{B}^*$;
(2) describe the basis $\mathcal{B}^* = \{\beta^1, \beta^2, \beta^3\}$;
(3) find the components of α with respect to $\widetilde{\mathcal{B}}^*$;
(4) describe the basis $\widetilde{\mathcal{B}}^* = \{\tilde{\beta}^1, \tilde{\beta}^2, \tilde{\beta}^3\}$;
(5) find the matrix of change of basis $L := L_{\widetilde{\mathcal{B}}\mathcal{B}}$ and compute $\Lambda = L^{-1}$;
(6) check the covariance of α;
(7) check the contravariance of $\mathcal{B}^*$.

(1) This is done in (3.6).

(2) This is done in (3.7).

(3) We proceed as in (3.6). Namely,

$$\begin{aligned}
\alpha_1 &= \alpha(\tilde{b}_1) = \alpha(x) = 2 - 1 = 1 \\
\alpha_2 &= \alpha(\tilde{b}_2) = \alpha(x-1) = 1 - 1 = 0 \\
\alpha_3 &= \alpha(\tilde{b}_3) = \alpha(x^2-1) = 3 - 4 = -1 \,,
\end{aligned}$$

so that

$$[\alpha]_{\widetilde{\mathcal{B}}^*} = \begin{pmatrix} 1 & -1 & -1 \end{pmatrix} \,.$$

[1] "co" is a prefix that in Latin means "joint".

(4) Since $\tilde{\beta}^i(v) = \tilde{v}^i$, to proceed as in (3.7) we first need to write the generic polynomial $p(x) = a + bx + cx^2$ as a linear combination of elements in $\widetilde{\mathcal{B}}$, namely we need to find $\tilde{a}, \tilde{b}$ and $\tilde{c}$ such that

$$p(x) = a + bx + cx^2 = \tilde{a}x + \tilde{b}(x-1) + \tilde{c}(x^2-1)\,.$$

By multiplying and collecting the terms, we obtain that

$$\begin{cases} -\tilde{b} - \tilde{c} = a \\ \tilde{a} + \tilde{b} = b \\ \tilde{c} = c \end{cases} \qquad \text{that is} \qquad \begin{cases} \tilde{a} = a + b + c \\ \tilde{b} = -a - c \\ \tilde{c} = c\,. \end{cases}$$

Hence

$$p(x) = a + bx + cx^2 = (a+b+c)x + (-a-c)(x-1) + c(x^2-1)\,,$$

so that it follows that

$$\begin{aligned} \beta^1(p(x)) &= a + b + c \\ \beta^2(p(x)) &= -a - c \\ \beta^3(p(x)) &= c\,, \end{aligned}$$

(5) The matrix of the change of bases is given by

$$L := L_{\widetilde{\mathcal{B}}\mathcal{B}} = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\,,$$

since for example $\tilde{b}_3$ can be written as a linear combination with respect to $\mathcal{B}$ as $\tilde{b}_3 = x^2 - 1 = -1b_1 + 0b_2 + 1b_3$, and hence its coordinates form the third column of L.

To compute $\Lambda = L^{-1}$ we can use the Gauss–Jordan elimination process

$$\left[\begin{array}{ccc|ccc} 0 & -1 & -1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \quad \leftrightsquigarrow \quad \dots \quad \leftrightsquigarrow \quad \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

Hence

$$\Lambda = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

(6) The linear form α is *covariant* since

$$\begin{pmatrix} \widetilde{\alpha}_1 & \widetilde{\alpha}_2 & \widetilde{\alpha}_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \begin{bmatrix} 0 & -1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{pmatrix} L$$

(7) The dual basis $\mathcal{B}^*$ is *contravariant* since

TABLE 1. Covariance and Contravariance

	The covariance of a tensor	The contravariace of a tensor
is characterized by	lower indices	upper indices
vectors are indicated as	row vectors	column vectors
the tensor transforms w.r.t. a change of basis $\mathcal{B} \to \widetilde{\mathcal{B}}$ by multiplication by	L on the right	L^{-1} on the left
(for later use) if a tensor is of type (p, q)	(p, q)	(p, q)

$$\begin{pmatrix} \tilde{\beta}^1 \\ \tilde{\beta}^2 \\ \tilde{\beta}^3 \end{pmatrix} = \Lambda \begin{pmatrix} \beta^1 \\ \beta^2 \\ \beta^3 \end{pmatrix},$$

as it can be verified by looking at an arbitrary vector $p(x) = a + bx + cx^2$

$$\begin{pmatrix} a+b+c \\ -a-c \\ c \end{pmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

□

In fact, the statement in Example 3.9(7) holds in general, namely:

CLAIM 3.13. Dual bases are **contravariant**.

PROOF. We will check that when bases $\mathcal{B}$ and $\widetilde{\mathcal{B}}$ are related by

$$\tilde{b}_j = L^i_j b_i$$

the corresponding dual bases $\mathcal{B}^*$ and $\widetilde{\mathcal{B}}^*$ of V^* are related by

$$\boxed{\tilde{\beta}^j = \Lambda^j_i \beta^i}. \tag{3.11}$$

It is enough to check that the $\Lambda^j_i \beta^i$ are *dual* of the $L^i_j b_i$. In fact, since $\Lambda L = I$, then

$$(\Lambda^k_\ell \beta^\ell)(L^i_j b_i) = \Lambda^k_\ell L^i_j \beta^\ell(b_i) = \Lambda^k_\ell L^i_j \delta^\ell_i = \Lambda^k_i L^i_j = \delta^k_j = \beta^j(\tilde{b}_j).$$

□

In Table 1 you will find a summary of the properties that characterize covariance and contravariance, while in Table 2 you can find a summary of the properties that bases and dual bases, coordinate vectors and coordinates of linear forms satisfy with respect to a change of coordinates and hence whether they are covariant or contravariant.

TABLE 2. Summary

V real vector space with $\dim V = n$ $\mathcal{B} := \{b_1, \dots, b_n\}$ basis of V $\widetilde{\mathcal{B}} := \{\tilde{b}_1, \dots, \tilde{b}_n\}$ another basis of V $L := L_{\widetilde{\mathcal{B}}\mathcal{B}}$ =matrix of the change of basis from $\mathcal{B}$ to $\widetilde{\mathcal{B}}$ $\tilde{b}_j = L^i_j b_i$ i.e. $\begin{pmatrix}\tilde{b}_1 & \dots & \tilde{b}_n\end{pmatrix} = \begin{pmatrix}b_1 & \dots & b_n\end{pmatrix} L$	$V^* = \{\alpha : V \to \mathbb{R}\}$ =linear forms = dual vector space $\mathcal{B}^* = \{\beta^1, \dots, \beta^n\}$ dual basis of V^* w.r.t $\mathcal{B}$ $\widetilde{\mathcal{B}}^* = \{\tilde{\beta}^1, \dots, \tilde{\beta}^n\}$ dual basis of V^* w.r.t $\widetilde{\mathcal{B}}^*$ $\Lambda = L^{-1}$ =matrix of the change of basis from $\widetilde{\mathcal{B}}$ to $\mathcal{B}$ $\tilde{\beta}^i = \Lambda^i_j \beta^j$ i.e. $\begin{pmatrix}\tilde{\beta}_1 \\ \vdots \\ \tilde{\beta}_n\end{pmatrix} = L^{-1}\begin{pmatrix}\beta_1 \\ \vdots \\ \beta_n\end{pmatrix}$
covariance of a basis	contravariance of the dual basis
If v is any vector in V then $v = v^i b_i = \tilde{v}^i \tilde{b}_i$ where $\tilde{v}^i = \Lambda^i_j v^j$ i.e. $[v]_{\widetilde{\mathcal{B}}} = L^{-1}[v]_{\mathcal{B}}$ or $\begin{pmatrix}\tilde{v}^1 \\ \vdots \\ \tilde{v}^n\end{pmatrix} = L^{-1}\begin{pmatrix}v^1 \\ \vdots \\ v^n\end{pmatrix}$	If α is any linear form in V^* then $\alpha = \alpha_j \beta^j = \widetilde{\alpha}_j \tilde{\beta}^j$ where $\widetilde{\alpha}_j = L^i_j \alpha_i$ i.e. $[\alpha]_{\widetilde{\mathcal{B}}} = [\alpha]_{\mathcal{B}} L$ or $\begin{pmatrix}\tilde{\alpha}_1 & \dots & \tilde{\alpha}_n\end{pmatrix} = \begin{pmatrix}\alpha_1 & \dots & \alpha_n\end{pmatrix} L$
contravariance of the coordinate vectors	covariance of linear forms

3.2. Bilinear Forms

3.2.1. Definition, Examples and Basis.

DEFINITION 3.14. A **bilinear form** on V is a function $\varphi : V \times V \to \mathbb{R}$ that is linear in each variable, that is

$$\begin{aligned}\varphi(u, \lambda v + \mu w) &= \lambda\varphi(u, v) + \mu\varphi(u, w)\\ \varphi(\lambda v + \mu w, u) &= \lambda\varphi(v, u) + \mu\varphi(w, u)\,,\end{aligned}$$

for every $\lambda, \mu \in \mathbb{R}$ and for every $u, v, w \in V$.

EXAMPLES 3.15. Let $V = \mathbb{R}^3$.

(1) The **scalar product**

$$\varphi(v, w) := v \bullet w = |v|\,|w|\cos\theta\,,$$

where θ is the angle between v and w is a bilinear form. It can be defined also for $n > 3$.

(2) Choose a vector $u \in \mathbb{R}^3$ and for any two vectors $v, w \in \mathbb{R}^3$, denote by $v \times w$ their **cross product**. The **scalar triple product**

$$\varphi_u(v, w) := u \bullet (v \times w) = \det \begin{bmatrix} u \\ v \\ w \end{bmatrix} \tag{3.12}$$

is a bilinear form in v and w, where $\begin{bmatrix} u \\ v \\ w \end{bmatrix}$ denotes the matrix with rows u, v and w. The quantity $\varphi_u(v, w)$ calculates the signed volume of the parallelepiped spanned by u, v, w: the sign of $\varphi_u(v, w)$ depends on the orientation of the triple u, v, w.

Since the cross product is defined only in $\mathbb{R}^3$, contrary to the scalar product, the scalar triple product cannot be defined in $\mathbb{R}^n$ with $n > 3$ (although there is a formula for an n dimensional paralleledipped involving some "generalization" of it).

□

Exercise 3.16. Verify the equality in (3.12) using the Leibniz formula for the determinant of a 3×3 matrix. Recall that

$$\begin{aligned} \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} =& a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} \\ & - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \\ =& \sum_{\sigma \in S_3} \operatorname{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} a_{3\sigma(3)}\,, \end{aligned}$$

where

$$\begin{aligned} \sigma = (\sigma(1), \sigma(2), \sigma(3)) \in S_3 :=& \ \{\text{permutations of 3 elements}\} \\ =& \ \{(1,2,3), (1,3,2), (2,3,1), (2,1,3), (3,1,2), (3,2,1)\}\,. \end{aligned}$$

Examples 3.17. Let $V = \mathbb{R}[x]_2$.

(1) Let $p, q \in \mathbb{R}[x]_2$. The function $\varphi(p, q) := p(\pi) q(33)$ is a bilinear form.
(2) Likewise,

$$\varphi(p, q) := p'(0) q(4) - 5p'(3) q''(\tfrac{1}{2})$$

is a bilinear form.

□

Exercise 3.18. Are the following functions bilinear forms?

(1) $V = \mathbb{R}^2$ and $\varphi(u, v) := \det \begin{bmatrix} u \\ v \end{bmatrix}$;
(2) $V = \mathbb{R}[x]_2$ and $\varphi(p, q) := \int_0^1 p(x) q(x) dx$;

(3) $V = M_{2\times 2}(\mathbb{R})$, the space of real 2×2 matrices, and $\varphi(L, M) := L^1_1 \operatorname{tr} M$, where L^1_1 it the (1,1)-entry of L and $\operatorname{tr} M$ is the trace of M;
(4) $V = \mathbb{R}^3$ and $\varphi(v, w) := v \times w$;
(5) $V = \mathbb{R}^2$ and $\varphi(v, w)$ is the area of the parallelogram spanned by v and w.

3.2.2. Tensor product of two linear forms on V. Let $\alpha, \beta \in V^*$ be two linear forms, $\alpha, \beta : V \to \mathbb{R}$, and define $\varphi : V \times V \to \mathbb{R}$, by

$$\varphi(v, w) := \alpha(v)\beta(w) .$$

Then φ is bilinear, is called the **tensor product** of α and β and is denoted by

$$\varphi = \alpha \otimes \beta .$$

NOTE 3.19. In general $\alpha \otimes \beta \neq \beta \otimes \alpha$, as there could be vectors v and w such that $\alpha(v)\beta(w) \neq \beta(v)\alpha(w)$.

EXAMPLE 3.20. Let $V = \mathbb{R}[x]_2$, let $\alpha(p) = p(2) - p'(2)$ and $\beta(p) = \int_3^4 p(x)dx$ be two linear forms. Then

$$(\alpha \otimes \beta)(p, q) = (p(2) - p'(2)) \int_3^4 q(x)dx$$

is a bilinear form. □

EXAMPLE 3.21. Let $\varphi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a function:
(1) $\varphi(x, y) := 2x - y$ is a *linear form* in $(x, y) \in \mathbb{R}^2$;
(2) $\varphi(x, y) := 2xy$ is *bilinear*, hence *linear* in $x \in \mathbb{R}$ and *linear* in $y \in \mathbb{R}$, but it is *not linear* in $(x, y) \in \mathbb{R}^2$.

□

Let

$$\operatorname{Bil}(V \times V, \mathbb{R}) := \{\text{all bilinear forms } \varphi : V \times V \to \mathbb{R}\} .$$

EXERCISE 3.22. Check that $\operatorname{Bil}(V \times V, \mathbb{R})$ is a vector space with the zero element equal to the bilinear form identically equal to zero.
Hint: It is enough to check that if $\varphi, \psi \in \operatorname{Bil}(V \times V, \mathbb{R})$, and $\lambda, \mu \in \mathbb{R}$, then $\lambda\varphi + \mu\psi \in \operatorname{Bil}(V \times V, \mathbb{R})$. Why? (Recall Example 2.3(4).)

Assuming Exercise 3.22, we are going to find a basis of $\operatorname{Bil}(V \times V, \mathbb{R})$ and determine its dimension. Let $\mathcal{B} := \{b_1, \dots, b_n\}$ be a basis of V and let $\mathcal{B}^* = \{\beta^1, \dots, \beta^n\}$ be the dual basis of V^* (that is $\beta^i(b_j) = \delta^i_j$).

PROPOSITION 3.23. *The bilinear forms* $\beta^1 \otimes \beta^j$, $i, j = 1, \dots, n$ *form a basis of* $\operatorname{Bil}(V \times V, \mathbb{R})$. *As a consequence* $\dim \operatorname{Bil}(V \times V, \mathbb{R}) = n^2$.

NOTATION. We denote

$$\boxed{\operatorname{Bil}(V \times V, \mathbb{R}) = V^* \otimes V^*}$$

the **tensor product** of V^* and V^*.

Proof of Proposition 3.23. The proof will be similar to the one of Proposition 3.8 for linear forms. We first check that the set of bilinear forms $\{\beta^1 \otimes \beta^j,\, i, j = 1, \dots, n\}$ span $\mathrm{Bil}(V \times V, \mathbb{R})$ and than that it consists of linearly independent elements.

To check that $\mathrm{span}\{\beta^i \otimes \beta^j,\, i, j = 1, \dots, n\} = \mathrm{Bil}(V \times V, \mathbb{R})$, we need to check that if $\varphi \in \mathrm{Bil}(V \times V, \mathbb{R})$, there exists $B_{ij} \in \mathbb{R}$ such that

$$\varphi = B_{ij} \beta^i \otimes \beta^j .$$

Because of (3.2), we obtain

$$\varphi(b_k, b_\ell) = B_{ij} \beta^i(b_k) \beta^j(b_\ell) = B_{ij} \delta^i_k \delta^j_\ell = B_{k\ell} ,$$

for every pair $(b_k, b_\ell) \in V \times V$. Hence we are forced to choose $B_{k\ell} := \varphi(b_k, b_\ell)$. Now we have to check that with this choice of $B_{k\ell}$ we have indeed

$$\varphi(v, w) = B_{ij} \beta^i(v) \beta^j(w)$$

for arbitrary $v = v^k b_k \in V$ and $w = w^\ell b_\ell \in V$.

On the one hand we have that

$$\varphi(v, w) = \varphi(v^k b_k, w^\ell b_\ell) = v^k w^\ell \varphi(b_k, b_\ell) = v^k w^\ell B_{k\ell} ,$$

where the next to the last equality follows from the bilinearity of φ and the last one from the definition of $B_{k\ell}$.

On the other hand,

$$\begin{aligned} B_{ij} \beta^i(v) \beta^j(w) &= B_{ij} \beta^i(v^k b_k) \beta^j(v^\ell b_\ell) \\ &= B_{ij} v^k \beta^i(b_k) v^\ell \beta^j(b_\ell) \\ &= B_{ij} v^k v^\ell \delta^i_k \delta^j_\ell \\ &= B_{k\ell} v^k w^\ell , \end{aligned}$$

where the second equality follows from the bilinearity of β^i and the next to the last from (3.2).

Now we need to check that the only linear combination of the $\beta^i \otimes \beta^j$ that gives the zero bilinear form is the trivial linear combination. Let $c_{ij} \beta^i \otimes \beta^j = 0$ be a linear combination of the $\beta^i \otimes \beta^j$. Then for all pairs of basis vectors (b_k, b_ℓ), with $k, \ell = 1, \dots, n$, we have

$$0 = c_{ij} \beta^i \otimes \beta^j (b_k, b_\ell) = c_{ij} \delta^i_k \delta^j_\ell = c_{k\ell} ,$$

thus showing the linear independence. □

3.2.3. Transformation of Bilinear Forms under a Change of Basis. If we summarize what we have done so far, we see that once we choose a basis $\mathcal{B} := \{b_1, \dots, b_n\}$ of V, we automatically have a basis $\mathcal{B}^* = \{\beta^1, \dots, \beta^n\}$ of V^* and a basis $\{\beta^1 \otimes \beta^j,\, i, j = 1, \dots, n\}$ of $V^* \otimes V^*$.

That is, any bilinear form $\varphi : V \times V \to \mathbb{R}$ can be represented by its components

$$B_{ij} = \varphi(b_i, b_j) \tag{3.13}$$

and these components can be arranged in a matrix

$$B := \begin{pmatrix} B_{11} & \dots & B_{1n} \\ \vdots & & \vdots \\ B_{n1} & \dots & B_{nn} \end{pmatrix}$$

called the **matrix of the bilinear form** φ with respect to the chosen basis $\mathcal{B}$. The natural question of course is: how does the matrix B change when we choose a different basis of V?

So, let us choose a different basis $\widetilde{\mathcal{B}} := \{\tilde{b}_1, \dots, \tilde{b}_n\}$ and corresponding bases $\widetilde{\mathcal{B}}^* = \{\tilde{\beta}^1, \dots, \tilde{\beta}^n\}$ of V^* and $\{\tilde{\beta}^i \otimes \tilde{\beta}^j,\ i, j = 1, \dots, n\}$ of $V^* \otimes V^*$, with respect to which φ will be represented by a matrix $\widetilde{B}$, whose entries are $\widetilde{B}_{ij} = \varphi(\tilde{b}_i, \tilde{b}_j)$.

To see the relation between B and $\widetilde{B}$, due to the change of basis from $\mathcal{B}$ to $\widetilde{\mathcal{B}}$, we start with the matrix of the change of basis $L := L_{\widetilde{\mathcal{B}}\mathcal{B}}$, according to which

$$\tilde{b}_j = L^i_j b_i \,. \tag{3.14}$$

Then

$$\widetilde{B}_{ij} = \varphi(\tilde{b}_i, \tilde{b}_j) = \varphi(L^k_i b_k, L^\ell_j b_\ell) = L^k_i L^\ell_j \varphi(b_k, b_\ell) = L^k_i L^\ell_j B_{k\ell} \,,$$

where the first and the last equality follow from (3.13), the second from (3.14) (after having renamed the dummy indices to avoid conflicts) and the remaining one from the bilinearity.

Exercise 3.24. Show that the formula of the transformation of the component of a bilinear form in terms of the matrices of the change of coordinates is

$$\widetilde{B} = {}^{\mathsf{t}}LBL \,, \tag{3.15}$$

where ${}^{\mathsf{t}}L$ denotes the transpose of the matrix L.

We hence say that a bilinear form φ is a **covariant 2-tensor** or a **tensor of type** $(0, 2)$.

3.3. Multilinear forms

We saw in § 3.1.2 that linear forms are covariant 1-tensors – or tensor of type $(0, 1)$ – and in § 3.2.3 that bilinear forms are covariant 2-tensors – or tensors of type $(0, 2)$.

Completely analogously to what was done until now, one can define **trilinear forms**, that is functions $T : V \times V \times V \to \mathbb{R}$ that are linear in each of the three variables. The space of trilinear forms is denoted by $V^* \otimes V^* \otimes V^*$, has basis $\{\beta^j \otimes \beta^j \otimes \beta^k,\ i, j, k = 1, \dots, n\}$ and hence dimension n^3.

Since the components of a trilinear form $T : V \times V \times V \to \mathbb{R}$ satisfy the following transformation with respect to a change of basis

$$\widetilde{T}_{ijk} = L^\ell_i L^p_j L^q_k T_{\ell pq} \,,$$

a trilinear form is a **covariant** 3**-tensor** or a **tensor of type** $(0, 3)$**.**

In fact, there is nothing special about $k = 1, 2$ or 3.

DEFINITION 3.25. A **multilinear form** if a function $f : V \times \cdots \times V \to \mathbb{R}$ from k-copies of V into $\mathbb{R}$, that is linear with respect to each variable.

A multilinear form is a **covariant k-tensor** or a **tensor of type** $(0, k)$. The vectors space of multilinear forms $V^* \otimes \cdots \otimes V^*$ has basis $\beta^{i_1} \otimes \beta^{i_2} \times \cdots \otimes \beta^{i_k}$, $i_1, \ldots, i_k := 1, \ldots, n$ and hence $\dim(V^* \otimes \cdots \otimes V^*) = n^k$.

3.4. Examples

3.4.1. A Bilinear Form.

EXAMPLE 3.26. We continue with the study of the *scalar triple product*, that was defined in Example 3.15. We want to find the components B_{ij} of φ_u with respect to the standard basis of $\mathbb{R}^3$. Let $u = \begin{bmatrix} u^1 \\ u^2 \\ u^3 \end{bmatrix}$ be the fixed vector. Recall the cross product in $\mathbb{R}^3$ is defined as

$$e_i \times e_j := \begin{cases} 0 & \text{if } i = j \\ e_k & \text{if } (i,j,k) \text{ is a cyclic permutation of } (1,2,3) \\ -e_k & \text{if } (i,j,k) \text{ is a non-cyclic permutation of } (1,2,3), \end{cases}$$

that is

$$\text{cyclic}\begin{cases} e_1 \times e_2 = e_3 \\ e_2 \times e_3 = e_1 \\ e_3 \times e_1 = e_2 \end{cases} \quad \text{and} \quad \text{non-cyclic}\begin{cases} e_2 \times e_1 = -e_3 \\ e_3 \times e_2 = -e_1 \\ e_1 \times e_3 = -e_2 \end{cases}$$

Since $u \bullet e_k = u^k$, then

$$B_{ij} = \varphi_u(e_i, e_j) = u \bullet (e_i \times e_j) = \begin{cases} 0 & \text{if } i = j \\ u^k & \text{if } (i,j,k) \text{ is a cyclic permutation of } (1,2,3) \\ -u^k & \text{if } (i,j,k) \text{ is a non-cyclic permutation of } (1,2,3) \end{cases}$$

Thus

$$\begin{aligned} &B_{12} = u^3 = -B_{21} \\ &B_{23} = u^1 = -B_{32} \\ &B_{II} = 0 \text{ (that is the diagonal components are zero)}, \end{aligned}$$

which can be written as a matrix

$$B = \begin{pmatrix} 0 & u^3 & -u^2 \\ -u^3 & 0 & u^1 \\ u^2 & -u^1 & 0 \end{pmatrix}.$$

We look now for the matrix of the scalar tripe product with respect to the basis

$$\widetilde{\mathcal{B}} := \left\{ \underbrace{\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}}_{\tilde{b}_1}, \underbrace{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}}_{\tilde{b}_2}, \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{\tilde{b}_3} \right\}.$$

The matrix of the change of coordinates from the standard basis to $\widetilde{\mathcal{B}}$ is

$$L = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix},$$

so that

$$\begin{aligned} \widetilde{B} &= \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{{}^{t}} \underbrace{\begin{bmatrix} 0 & u^3 & -u^2 \\ -u^3 & 0 & u^1 \\ u^2 & -u^1 & 0 \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}}_{L} \\ &= \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{{}^{t}L} \underbrace{\begin{bmatrix} u^3 & -u^2 & -u^2 \\ 0 & u^1 - u^3 & u^1 \\ -u^1 & u^2 & 0 \end{bmatrix}}_{BL} = \begin{bmatrix} 0 & u^1 - u^3 & u^1 \\ u^3 - u^1 & 0 & -u^2 \\ -u^1 & u^2 & 0 \end{bmatrix}. \end{aligned}$$

It is easy to check that $\widetilde{B}$ is antisymmetric just like B is, and to check that the components of $\widetilde{B}$ are correct by using the formula for φ. In fact

$$\begin{aligned} \widetilde{B}_{12} &= \varphi(\tilde{b}_1, \tilde{b}_2) = u \bullet (e_2 \times (e_1 + e_3)) = u^1 - u^3 \\ \widetilde{B}_{13} &= \varphi(\tilde{b}_2, \tilde{b}_3) = u \bullet ((e_2) \times e_3) = u^1 \\ \widetilde{B}_{23} &= \varphi(\tilde{b}_2, \tilde{b}_3) = u \bullet ((e_1 + e_3) \times e_3) = -u^2 \\ \widetilde{B}_{11} &= \varphi(\tilde{b}_1, b_1) = u \bullet (e_2 \times e_2) = 0 \\ \widetilde{B}_{22} &= \varphi(\tilde{b}_2, b_2) = u \bullet ((e_1 + e_3) \times (e_1 + e_3)) = 0 \\ \widetilde{B}_{33} &= \varphi(\tilde{b}_3, b_3) = u \bullet (e_3 \times e_3) = 0 \end{aligned}$$

□

3.4.2. A Trilinear Form.

EXAMPLE 3.27. If in the definition of the scalar triple product instead of fixing a vector $a \in \mathbb{R}$, we let the vector vary, we have a function $\varphi : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, defined by

$$\varphi(u, v, w) := u \bullet (v \times w) = \det \begin{bmatrix} u \\ v \\ w \end{bmatrix}.$$

One can verify that such function is trilinear, that is linear in each of the three variables separately.

3.5. Basic Operation on Multilinear Forms

Let $T : \underbrace{V \times \cdots \times V}_{k \text{ times}} \to \mathbb{R}$ and $U : \underbrace{V \times \cdots \times V}_{\ell \text{ times}} \to \mathbb{R}$ be respectively a k-linear and an ℓ-linear form. Then the **tensor product** of T and U

$$T \otimes U : \underbrace{V \times \cdots \times V}_{k+\ell \text{ times}} \to \mathbb{R}\,,$$

defined by

$$T \otimes U(v_1, \dots, v_{k+\ell}) := T(v_1, \dots, v_k)U(v_{k+1}, \dots, v_{k+\ell})$$

is a $(k+\ell)$-linear form.

Likewise, one can take the tensor product of a tensor of type $(0,k)$ and a tensor of type $(0,\ell)$ to obtain a tensor of type $(0,k+\ell)$.

CHAPTER 4

Inner Products

4.1. Definitions and First Properties

Inner products are a special case of bilinear forms. They add an important structure to a vector space, as for example they allow to compute the length of a vector. Moreover, they provide a canonical identification between the vector space V and its dual V^*.

DEFINITION 4.1. An **inner product** $g : V \times V \to \mathbb{R}$ on a vector space V is a *bilinear form* on V that is

(1) *symmetric*, that is $g(v, w) = g(w, v)$ for all $v, w \in V$ and
(2) *positive definite*, that is $g(v, v) \geq 0$ for all $v \in V$, and $g(v) = 0$ if and only if $v = 0$.

EXERCISE 4.2. Let $V = \mathbb{R}^3$. Determine whether the following bilinear forms are inner products, by verifying whether they are symmetric and positive definite:

(1) the scalar or dot product $\varphi(v, w) := v \bullet w$, defined as

$$v \bullet w = v^i w^i ,$$

where $v = \begin{bmatrix} v^1 \\ v^2 \\ v^3 \end{bmatrix}$ and $w = \begin{bmatrix} w^1 \\ w^2 \\ w^3 \end{bmatrix}$;

(2) $\varphi(v, w) := -v \bullet w$, for all $v, w \in V$;
(3) $\varphi(v, w) = v \bullet w + 2v^1 w^2$, for $v, w \in V$;
(4) $\varphi(v, w) = v \bullet 3w$, for $v, w \in V$.

EXERCISE 4.3. Let $V := \mathbb{R}[x]_2$ be the vector space of polynomials of degree ≤ 2. Determine whether the following bilinear forms are inner products, by verifying whether they are symmetric and positive definite:

(1) $\varphi(p, q) = \int_0^1 p(x)q(x)dx$;
(2) $\varphi(p, q) = \int_0^1 p'(x)q'(x)dx$;
(3) $\varphi(p, q) = \int_3^\pi e^x p(x)q(x)dx$;
(4) $\varphi(p, q) = p(1)q(1) + p(2)q(2)$;
(5) $\varphi(p, q) = p(1)q(1) + p(2)q(2) + p(3)q(3)$.

DEFINITION 4.4. Let $g : V \times V \to \mathbb{R}$ be an inner product on V.

(1) The **norm** $\|v\|$ of a vector $v \in V$ is defined as
$$\|v\| := \sqrt{g(v,v)}\,.$$
(2) A vector $v \in V$ is **unit vector** if $\|v\| = 1$;
(3) Two vectors $v, w \in V$ are **orthogonal** (that is **perpendicular or** $v \perp w$), if $g(v,w) = 0$;
(4) Two vectors $v, w \in V$ are **orthonormal** if they are orthogonal and $\|v\| = \|w\| = 1$;
(5) A basis $\mathcal{B}$ of V is an **orthonormal basis** if $b_1, \dots, b_n$ are pairwise orthonormal vectors, that is

$$g(b_i, b_j) = \delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j\,, \end{cases} \tag{4.1}$$

for all $i, j = 1 \dots, n$. The condition for $i = j$ implies that an orthonormal basis consists of unit vectors, while the one for $i \neq j$ implies that it consists of pairwise orthogonal vectors.

Example 4.5. (1) Let $V = \mathbb{R}[x]_2$ and g the standard inner product. The standard basis $\mathcal{B} = \{e_1, \dots, e_n\}$ is an orthonormal basis with respect to the standard inner product.
(2) Let $V = \mathbb{R}[x]_2$ and let $g(p,q) := \int_{-1}^{1} p(x)q(x)dx$. Check that the basis
$$\mathcal{B} = \{p_1, p_2, p_3\}\,,$$
where
$$p_1(x) := \frac{1}{\sqrt{2}}, \quad p_2(x) := \sqrt{\frac{3}{2}}x, \quad p_3(x) := \sqrt{\frac{5}{8}}(3x^2 - 1)\,,$$
is an orthonormal basis with respect to the inner product g.

Remark 4.6. p_1, p_2, p_3 are the first three **Legendre polynomials** up to scaling.

An inner product g on a vector space V is also called a **metric** on V.

4.1.1. Correspondence Between Inner Products and Symmetric Positive Definite Matrices. Recall that a matrix $S \in M_{n\times n}(\mathbb{R})$ is **symmetric** if $S = {}^{t}S$, that is if

$$S = \begin{bmatrix} * & A & B & \iddots \\ A & * & C & \iddots \\ B & C & * & \iddots \\ \iddots & \iddots & & * \end{bmatrix}.$$

Moreover if S is symmetric, then

(1) S is **positive definite** if ${}^{t}vSv > 0$ for all $v \in \mathbb{R}^n$;
(2) S is **negative definite** if ${}^{t}vSv < 0$ for all $v \in \mathbb{R}^n$;
(3) S is **positive semidefinite** if ${}^{t}vSv \geq 0$ for all $v \in \mathbb{R}^n$;
(4) S is **positive semidefinite** if ${}^{t}vSv \leq 0$ for all $v \in \mathbb{R}^n$;
(5) S is **indefinite** if $v^{t}Sv$ takes both positive and negative values for different $v \in \mathbb{R}^n$.

Definition 4.7. A **quadratic form** $Q : \mathbb{R}^n \to \mathbb{R}$ is a homogeneous quadratic polynomial in n variables.

Any symmetric matrix S correspond to a **quadratic form** as follow:

$$S \mapsto Q_S \,,$$

where $Q_S : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$Q_S(v) = {}^{t}vSv = \underbrace{\begin{bmatrix} v^1 & \dots & v^n \end{bmatrix} S \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix}}_{\text{matrix notation}} S_{ij} = \underbrace{v^i v^j S_{ij}}_{\text{Einstein notation}} \quad . \tag{4.2}$$

Note that Q is *not* linear in v.

Let S be a symmetric matrix and Q_S be the corresponding quadratic form. The notion of positive definiteness, etc. for S can be translated into corresponding properties for Q_S, namely:

(1) Q is **positive definite** if $Q(v) > 0$ for all $v \in V$;
(2) Q is **negative definite** if $Q(v) < 0$ for all $v \in V$;
(3) Q is **positive semidefinite** if $Q(v) \geq 0$ for all $v \in V$;
(4) Q is **negative semidefinite** if $Q(v) \leq 0$ for all $v \in V$;
(5) Q is **indefinite** if $Q(v)$ takes both positive and negative values.

To find out the type of a symmetric matrix S (or, equivalently of a quadratic form Q_S) it is enough to look at the eigenvalues of S, namely:

(1) S and Q_S are *positive definite* if all eigenvalues of S are positive:
(2) S and Q_S are *negative definite* if all eigenvalues of S are negative;
(3) S and Q_S are *positive semidefinite* if all eigenvalues of S are non-negative;
(4) S and Q_S are *negative semidefinite* if all eigenvalues of S are non-positive;
(5) S and Q_S are *indefinite* if S has both positive and negative eigenvalues.

The reason this makes sense is the same reason for which we need to restrict our attention to symmetric matrices and lies in the so-called Spectral Theorem:

Theorem 4.8. *[Spectral Theorem] Any symmetric matrix S has the following properties:*

(1) *it has only* real *eigenvalues;*
(2) *it is* diagonalizable*;*

(3) *it admits an* orthonormal eigenbasis, *that is a basis* $\{b_1, \dots, b_n\}$ *such that the* b_j *are orthonormal and are eigenvectors of* S.

4.1.1.1. *From Inner Products to Symmetric Positive Definite Matrices.* Let $\mathcal{B} := \{b_1, \dots, b_n\}$ be a basis of V. The **components of g with respect to $\mathcal{B}$** are

$$g_{ij} = g(b_i, b_j)\,. \tag{4.3}$$

Let G be the matrix with entries g_{ij}

$$G = \begin{bmatrix} g_{11} & \cdots & g_{1n} \\ \vdots & \ddots & \vdots \\ g_{n1} & \cdots & g_{nn} \end{bmatrix}. \tag{4.4}$$

We claim that G is symmetric and positive definite. In fact:

(1) Since g is *symmetric*, then for $1 \leq i, j \leq n$,

$$g_{ij} = g(b_i, b_j) = g(b_j, b_i) = g_{ji} \Rightarrow G \text{ is a } \textit{symmetric} \text{ matrix;}$$

(2) Since g is *positive definite*, then G is *positive definite* as a symmetric matrix. In fact, let $v = v^i b_i, w = w^j b_j \in V$ be two vectors. Then, using the bilinearity of g in (1), (4.3) and with the Einstein notation, we have:

$$g(v, w) = g(v^i b_i, w^j b_j) \overset{(1)}{=} v^i w^j \underbrace{g(b_i, b_j)}_{g_{ij}} = v^i w^j g_{ij}$$

or, in matrix notation,

$$g(v, w) = {}^{\mathtt{t}}[v]_{\mathcal{B}} G [w]_{\mathcal{B}} = \begin{bmatrix} v^1 & \dots & v^n \end{bmatrix} G \begin{bmatrix} w^1 \\ \vdots \\ w^n \end{bmatrix}.$$

4.1.1.2. *From Symmetric Positive Definite Matrices to Inner Products.* If S is a symmetric positive definite matrix, then the assignment

$$(v, w) \mapsto {}^{\mathtt{t}}vSv$$

defines a map that is easily seen to be bilinear, symmetric and positive definite and is hence an inner product.

4.1.2. Orthonormal Basis. Suppose that there in a basis $\mathcal{B} := \{b_1, \dots, b_n\}$ of V consisting of orthonormal vectors with respect to g, so that

$$g_{ij} = \delta_{ij}\,,$$

because of Definition 4.4(5) and of (4.3). In other words the symmetric matrix corresponding to the inner product g in the basis consisting of orthonormal vectors is the identity matrix. Moreover

$$g(v, w) = v^i w^j g_{ij} = v^i w^j \delta_{ij} = v^i w^i\,,$$

so that, if $v = w$,

$$\|v\|^2 = g(v,v) = v^i v^i = (v^1)^2 + \cdots + (v^n)^2 .$$

We deduce the following important

FACT 4.9. *Any inner product g can be expressed in the standard form*

$$g(v,w) = v^i w^i ,$$

as long as $[v]_{\mathcal{B}} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$ *and* $[w]_{\mathcal{B}} = \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix}$ *are the coordinates of v and w with respect to an orthonormal basis $\mathcal{B}$ for g.*

EXAMPLE 4.10. Let g be an inner product of $\mathbb{R}^3$ with respect to which

$$\widetilde{\mathcal{B}} := \Bigg\{ \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{\tilde{b}_1}, \underbrace{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}_{\tilde{b}_2}, \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{\tilde{b}_3} \Bigg\}$$

is an orthonormal basis. We want to express g with respect to the standard basis $\mathcal{E}$ of $\mathbb{R}^3$

$$\mathcal{E} := \Bigg\{ \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{e_1}, \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{e_2}, \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{e_3} \Bigg\} .$$

The matrices of the change of basis are

$$L := L_{\widetilde{\mathcal{B}}\mathcal{E}} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{and } \Lambda = L^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} .$$

Since g is a bilinear form, we saw in (3.15) that its matrices with respect to a change of basis are related by the formula

$$\widetilde{G} = {}^{\mathtt{t}}LGL .$$

Since the basis $\widetilde{\mathcal{B}}$ is orthonormal with respect to g, the associated matrix $\widetilde{G}$ is the identity matrix, so that

$$\begin{aligned} G &= {}^{\mathtt{t}}\Lambda \widetilde{G} \Lambda = {}^{\mathtt{t}}\Lambda\Lambda \\ &= \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} . \end{aligned} \tag{4.5}$$

It follows that, with respect to the standard basis, g is given by

$$\begin{aligned} g(v,w) &= \begin{pmatrix} v^1 & v^2 & v^3 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} w^1 \\ w^2 \\ w^3 \end{pmatrix} \\ &= v^1w^1 - v^1w^2 - v^2w^1 + 2v^2w^2 \\ &\qquad -v^2w^3 - w^3v^2 + 2v^3w^3 . \end{aligned} \tag{4.6}$$

□

EXERCISE 4.11. Verify the formula (4.6) for the inner product g in the coordinates of the basis $\widetilde{\mathcal{B}}$ by applying the matrix of the change of coordinate directly on the coordinates vectors $[v]_{\mathcal{E}}$.

REMARK 4.12. Norm and inner product of vectors depend *only* on the choice of g, but *not* on the choice of basis: different coordinate expressions yield the same result.

EXAMPLE 4.13. We verify the assertion of the previous remark with the inner product in Example 4.10. Let $v, w \in \mathbb{R}^3$ such that

$$[v]_{\mathcal{E}} = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \quad \text{and} \quad [v]_{\widetilde{\mathcal{B}}} = \begin{pmatrix} \tilde{v}^1 \\ \tilde{v}^2 \\ \tilde{v}^3 \end{pmatrix} = L^{-1} \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}$$

and

$$[w]_{\mathcal{E}} = \begin{pmatrix} w^1 \\ w^2 \\ w^3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{and} \quad [w]_{\widetilde{\mathcal{B}}} = \begin{pmatrix} \tilde{w}^1 \\ \tilde{w}^2 \\ \tilde{w}^3 \end{pmatrix} = L^{-1} \begin{pmatrix} w^1 \\ w^2 \\ w^3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 3 \end{pmatrix} .$$

Then with respect to the basis $\widetilde{\mathcal{B}}$ we have that

$$g(v,w) = 1 \cdot (-1) + 1 \cdot (-1) + 1 \cdot 3 = 1 ,$$

and also with respect to the basis $\mathcal{E}$

$$g(v,w) = 3 \cdot 1 - 3 \cdot 2 - 2 \cdot 1 + 2 \cdot 2 \cdot 2 - 2 \cdot 3 - 1 \cdot 2 + 2 \cdot 1 \cdot 3 = 1 .$$

□

EXERCISE 4.14. Verify that $\|v\| = \sqrt{3}$ and $\|w\| = \sqrt{11}$, when computed with respect of both bases.

Let $\mathcal{B} := \{b_1, \dots, b_n\}$ be an orthonormal basis and let $v = v^i b_i$ be a vector in V. We want to compute the coordinates v^i of v with respect of the metric g and of the elements of the basis. In fact

$$g(v, b_j) = g(v^i b_i, b_j) = v^i g(b_i, b_j) = v^i \delta_{ij} = v^j ,$$

that is the coordinates of a vector with respect to an orthonormal basis are the inner product of the vector with the basis vectors. This is particularly nice, so that we have to make sure that we remember how to construct an orthonormal basis from a given arbitrary basis.

Recall (Gram–Schmidt orthogonalization process). The Gram–Schmidt orthogonalization process is a recursive process that allows us to obtain an orthonormal basis starting from an arbitrary one. Let $\mathcal{B} := \{b_1, \dots, b_n\}$ be an arbitrary basis, let $g : V \times V \to \mathbb{R}$ be an inner product and $\| \cdot \|$ the corresponding norm.

We start by defining

$$u_1 := \frac{1}{\|b_1\|} b_1 \, .$$

Next, observe that $g(b_2, u_1)u_1$ is the projection of the vector b_2 in the direction of u_1. It follows that

$$b_2^{\perp} := b_2 - g(b_2, u_1)u_1$$

is a vector orthogonal to u_1, but not necessarily of unit norm. Hence we set

$$u_2 := \frac{1}{\|b_2\|} b_2^{\perp} \, .$$

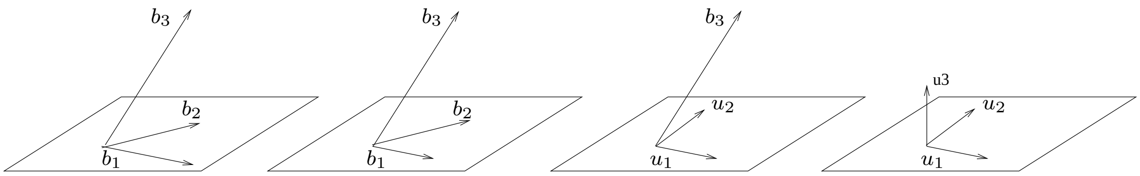
Likewise $g(b_3, u_1)u_1 + (b_3, u_2)u_2$ is the projection of b_3 on the plane generated by u_1 and u_2, so that

$$b_3^{\perp} := b_3 - g(b_3, u_1)u_1 - g(b_3, u_2)u_2$$

is orthogonal both to u_1 and to u_2. Set

$$u_3 := \frac{1}{\|b_3\|} b_3^{\perp} \, .$$

Continuing until we have exhausted all elements of the basis $\mathcal{B}$, we obtain an orthonormal basis $\{u_1, \dots, u_n\}$.



Example 4.15. Let V be the subspace of $\mathbb{R}^4$ spanned by

$$b_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \quad b_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \\ 0 \end{bmatrix} \quad b_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} .$$

(One can check that b_1, b_2, b_3 are linearly independent and hence form a basis of V.) We look for an orthonormal basis of V with respect to the standard inner product $\langle \cdot \, , \cdot \rangle$. Since

$$\|b_1\| = (1^1 + 1^2 + (-1)^2 + (-1)^2)^{1/2} = 2 \Rightarrow u_1 := \frac{1}{2} b_1 \, .$$

Moreover

$$\langle b_2, u_1 \rangle = \frac{1}{2}(1+1) = 2 \qquad \Longrightarrow \qquad b_2^\perp := b_2 - \langle b_2, u_1 \rangle u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},$$

so that

$$\|b_2\| = 2 \qquad \text{and} \qquad u_2 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Finally,

$$\langle b_3, u_1 \rangle = \frac{1}{2}(1+1-1) = \frac{1}{2} \qquad \text{and} \qquad \langle b_3, u_2 \rangle = \frac{1}{2}(1+1+1) = \frac{3}{2}$$

implies that

$$b_3^\perp := b_3 - \langle b_3, u_1 \rangle u_1 - \langle b_3, u_2 \rangle u_2 = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}.$$

Since

$$\|b_3^\perp\| = \frac{\sqrt{2}}{2} \qquad \Longrightarrow \qquad u_3 := \frac{\sqrt{2}}{2}\begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

4.2. Reciprocal Basis

Let $g : V \times V \to \mathbb{R}$ be an inner product and $\mathcal{B} := \{b_1, \dots, b_n\}$ any basis of V. From g and $\mathcal{B}$ we can define another basis of V, denoted by

$$\mathcal{B}^g = \{b^1, \dots, b^n\}$$

and satisfying

$$\boxed{g(b^i, b_j) = \delta^i_j}. \tag{4.7}$$

The basis $\mathcal{B}^g$ is called the **reciprocal basis** of V with respect to g and $\mathcal{B}$.

Note that, strictly speaking, we are very imprecise here. In fact, while it is certainly possible to define a set of $n = \dim V$ vectors as in (4.7), we should justify the fact that we call it a *basis*. This will be done in Claim 4.18.

Remark 4.16. In general $\mathcal{B}^g \neq \mathcal{B}$ and in fact, because of Definition 4.4(5),

$$\mathcal{B} = \mathcal{B}^g \Longleftrightarrow \mathcal{B} \text{ is an orthonormal basis.}$$

EXAMPLE 4.17. Let g be the inner product in (4.6) in Example 4.10 and let $\mathcal{E}$ the standard basis of $\mathbb{R}^3$. We want to find the reciprocal basis $\mathcal{E}^g$, that is we want to find $\mathcal{E}^g := \{b^1, b^2, b^3\}$ such that

$$g(b^i, e_j) = \delta^i_j \,.$$

If G is the matrix of the inner product in (4.5), using the matrix notation and considering b^j as a row vector and e_i as a column vector for $i, j = 1, 2, 3$,

$$\begin{bmatrix} — & {}^{\mathtt{t}}b^i & — \end{bmatrix} G \begin{bmatrix} | \\ e_j \\ | \end{bmatrix} = \delta^i_j \,.$$

Letting i and j vary from 1 to 3, we obtain

$$\begin{bmatrix} — & {}^{\mathtt{t}}b^1 & — \\ — & {}^{\mathtt{t}}b^2 & — \\ — & {}^{\mathtt{t}}b^3 & — \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} | & | & | \\ e_1 & e_2 & e_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} ,$$

from which we conclude that

$$\begin{aligned} \begin{bmatrix} — & {}^{\mathtt{t}}b^1 & — \\ — & {}^{\mathtt{t}}b^2 & — \\ — & {}^{\mathtt{t}}b^3 & — \end{bmatrix} &= \begin{bmatrix} | & | & | \\ e_1 & e_2 & e_3 \\ | & | & | \end{bmatrix}^{-1} \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} | & | & | \\ e_1 & e_2 & e_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} . \end{aligned}$$

Hence

$$b^1 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} , \quad b^2 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} , \quad b^3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} . \tag{4.8}$$

Observe that in order to compute G^{-1} we used the Gauss–Jordan elimination method

$$\begin{aligned} \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right] &\rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right] \\ &\rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 2 & 1 & 0 \\ 0 & 1 & -1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right] \\ &\rightsquigarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 2 & 1 \\ 0 & 1 & 0 & 2 & 2 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right] \end{aligned}$$

4.2.1. Properties of Reciprocal Bases.

CLAIM 4.18. Given a vector space V with a basis $\mathcal{B}$ and an inner product $g : V\times V \to \mathbb{R}$, a reciprocal basis *exists* and is *unique*.

As we pointed out right after the definition of reciprocal basis, what this claim really says is that there is a set of vectors $\{b^1, \dots, b^n\}$ in V that satisfy (4.7), that form a basis and that this basis is unique.

PROOF. Let $\mathcal{B} := \{b_1, \dots, b_n\}$ be the given basis. Any other basis $\{b^1, \dots, b^n\}$ is related to $\mathcal{B}$ by the relation

$$b^i = M^{ij} b_j \tag{4.9}$$

for some *invertible* matrix M. We want to show that there exists a *unique* matrix M such that, when (4.9) is plugged into $g(b^i, b_j)$, we have

$$g(b^i, b_j) = \delta^i_j \,. \tag{4.10}$$

From (4.9) and (4.10) we obtain

$$\delta^i_j = g(b^i, b_j) = g(M^{ik} b_k, b_j) = M^{ik} g(b_k, b_j) = M^{ik} g_{kj} \,,$$

which, in matrix notation becomes

$$I = MG \,,$$

where G is the matrix of g with respect to $\mathcal{B}$ whose entries are g_{ij} as in (4.4). Since G is invertible because it is positive definite, then $M = G^{-1}$ exists and is unique. □

REMARK 4.19. Note that in the course of the proof we have found that, since $M = L_{\mathcal{B}^g\mathcal{B}}$, then

$$\boxed{G = (L_{\mathcal{B}^g\mathcal{B}})^{-1} = L_{\mathcal{B}\mathcal{B}^g}} \,.$$

We denote with g^{ij} the entries of $M = G^{-1}$. From the above discussion, it follows that with this notation

$$\boxed{g^{ik} g_{kj} = \delta^i_j} \tag{4.11}$$

as well as

$$\boxed{b^i = g^{ij} b_j} \,, \tag{4.12}$$

or

$$\boxed{\begin{pmatrix} b^1 & \dots & b^n \end{pmatrix} = \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix} G^{-1}} \,. \tag{4.13}$$

(check for example the dimensions and the indices to understand why G^{-1} has to be multiplied on the right). We can now compute $g(b^i, b^j)$

$$\begin{aligned} g(b^i, b^j) &\overset{(4.12)}{=} g(g^{ik} b_k, g^{j\ell} b_\ell) = g^{ik} g^{j\ell} g(b_k, b_\ell) \\ &\overset{(4.1)}{=} g^{j\ell} \delta^i_\ell \overset{(4.11)}{=} g^{ji} = g^{ij} \,, \end{aligned}$$

where we used in the second equality the bilinearity of g and in the last its symmetry. Thus, similarly to (4.1), we have

$$\boxed{g^{ij} = g(b^i, b^j)}\,. \tag{4.14}$$

Given that we just proved that reciprocal basis are unique, we can talk about *the* reciprocal basis (of a fixed vector space V associated to a basis and an inner product).

Claim 4.20. The reciprocal basis is **contravariant**.

Proof. Let $\mathcal{B}$ and $\widetilde{\mathcal{B}}$ be two bases of V and $L := L_{\widetilde{\mathcal{B}}\mathcal{B}}$ be the corresponding matrix of the change of bases, with $\Lambda = L^{-1}$. Recall that this means that

$$\tilde{b}_i = L^j_i b_j\,.$$

We have to check that if $\mathcal{B}^g = \{b^1, \dots, b^n\}$ is a reciprocal basis for $\mathcal{B}$, then the basis $\{\tilde{b}^1, \dots, \tilde{b}^n\}$ defined by

$$\tilde{b}^i = \Lambda^i_k b^k \tag{4.15}$$

is a reciprocal basis for $\widetilde{\mathcal{B}}$. Then the assertion will be proven, since $\{\tilde{b}^1, \dots, \tilde{b}^n\}$ is contravariant by construction.

To check that $\{\tilde{b}^1, \dots, \tilde{b}^n\}$ is the reciprocal basis, we need with check that with the choice of $\tilde{b}^i$ as in (4.15), the property (4.1) of the reciprocal basis is verified, namely that

$$g(\tilde{b}^i, \tilde{b}_j) = \delta^i_j\,.$$

But in fact,

$$g(\tilde{b}^i, \tilde{b}_j) \overset{(4.15)}{=} g(\Lambda^i_k b^k, L^\ell_j b_\ell) = \Lambda^i_k L^\ell_j g(b^k, b_\ell) \overset{(4.10)}{=} \Lambda^i_k L^\ell_j \delta^k_\ell = \Lambda^i_k L^k_j = \delta^i_j\,,$$

where the second equality comes from the bilinearity of g, the third from the property (4.7) defining reciprocal basis and the last from the fact that $\Lambda = L^{-1}$. ☐

Suppose now that V is a vector space with a basis $\mathcal{B}$ and that $\mathcal{B}^g$ is the reciprocal basis of V with respect to $\mathcal{B}$ and to a fixed inner product $g : V \times V \to \mathbb{R}$. Then there are two ways of writing a vector $v \in V$, namely

$$v = \underbrace{v^i b_i}_{\text{with respect to } \mathcal{B}} = \underbrace{v_j b^j}_{\text{with respect to } \mathcal{B}^g}\,.$$

Recall that the (ordinary) coordinates of v with respect to $\mathcal{B}$ are *contravariant* (see Example 1.2).

Claim 4.21. Vector coordinates with respect to the reciprocal basis are **covariant**.

PROOF. This will follow from the fact that the reciprocal basis is contravariant and the idea of the proof is the same as in Claim 4.20.

Namely, let $\mathcal{B}, \widetilde{\mathcal{B}}$ be two bases of V, $L := L_{\widetilde{\mathcal{B}}\mathcal{B}}$ the matrix of the change of basis and $\Lambda = L^{-1}$. Let $\mathcal{B}^g$ and $\widetilde{\mathcal{B}}^g$ be the corresponding reciprocal bases and $v = v_j b^j$ a vector with respect to $\mathcal{B}^g$.

It is enough to check that the numbers

$$\tilde{v}_i := L^j_i v_j$$

are the coordinates of v with respect to $\widetilde{\mathcal{B}}^g$, because in fact these coordinates are covariant by definition. But in fact, using this and (4.15), we obtain

$$\tilde{v}_i \tilde{b}^i = (L^j_i v_j)(\Lambda^i_k b^k) = \underbrace{L^j_i \Lambda^i_k}_{\delta^j_k} v_j b^k = v_j b^j = v$$

□

DEFINITION 4.22. The coordinates v_i of a vector $v \in V$ with respect to the reciprocal basis $\mathcal{B}^g$ are called the **covariant coordinates** of v.

4.2.2. Change of basis from a basis $\mathcal{B}$ to its reciprocal basis $\mathcal{B}^g$. We want to look now at the direct relationship between the covariant and the contravariant coordinates of a vector v. Recall that we can write

$$\underbrace{v^i b_i}_{\text{with respect to } \mathcal{B}} = v = \underbrace{v_j b^j}_{\text{with respect to } \mathcal{B}^g} .$$

from which we obtain

$$(v^i g_{ij}) b^j = v^i (g_{ij} b^j) = v^i b_i = v = v_j b^j ,$$

and hence

$$\boxed{v_j = v^i g_{ij}} \quad \text{or} \quad \boxed{{}^{\mathsf{t}}[v]_{\mathcal{B}^g} = G[v]_{\mathcal{B}}}, . \tag{4.16}$$

Likewise, from

$$v^i b_i = v = v_j b^j = v_j (g^{ji} b_i) = (v_j g^{ji}) b_i$$

it follows that

$$\boxed{v^i = v_j g^{ji}} \quad \text{or} \quad \boxed{[v]_{\mathcal{B}} = G^{-1\,\mathsf{t}}[v]_{\mathcal{B}^g}} . \tag{4.17}$$

EXAMPLE 4.23. Let $\mathcal{B} = \{e_1, e_2, e_3\}$ be the standard basis of $\mathbb{R}^3$ and let

$$G = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

be the matrix of g with respect to $\mathcal{B}$. In (4.8) Example 4.17 we saw that

$$\mathcal{B}^g = \left\{ b^1 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}, b^2 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, b^3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

is the reciprocal basis. We find the covariant coordinates of $v = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ with respect to $\mathcal{B}^g$ using (4.16), namely

$$[v]_{\mathcal{B}^g} = G[v]_{\mathcal{B}} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 7 \end{pmatrix}.$$

In fact,

$$v_i b^i = (-1) \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} + 0 \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} + 7 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}.$$

□

Example 4.24. Let $V := \mathbb{R}[x]_1$ be the vector space of polynomials of degree ≤ 1 (that is "linear" polynomial, or of the form $a + bx$). Let $g : V \times V \to \mathbb{R}$ be defined by

$$g(p, q) := \int_0^1 p(x)q(x)dx ,$$

and let $\mathcal{B} := \{1, x\}$ be a basis of V. Determine:

(1) the matrix G;
(2) the matrix G^{-1};
(3) the reciprocal basis $\mathcal{B}^g$;
(4) the contravariant coordinates of $p(x) = 6x$ (that is the coordinates of $p(x)$ with respect to $\mathcal{B}$);
(5) the covariant coordinates of $p(x) = 6x$ (that is the coordinates of $p(x)$ with respect to $\mathcal{B}$).

(1) The matrix G has entries $g_{ij} = g(b_i, b_i)$, that is

$$\begin{aligned} g_{11} &= g(b_1, b_1) = \int_0^1 (b_1)^2 dx = \int_0^1 dx = 1 \\ g_{12} &= g(b_1, b_2) = \int_0^1 b_1 b_2 dx = \int_0^1 x = \frac{1}{2} \\ g_{21} &= g(b_2, b_1) = \int_0^1 b_2 b_1 dx = \frac{1}{2} \\ g_{22} &= g(b_2, b_2) = \int_0^1 (b_2)^2 dx = \int_0^1 x^2 dx = \frac{1}{3}, \end{aligned}$$

so that

$$G = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}.$$

(2) Since $\det G = 1 \cdot \frac{1}{3} - \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{12}$, then by using (2.7), we get

$$G^{-1} = \begin{pmatrix} 4 & -6 \\ -6 & 12 \end{pmatrix},.$$

(3) Using (4.13), we obtain that

$$\begin{pmatrix} b^1 & b^2 \end{pmatrix} = \begin{pmatrix} 1 & x \end{pmatrix} G^{-1} = \begin{pmatrix} 1 & x \end{pmatrix} \begin{pmatrix} 4 & -6 \\ -6 & 12 \end{pmatrix} = \begin{pmatrix} 4 - 6x & -6 + 12x \end{pmatrix},$$

so that $\mathcal{B}^g = \{4 - 6x, -6 + 12x\}$.

(4) $p(x) = 6x = 0 \cdot 1 + 6 \cdot x$, so that $p(x)$ has contravariant coordinates $[p(x)]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 6 \end{pmatrix}$.

(5) From (4.16) it follows that if $v = p(x)$, then

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = G \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 0 \\ 6 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

In fact, one can easily check that

$$v_1 b^1 + v_2 b^2 = 3 \cdot (4 - 6x) + 2 \cdot (-6 + 12x) = 6x.$$

4.2.3. Isomorphisms Between a Vector Space and its Dual. We saw already in Proposition 3.8 that If V is a vector space and V^* is its dual, then $\dim V = \dim V^*$. In particular this means that V and V^* can be identified, once we choose a basis $\mathcal{B}$ of V and a basis $\mathcal{B}^*$ of V^*. In fact, the basis $\mathcal{B}^*$ of V^* is given once we choose the basis $\mathcal{B}$ of V, as the dual basis of V^* with respect to $\mathcal{B}$. Then there is the following correspondence:

$$v \in V \leftrightsquigarrow \alpha \in V^*,$$

TABLE 1. Summary of covariance and contravariance of vector coordinates

	$\mathcal{B} := \{b_1, \dots, b_n\}$ basis	$\mathcal{B}^g = \{b^1, \dots, b^n\}$ reciprocal basis
they are related by	$g(b^i, b_j) = \delta^i_j$	
	$v = v^i b_i$ contravariant coordinates	$v = v_i b^i$ covariant coordinates
the matrices of g are	$g_{ij} = g(b_i, b_j)$	$g^{ij} = g(b^i, b^j)$
the matrices are inverse of each other, that is	$g^{ik} g_{kj} = \delta^i_j$	
the relation between the basis and the reciprocal basis is	$b_i = g_{ij} b^j$	$b^i = g^{ij} b_j$
the relation between covariant coordinates and contravariant coordinates is	$v^i = g_{ij} v_j$	$v_i = g_{ij} v^j$

exactly when v and α have the same coordinates, respectively with respect to $\mathcal{B}$ and $\mathcal{B}^*$, However this correspondence depends on the choice of the basis $\mathcal{B}$ and hence *not canonical*.

If however V is endowed with an inner product, then there is a **canonical identification** of V with V^* (that is an identification that does not depend on the basis $\mathcal{B}$ ov V). In fact, let $g : V \times V \to \mathbb{R}$ be an inner product and let $v \in V$. Then

$$\begin{aligned} g(v, \cdot\,) : V &\longrightarrow \mathbb{R} \\ w &\longmapsto g(v, w) \end{aligned}$$

is a linear form and hence we have the following *canonical* identification given by the metric

$$\begin{aligned} V &\longleftrightarrow V^* \\ v &\longleftrightarrow v^* := g(v, \cdot\,)\,. \end{aligned} \tag{4.18}$$

Note that the isomorphism sends the zero vector to the linear form identically equal to zero, since $g(v, \cdot\,) \equiv 0$ if and only if $v = 0$, since g is positive definite.

So far, we have two bases of the vector space V, namely the basis $\mathcal{B}$ and the reciprocal basis $\mathcal{B}^g$ and we have also the dual basis of the dual vector space V^*. In fact, under the isomorphism (4.18), the reciprocal basis of V and the dual basis of V^* correspond to each other. This is easily seen because, under the isomorphism (4.18) an element of the reciprocal basis b^i correspond to the linear form $g(b^i, \cdot\,)$

$$b^i \longleftrightarrow g(b^i, \cdot\,)$$

and the linear form $g(b^i, \cdot\,) : V \to \mathbb{R}$ has the property that

$$g(b^i, b_j) = \delta^i_j\,.$$

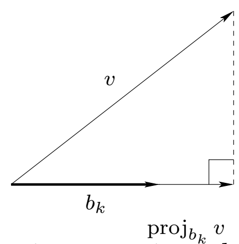
Hence

$$g(b^i, \cdot\,) \equiv \beta^i\,,$$

and under the canonical identification between V and V^* the reciprocal basis of V corresponds to the dual basis of V^*.

4.2.4. Geometric Interpretation. Let $g : V \times V \to \mathbb{R}$ be an inner product and $\mathcal{B} := \{b_1, \dots, b_n\}$ a basis of V. The orthogonal projection of a vector $\in V$ onto b_k is defined as

$$\operatorname{proj}_{b_k} v = \frac{g(v, b_k)}{g(b_k, b_k)} b_k\,. \tag{4.19}$$



In fact, $\operatorname{proj}_{b_k} v$ is obviously parallel to b_k and the following exercises shows that the component $v - \operatorname{proj}_{b_k} v$ is orthogonal to b_k.

EXERCISE 4.25. With $\operatorname{proj}_{b_k} v$ defined as in , we have

$$v - \operatorname{proj}_{b_k} v \perp b_k\,,$$

where the orthogonality is meant with respect to the inner product g.

Now let $v = v_i b^i \in V$ be a vector written in terms of its covariant coordinates (that is the coordinates with respect to the reciprocal basis). Then

$$g(v, b_k) = g(v_i b^i, b_k) = v_i \underbrace{g(b^i, b_k)}_{\delta^i_k} = v_k\,,$$

so that (4.19) becomes

$$\operatorname{proj}_{b_k} v = \frac{v_k}{g(b_k, b_k)} b_k\,.$$

If we assume that the elements of the basis $\mathcal{B} := \{b_1, \dots, b_n\}$ are unit vectors, then (4.19) further simplifies to give

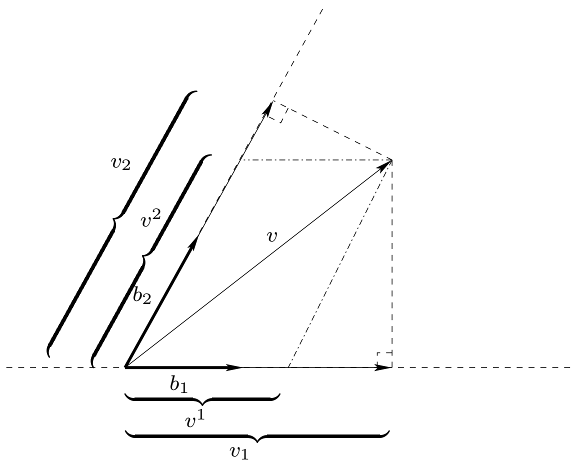
$$\operatorname{proj}_{b_k} v = v_k b_k\,. \tag{4.20}$$

This equation shows the following:

FACT 4.26. *The covariant coordinates of v give the orthogonal projection of v onto $b_1, \dots, b_n$.*

Likewise, the following holds basically by definition:

Fact 4.27. *The contravariant coordinates of v give the "parallel" projection of v onto $b_1, \dots, b_n$.*



CHAPTER 5

Tensors

5.1. Generalities

Let V be a vector space. Up to now we saw several objects related to V, that we said were "tensors". We summarize them in Table 1. So, we seem to have a good candidate for the definition of a tensor of type $(0,q)$ for all $q \in \mathbb{N}$, but we cannot say the same for a tensor of type $(p,0)$ for all $p \in \mathbb{N}$. The next discussion will lead us to that point, and in the meantime we will discuss an important point.

5.1.1. Canonical isomorphism between V and $(V^*)^*$. We saw in § 4.2.3 that any vector space is isomorphic to its dual, but the the isomorphism is not canonical (that is, it depends on the choice of basis). We also saw that if there is an inner product on V, then there is a canonical isomorphism. The point of this section is to show that, even without an inner product, there is a always a canonical isomorphism between V and its **bidual** $(V^*)^*$, that is the dual of its dual.

To see this, let us observe first of all that

$$\dim V = \dim (V^*)^* . \tag{5.1}$$

If fact, for any vector space W, we saw in Proposition 3.8 that $\dim W = \dim W^*$. If we apply this equality both to $W = V$ and to $W = V^*$, we obtain

$$\dim V = \dim V^* \quad \text{and} \quad \dim V^* = \dim (V^*)^* ,$$

from which (5.1) follows immediately. From (4.2.3) we deduce immediately that V and $(V^*)^*$ are isomorphic, and we only have to see that the isomorphism is canonical.

TABLE 1. Covariance and Contravariance

Tensor	Components	Behavior under a change of basis	Type
vectors in V	v^i	contravariant tensor	(1,0)
linear forms $V \to \mathbb{R}$	α_j	covariant tensor	(0,1)
linear transformations $V \to V$	A^i_j	mixed: {contravariant, covariant} tensor	(1,1)
bilinear forms[1] $V \times V \to \mathbb{R}$	B_{ij}	covariant 2-tensor	(0,2)
k-linear forms $\underbrace{V \times \cdots \times V}_{k} \to \mathbb{R}$	$F_{i_1 i_2 \dots i_k}$	covariant k-tensor	(0,k)

To this end, observe that a vector $v \in V$ gives rise to a linear form on V^* defined by

$$\begin{aligned} \varphi_v : V^* &\longrightarrow \mathbb{R} \\ \alpha &\longmapsto \alpha(v) \,. \end{aligned}$$

Then we can define a linear map as follows:

$$\begin{aligned} \Phi : V &\longrightarrow (V^*)^* \\ v &\mapsto \varphi_v\{\alpha \mapsto \alpha(v)\} \end{aligned} \tag{5.2}$$

Since for any linear map $T : W \to W$ between vector spaces

$$\dim W = \dim \operatorname{Range}(T) + \dim \ker(T) \,,$$

it will be enough to show that $\ker \Phi = \{0\}$, because then

$$\dim (V^*)^* = \dim V = \dim \operatorname{Range}(\Phi) \,,$$

that is Φ is an isomorphism. Notice that the important fact is that we have *not* chosen a basis to define the isomorphism Φ.

To see that $\ker \Phi = \{0\}$, observe that the kernel consists of all vectors $v \in V$ such that $\alpha(v) = 0$ for all $\alpha \in V^*$. We want to see that the only vector $v \in V$ for which this happens is the zero vector. In fact, if $0 \neq v \in V$ and $\mathcal{B} := \{b_1, \dots, b_n\}$ is *any* basis of V, then we can write $v = v^i b_i$, where at least one $v^j \neq 0$. But then, if $\mathcal{B}^* = \{\beta_1, \dots, \beta_n\}$ is the dual basis, $0 \neq v^j = \beta_j(v)$. Notice again that the dimension of the kernel of a linear map is invariant under a change of basis, and hence $\ker \Phi = \{0\}$ no matter what the basis $\mathcal{B}$ here was.

We record this fact as follows:

Fact 5.1. *Let V be a vector space and V^* its dual. The dual $(V^*)^*$ of V^* is canonically isomorphic to V.*

5.1.2. Towards general tensors.

Recall that

$$V^* := \{\alpha : V \to \mathbb{R} : \alpha \text{ is a linear form}\} = \{(0, 1)\text{-tensors}\} \,.$$

Applying this formula to the vector space V^* we obtain

$$(V^*)^* := \{\alpha : V^* \to \mathbb{R} : \alpha \text{ is a linear form}\} \,.$$

Using the isomorphism $(V^*)^* \cong V$ and the fact that coordinate vectors are contravariant, we conclude that

$$\{\alpha : V^* \to \mathbb{R} : \alpha \text{ is a linear form}\} = (V^*)^* = V = \{(1, 0)\text{-tensors}\} \,.$$

So changing the vector space from V to its dual V^* seems to have had the effect to transform a covariant tensor of type $(0, 1)$ into a contravariant one of type $(1, 0)$. We are going to apply this procedure to try to transform a covariant of type $(0, 2)$ into a contravariant one of type $(2, 0)$.

Recall that

$$\{\varphi : V \times V \to \mathbb{R} : \text{ bilinear } \} = \{(0, 2)\text{-tensors}\} \,.$$

and consider

$$\{\varphi : V^* \times V^* \to \mathbb{R} : \text{ bilinear}\} .$$

We can hence give the following definition:

Definition 5.2. A **tensor of type** $(2,0)$ is a bilinear form on V^*, that is a bilinear function $\sigma : V^* \times V^* \to \mathbb{R}$.

Let

$$\mathrm{Bil}(V^* \times V^*, \mathbb{R}) = \{\sigma : V^* \times V^* \to \mathbb{R} : \sigma \text{ is a bilinear form}\} = \{(2,0)\text{-tensors}\} .$$

Exercise 5.3. Check that $\mathrm{Bil}(V^* \times V^*, \mathbb{R})$ is a vector space, that is that if $\sigma, \tau \in \mathrm{Bil}(V^* \times V^*, \mathbb{R})$ and $\lambda, \mu \in \mathbb{R}$, then $\lambda\sigma + \mu\tau \in \mathrm{Bil}(V^* \times V^*, \mathbb{R})$.

5.1.3. Tensor product of $(1,0)$-tensors on V^*. If $v, w \in V$ are two vectors (that is two $(1,0)$-tensors), we define

$$\sigma_{v,w} : V^* \times V^* \to \mathbb{R}$$

by

$$\sigma_{v,w}(\alpha, \beta) := \alpha(v)\beta(w) .$$

Since α and β are two linear forms, then σ

$$\boxed{\sigma_{v,w} =: v \otimes w}$$

is *bilinear* and called the **tensor product** of v and w. Hence σ is a $(2,0)$**-tensor**.

Note 5.4. In general

$$\boxed{v \otimes w \neq w \otimes v},$$

as there can be linear forms α, β such that $\alpha(v)\beta(w) \neq \alpha(w)\beta(v)$.

Similar to what we saw in § 3.2.2, we can define a basis for the space of $(2,0)$-tensors by considering the $(2,0)$-tensors defined by $b_i \otimes b_j$, where $\mathcal{B} := \{b_1, \dots, b_n\}$ is a basis of V.

Proposition 5.5. *The elements* $b_j \otimes b_j$, $i, j = 1, \dots, n$ *form a basis of* $\mathrm{Bil}(V^* \times V^*, \mathbb{R})$. *Thus* $\dim \mathrm{Bil}(V^* \times V^*, \mathbb{R}) = n^2$.

We will not prove the proposition here, as the proof is be completely analogous to the one of Proposition 3.23.

Notation. We write

$$\boxed{\mathrm{Bil}(V^* \times V^*, \mathbb{R}) = V \otimes V}.$$

5.1.4. Components of a $(2,0)$-tensor and their contravariance. Let $\sigma : V^* \times V^* \to \mathbb{R}$ be a bilinear form on V^*, that is a $(2,0)$-tensor, as we just saw. We want to verify that it behaves as we expect with respect to a change of basis. After choosing a basis $\mathcal{B} := \{b_1, \dots, b_n\}$ of V, we have the dual basis $\mathcal{B}^* = \{\beta^1, \dots, \beta^n\}$ of V^* and the basis $\{b_i \otimes b_j : i, j = 1, \dots, n\}$ of the space of $(2,0)$-tensors.

The $(2,0)$-tensor σ is represented by its components

$$\boxed{S^{ij} = \sigma(\beta^i, \beta^j)},$$

that is

$$\sigma = S^{ij} b_j \otimes b_j ,$$

and the components S^{ij} can be arranged into a matrix

$$S = \begin{pmatrix} S^{11} & \dots & S^{1n} \\ \vdots & \ddots & \vdots \\ S^{n1} & \dots & S^{nn} \end{pmatrix}$$

called the **matrix of the $(2,0)$-tensor with respect to the chosen basis of V.**

We look now at how the components of a $(2,0)$-tensor change with a change of basis. Let $\mathcal{B} := \{b_1, \dots, b_n\}$ and $\widetilde{\mathcal{B}} := \{\tilde{b}_1, \dots, \tilde{b}_n\}$ be two basis of V and let $\mathcal{B}^* := \{\beta^1, \dots, \beta^n\}$ and $\widetilde{\mathcal{B}}^* := \{\tilde{\beta}^1, \dots, \tilde{\beta}^n\}$ be the corresponding dual basis of V^*. Let $\sigma : V^* \times V^* \to \mathbb{R}$ be a $(2,0)$-tensor with components

$$S^{ij} = \sigma(\beta^i, \beta^j) \quad \text{and} \quad \widetilde{S}^{ij} = \sigma(\tilde{\beta}^i, \tilde{\beta}^j)$$

with respect to $\mathcal{B}^*$ and $\widetilde{\mathcal{B}}^*$ respectively. Let $L := L_{\widetilde{\mathcal{B}}\mathcal{B}}$ be the matrix of the change of basis from $\mathcal{B}$ to $\widetilde{\mathcal{B}}$, and let $\Lambda = L^{-1}$. Then, as seen in (1.2) and (3.11) we have that

$$\tilde{b}_j = L^i_j b_i \quad \text{and} \quad \tilde{\beta}^i = \Lambda^i_j \beta^j .$$

It follows that

$$\widetilde{S}^{ij} = \sigma(\tilde{\beta}^i, \tilde{\beta}^j) = \sigma(\Lambda^i_k \beta^k, \Lambda^j_\ell \beta^\ell) = \Lambda^i_k \Lambda^j_\ell \sigma(\beta^k, \beta^\ell) = \Lambda^i_k \Lambda^j_\ell S^{k\ell} ,$$

where the first and the last equality follow from the definition of $\widetilde{S}^{ij}$ and of $S^{k\ell}$ respectively, the second from the change of bases and the third from the bilinearity of σ. We conclude that

$$\boxed{\widetilde{S}^{ij} = \Lambda^i_k \Lambda^j_\ell S^{kl}}. \tag{5.3}$$

Hence σ is a **contravariant** 2-tensor..

EXERCISE 5.6. Verify that in terms of matrices (5.3) translates into

$$\boxed{\widetilde{S} = {}^{t}\Lambda S \Lambda}.$$

(Compare with (3.15).)

5.2. Tensors of type (p,q)

Definition 5.7. A **tensor of type** (p,q) or (p,q)**-tensor** is a multilinear form

$$T : \underbrace{V^* \times \dots V^*}_{p} \times \underbrace{V \times \dots \times V}_{q} \longrightarrow \mathbb{R}\,.$$

Let T be a (p,q)-tensor, $\mathcal{B} := \{b_1, \dots, b_n\}$ a basis of V and $\mathcal{B}^* = \{\beta_1, \dots, \beta_n\}$ the corresponding dual basis of V^*. The components of T with respect to these bases are

$$\boxed{T^{i_1,\dots,i_p}_{j_1,\dots,j_q} = T(\beta^{i_1}, \dots, \beta^{i_p}, b_{j_1}, \dots, b_{j_q})}\,.$$

If moreover $\widetilde{\mathcal{B}} := \{\tilde{b}_1, \dots, \tilde{b}_n\}$ is another basis, $\widetilde{\mathcal{B}}^* = \{\tilde{\beta}_1, \dots, \tilde{\beta}_n\}$ is the coresponding dual basis of V^* and $L := L_{\widetilde{\mathcal{B}}\mathcal{B}}$ is the matrix of the change of basis with inverse $\Lambda := L^{-1}$, then the components of T with respect to these new bases are

$$\boxed{\widetilde{T}^{i_1,\dots,i_p}_{j_1,\dots,j_q} = \Lambda^{i_1}_{k_1} \dots \Lambda^{i_p}_{k_p} L^{\ell_1}_{j_1} \dots L^{\ell_q}_{j_q} T^{k_1,\dots,l_p}_{\ell_1,\dots,\ell_q}}\,.$$

5.3. Tensor product

We saw already in § 3.2.2 and § 3.5 the tensor product of two multilinear forms. Since multilinear forms are covariant tensors, we said that this corresponds to the tensor product of two covariant tensors. More generally, we can define the tensor product of any two tensors as follows:

Definition 5.8. Let

$$T : \underbrace{V^* \times \dots V^*}_{p} \times \underbrace{V \times \dots \times V}_{q} \longrightarrow \mathbb{R}$$

be a (p,q)-tensor and

$$U : \underbrace{V^* \times \dots V^*}_{k} \times \underbrace{V \times \dots \times V}_{\ell} \longrightarrow \mathbb{R}$$

a (k,ℓ) tensor. The **tensor product** $T \otimes U$ of T and U is a $(p+k, q+\ell)$-tensor

$$T \otimes U : \underbrace{V^* \times \dots V^*}_{p+k} \times \underbrace{V \times \dots \times V}_{q+\ell} \longrightarrow \mathbb{R}$$

defined by

$$\begin{aligned}(T \otimes U)(\alpha_1, \dots, \alpha_{p+k}, v_1, \dots, v_{q+\ell}) :=\\ T(\alpha_1, \dots, \alpha_p, v_1, \dots, v_q)U(\alpha_{p+1}, \dots, \alpha_{p+k}, v_{q+1}, \dots, v_{q+\ell})\,.\end{aligned}$$

Note that both $T \otimes U$ and $U \otimes T$ are tensors of the same type $(p+k, q+\ell)$, but in general

$$T \otimes U \neq U \otimes T\,.$$

The set of all tensors of type (p,q) on a vector space V is denoted by

$$\boxed{\mathcal{T}_q^p(V) := \{\text{all } (p,q)\text{-tensors on } V\}}\,.$$

Analogously to how we proceeded in the case of $(0,2)$-tensors, we compute the dimension of $\mathcal{T}_q^p(V)$. If $\mathcal{B} := \{b_1,\dots,b_n\}$ is a basis of V and $\mathcal{B}^* := \{\beta^1,\dots,\beta^n\}$ is the corresponding dual basis of V^*. Just like we saw in Proposition 5.5 in the case of $(0,2)$-tensors, a basis of $\mathcal{T}_q^p(V)$ is

$$\{b_{i_1}\otimes b_{i_2}\otimes\cdots\otimes b_{i_p}\otimes\beta^{j_1}\otimes\beta^{j_2}\otimes\cdots\otimes\beta^{j_q} : 1\le i_1,\dots,i_p\le n,\ 1\le j_1,\dots,j_q\le n\}\,.$$

Since there are $\underbrace{n\times\cdots\times n}_{p}\times\underbrace{n\times\cdots\times n}_{q} = n^{p+q}$ elements in this basis (corresponding to the possible choices of b_{i_k} and β^{j_ℓ}), we deduce that

$$\boxed{\dim\mathcal{T}_q^p(V) = n^{p+q}\,.}$$

We now define the tensor product of two vector spaces:

Definition 5.9. Let V and W be two finite dimensional vector spaces, with $\dim V = n$ and $\dim W = m$. Choose $\{b_1,\dots,b_n\}$ a basis of V and $\{a_1,\dots,a_m\}$ a basis of W. Then the **tensor product $V\otimes W$ of V and W** is an $(n\cdot m)$-dimensional vector space with basis

$$\{b_i\otimes a_j : 1\le i\le n,\ 1\le j\le m\}\,.$$

We remark that there is no reason to restrict oneself to the tensor product of only two factors. One can equally define the tensor product $V_1\otimes\cdots\otimes V_k$, and obtain a vector space of dimension $\dim V_1\times\cdots\times\dim V_k$.

There is hence an identification

$$\mathcal{T}_q^p(V)\cong\underbrace{V\otimes\cdots\otimes V}_{p}\otimes\underbrace{V^*\otimes\cdots\otimes V^*}_{q},$$

that follows from the fact that both spaces have the same basis. The following proposition gives other useful identifications.

Proposition 5.10. *Let V and W be two finite dimensional vector spaces, with $\dim V = n$ and $\dim W = m$ and let us denote by*

$$\operatorname{Lin}(V,W^*) := \{\textit{linear maps } V\to W^*\}\,.$$

Then

$$\begin{aligned}\operatorname{Bil}(V\times W,\mathbb{R}) &\cong \operatorname{Lin}(V,W^*)\\ &\cong \operatorname{Lin}(W,V^*)\\ &\cong V^*\otimes W^*\\ &\cong (V\otimes W)^*\\ &\cong \operatorname{Lin}(V\otimes W,\mathbb{R})\,.\end{aligned}$$

PROOF. Here is the idea behind this chain of identifications. Let $f \in \mathrm{Bil}(V \times W, \mathbb{R})$, that is a bilinear function $f : V \times W \to \mathbb{R}$. This means that f takes two vectors, $v \in V$ and $w \in W$, as input and gives back a real number $f(v, w) \in \mathbb{R}$. If on the other hand we only feed f one vector $v \in V$, then there is a remaining spot waiting for a vector $w \in W$ to produce a real number. Since f is linear on V and on W, the map $f(v, \cdot\,) : W \to \mathbb{R}$ is a linear form, so $f(v, \cdot\,) \in W^*$. In other words, we can view $f \in \mathrm{Lin}(V, W^*)$. It follows that there is a linear map

$$\begin{array}{ccc} \mathrm{Bil}(V \times W, \mathbb{R}) & \longrightarrow & \mathrm{Lin}(V, W^*) \\ f & \longmapsto & T_f \,, \end{array}$$

where

$$T_f(v)(w) := f(v, w)\,.$$

Conversely, any $T \in \mathrm{Lin}(V, W^*)$ can be identified with a bilinear map $f_T \in \mathrm{Bil}(V \times W, \mathbb{R})$ defined by

$$f_T(v, w) := T(v)(w)\,.$$

Since $f_{T_f} = f$ and $T_{f_T} = T$, we have proven the first identification in the proposition.

Analogously, if the input is only a vector $w \in W$, then $f(\,\cdot\,, w) : V \to \mathbb{R}$ is a linear map and hence $f \in \mathrm{Bil}(V \times W, \mathbb{R})$ defines a linear map $T_f \in \mathrm{Lin}(W, V^*)$. The same reasoning as in the previous paragraph, shows that $\mathrm{Bil}(V \times W, \mathbb{R}) \cong \mathrm{Lin}(W, V^*)$.

To proceed with the identifications, observe that, because of our definition of $V^* \otimes W^*$, we have

$$\mathrm{Bil}(V \times W, \mathbb{R}) \cong V^* \otimes W^*\,,$$

since these spaces both have basis[2]

$$\{\beta^i \otimes \alpha^j : 1 \leq i \leq n,\, 1 \leq j \leq m\}\,,$$

where $\{b_1, \dots, b_n\}$ is a basis of V with corresponding dual basis $\{\beta^1, \dots, \beta^n\}$ of V^*, and $\{a_1, \dots, a_n\}$ is a basis of W with corresponding dual basis $\{\alpha^1, \dots, \alpha^n\}$ of W^*.

Finally, an element $D_{ij}\beta^i \otimes \alpha^j \in V^* \otimes W^*$ may be viewed as a linear map $V \otimes W \to \mathbb{R}$, that is as an element of $(V \otimes W)^*$ by

$$\begin{array}{ccc} V \otimes W & \longrightarrow & \mathbb{R} \\ C^{k\ell} b_k \otimes a_\ell & \longmapsto & D_{ij} C^{k\ell} \underbrace{\beta^i(b_k)}_{\delta^i_k} \underbrace{\alpha^j(a_\ell)}_{\delta^j_\ell} = D_{ij} C^{k\ell}\,. \end{array}$$

□

Because of the identification $\mathrm{Bil}(V \times W, \mathbb{R}) \cong \mathrm{Lin}(V \otimes W, \mathbb{R})$, sometimes one says that “the tensor product linearizes what was bilinear or multilinear”.

[2]There is a way of defining the tensor product of vector spaces without involving bases, but we will not do it here.

CHAPTER 6

Applications

6.1. Inertia tensor

6.1.1. Moment of inertia with respect to the axis determined by the angular velocity. Let M be a rigid body fixed at a point O. The motion of this rigid body at time t is by *rotation* by an angle θ with *angular velocity* ω about some axis through O. The angular velocity has magnitude $|\cdot|$ or $\|\cdot\|$

$$|\omega| = \left|\frac{d\theta}{dt}\right|,$$

direction given by the axis of rotation and orientation by the right-hand rule. The *position vector* of a point P in the body M relative to the origin O is

$$\mathbf{x} = \overrightarrow{OP}$$

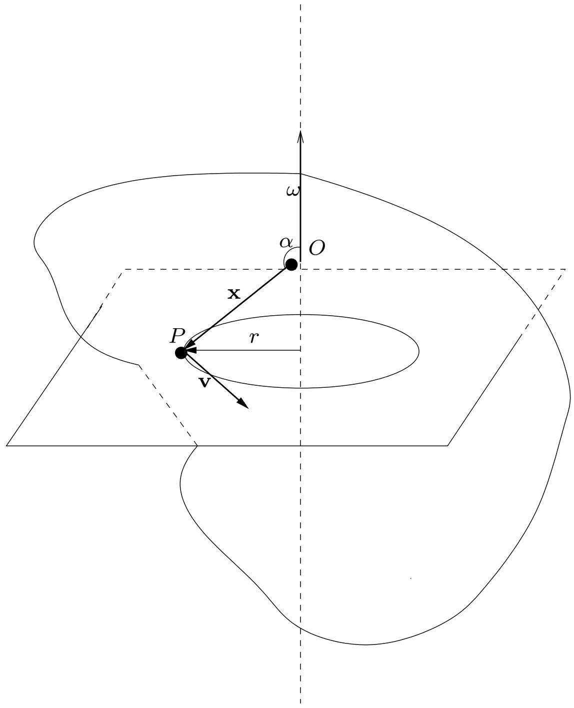
while the *linear velocity* of a point P is

$$\mathbf{v} = \omega \times \mathbf{x}.$$

The linear velocity $\mathbf{v}$ has magnitude

$$|\mathbf{v}| = \underbrace{|\omega|}_{\left|\frac{d\theta}{dt}\right|} \underbrace{|\mathbf{x}|\sin\alpha}_{r}$$

and direction tangent at P to the circle of radius r perpendicular to the axis of rotation.



The *kinetic energy* of an infinitesimal region dM of M around P is

$$dE = \frac{1}{2}\mathbf{v}^2 dm\,,$$

where $\mathbf{v}^2 = \mathbf{v} \bullet \mathbf{v}$ and dm is the mass of dM. The *total kinetic energy* of M is

$$E = \frac{1}{2}\int_M \mathbf{v}^2 dm = \frac{1}{2}\int_M (\omega \times \mathbf{x})^2 dm\,.$$

Note that, depending on the type of the rigid body, we might take here the sum of the integral, and what type of integral depends again from the kind of rigid body we have. More precisely:

(1) If M is a solid in 3-dimensional space, then

$$E = \frac{1}{2}\iiint_M (\omega_P \times \mathbf{x}_P)^2 \rho_P dx^1 dx^2 dx^3\,,$$

where $(\omega \times \mathbf{x}_P)^2 \rho_P$ is a function of $P(x^1, x^2, x^3)$.

(2) If M is a flat sheet in 3-dimensional space, then

$$E = \frac{1}{2}\iint_M (\omega_P \times \mathbf{x}_P)^2 \rho_P dx^1 dx^2\,.$$

(3) If M is a surface in 3-dimensional space, then

$$E = \frac{1}{2}\iiint_M (\omega_P \times \mathbf{x}_P)^2 \rho_P d\sigma\,,$$

where $d\sigma$ is the infinitesimal element of the surface for a surface integral.

(4) If M is a wire in 3-dimensional space, then

$$E = \frac{1}{2}\int_M (\omega_P \times \mathbf{x}_P)^2 \rho_P ds\,,$$

where ds is the infinitesimal element of length for a line integral.

(5) If M is a finite set of point masses with rigid relative positions, then

$$E = \frac{1}{2}\sum_{i=1}^{N} (\omega_P \times \mathbf{x_i})^2 m_i\,.$$

In any case we need to work out the quantity

$$(\omega \times \mathbf{x})^2$$

for vectors ω and $\mathbf{x_i}$ in 3-dimensional space.

To this purpose we use the **Lagrange identity**[1], according to which

$$(a \times b)\cdot(c \times d) = \det \begin{bmatrix} a\cdot c & a\cdot d \\ b\cdot c & b\cdot d \end{bmatrix}\,. \tag{6.1}$$

Applying (6.1) with $a = c = \omega$ and $b = d = \mathbf{x}$, we obtain

$$(\omega \times \mathbf{x})^2 = (\omega \times \mathbf{x})\cdot(\omega \times \mathbf{x}) = \det \begin{bmatrix} \omega\cdot\omega & \omega\cdot\mathbf{x} \\ \mathbf{x}\cdot\omega & \mathbf{x}\cdot\mathbf{x} \end{bmatrix} = \omega^2\mathbf{x}^2 - (\omega\cdot\mathbf{x})^2\,.$$

Let now $\mathcal{B} = \{e_1, e_2, e_3\}$ be an orthonormal[2] basis of $\mathbb{R}$, so that

$$\omega = \omega^i e_i \quad \text{and} \quad \mathbf{x} = x^i e_i\,.$$

Then

$$\begin{aligned} \omega^2 &= \omega\cdot\omega = \delta_{ij}\omega^i\omega^j = \omega^1\omega^1 + \omega^2\omega^2 + \omega^3\omega^3 \\ \mathbf{x}^2 &= \mathbf{x}\cdot\mathbf{x} = \delta_{k\ell}x^k x^\ell = x^1x^1 + x^2x^2 + x^3x^3 \\ \omega\cdot\mathbf{x} &= \delta_{ik}\omega^i x^k \end{aligned}$$

so that

$$\begin{aligned} (\omega \times \mathbf{x})^2 &= \omega^2\mathbf{x}^2 - (\omega\cdot\mathbf{x})^2 \\ &= (\delta_{ij}\omega^i\omega^j)(\delta_{k\ell}x^k x^\ell) - (\delta_{ik}\omega^i x^k)(\delta_{jl}\omega^j x^\ell) \\ &= (\delta_{ij}\delta_{k\ell} - \delta_{ik}\delta_{j\ell})\omega^i\omega^j x^k x^\ell\,. \end{aligned}$$

Therefore the total kinetic energy is

$$E = \frac{1}{2}(\delta_{ij}\delta_{k\ell} - \delta_{ik}\delta_{j\ell})\omega^i\omega^j \int_M x^k x^\ell dm$$

and it depends only on $\omega^1, \omega^2, \omega^3$ (since we have integrated over the x^1, x^2, x^3).

[1] The Lagrange identity can easily be proven in coordinates.

[2] We could use any basis of $\mathbb{R}^3$. Then, instead of the δ_{ij}, the formula would have involved the metric tensor g_{ij}. However computations with orthonormal bases are simpler. In addition in this case we will see that the metric tensor is symmetric, and hence it admits an orthonormal eigenbasis.

DEFINITION 6.1. The **inertia tensor** is the tensor whose components with respect to an orthonormal basis $\mathcal{B}$ are

$$\boxed{I_{ij} = (\delta_{ij}\delta_{k\ell} - \delta_{ik}\delta_{j\ell}) \int_M x^k x^\ell dm}.$$

Then the kinetic energy of the rotating rigid body is

$$E = \underbrace{\frac{1}{2} I_{ij}\omega^i\omega^j}_{\substack{\text{Einstein}\\ \text{notation}}} = \underbrace{\frac{1}{2}\omega \cdot I\omega}_{\substack{\text{matrix}\\ \text{notation}}}.$$

One can check that

$$\begin{aligned}
I_{11} &= \int_M (x^2x^2 + x^3x^3)dm \\
I_{22} &= \int_M (x^1x^1 + x^3x^3)dm \\
I_{33} &= \int_M (x^1x^1 + x^2x^2)dm \\
I_{23} = I_{32} &= -\int_M x^2x^3 dm \\
I_{31} = I_{13} &= -\int_M x^1x^3 dm \\
I_{12} = I_{21} &= -\int_M x^1x^2 dm\,,
\end{aligned}$$

so that with respect to the basis $\mathcal{B}$, the metric tensor is represented by the symmetric matrix

$$I = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix}.$$

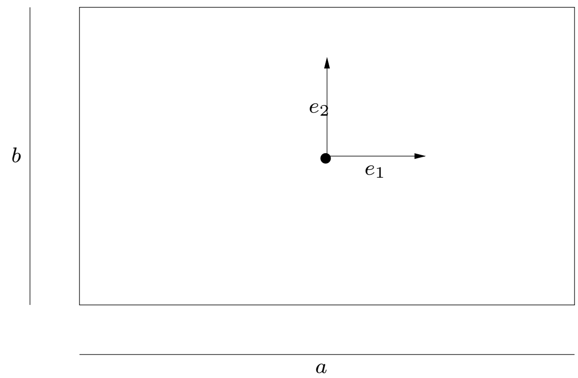
We check only the formula for I_{11}. In fact,

$$I_{11} = (\underbrace{\delta_{11}\delta_{k\ell}}_{\substack{=0\\ \text{unless}\\ k=\ell}} - \underbrace{\delta_{1k}\delta_{1\ell}}_{\substack{=0\\ \text{unless}\\ k=\ell=1}}) \int_M x^k x^\ell dm$$

If $k = \ell = 1$, then $\delta_{11}\delta_{11} - \delta_{11}\delta_{11} = 0$, so that the non-vanishing terms have $k = \ell \neq 1$.

So I_{11}, I_{22}, I_{33} are the **moments of inertia** of the rigid body M with respect to the axes Ox_1, Ox_2, Ox_3 respectively; I_{12}, I_{23}, I_{31} are the **polar moments of inertia** or the **products of inertia** of the rigid body M.

EXAMPLE 6.2. Find the inertia tensor of a homogeneous rectangular plate with sides a and b and total mass m, assuming that the center of rotation O coincides with the center of inertia. We choose an orthonormal basis with e_1 aligned with the side of length a, e_2 aligned with the side of length b and e_3 perpendicular to the plate.



Since the plate is assumed to be homogeneous, it has a constant *mass density* equal to

$$\rho = \frac{\text{total mass}}{\text{area}} = \frac{m}{ab}\,.$$

Denote by x, y and z the coordinates. Then

$$\begin{aligned}\underbrace{I_{11}}_{I_{xx}} &= \int_{-\frac{a}{2}}^{\frac{a}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} (y^2 + \underbrace{z^2}_{=0}) \underbrace{\rho}_{\frac{m}{ab}} \, dydx \\ &= \frac{m}{ab} a \int_{-\frac{b}{2}}^{\frac{b}{2}} y^2 dy \\ &= \frac{m}{b} \left[\frac{y^3}{3} \right]_{-\frac{b}{2}}^{\frac{b}{2}} \\ &= \frac{m}{12} b^2\,.\end{aligned}$$

Similarly

$$\underbrace{I_{22}}_{I_{yy}} = \frac{m}{12} a^2\,,$$

and

$$\underbrace{I_{33}}_{I_{zz}} = \int_{-\frac{a}{2}}^{\frac{a}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} (x^2 + y^2) \rho \, dydx = \frac{m}{12}(a^2 + b^2)\,,$$

which turns out to be just the sum of I_{11} and I_{22}.

Furthermore,

$$I_{23} = I_{32} = -\int_{-\frac{a}{2}}^{\frac{a}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} y \underbrace{z}_{=0} \rho dy dx = 0 \,,$$

and similarly $I_{31} = I_{13} = 0$. Finally

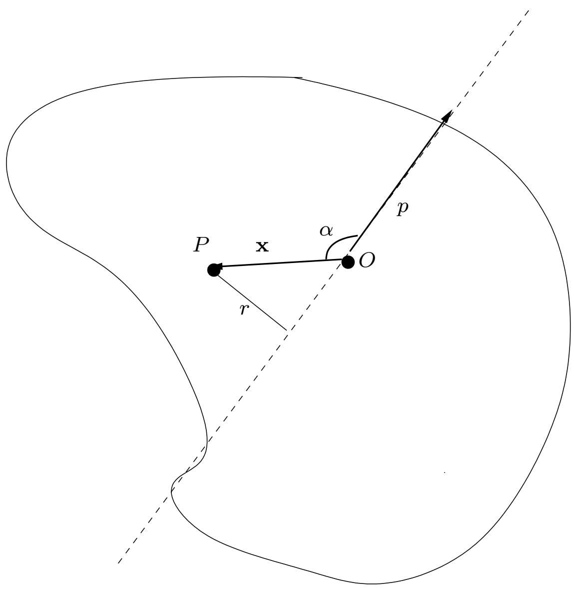
$$I_{21} = I_{21} = -\int_{-\frac{a}{2}}^{\frac{a}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} xy\rho dy dx = -\frac{m}{ab} \underbrace{\underbrace{\left(\int_{-\frac{a}{2}}^{\frac{a}{2}} x\,dx\right)}_{=0} \underbrace{\left(\int_{-\frac{b}{2}}^{\frac{b}{2}} y\,dy\right)}_{=0}}_{\substack{\text{because the integral of an odd function} \\ \text{on a symmetric interval is 0}}} \,.$$

We conclude that the inertia tensor is given by the matrix

$$\frac{m}{12}\begin{pmatrix} b^2 & 0 & 0 \\ 0 & a^2 & 0 \\ 0 & 0 & a^2+b^2 \end{pmatrix} . \quad \square$$

Exercise 6.3. Compute the inertia tensor of the same plate but now with center of rotation O coinciding with a vertex of the rectangular plate.

6.1.2. Moment of inertia about any axis through the fixed point. We compute the moment of inertia of the body M about an axis through O. Let p be a unit vector defining an axis through O.



The **moment of inertia** of an infinitesimal region of M around P is

$$dI = \underbrace{r^2}_{\substack{r \text{ is the distance} \\ \text{from } P \text{ to the axis}}} \underbrace{dm}_{\substack{\text{infinitesimal} \\ \text{mass}}} = \|p \times x\|^2 dm \,,$$

where the last equality follows from the fact that $\|p \times x\| = \|p\| \, \|x\| \sin\alpha = r$, since p is a unit vector. Hence the **total moment of inertia** of M with respect to the axis given by p is

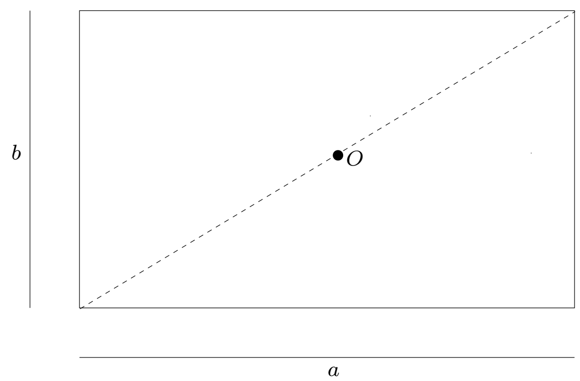
$$I = \int_M \|p \times x\|^2 dm \geq 0\,.$$

(This is very similar to the total kinetic energy E: just replace ω by p and omit the factor $\frac{1}{2}$.) By the earlier computations, we conclude that

$$I = I_{ij} p^i p^j\,,$$

where I_{ij} is the inertia tensor. This formula shows that *the total moment of inertia of the rigid body M with respect to an arbitrary axis passing through the point O is determined only by the inertia tensor of the rigid body.*

Example 6.4. For the rectangular plate in Example 6.2, compute the moment of inertia with respect to the diagonal of the plate.



Choose $p = \frac{1}{\sqrt{a^2+b^2}}(ae_1 + be_2)$ (the other possibility is the negative of this vector). So

$$p^1 = \frac{a}{\sqrt{a^2+b^2}}\,, \quad p^2 = \frac{b}{\sqrt{a^2+b^2}} \quad p^3 = 0\,.$$

The moment of inertia is

$$\begin{aligned} I =& I_{ij} p^i p^j \\ =& \begin{pmatrix} \frac{a}{\sqrt{a^2+b^2}} & \frac{b}{\sqrt{a^2+b^2}} & 0 \end{pmatrix} \begin{pmatrix} \frac{m}{12}b^2 & 0 & 0 \\ 0 & \frac{m}{12}a^2 & 0 \\ 0 & 0 & \frac{m}{12}(a^2+b^2) \end{pmatrix} \begin{pmatrix} \frac{a}{\sqrt{a^2+b^2}} \\ \frac{b}{\sqrt{a^2+b^2}} \\ 0 \end{pmatrix} \\ =& \frac{m}{6}\frac{a^2b^2}{a^2+b^2}\,. \quad \square \end{aligned}$$

6.1.3. Moment of inertia with respect to an eigenbasis of the inertia tensor. Observe that the inertia tensor is *symmetric* and recall the **spectral theorem** for symmetric matrices that we saw in Theorem 4.8. Let $(\tilde{e}_1, \tilde{e}_2, \tilde{e}_3)$ be an orthonormal basis for the inertia tensor of a rigid body M. Let I_1, I_2, I_3 be the corresponding eigenvalues of the inertia tensor. The matrix representing the inertia tensor with respect to this eigenbasis is

$$\begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} .$$

The axes of the eigenvectors $\tilde{e}_1, \tilde{e}_2, \tilde{e}_3$ are called the **principal axes of inertia** of the rigid body M. The eigenvalues I_i are called the **principal moments of inertia**.

The *principal moments of inertia* are the moments if inertia with respect to the *principal axes of inertia*, hence they are non-negative

$$I_1, \, , I_2, \, , I_3 \geq 0 \, .$$

A rigid body is called

(1) an **asymmetric top** if $I_1 \neq I_2 \neq I_3 \neq I_1$;
(2) a **symmetric top** if $I_1 = I_2 \neq I_3$: any axis passing through the plane determined by e_1 and e_2 is a principal axis of inertia;
(3) a **spherical top** if $I_1 = I_2 = I_3$: any axis passing through O is a principal axis of inertia.

With respect to the eigenbasis $\{\tilde{e}_1, \tilde{e}_2, \tilde{e}_3\}$ the *kinetic energy* is

$$E = \frac{1}{2}(I_1(\tilde{\omega}^1)^2 + I_2(\tilde{\omega}^2)^2 + I_3(\tilde{\omega}^3)^2) \, ,$$

where $\omega = \tilde{\omega}^i \tilde{e}_i$, with $\tilde{\omega}^i$ the components of the angular velocity with respect to the basis $\{\tilde{e}_1, \tilde{e}_2, \tilde{e}_3\}$.

The surface determined by the equation (with respect to the coordinates x, y, z)

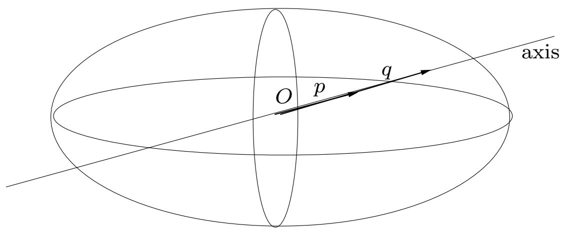
$$I_1 x^2 + I_2 y^2 + I_3 z^2 = 1$$

is called the **ellipsoid of inertia**. The symmetry axes of the ellipsoid coincide with the principal axes of inertia. Note that for a spherical top, the ellipsoid of inertia is actually a sphere.

The ellipsoid of inertia gives the moment of inertia with respect to any axis as follows: Consider an axis given by the unit factor p and let q be a vector of intersection of the axis with the ellipsoid of inertia.

$$q = cp$$

where c is the (signed) distance to O of the intersection of the axis with the ellipsoid of inertia.

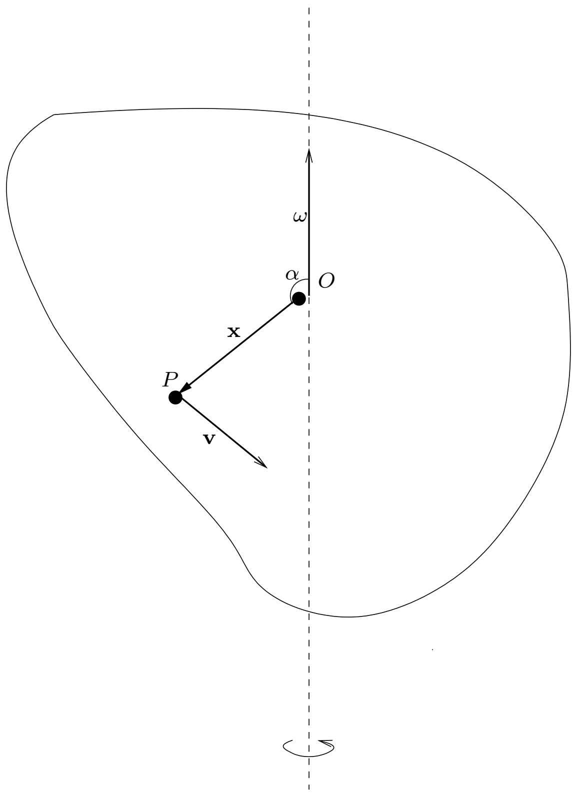


The moment of inertia with respect to this axis if

$$I = I_{ij}p^ip^j = \frac{1}{c^2}I_{ij}q^iq^j = \frac{1}{c^2},$$

where the last equality follows from the fact that, since q is on the ellipsoid, then $I_{ij}q^iq^j = 1$.

6.1.4. Angular momentum. Let M be a body rotating with angular velocity ω about an axis through the point O. Let $\mathbf{x} = \overrightarrow{OP}$ be the position vector of a point P and $\mathbf{v} = \omega \times \mathbf{x}$ the linear velocity of P.



Then the **angular momentum** of an infinitesimal region of M around P is

$$dL = (\mathbf{x} \times \mathbf{v})dm\,,$$

so that the **total angular momentum** of M is

$$L = \int_M (\mathbf{x} \times (\omega \times \mathbf{x}))dm\,.$$

We need to work out $\mathbf{x} \times (\omega \times \mathbf{x})$ for vectors $\mathbf{x}$ and ω in three dimensional space. It is easy to prove, using coordinates[3], the equality

$$\mathbf{x} \times (\omega \times \mathbf{x}) = \omega(\mathbf{x} \bullet \mathbf{x}) - \mathbf{x}(\omega \bullet \mathbf{x})\,. \tag{6.2}$$

Let $\mathcal{B} = \{e_1, e_2, e_3\}$ be an orthonormal basis of $\mathbb{R}^3$. Then, replacing the following equalities

$$\omega = \omega^i e_i = \delta^i_j \omega^j e_i \tag{6.3}$$

$$\mathbf{x} = x^i e_i = \delta^i_k x^k e_i \tag{6.4}$$

$$\mathbf{x} \bullet \mathbf{x} = \delta_{k\ell} x^k x^\ell \tag{6.5}$$

$$\omega \bullet \mathbf{x} = \delta_{j\ell} \omega^j x^\ell \tag{6.6}$$

into (6.2), we obtain

$$\mathbf{x} \times (\omega \times \mathbf{x}) = \underbrace{\delta^i_j \omega^j e_i}_{(6.3)}(\underbrace{\delta_{k\ell} x^k x^\ell}_{(6.5)}) - \underbrace{\delta^i_k x^k e_i}_{(6.4)}(\underbrace{\delta_{j\ell} \omega^j x^\ell}_{(6.6)}) = (\delta^i_j \delta_{k\ell} - \delta^i_k \delta_{j\ell})\omega^j x^k x^\ell e_i\,.$$

Therefore the total angular momentum is

$$L = L^i e_i\,,$$

where the components L^i are

$$L^i = (\delta^i_j \delta_{k\ell} - \delta^i_k \delta_{j\ell})\omega^j \int_M x^k x^\ell dm\,.$$

Since we restrict to orthonormal bases, we have always

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

Hence the above expression for L^i can be written in terms of the inertia tensor I_{ij} as

$$L^i = I_{ij}\omega^j\,.$$

[3]Consider only the case in which $\mathbf{x}$ is a basis vector and use the linearity in ω.

Example 6.5. Suppose the rectangular plate in the previous examples is rotating about an axis through O with angular velocity

$$\omega = e_1 + 2e_2 + 3e_3\,, \qquad \text{or} \qquad \begin{pmatrix}\omega^1\\ \omega^2\\ \omega^3\end{pmatrix} = \begin{pmatrix}1\\2\\3\end{pmatrix}.$$

Compute its angular momentum.

The inertia tensor is given by the matrix I_{ij}

$$\begin{pmatrix}\frac{m}{12}b^2 & 0 & 0\\ 0 & \frac{m}{12}a^2 & 0\\ 0 & 0 & \frac{m}{12}(a^2+b^2)\end{pmatrix}.$$

The total angular momentum has components given by

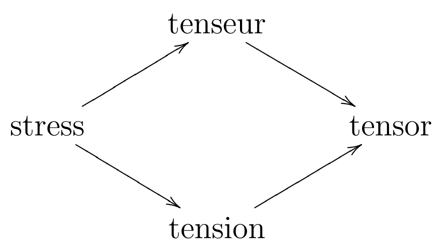
$$\begin{pmatrix}\frac{m}{12}b^2 & 0 & 0\\ 0 & \frac{m}{12}a^2 & 0\\ 0 & 0 & \frac{m}{12}(a^2+b^2)\end{pmatrix}\begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}\frac{m}{12}b^2\\ \frac{m}{6}a^2\\ \frac{m}{4}(a^2+b^2)\end{pmatrix} = \begin{pmatrix}L^1\\L^2\\L^3\end{pmatrix},$$

so that

$$L = \frac{m}{12}b^2e_1 + \frac{m}{6}a^2e_2 + \frac{m}{4}(a^2+b^2)e_3\,.$$

6.2. Stress tensor (Spannung)

It was the concept of **stress** in mechanics that originally led to the invention of tensors



Let us consider a rigid body M acted upon by external forces but in *static equilibrium*, and let us consider an infinitesimal region dM around a point P. There are two types of external forces:

(1) The **body forces**, that is forces whose magnitude is proportional to the volume/mass of the region. For instance, *gravity*, *attractive force* or the *centrifugal force*.
(2) The **surface forces**, that is forces exerted on the surface of the element by the material surrounding it. They are forces whose magnitude is proportional to the area of the region in consideration.

The *surface force per unit area* is called the **stress**. We will concentrate on **homogeneous stress**, that is stress that does not depend on the location of the element in the body, but depends only on the orientation of the surface/plane. Moreover, we assume that the body in consideration is in *static equilibrium*.

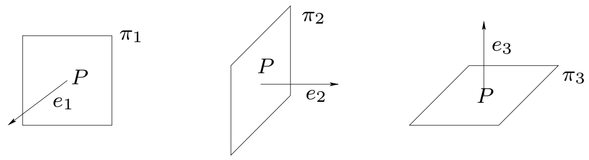
Choose an orthonormal basis $\{e_1, e_2, e_3\}$ and the plane π through P parallel to the e_2e_3 coordinate plane. The normal to the plane is the vector e_1. Let ΔA_1 be the area of the slice of the infinitesimal region around P cut by the plane and let ΔF be the force acting on that slice. We write ΔF in terms of its components

$$\Delta F = \Delta F^1 e_1 + \Delta F^2 e_e + \Delta F^3 e_3$$

and, since we defined the stress to be the surface force per unit area, we define, for $j = 1, 2, 3$,

$$\sigma^{1j} := \lim_{\Delta A_1 \to 0} \frac{\Delta F^j}{\Delta A_1} .$$

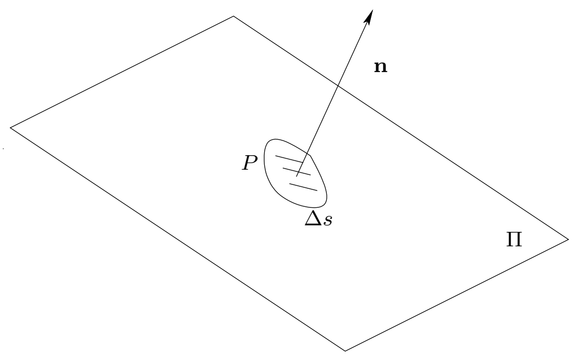
Similarly we can consider planes parallel to the other coordinate planes



and define

$$\sigma^{ij} := \lim_{\Delta A_i \to 0} \frac{\Delta F^j}{\Delta A_i} .$$

It turns out that the resulting nine numbers σ^{ij} form a contravariant 2-tensor called the **stress tensor**. To see this, we compute the stress tensor across *other* slices through P, that is other planes with other normal vectors. Let Π be a plane passing through P, $\mathbf{n}$ the unit vector through P perpendicular to the plane π, $\Delta s = \pi \cap dM$ the area of a small element of the plane Π containing P and ΔF the force acting on that element.

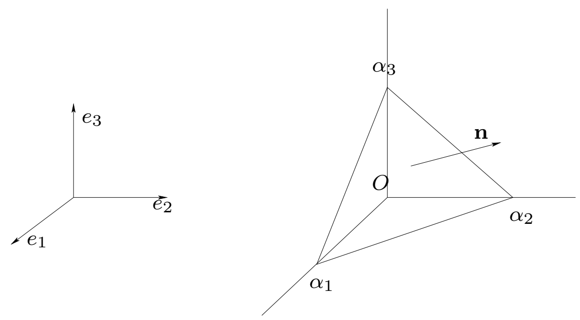


Claim 6.6. The stress at P across the surface perpendicular to $\mathbf{n}$ is

$$\sigma(\mathbf{n}) \stackrel{\text{def}}{=} \lim_{\Delta s \to 0} \frac{\Delta F}{\Delta s} = \sigma^{ij}(\mathbf{n} \bullet e_i) e_j .$$

It follows from the claim that the stress σ is a vector valued function that depends linearly on the normal $\mathbf{n}$ to the surface element, and we will see in § 6.2.2 that the matrix σ^{ij} of this linear vector valued function forms a second-order tensor.

PROOF. Consider the tetrahedron $OA_1A_2A_3$ formed by the *triangular slice* on the plane Π having area Δs and three triangles on planes parallel to the coordinate planes



Consider all forces acting on this tetrahedron as a volume element of the rigid body. There can be two types of forces:

(1) *Body forces*$= f \bullet \Delta v$, where f is the force per unit of volume and Δv is the volume of the tetrahedron. We actually do not know these forces, but we will see later that this is not relevant.
(2) *Surface forces*, that is the sum of the forces on each of the four sides of the tetrahedron.

We want to assess each of the four surface contributions due to the surface forces. If Δs is the area of the slice on the plane Π, the contribution of that slice is

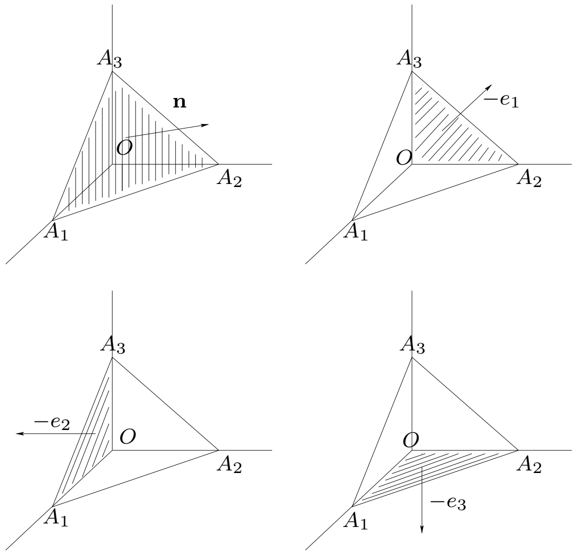
$$\sigma(\mathbf{n})\Delta s\,.$$

If Δs_1 is the area of the slice on the plane with normal $-e_1$, the contribution of that slice is

$$-\sigma^{1j}e_j\Delta s_1\,,$$

and similarly the contributions of the other two slices are

$$-\sigma^{2j}e_j\Delta s_2 \qquad \text{and} \qquad -\sigma^{3j}e_j\Delta s_3\,.$$



Note that the minus sign comes from the fact that we use everywhere outside pointing normals.

So the total surface force is

$$\sigma(\mathbf{n})\Delta s - \sigma^{1j} e_j \Delta s_1 - \sigma^{2j} e_j \Delta s_2 - \sigma^{3j} e_j \Delta s_3 \,.$$

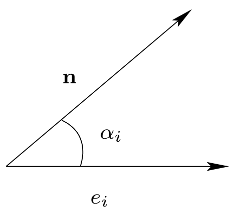
Since there is static equilibrium the sum of all (body and surface) forces must be zero

$$f\Delta v + \sigma(\mathbf{n})\Delta s - \sigma^{ij} e_j \Delta s_i = 0 \,.$$

The term $f\Delta v$ can be neglected when Δs is small, as it contains terms of higher order (in fact $\Delta v \to 0$ faster than $\Delta s \to 0$). We conclude that

$$\sigma(\mathbf{n})\Delta s = \sigma^{ij} e_j \Delta s_i \,.$$

It remains to relate Δs to $\Delta s_1, \Delta s_2, \Delta s_3$. The side with area Δs_i is the orthogonal projection of the side with area Δs onto the plane with normal e_i. The scaling factor for the area under projection is $\cos\alpha_i$, where α_i is the convex angle between the plane normal vectors



$$\frac{\Delta s_i}{\Delta s} = \cos\alpha_i = \cos\alpha_i \|\mathbf{n}\| \, \|e_i\| = \mathbf{n} \bullet e_i \,.$$

Therefore

$$\sigma(\mathbf{n})\Delta s = \sigma^{ij} e_j (\mathbf{n} \bullet e_i)\Delta s$$

or, equivalently,

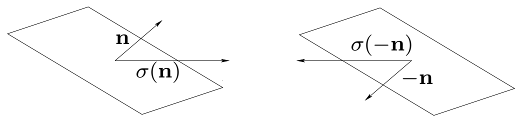
$$\sigma(\mathbf{n}) = \sigma^{ij}(\mathbf{n} \bullet e_i)e_j\,.$$

□

REMARK 6.7. (1) For homogeneous stress, the stress tensor σ^{ij} does not depend on the point P. However, when we flip the orientation of the normal to the plane, the stress tensor changes sign. In other words, if $\sigma(\mathbf{n})$ is the stress across a surface with normal $\mathbf{n}$, then

$$\sigma(-\mathbf{n}) = -\sigma(\mathbf{n})\,.$$

The stress considers orientation as if the forces on each side of the surface have to balance each other in static equilibrium.

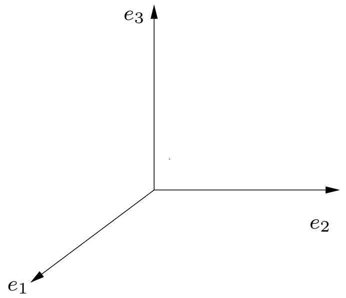


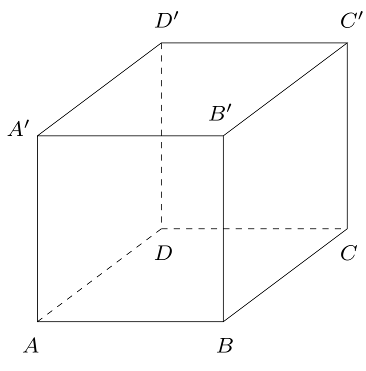
(2) In the formula $\sigma(\mathbf{n}) = \sigma^{ij}(\mathbf{n} \bullet e_i)e_j$, the quantities $\mathbf{n} \bullet e_i$ are the coordinates of $\mathbf{n}$ with respect to the orthonormal basis $\{e_1, e_2, e_3\}$, namely

$$\mathbf{n} = (\mathbf{n} \bullet e_1)e_1 + (\mathbf{n} \bullet e_2)e_2 + (\mathbf{n} \bullet e_3)e_3 = n^1 e_1 + n^2 e_2 + n^3 e_3\,.$$

CLAIM 6.8. The stress tensor is a *symmetric tensor*, that is $\sigma^{ij} = \sigma^{ji}$.

In fact, let us consider an infinitesimal cube of side $\Delta\ell$ surrounding P and with faces parallel to the coordinate planes.





The force acting on each of the six faces of the cube are:

- $\sigma^{1j}\Delta A_1 e_j$ and $-\sigma^{1j}\Delta A_1 e_j$, respectively for the front and the back faces, $ABB'A'$ and $DCC'D'$;

- $\sigma^{2j}\Delta A_2 e_j$ and $-\sigma^{2j}\Delta A_2 e_j$, respectively for the right and the left faces $BCC'B'$ and $ADD'A'$;
- $\sigma^{3j}\Delta A_3 e_j$ and $-\sigma^{3j}\Delta A_3 e_j$, respectively for the top and the bottom faces $ABCD$ and $A'B'C'D'$,

where $\Delta A_1 = \Delta A_2 = \Delta A_3 = \Delta s = (\Delta\ell)^2$ is the common face area. We compute now the torque μ, assuming the forces are applied at the center of the faces (whose distance is $\frac{1}{2}\Delta\ell$ to the center point P). Recall that the torque is the tendency of a force to twist or rotate an object.

$$\begin{aligned}
\mu =& \frac{\Delta\ell}{2} e_1 \times \sigma^{1j}\Delta s\, e_j + \left(-\frac{\Delta\ell}{2} e_1\right) \times (-\sigma^{1j}\Delta s\, e_j\,) \\
&+ \frac{\Delta\ell}{2} e_2 \times \sigma^{2j}\Delta s\, e_j + \left(-\frac{\Delta\ell}{2} e_2\right) \times (-\sigma^{2j}\Delta s\, e_j\,) \\
&+ \frac{\Delta\ell}{2} e_2 \times \sigma^{3j}\Delta s\, e_j + \left(-\frac{\Delta\ell}{2} e_3\right) \times (-\sigma^{3j}\Delta s\, e_j\,) \\
=& \Delta\ell\Delta s\,(e_i \times \sigma^{ij} e_j) = \\
=& \Delta\ell\Delta s\big((\sigma^{23}-\sigma^{32})e_1 + (\sigma^{31}-\sigma^{13})e_2 + (\sigma^{12}-\sigma^{21})e_3\big)\,.
\end{aligned}$$

Since the equilibrium is static, then $L = 0$, so that $\sigma^{ij} = \sigma^{ji}$.

We can hence write

$$\sigma = \begin{pmatrix} \sigma^{11} & \sigma^{12} & \sigma^{13} \\ \sigma^{12} & \sigma^{22} & \sigma^{23} \\ \sigma^{13} & \sigma^{23} & \sigma^{33} \end{pmatrix},$$

where the diagonal entries σ^{11}, σ^{22} and σ^{33} are the **normal components**, that is the components of the forces perpendicular to the coordinate planes and the remaining entries σ^{12}, σ^{13} and σ^{23} are the **shear components**, that is the components of the forces parallel to the coordinate planes.

Since the stress tensor is symmetric, it can be orthogonally diagonalized, that is

$$\sigma = \begin{pmatrix} \sigma^{1} & 0 & 0 \\ 0 & \sigma^{2} & 0 \\ 0 & 0 & \sigma^{3} \end{pmatrix},$$

where now σ^1, σ^2 and σ^3 are the **principal stresses**, that is the *eigenvalues of* σ. The eigenspaces of σ are the **principal directions** and the shear components disappear for the **principal planes**.

6.2.1. Special forms of the stress tensor (written with respect to an orthonormal eigenbasis or another special basis).

- **Uniaxial stress** with stress tensor given by

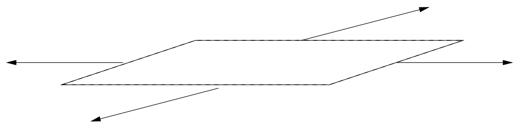
$$\begin{pmatrix} \sigma & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

EXAMPLE 6.9. This is the stress tensor in a long vertical rod loaded by hanging a weight on the end.

- **Plane stressed state** or **biaxial stress** with stress tensor given by

$$\begin{pmatrix} \sigma^1 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

EXAMPLE 6.10. This is the stress tensor in plate on which forces are applied as in the picture.



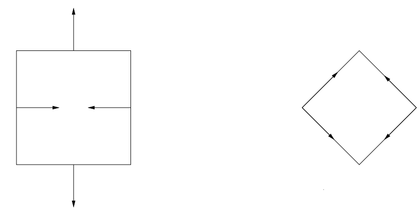
- **Pure shear** with stress tensor given by

$$\begin{pmatrix} -\sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & \sigma & 0 \\ \sigma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{6.7}$$

This is special case of the biaxial stress, in the case in which $\sigma^1 = \sigma^2$. In (6.7) the first is the stress tensor written with respect to an eigenbasis, while the second is the stress tensor written with respect to an orthonormal basis obtained by rotating an eigenbasis by 45° about the third axis. In fact

$$\begin{pmatrix} 0 & \sigma & 0 \\ \sigma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{{}^{t}L} \begin{pmatrix} -\sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{pmatrix} \underbrace{\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{L}$$

where L is the matrix of the change of coordinates.



- **Shear deformation** with stress tensor given by

$$\begin{pmatrix} 0 & \sigma^{12} & \sigma^{13} \\ \sigma^{12} & 0 & \sigma^{23} \\ \sigma^{13} & \sigma^{23} & 0 \end{pmatrix}$$

with respect to some orthonormal basis.

FACT 6.11. *The stress tensor σ is a shear deformation if and only if its trace is zero.*

EXAMPLE 6.12. The stress tensor

$$\begin{pmatrix} 2 & -4 & 0 \\ -4 & 0 & 4 \\ 0 & 4 & -2 \end{pmatrix}$$

represents a shear deformation. In fact one can check that

$$\underbrace{\begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{pmatrix}}_{{}^{t}L} \begin{pmatrix} 2 & -4 & 0 \\ -4 & 0 & 4 \\ 0 & 4 & -2 \end{pmatrix} \underbrace{\begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{pmatrix}}_{L} = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 4\sqrt{2} \\ -2 & 4\sqrt{2} & 0 \end{pmatrix}$$

- **Hydrostatic pressure** with stress tensor given by

$$\begin{pmatrix} -p & 0 & 0 \\ 0 & -p & 0 \\ 0 & 0 & -p \end{pmatrix},$$

where $p \neq 0$ is the pressure. Here all eigenvalues are equal to $-p$.

EXAMPLE 6.13. Pressure of a fluid on a bubble.

EXERCISE 6.14. Any stress tensor can be written as the sum of a hydrostatic pressure and a shear deformation. *Hint:* look at the trace.

6.2.2. Contravariance of the stress tensor. Let $\mathcal{B} = \{e_1, e_2, e_3\}$ and $\widetilde{\mathcal{B}} = \{\tilde{e}_1, \tilde{e}_2, \tilde{e}_3\}$ be two basis, and let

$$\tilde{e}_i = L_i^j e_j \qquad \text{and} \qquad e_i = \Lambda_i^j \tilde{e}_j\,, \tag{6.8}$$

where $L := L_{\widetilde{\mathcal{B}}\mathcal{B}}$ is the matrix of the change of basis and $\Lambda = L^{-1}$ is the inverse. Let $\mathbf{n}$ be a given unit vector and S the stress across a surface perpendicular to $\mathbf{n}$. Then S can be expressed in two way, respectively with respect to $\mathcal{B}$ and to $\widetilde{\mathcal{B}}$

$$S = \sigma^{ij}(\mathbf{n} \bullet e_i)e_j \tag{6.9}$$

$$S = \tilde{\sigma}^{ij}(\mathbf{n} \bullet \tilde{e}_i)\tilde{e}_j\,, \tag{6.10}$$

and we want to relate σ^{ij} to $\tilde{\sigma}^{ij}$. We start with the expression for S in (6.9) and rename the indices for later convenience.

$$S = \sigma^{km}(\mathbf{n} \bullet e_k)e_m = \sigma^{km}(\mathbf{n} \bullet \Lambda_k^i \tilde{e}_i)(\Lambda_m^j \tilde{e}_j) = \sigma^{km}\Lambda_k^i \Lambda_m^j (\mathbf{n} \bullet \tilde{e}_i)\tilde{e}_j\,,$$

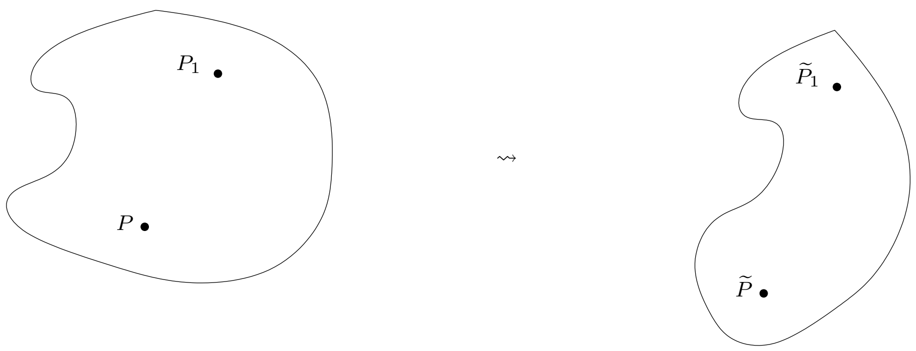
where in the second equality we used (6.8), and in the third we used linearity. Comparing the last expression with the expression in (6.10) we obtain

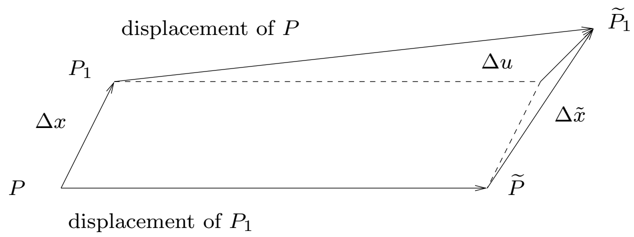
$$\tilde{\sigma}^{ij} = \sigma^{km}\Lambda_k^i \Lambda_m^j\,,$$

thus showing that σ is a **contravariant 2-tensor** or a **tensor of type** $(0, 2)$.

6.3. Strain tensor (Verzerrung)

Consider a slightly deformation of a body





We have

$$\Delta\tilde{x} = \Delta x + \Delta u\,,$$

where Δx is the old relative position of P and P_1, $\Delta\tilde{x}$ is their new relative position and Δu is the difference of the displacement, which hence measures the deformation.

Assume that we have a small homogeneous deformation, that is

$$\Delta u = f(\Delta x)\,;$$

in other words f is a small linear function independent of the point P. If we write the components of Δu and Δx with respect to an orthonormal basis $\{e_1, e_2, e_3\}$, the function f will be represented by a matrix with entries that we denote by f_{ij},

$$\Delta u_i = f_{ij}\Delta x^j\,.$$

The matrix (f_{ij}) can be written as a sum of a symmetric and an antisymmetric matrix as follows:

$$f_{ij} = \epsilon_{ij} + \omega_{ij}\,,$$

where

$$\epsilon_{ij} = \frac{1}{2}(f_{ij} + f_{ji})$$

is a symmetric matrix and is called the **strain tensor** or **deformation tensor** and

$$\omega_{ij} = \frac{1}{2}(f_{ij} - f_{ji})$$

is an antisymmetric matrix called the **rotation tensor**. We try to understand now where these names come from.

REMARK 6.15. First we verify that a (small) antisymmetric 3×3 matrix represents a (small) rotation in 3-dimensional space.

FACT 6.16. *Let V be a vector space with orthonormal basis $\mathcal{B} = \{e_1, e_2, e_3\}$, and let $\omega = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$. The matrix of the linear map $V \to V$ defined by $v \mapsto \omega \times v$ with respect to the basis $\mathcal{B}$ is*

$$\begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}.$$

In fact

$$\begin{aligned} \omega \times v = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \det \begin{pmatrix} e_1 & e_2 & e_3 \\ a & b & c \\ x & y & z \end{pmatrix} \\ = \begin{pmatrix} bz - cy \\ cx - az \\ ay - bx \end{pmatrix} &= \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \end{aligned}$$

Note that the matrix $(\omega_{ij}) = \begin{pmatrix} 0 & \omega_{12} & -\omega_{13} \\ -\omega_{12} & 0 & \omega_{23} \\ \omega_{13} & -\omega_{23} & 0 \end{pmatrix}$ corresponds to the cross product with the vector $\omega = \begin{pmatrix} -\omega_{23} \\ -\omega_{13} \\ -\omega_{12} \end{pmatrix}$. □

The antisymmetric case. Suppose that the matrix (f_{ij}) was already *antisymmetric*, so that

$$\omega_{ij} = f_{ij} \qquad \text{and} \qquad \epsilon_{ij} = 0\,.$$

By the Fact 6.16, the relation

$$\Delta u_i = f_{ij} \Delta x^j \tag{6.11}$$

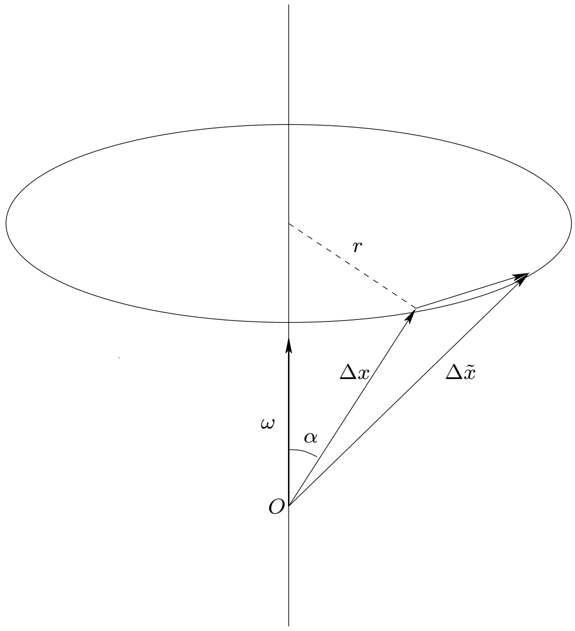
is equivalent to

$$\Delta u = \omega \times \Delta x\,,$$

so that

$$\Delta \tilde{x} = \Delta x + \Delta u = \Delta x + \omega \times \Delta x\,.$$

When ω is small, this represents an infinitesimal rotation of an angle $\|\omega\|$ about the axis $O\omega$.



In fact, since $\omega \times \Delta x$ is orthogonal to the plane determined by ω and by Δx, it is tangent to the circle with center along the axis $O\omega$ and radius determined by Δx. Moreover,

$$\|\Delta u\| = \|\omega \times \Delta x\| = \|\omega\| \underbrace{\|\Delta x\| \sin\alpha}_{r},$$

and hence, since the length of an arc of a circle of radius r corresponding to an angle θ is $r\theta$, infinitesimally this represents a rotation by an angle $\|\omega$

.

The symmetric case. The opposite extreme case is when the matrix f_{ij} was already *symmetric*, so that

$$\epsilon_{ij} = f_{ij} \qquad \text{and} \qquad \omega_{ij} = 0\,.$$

We will see that it is ϵ_{ij} that encodes the changes in the distances: in fact,

$$\begin{aligned} \|\Delta\tilde{x}\|^2 &= \Delta\tilde{x} \bullet \Delta\tilde{x} = (\Delta x + \Delta u) \bullet (\Delta x + \Delta u) \\ &= \Delta x \bullet \Delta x + 2\Delta x \bullet \Delta u + \Delta u \bullet \Delta u \\ &\simeq \|\Delta x\|^2 + 2\epsilon_{ij}\Delta x^i \Delta x^j\,, \end{aligned} \tag{6.12}$$

where in the last step we neglected the term $\|\Delta u\|^2$ since it is small compared to Δu when $\Delta u \to 0$ and used (6.11).

Remark 6.17. Even when f_{ij} is not purely symmetric, only the symmetric part ϵ_{ij} is relevant for the distortion of the distances. In fact, if ω_{ij} is antisymmetric, the term $2\omega_{ij}\Delta x^i \Delta x^j = 0$, so that

$$\|\Delta\tilde{x}\|^2 \simeq \|\Delta x\|^2 + 2f_{ij}\Delta x^i \Delta x^j = \|\Delta x\|^2 + 2\epsilon_{ij}\Delta x^i \Delta x^j \,. \quad \square$$

Recall that a metric tensor (or inner product) encodes the distances among points. It follows that a deformation changes the metric tensor. Let us denote by g the metric before the deformation and by $\tilde{g}$ the metric after the deformation. By definition we have

$$\|\Delta\tilde{x}\|^2 \stackrel{\text{def}}{=} \tilde{g}(\Delta\tilde{x}, \Delta\tilde{x}) = \tilde{g}_{ij}\Delta\tilde{x}^i \Delta\tilde{x}^j = \tilde{g}_{ij}(\Delta x^i + \Delta u^i)(\Delta x^j + \Delta u^j) \tag{6.13}$$

and

$$\|\Delta x\|^2 \stackrel{\text{def}}{=} g(\Delta x, \Delta x) = g_{ij}\Delta x^i \Delta x^j \,. \tag{6.14}$$

For infinitesimal deformations (that is if $\Delta u \sim 0$), (6.13) becomes

$$\|\Delta\tilde{x}\|^2 = \tilde{g}_{ij}\Delta x^i \Delta x^j \,.$$

This, together with (6.14) and (6.12), leads to

$$\tilde{g}_{ij}\Delta x^i \Delta x^j \simeq g_{ij}\Delta x^i \Delta x^j + 2\epsilon_{ij}\Delta x^i \Delta x^j$$

and hence

$$\epsilon_{ij} \simeq \frac{1}{2}(\tilde{g}_{ij} - g_{ij}) \,,$$

that is ϵ_{ij} measures the change in the metric.

By definition the strain tensor ϵ_{ij} is symmetric

$$\mathcal{E} = \begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{12} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{13} & \epsilon_{23} & \epsilon_{33} \end{pmatrix} ,$$

where the terms on the diagonal (in green) determine the elongation or the contraction of the body along the coordinate directions e_1, e_2, e_3, and the terms above the diagonal (in orange) are the **shear components** of the strain tensor; that is ϵ_{ij} is the movement of a line element parallel to Oe_j towards Oe_i. Since it is a symmetric tensor it can be orthogonally diagonalized

$$\begin{pmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_2 & 0 \\ 0 & 0 & \epsilon_3 \end{pmatrix} ,$$

The eigenvalues of $\mathcal{E}$ are the **principal coefficients** of the deformation and the eigenspaces are the **principal directions** of the deformation.

6.3.1. Special forms of the strain tensor.

(1) **Shear deformation** when $\mathcal{E}$ is traceless

$$\operatorname{tr}\mathcal{E} = \epsilon_{11} + \epsilon_{22} + \epsilon_{33} = 0\,.$$

(2) **Uniform compression** when the principal coefficients of $\mathcal{E}$ are equal (and nonzero)

$$\begin{pmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{pmatrix}$$

EXERCISE 6.18. Any strain tensor can be written as the sum of a uniform compression and a shear deformation.

6.4. Elasticity tensor

The stress tensor represents an *external exertion* on the material, while the strain tensor represents the *material reaction* to that exertion. In crystallography these are called **field tensors** because they represent imposed conditions, opposed to **matter tensors**, that represents material properties.

Hooke's law says that, for small deformations, stress is related to strain by a matter tensor called **elasticity tensor** or **stiffness tensor**

$$\sigma^{ij} = E^{ijk\ell}\epsilon_{kl}\,,$$

while the tensor relating strain to stress is the **compliance tensor**

$$\epsilon_{k\ell} = S_{ijk\ell}\sigma^{ij}\,.$$

The elasticity tensor has rank 4, and hence in 3-dimensional space it has $3^4 = 81$ components. However symmetry reduces the number of independent components for $E^{ijk\ell}$.

(1) *Minor symmetries:* The symmetry of the stress tensor

$$\sigma^{ij} = \sigma^{ji}$$

implies that

$$E^{ijk\ell} = E^{jik\ell} \qquad \text{for each } k, \ell\,;$$

it follows that for each k, ℓ fixed there are only 6 independent components $E^{ijk\ell}$

$$\begin{pmatrix} E^{11k\ell} & E^{12k\ell} & E^{13k\ell} \\ E^{12k\ell} & E^{22k\ell} & E^{23k\ell} \\ E^{13k\ell} & E^{23k\ell} & E^{33k\ell}\,. \end{pmatrix}$$

Having taken this in consideration, the number of independent components decreases to 6×3^2 at the most. However the symmetry also of the strain tensor

$$\epsilon_{k\ell} = \epsilon_{\ell k}$$

implies that

$$E^{ijk\ell} = E^{ij\ell k} \qquad \text{for each } i, j\,.$$

This means that for each i, j fixed there are also only 6 independent components $E^{ijk\ell}$, so that $E^{ijk\ell}$ has at most $6^2 = 36$ independent components.

(2) *Major symmetries:* Since (under appropriate conditions) partial derivatives commute, if follows from the existence of a *strain energy density functional* U satisfying

$$\frac{\partial^2 U}{\partial \epsilon_{ij} \partial \epsilon_{k\ell}} = E^{ijk\ell}$$

that

$$E^{ijk\ell} = E^{k\ell ij}\,,$$

that means the matrix with rows labelled by (i, j) and columns labelled by (k, ℓ) is symmetric. Since from (1) that there are only 6 entries (i, j) for a fixed (k, ℓ), $E^{ijk\ell}$ can be written in a 6×6 matrix with rows labelled by (i, j) and columns labelled by (k, ℓ)

$$\begin{pmatrix} * & * & * & * & * & * \\ & * & * & * & * & * \\ & & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \\ & & & & & * \end{pmatrix}$$

so that $E^{ijk\ell}$ has in fact only $6 + 5 + 4 + 3 + 2 + 1 = 21$ components.

6.5. Conductivity tensor

Consider a homogeneous continuous crystal. Its properties can be divided into two classes:

- Properties that *do not depend* on a direction, and are hence described by *scalars*. Examples are density and heat capacity.
- Properties that *depends* on a direction, and are hence described by *tensors*. Examples are **elasticity**, **electrical conductivity** and **heat conductivity**. We say that a crystal is **anisotropic** when it has such "tensorial" properties.

6.5.1. Electrical conductivity.

Let E be the *electric field* and J the *electrical current density*. We assume that these are constant throughout the crystal. At each point of the crystal:

(1) E gives the **electric force** (in Volts/m) that would be exerted on a positive test charge (of 1 Coulomb) placed at the point;

(2) J (in Amperes/m^2) gives the direction the charge carriers move and the **rate of electric current** across an infinitesimal surface perpendicular to that direction.

J is a function of E,

$$J = f(E)\,.$$

Consider a small increment ΔJ in J caused by a small increment ΔE in E, and write these increments in terms of their components with respect to a chosen orthonormal basis $\{e_1, e_2, e_3\}$.

$$\Delta J = \Delta J^i e_i \qquad \text{and} \qquad \Delta E = \Delta E^i e_i\,.$$

The increments are related by

$$\Delta J^i = \frac{\partial f^i}{\partial E^j}\Delta E^j + \text{higher order terms in } (\Delta E^j)^2, (\Delta E^j)^2, \dots$$

If the quantities ΔE^j are small, we can assume that

$$\Delta J^i = \frac{\partial f^i}{\partial E^j}\Delta E^j \tag{6.15}$$

If we assume that $\frac{\partial f^i}{\partial E^j}$ is independent of the point of the crystal,

$$\frac{\partial f^i}{\partial E^j} = \sigma^i_j \in \mathbb{R}$$

we obtain the relation

$$\Delta J^i = \sigma^i_j \Delta E^j$$

or simply

$$\Delta J = \sigma\, \Delta E\,,$$

where σ if the **electrical conductivity tensor**. This is a $(1,1)$-tensor and may depend[4] on the initial value of E, that is the electrical conductivity may be different for small and large electric forces. If initially $E = 0$ and σ^0 is the corresponding electrical conductivity tensor, we obtain the relation

$$J = \sigma^0\, E$$

that is called the *generalized Ohm law*. This is always under the assumption that ΔE and ΔJ are small and that the relation is linear.

The **electrical resistivity tensor** is

$$\rho = K^{-1}\,,$$

that is, it is the $(1,1)$-tensor such that

$$\rho^j_i K^\ell_j = \delta^\ell_i\,.$$

[4]Typically if the dependence between E and J is linear for any value, and not only for small ones, the tensor ail not depend on the initial value of E.

The electrical conductivity measures the material's ability to conduct an electrical current, while the electrical resistivity quantifies the ability of the material to oppose the flow of the electrical current.

For an *isotropic* crystal, all directions are equivalent and these tensors are spherical

$$\sigma_i^j = k\delta_i^j \qquad \text{and} \qquad \rho_i^j = \frac{1}{k}\delta_i^j\,, \tag{6.16}$$

where k is a scalar, called the **electrical conductivity** of the crystal. Equation (6.16) can also be written as

$$\begin{pmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} \frac{1}{k} & 0 & 0 \\ 0 & \frac{1}{k} & 0 \\ 0 & 0 & \frac{1}{k} \end{pmatrix}.$$

In general, σ_i^j is neither symmetric nor antisymmetric (and actually *symmetry* does not even make sense for a $(1,1)$ tensor unless a metric is fixed, since it does require a canonical identification of V with V^*).

6.5.2. Heat conductivity. Let T be the *temperature* and H the *heat flux vector*. For a homogeneous crystal and constant H and for a constant gradient of T, *Fourier heat conduction law* says that

$$H = -K \operatorname{grad} T\,. \tag{6.17}$$

At each point of the crystal:

(1) $\operatorname{grad} T$ points in the direction of the highest ascent of the temperature and measures the rate of increase of T in that direction. The minus sign in (6.17) comes from the fact that the heat flows in the direction of the decreasing temperature.
(2) H measure the amount of heat passing per unit area perpendicular to its direction per unit time.

Here K is the **heat conductivity tensor** or **thermal conductivity tensor**. In terms of components with respect to a chosen orthonormal basis

$$H^i = -K^{ij}(\operatorname{grad} T)_j\,.$$

EXERCISE 6.19. Verify that the gradient of a real function is a **covariant 1-tensor**.

The heat conductivity tensor is a contravariant 2-tensor and experiments show that it is symmetric and hence can be orthogonally diagonalized. The **heat resistivity tensor** is

$$r = K^{-1}\,,$$

and hence is also symmetric. With respect to an orthonormal basis, K is represented by

$$\begin{pmatrix} K_1 & 0 & 0 \\ 0 & K_2 & 0 \\ 0 & 0 & K_3 \end{pmatrix},$$

where the eigenvalues of K are called the **principal coefficients** of heat conductivity.

Physical considerations (that is the fact that heat flows always in the direction of decreasing temperature) show that the eigenvalues are positive

$$K_i > 0 \,.$$

The eigenspaces of K are called the **principal directions** of heat conductivity.

CHAPTER 7

Solutions

Exercise 2.4 (1) yes; (2) no, (3) no, (4) yes, (5) no, (6) yes.

Exercise 2.14

(1) The vectors in $\mathcal{B}$ span V since

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} .$$

Moreover they are linearly independent since

$$a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

if and only if

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

that is if and only if $a = b = c = 0$.

(2) $\mathcal{B}$ is a basis of V, hence $\dim V = 3$. Since $\widetilde{\mathcal{B}}$ has three elements, it is enough to check either that it spans V or that it consists of linearly independent vectors. We will check this last condition. In fact

$$a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \iff \begin{bmatrix} a & c-b \\ b+c & -a \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

that is

$$\begin{cases} a = 0 \\ b + c = 0 \\ c - b = 0 \end{cases} \iff \begin{cases} a = 0 \\ b = 0 \\ c = 0 \end{cases}$$

(3)

$$\begin{bmatrix} 2 & 1 \\ 7 & -2 \end{bmatrix} = 2 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + 1 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 7 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} ,$$

therefore

$$[v]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 1 \\ 7 \end{pmatrix} .$$

To compute the coordinates of v with respect to $\widetilde{\mathcal{B}}$ we need to find $a, b, c \in \mathbb{R}$ such that

$$\begin{bmatrix} 2 & 1 \\ 7 & -2 \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} .$$

Similar calculations to the above ones yield

$$[v]_{\widetilde{\mathcal{B}}} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} .$$

Exercise 3.2: (1) no; (2) no; (3) yes.

Exercise 3.3: (1) yes; (2) yes; (3) no.

Exercise 3.11:

Exercise 3.18: (1) yes; (2) yes; (3) yes; (4) no, because $v \times w$ is not a real number; (5) yes.

Exercise 4.2: (1) yes, this is the **standard inner product**; (2) no, as φ is **negative definite**, that is $\varphi(v, v) < 0$ if $v \in V$, $v \neq 0$; (3) no, as φ is not symmetric; (4) yes.

Exercise 4.3:

(1) Yes, in fact:
   (a) $\int_0^1 p(x)q(x)dx = \int_0^1 q(x)p(x)dx$ because $p(x)q(x) = q(x)p(x)$;
   (b) $\int_0^1 (p(x))^2 dx \geq 0$ for all $p \in \mathbb{R}[x]_2$ because $(p(x))^2 \geq 0$, and $\int_0^1 (p(x))^2 dx = 0$ only when $p(x) = 0$ for all $x \in [0, 1]$, that is only if $p \equiv 0$.

(2) No, since $\int_0^1 (p'(x))^2 dx = 0$ implies that $p'(x) = 0$ for all $x \in [0, 1]$, but p is not necessarily the zero polynomial.

(3) Yes

(4) No. Is there $p \in \mathbb{R}[x]_2$, $p \neq 0$ such that $(p(1))^2 + (p(2))^2 = 0$?

(5) Yes. Is there a non-zero polynomial of degree 2 with 3 distinct zeros?

Exercise 4.11. We write

$$[v]_{\widetilde{\mathcal{B}}} = \begin{pmatrix} \tilde{v}^1 \\ \tilde{v}^2 \\ \tilde{v}^3 \end{pmatrix} \quad \text{and} \quad [w]_{\widetilde{\mathcal{B}}} = \begin{pmatrix} \tilde{w}^1 \\ \tilde{w}^2 \\ \tilde{w}^3 \end{pmatrix}$$

and we know that g with respect to the basis $\widetilde{\mathcal{B}}$ has the standard form $g(v, w) = \tilde{v}^i \tilde{w}^i$ and we want to verify (4.6) using the matrix of the change of coordinates $L^{-1} = \Lambda$. If

$$[v]_{\mathcal{B}} = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} \quad \text{and} \quad [w]_{\mathcal{B}} = \begin{pmatrix} w^1 \\ w^2 \\ w^3 \end{pmatrix}$$

then we have that

$$\begin{pmatrix} \tilde{v}^1 \\ \tilde{v}^2 \\ \tilde{v}^3 \end{pmatrix} = \Lambda \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = \begin{pmatrix} v^1 - v^2 \\ v^2 - v^3 \\ v^3 \end{pmatrix}$$

and

$$\begin{pmatrix} \tilde{w}^1 \\ \tilde{w}^2 \\ \tilde{w}^3 \end{pmatrix} = \Lambda \begin{pmatrix} w^1 \\ w^2 \\ w^3 \end{pmatrix} = \begin{pmatrix} w^1 - w^2 \\ w^2 - w^3 \\ w^3 \end{pmatrix}$$

It follows that

$$\begin{aligned} g(v, w) = \tilde{v}^i \tilde{w}^i &= (v^1 - v^2)(w^1 - w^2) + (v^2 - v^3)(w^2 - w^3) + v^3 w^3 \\ &= v^1 w^1 - v^1 w^2 - v^2 w^1 + 2v^2 w^2 - v^2 w^3 - w^3 v^2 + 2v^3 w^3 \,. \end{aligned}$$

Exercise 4.14 With respect to $\widetilde{\mathcal{B}}$, we have

$$\begin{aligned} \|v\| &= (1^2 + 1^2 + 1^2)^{1/2} = \sqrt{3} \\ \|w\| &= ((-1)^2 + (-1)^2 + 3^2)^{1/2} = \sqrt{11} \end{aligned}$$

and with respect to $\mathcal{E}$

$$\begin{aligned} \|v\| &= (3 \cdot 3 - 3 \cdot 2 - 2 \cdot 3 + 2 \cdot 2 \cdot 2 - 2 \cdot 1 - 1 \cdot 2 + 2 \cdot 1 \cdot 1)^{1/2} = \sqrt{3} \\ \|w\| &= (1 \cdot 1 - 1 \cdot 2 - 2 \cdot 1 + 2 \cdot 2 \cdot 2 - 2 \cdot 3 - 3 \cdot 2 + 2 \cdot 3 \cdot 3)^{1/2} = \sqrt{11} \,. \end{aligned}$$

Exercise: 4.25. Saying that the orthogonality is meant with respect to g, means that we have to show that $g(v - \text{proj}_{b_k} v, b_k) = 0$. In fact,

$$g(v - \text{proj}_{b_k} v, b_k) = g(v - \frac{g(v, b_k)}{g(b_k, b_k)} b_k, b_k) = g(v, b_k) - \frac{g(v, b_k)}{\cancel{g(b_k, b_k)}} \cancel{g(b_k, b_k)} = 0$$