

# Chapter 13

## Quadratic Forms



### 13.1 Quadratic Forms on Real Vector Spaces

In Sect. 3.1 the notion of scalar product on a finite dimensional real vector space has been introduced as a bilinear symmetric map  $\cdot : V \times V \rightarrow \mathbb{R}$  with additional properties. Such additional properties are that  $v \cdot v \geq 0$  for  $v \in V$ , with  $v \cdot v = 0 \Leftrightarrow v = 0_V$ . This is referred to as positive definiteness.

We start by introducing the more general notion of *quadratic form*.

**Definition 13.1.1** Let  $V$  be a finite dimensional real vector space. A *quadratic form* on  $V$  is a map

$$Q : V \times V \longrightarrow \mathbb{R} \quad (v, w) \mapsto Q(v, w)$$

that fulfils the following properties. For any  $v, w, v_1, v_2 \in V$  and  $a_1, a_2 \in \mathbb{R}$  it holds that:

- (Q1)  $Q(v, w) = Q(w, v)$ ,
- (Q2)  $Q((a_1 v_1 + a_2 v_2), w) = a_1 Q(v_1, w) + a_2 Q(v_2, w)$ .

When a quadratic form is positive definite, that is for any  $v \in V$  the additional conditions

- (E1)  $Q(v, v) \geq 0$ ;
- (E2)  $Q(v, v) = 0 \Leftrightarrow v = 0_V$ .

are satisfied, then  $Q$  is a scalar product, and we say that  $V$  is an euclidean space.

With respect to a basis  $\mathcal{B} = (u_1, \dots, u_n)$  for  $V$ , the conditions Q1 and Q2 are clearly satisfied if and only if there exists a symmetric matrix  $F = (F_{ab}) \in \mathbb{R}^{n,n}$  such that

$$Q(v, w) = Q((v_1, \dots, v_n)_{\mathcal{B}}, (w_1, \dots, w_n)_{\mathcal{B}}) = \sum_{a,b=1}^n F_{ab} v_a w_b.$$

This expression can be also written as

$$\mathcal{Q}(v, w) = (v_1 \cdots v_n) \begin{pmatrix} F_{11} & \cdots & F_{1n} \\ \vdots & & \vdots \\ F_{n1} & \cdots & F_{nn} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

Not surprisingly, the matrix representing the action of the quadratic form  $\mathcal{Q}$  depends on the basis considered in  $V$ . Under a change of basis  $\mathcal{B} \rightarrow \mathcal{B}'$  with  $\mathcal{B}' = (u'_1, \dots, u'_n)$  and corresponding matrix  $M^{\mathcal{B}', \mathcal{B}}$ , as we know, the components of the vectors  $v, w$  are transformed as

$$\begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = M^{\mathcal{B}', \mathcal{B}} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

and analogously for  $w$ . So we write the action of the quadratic form  $\mathcal{Q}$  as

$$\mathcal{Q}(v, w) = (v'_1 \cdots v'_n) ({}^t M^{\mathcal{B}', \mathcal{B}} F M^{\mathcal{B}', \mathcal{B}}) \begin{pmatrix} w'_1 \\ \vdots \\ w'_n \end{pmatrix}.$$

If we write the dependence on the basis as  $\mathcal{Q} \rightarrow F^{\mathcal{B}}$ , we have then shown the following result.

**Proposition 13.1.2** *Given a quadratic form  $\mathcal{Q}$  on the finite dimensional real vector space  $V$ , with  $F^{\mathcal{B}}$  and  $F^{\mathcal{B}'}$  the matrices representing  $\mathcal{Q}$  on  $V$  with respect to the bases  $\mathcal{B}$  and  $\mathcal{B}'$ , it holds that*

$$F^{\mathcal{B}'} = {}^t M^{\mathcal{B}', \mathcal{B}} F^{\mathcal{B}} M^{\mathcal{B}', \mathcal{B}}.$$

**Corollary 13.1.3** *Since the matrix  $F^{\mathcal{B}}$  associated with the quadratic form  $\mathcal{Q}$  on  $V$  for the basis  $\mathcal{B}$  is symmetric, it is evident from the Proposition 4.1.20 that the matrix  $F^{\mathcal{B}'}$  associated with  $\mathcal{Q}$  with respect to any other basis  $\mathcal{B}'$  is symmetric as well.*

The Proposition 13.1.2 is the counterpart of the Proposition 7.9.9 which related the matrices of a linear maps in different bases. This transformation is not the same as the one for the matrix of an endomorphism as described at the beginning of Chap. 9. To parallel the definition there, one is led to the following definition.

**Definition 13.1.4** The symmetric matrices  $A, B \in \mathbb{R}^{n,n}$  are called *quadratically equivalent* (or simply *equivalent*) if there exists a matrix  $P \in \text{GL}(n)$ , such that  $B = {}^t P A P$ . Analogously, the quadratic forms  $\mathcal{Q}$  and  $\mathcal{Q}'$  defined on a real finite dimensional vector space  $V$  are called *equivalent* if their representing matrices are (quadratically) equivalent.

**Exercise 13.1.5** Let us consider the symmetric matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}.$$

They are *not* similar, since for example  $\det(A) = 2 \neq \det(B) = 3$  (recall that if two matrices are similar, then their determinants must coincide, from the Binet Theorem 5.1.16). They are indeed *quadratically equivalent*: the matrix

$$P = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{\frac{3}{2}} \end{pmatrix}$$

gives  ${}^tPAP = B$ .

In parallel with the Remark 9.1.4 concerning similarity of matrices, it is easy to show that the quadratic equivalence is an equivalence relation within the collection of symmetric matrices in  $\mathbb{R}^{n,n}$ . It is then natural to look for a canonical representative in any equivalence class.

**Proposition 13.1.6** *Any quadratic form  $\mathcal{Q}$  is equivalent to a diagonal quadratic form, that is one whose representing matrix is diagonal.*

*Proof* This is just a consequence of the fact that symmetric matrices are orthogonally diagonalisable. From the Proposition 10.5.1 we know that for any symmetric matrix  $A \in \mathbb{R}^{n,n}$  there exists a matrix  $P$  which is orthogonal, that is  $P^{-1} = {}^tP$ , such that

$${}^tPAP = \Delta_A$$

where  $\Delta_A$  is a diagonal matrix whose entries are the eigenvalues of  $A$ . □

Without any further requirements on the quadratic form, the matrix  $\Delta_A$  may have a number  $\mu$  of positive eigenvalues, a number  $\nu$  of negative eigenvalues, and also the zero eigenvalue with multiplicity  $m_0 = m_{\lambda=0}$ . We can order the eigenvalues as follows

$$\Delta_A = \text{diag}(\lambda_{p_1}, \dots, \lambda_{p_\mu}, \lambda_{n_1}, \dots, \lambda_{n_\nu}, 0, \dots, 0)$$

As in the Exercise 13.1.5, we know that the diagonal matrix

$$Q = \text{diag}\left(\frac{1}{\sqrt{\lambda_{p_1}}}, \dots, \frac{1}{\sqrt{\lambda_{p_\mu}}}, \frac{1}{\sqrt{|\lambda_{n_1}|}}, \dots, \frac{1}{\sqrt{|\lambda_{n_\nu}|}}, 1, \dots, 1\right)$$

is such that

$${}^tQ\Delta_A Q = \text{diag}(1, \dots, 1, -1, \dots, -1, 0, \dots, 0) = \mathcal{D}_A$$

with the expected multiplicities  $\mu$  for  $+1$ ,  $\nu$  for  $-1$  and  $m_0$  for  $0$ . Since we are considering only transformations between real basis, these multiplicities are constant in each equivalence class of symmetric matrices.

For quadratic forms, this means that any quadratic form  $\mathcal{Q}$  on  $V$  is equivalent to a diagonal one whose diagonal matrix has a number of  $\mu$  times  $+1$ , a number of  $\nu$  times  $-1$  and a number of  $m_0 = \dim(V) - \mu - \nu$  times  $0$ . The multiplicities  $\mu$  and  $\nu$  depend only on the equivalence class. Equivalently, for a quadratic form  $\mathcal{Q}$  on  $V$ , there is a basis for  $V$  with respect to which the matrix representing  $\mathcal{Q}$  is diagonal, with diagonal entries given  $+1$  repeated  $\mu$  times,  $-1$  repeated  $\nu$  times and  $m_0$  multiplicity of  $0$ .

**Definition 13.1.7** Given a symmetric matrix  $A$  on  $\mathbb{R}^{n,n}$ , we call  $\mathcal{D}_A$  its *canonical form* (or *reduced form*). If  $\mathcal{Q}$  is a quadratic form on  $\mathbb{R}^n$  whose matrix  $F^{\mathcal{B}}$  is canonical, then one has

$$\mathcal{Q}(v, w) = v_{p_1}w_{p_1} + \cdots + v_{p_\mu}w_{p_\mu} - (v_{n_1}w_{n_1} + \cdots + v_{n_\nu}w_{n_\nu})$$

with  $v = (v_{p_1}, \dots, v_{p_\mu}, v_{n_1}, \dots, v_{n_\nu}, \tilde{v}_1, \dots, \tilde{v}_{m_0})$  and analogously for  $w$ . This is the *canonical form* for the quadratic form  $\mathcal{Q}$ . The triple  $\text{sign}(\mathcal{Q}) = (\mu, \nu, m_0)$  is called the *signature* of the quadratic form  $\mathcal{Q}$ . In particular, the quadratic form  $\mathcal{Q}$  is called *positive definite* if  $\text{sign}(\mathcal{Q}) = (\mu = n, 0, 0)$ , and *negative definite* if  $\text{sign}(\mathcal{Q}) = (0, \nu = n, 0)$ .

**Exercise 13.1.8** On  $V = \mathbb{R}^3$  consider the quadratic form

$$\mathcal{Q}(v, w) = v_1w_2 + v_2w_1 + v_1w_3 + v_3w_1 + v_2w_3 + v_3w_2$$

where  $v = (v_1, v_2, v_3)_B$  and  $w = (w_1, w_2, w_3)_B$  with respect to a given basis  $(u_1, u_2, u_3)$ . Its action is represented by the matrix

$$F^{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

To diagonalise it, we compute its eigenvalues from the characteristic polynomial,

$$p_{F^{\mathcal{B}}}(T) = -T^3 + 3T + 2 = (2 - T)(1 + T)^2.$$

The eigenvalue  $\lambda = 2$  is simple, with eigenspace  $V_{\lambda=2} = \mathcal{L}((1, 1, 1))$ , while the eigenvalue  $\lambda = -1$  has multiplicity  $m_{\lambda=-1} = 2$ , with corresponding eigenspace  $V_{\lambda=-1} = \mathcal{L}((1, -1, 0), (1, 1, -2))$ . If we define

$$P = M^{\mathcal{B}, \mathcal{B}} = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & \sqrt{3} & 1 \\ \sqrt{2} & -\sqrt{3} & 1 \\ \sqrt{2} & 0 & -2 \end{pmatrix}$$

we see that

$${}^t P F^{\mathcal{B}} P = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \Delta_A = F^{\mathcal{B}'}$$

with respect to the basis  $\mathcal{B}' = (u'_1, u'_2, u'_3)$  of eigenvectors given explicitly by

$$\begin{aligned} u'_1 &= \frac{1}{\sqrt{3}}(u_1 + u_2 + u_3), \\ u'_2 &= \frac{1}{\sqrt{2}}(u_1 - u_2), \\ u'_3 &= \frac{1}{\sqrt{6}}(u_1 + u_2 - 2u_3). \end{aligned}$$

With respect to the basis  $\mathcal{B}'$  the quadratic form is written as

$$\mathcal{Q}(v, w) = 2v'_1 w'_1 - (v'_2 w'_2 + v'_3 w'_3).$$

Motivated by the Exercise 13.1.5, with the matrix

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

we have that

$${}^t Q F^{\mathcal{B}'} Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = F^{\mathcal{B}''}$$

on the basis  $\mathcal{B}'' = (u''_1 = \frac{1}{\sqrt{2}}u'_1, u''_2 = u'_2, u''_3 = u'_3)$ . With respect to  $\mathcal{B}''$  the quadratic form is

$$\mathcal{Q}(v, w) = v''_1 w''_1 - v''_2 w''_2 - v''_3 w''_3,$$

in terms of the components of  $v, w$  in the basis  $\mathcal{B}''$ . Its signature is  $\text{sign}(Q) = (1, 2, 0)$ .

**Exercise 13.1.9** On the vector space  $\mathbb{R}^4$  with canonical basis  $\mathcal{E}$ , consider the quadratic form

$$\mathcal{Q}(v, w) = u_1 w_1 + u_2 w_2 + u_1 w_2 + u_2 w_1 + u_3 w_4 + u_4 w_3 - u_3 w_3 - u_4 w_4,$$

for any two vectors  $v, w$  in  $\mathbb{R}^4$ . Its representing matrix is

$$F^{\mathcal{E}} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix},$$

which has been already studied in the Exercise 10.5.3. We can then immediately write

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{pmatrix}, \quad {}^t P F^{\mathcal{E}} P = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = F^{\mathcal{E}'},$$

with the basis  $\mathcal{E}' = (e'_1, e'_2, e'_3, e'_4)$  given by

$$\begin{aligned} e'_1 &= \frac{1}{\sqrt{2}}(e_1 - e_2), \\ e'_2 &= \frac{1}{\sqrt{2}}(e_3 + e_4), \\ e'_3 &= \frac{1}{\sqrt{2}}(e_1 + e_2), \\ e'_4 &= \frac{1}{\sqrt{2}}(e_3 - e_4). \end{aligned}$$

With respect to the basis  $\mathcal{E}'' = (e''_1 = e'_1, e''_2 = e'_2, e''_3 = \frac{1}{\sqrt{2}}e'_3, e''_4 = \frac{1}{\sqrt{2}}e'_4)$  it is clear that the matrix representing the action of  $\mathcal{Q}$  is  $F^{\mathcal{E}''} = \text{diag}(0, 0, -1, 1)$ , so that the canonical form of the quadratic form  $\mathcal{Q}$  reads

$$\mathcal{Q}(v, w) = -v''_3 w''_3 + v''_4 w''_4$$

with  $v = (v''_1, v''_2, v''_3, v''_4)_{\mathcal{E}''}$  and analogously for  $w$ . Its signature is  $\text{sign}(\mathcal{Q}) = (1, 1, 2)$

*Remark 13.1.10* Once the dimension  $n$  of the real vector space  $V$  is fixed, the collection of inequivalent quadratic forms, that is the quotient of the symmetric matrices by the quadratic equivalence relation of the Definition 13.1.7, is labelled by the possible signatures of the quadratic forms, or equivalently by the signatures of the symmetric matrices, written as  $\text{sign}(\mathcal{Q}) = (\mu, \nu, n - \mu - \nu)$ .

Finally, we state the conditions for a quadratic form to provides a scalar product for a finite dimensional real vector space  $V$ . Since we have discussed the topics at length, we omit the proof of the following proposition.

**Proposition 13.1.11** *A quadratic form  $\mathcal{Q}$  on a finite dimensional real vector space  $V$  provides a scalar product if and only if it is positive definite. In such a case we denote the scalar product by*

$$v \cdot w = \mathcal{Q}(v, w).$$

**Exercise 13.1.12** With respect to the canonical basis  $\mathcal{E}$  on  $\mathbb{R}^2$  we consider the quadratic form

$$\mathcal{Q}(v, w) = av_1 w_1 + v_1 w_2 + v_2 w_1, \quad \text{with } a \in \mathbb{R},$$

for  $v = (v_1, v_2)$  and  $w = (w_1, w_2)$ . The matrix representing  $\mathcal{Q}$  is given by

$$F^{\mathcal{E}} = \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix}$$

and its characteristic polynomial,  $p_{F^{\mathcal{E}}}(T) = T^2 - aT - 1$ , gives eigenvalues

$$\lambda_{\pm} = \frac{1}{2}(a \pm \sqrt{a^2 + 4}).$$

Since for any real value of  $a$  there is one positive eigenvalue and one negative eigenvalue, we conclude that the signature of the quadratic form is  $\text{sign}(\mathcal{Q}) = (1, 1, 0)$ .

**Exercise 13.1.13** Consider, from the Exercise 11.1.11, the three dimensional vector space  $V$  of antisymmetric matrices in  $\mathbb{R}^{3,3}$ . If we set

$$\mathcal{Q}(L, L') = -\frac{1}{2} \text{tr}(LL')$$

with  $L, L' \in V$ , it is immediate to verify that  $\mathcal{Q}$  is a quadratic form. Also, the basis elements  $L_a$  given in the Exercise 11.1.11 are orthonormal,

$$\mathcal{Q}(L_a, L_b) = \delta_{ab}.$$

Then, the space of real antisymmetric  $3 \times 3$  matrices is an euclidean space for this scalar product.

**Exercise 13.1.14** On  $\mathbb{R}^2$  again with the canonical basis, we consider the quadratic form

$$\mathcal{Q}(v, w) = v_1w_1 + v_2w_2 + a(v_1w_2 + v_2w_1), \quad \text{with } a \in \mathbb{R},$$

whose representing matrix is

$$F^{\mathcal{E}} = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}.$$

Its characteristic polynomial is  $p_{F^{\mathcal{E}}} = (1 - T)^2 - a^2 = (1 - T - a)(1 - T + a)$ , with eigenvalues

$$\lambda_{\pm} = 1 \pm a.$$

We have the following cases:

- for  $a > 1$ , it is  $\text{sign}(\mathcal{Q}) = (1, 1, 0)$ ;
- for  $a = \pm 1$ , it is  $\text{sign}(\mathcal{Q}) = (1, 0, 1)$ ;
- for  $a < -1$ , it is  $\text{sign}(\mathcal{Q}) = (1, 1, 0)$ ;
- for  $-1 < a < 1$ , it is  $\text{sign}(\mathcal{Q}) = (2, 0, 0)$ .

In this last case, the quadratic form endows  $\mathbb{R}^2$  with a scalar product. The eigenspaces are

$$\begin{aligned} \lambda_- = (1 - a), & \quad V_{\lambda_-} = \mathcal{L}((1, -1)), \\ \lambda_+ = (1 + a), & \quad V_{\lambda_+} = \mathcal{L}((1, 1)), \end{aligned}$$

so we can define the matrix

$$M^{\mathcal{E}', \mathcal{E}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

which gives

$${}^t M^{\mathcal{E}', \mathcal{E}} F^{\mathcal{E}} M^{\mathcal{E}', \mathcal{E}} = \begin{pmatrix} 1-a & 0 \\ 0 & 1+a \end{pmatrix}.$$

With respect to the basis  $\mathcal{E}' = \frac{1}{\sqrt{2}}(e'_1 = (e_1 - e_2), e'_2 = \frac{1}{\sqrt{2}}(e_1 + e_2))$  the quadratic form is

$$Q(v, w) = (1-a)v'_1 w'_1 + (1+a)v'_2 w'_2.$$

We obtain the canonical form for  $Q$  if we consider the basis  $\mathcal{E}''$  given by

$$e''_1 = \frac{1}{\sqrt{1-a}} e'_1, \quad e''_2 = \frac{1}{\sqrt{1+a}} e'_2.$$

The basis  $\mathcal{E}''$  is orthonormal with respect to the scalar product defined by  $Q$ .

**Exercise 13.1.15** This exercise puts the results of the previous one in a more general context.

(a) From Exercise 13.1.14 we know that the symmetric matrix

$$S = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix},$$

with  $a \in \mathbb{R}$ , is quadratically equivalent to the diagonal matrix

$$S' = \begin{pmatrix} 1-a & 0 \\ 0 & 1+a \end{pmatrix}.$$

Let us consider  $S$  and  $S'$  as matrices in  $\mathbb{C}^{2,2}$  with real entries (recall that  $\mathbb{R}$  is a subfield of  $\mathbb{C}$ ). We can then write

$$\begin{pmatrix} (1-a)^{-1/2} & 0 \\ 0 & (1+a)^{-1/2} \end{pmatrix} \begin{pmatrix} 1-a & 0 \\ 0 & 1+a \end{pmatrix} \begin{pmatrix} (1-a)^{-1/2} & 0 \\ 0 & (1+a)^{-1/2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

for any  $a \in \mathbb{R}$ . This means that, by complexifying the entries of the real symmetric matrix  $S$ , there exists a transformation

$$S \mapsto {}^t P S P = I_2$$

with  $P \in \text{GL}(n, \mathbb{C})$  (the group of invertible  $n \times n$  complex matrices), which transforms  $S$  to  $I_2$ .

(b) From the Exercise 13.1.8 we know that the symmetric matrix

$$S = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

is quadratically equivalent to

$$S' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

By again considering them as complex matrices, we can write

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & i \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & i \end{pmatrix}.$$

Thus,  $S$  is quadratically equivalent to  $I_3$  via an invertible matrix  $P \in \mathbb{C}^{n,n}$ .

If  $A$  is a symmetric matrix with real entries, from the Proposition 13.1.6 we know that it is quadratically equivalent to

$$\Delta_A = \text{diag}(\lambda_{p_1}, \dots, \lambda_{p_\mu}, \lambda_{n_1}, \dots, \lambda_{n_\nu}, 0, \dots, 0),$$

with  $\lambda_{p_j} > 0$  and  $\lambda_{n_j} < 0$ . Given the invertible matrix

$$P = \text{diag}\left(\frac{1}{\sqrt{\lambda_{p_1}}}, \dots, \frac{1}{\sqrt{\lambda_{p_\mu}}}, \frac{i}{\sqrt{|\lambda_{n_1}|}}, \dots, \frac{i}{\sqrt{|\lambda_{n_\nu}|}}, 1, \dots, 1\right)$$

in  $\mathbb{C}^{n,n}$ , one finds that

$${}^t P \Delta_A P = \text{diag}(1, \dots, 1, 1, \dots, 1, 0, \dots, 0) = \tilde{D}_A,$$

where the number of non zero terms  $+1$  is given by the rank of  $A$ .

If we now define that two symmetric matrices  $A, B \in \mathbb{C}^{n,n}$  are quadratically equivalent if there exists a matrix  $P \in \text{GL}(n, \mathbb{C})$  such that  $B = {}^t P A P$ , we can conclude that any real symmetric matrix  $A$  is quadratically equivalent to a diagonal matrix  $\tilde{D}_A$  as above.

The diagonal matrix  $\tilde{D}_A$  above gives a canonical form for  $A$  with respect to quadratic equivalence *after complexification*. Notice that, since  $(iI_n)A(iI_n) = -A$ , we have that  $A$  is quadratically equivalent to  $-A$ . This means that a notion of *complex* signature does not carry much information since it cannot measure the signs of the eigenvalues of  $A$ , but only its rank. If  $A = {}^t A = \bar{A}$ , then we set  $\text{sign}(A) = (\text{rk}(A), \dim \ker(A))$ .

We conclude by observing that what we have sketched above gives the main properties of a *real quadratic form on a complex* finite dimensional vector space, whose definition is as follows.

**Definition 13.1.16** A *real quadratic form* on a *complex* finite dimensional vector spaces is a map

$$\mathcal{S} : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}, \quad (v, w) \mapsto \mathcal{S}(v, w)$$

such that, for any  $v, w, v_1, v_2 \in \mathbb{C}^n$  and  $a_1, a_2 \in \mathbb{C}$  it holds that:

- (S1)  $\mathcal{S}(v, w) = \mathcal{S}(w, v)$ ,
- (S2)  $\mathcal{S}(v, w) \in \mathbb{R}$  if and only if  $v = \bar{v}$  and  $w = \bar{w}$ ,
- (S3)  $\mathcal{S}((a_1v_1 + a_2v_2), w) = a_1\mathcal{S}(v_1, w) + a_2\mathcal{S}(v_2, w)$ .

It is clear that  $\mathcal{S}$  is a real quadratic form on  $\mathbb{C}^n$  if and only if there exists a real basis  $\mathcal{B}$  for  $\mathbb{C}^n$ , that is a basis which is invariant under complex conjugation, with respect to which the matrix  $S^{\mathcal{B}} \in \mathbb{C}^{n,n}$  representing  $\mathcal{S}$  is symmetric with real entries.

In order to have a more elaborate notion of signature for a bilinear form on complex vector spaces, one needs the notion of hermitian form as explained in the next section.

## 13.2 Quadratic Forms on Complex Vector Spaces

It is straightforward to generalise to  $\mathbb{C}^n$  the main results of the theory of quadratic forms on  $\mathbb{R}^n$ . The following definition comes naturally after Sects. 3.4 and 8.2.

**Definition 13.2.1** Let  $V$  be a finite dimensional complex vector space. A *hermitian form* on  $V$  is a map

$$\mathcal{H} : V \times V \longrightarrow \mathbb{C}, \quad (v, w) \mapsto \mathcal{H}(v, w)$$

that fulfils the following properties. For any  $v, w, v_1, v_2 \in V$  and  $a_1, a_2 \in \mathbb{C}$  it holds that:

- (H1)  $\mathcal{H}(v, w) = \overline{\mathcal{H}(w, v)}$ ,
- (H2)  $\mathcal{H}((a_1v_1 + a_2v_2), w) = \bar{a}_1\mathcal{H}(v_1, w) + \bar{a}_2\mathcal{H}(v_2, w)$ .

When a hermitian form is positive definite, that is for any  $v \in V$  the additional conditions

- (E1)  $\mathcal{H}(v, v) \geq 0$ ;
- (E2)  $\mathcal{H}(v, v) = 0 \iff v = 0_V$ .

are satisfied, then  $\mathcal{H}$  is a hermitian product, and we say that  $V$  is a hermitian space.

We list the properties of hermitian forms in parallel with those of the real case.

- (a) With respect to any given basis  $\mathcal{B} = (u_1, \dots, u_n)$  of  $V$ , the conditions H1 and H2 are satisfied if and only if there exists a selfadjoint matrix  $H = (H_{ab}) \in \mathbb{C}^{n,n}$ ,  $H = H^\dagger$ , such that

$$\mathcal{H}(v, w) = \sum_{a,b=1}^n H_{ab} \bar{v}_a w_b.$$

If we denote by  $H^\mathcal{B}$  the dependence on the basis of  $V$  for the matrix giving the action of  $\mathcal{H}$ , under a change of bases  $\mathcal{B} \rightarrow \mathcal{B}'$  we have

$$H^{\mathcal{B}'} = (M^{\mathcal{B}',\mathcal{B}})^\dagger H^\mathcal{B} M^{\mathcal{B},\mathcal{B}'} = (H^{\mathcal{B}'})^\dagger. \tag{13.1}$$

- (b) Two selfadjoint matrices  $A, B \in \mathbb{C}^{n,n}$  are defined to be *equivalent* if there exists an invertible matrix  $P$  such that  $B = P^\dagger A P$ . This is an equivalence relation within the set of selfadjoint matrices. Analogously, two hermitian forms  $\mathcal{H}$  and  $\mathcal{H}'$  on  $\mathbb{C}^n$  are defined to be equivalent if their representing matrices are equivalent.
- (c) From the spectral theory for selfadjoint matrices it is clear that any hermitian form  $\mathcal{H}$  is equivalent to a hermitian form whose representing matrix is diagonal. Referring to the relation (13.1), there exists a unitary matrix  $U = M^{\mathcal{B}',\mathcal{B}}$  of the change of basis from  $\mathcal{B}$  to  $\mathcal{B}'$  such that  $H^{\mathcal{B}'} = \text{diag}(\lambda_1, \dots, \lambda_n)$ , with  $\lambda_j \in \mathbb{R}$  giving the spectrum of  $H^\mathcal{B}$ .
- (d) The matrix  $H^{\mathcal{B}'}$  is further reduced to its *canonical form* via the same conjugation operation described for the real case after the Proposition 13.1.6.

Since, as in real case, no conjugation as in (13.1) can alter the signs of the eigenvalues of a given selfadjoint matrix, the notion of signature is meaningful for hermitian forms. Such a signature characterises equivalence classes of selfadjoint matrices (and then of hermitian forms) via the equivalence relation we are considering.

- (e) A hermitian form  $\mathcal{H}$  equips  $\mathbb{C}^n$  with a hermitian product if and only if it is positive definite.

**Exercise 13.2.2** On  $\mathbb{C}^2$  we consider the basis  $\mathcal{B} = (u_1, u_2)$  and the hermitian form

$$\mathcal{H}(v, w) = a(v_1 w_1 + v_2 w_2) + i b(v_1 w_2 - v_2 w_1), \quad \text{with } a, b \in \mathbb{R}$$

for  $v = (v_1, v_2)_\mathcal{B}$  and  $w = (w_1, w_2)_\mathcal{B}$ . The hermitian form is represented by the matrix

$$H^\mathcal{B} = \begin{pmatrix} a & i b \\ -i b & a \end{pmatrix} = (H^\mathcal{B})^\dagger.$$

The spectral resolution of this matrix gives

$$\lambda_\pm = a \pm b, \quad V_{\lambda_\pm} = \mathcal{L}(u_\pm)$$

with normalised eigenvectors

$$u_{\pm} = \frac{1}{\sqrt{2}} (\pm i, 1)_{\mathcal{B}},$$

and with respect to the basis  $\mathcal{B}' = (b'_1 = u_+, b'_2 = u_-)$  one finds

$$H^{\mathcal{B}'} = \begin{pmatrix} a+b & 0 \\ a-b & 0 \end{pmatrix}.$$

We reduce the hermitian form  $\mathcal{H}$  to its canonical form by defining a basis

$$\mathcal{B}'' = \left( \frac{1}{\sqrt{|a+b|}} b'_1, \frac{1}{\sqrt{|a-b|}} b'_2 \right)$$

so to have

$$M^{\mathcal{B}''} = \begin{pmatrix} \frac{a+b}{|a+b|} & 0 \\ 0 & \frac{a-b}{|a-b|} \end{pmatrix}.$$

We see that the signature of  $\mathcal{H}$  depends on the relative moduli of  $a$  and  $b$ . It endows  $\mathbb{C}^2$  with a hermitian product if and only if  $|a| > |b|$ , with  $\mathcal{B}''$  giving an orthonormal basis for it.

### 13.3 The Minkowski Spacetime

We now describe the quadratic form used for a geometrical description of the electromagnetism and for the special theory of relativity.

Let  $V$  be a four dimensional real vector space equipped with a quadratic form  $\mathcal{Q}$  with signature  $\text{sign}(\mathcal{Q}) = (3, 1, 0)$ . From the theory we have developed in Sect. 13.1 there exists a (canonical) basis  $\mathcal{E} = (e_0, e_1, e_2, e_3)$  with respect to which the action of  $\mathcal{Q}$  is given by<sup>1</sup>

$$\mathcal{Q}(v, w) = -v_0 w_0 + v_1 w_1 + v_2 w_2 + v_3 w_3$$

with  $v = (v_0, v_1, v_2, v_3)$  and  $w = (w_0, w_1, w_2, w_3)$ .

**Definition 13.3.1** The equivalence class of quadratic forms on  $\mathbb{R}^4$  characterised by the signature  $(3, 1, 0)$  is said to provide  $\mathbb{R}^4$  a *Minkowski* quadratic form, that we denote by  $\eta$ . The datum  $(\mathbb{R}^4, \eta)$  is called the Minkowski spacetime, using the name from physics. We shall denote it by  $M^4$  and with a slight abuse of terminology, we shall also denote the action of  $\eta$  as a scalar product

$$v \cdot w = \eta(v, w)$$

and refer to it as the (*Minkowski*) scalar product in  $M^4$ .

---

<sup>1</sup>The reason why we denote the first element by  $e_0$  and the corresponding component of a vector  $v$  by  $v_0$  comes from physics, since such components is identified with the *time* coordinate of an event.

**Definition 13.3.2** We list the natural generalisations to  $M^4$  of well known definitions in  $E^n$ .

- (a) For any  $v \in (\mathbb{R}^4, \eta)$ , the quantity  $\|v\|^2 = v \cdot v$  is the square of the (Minkowski) norm of  $v \in \mathbb{R}^4$ ;

the vector  $v$  is called *space-like* if  $\|v\|^2 > 0$ ,

the vector  $v$  is called *light-like* if  $\|v\| = 0$ ,

the vector  $v$  is called *time-like* if  $\|v\|^2 < 0$ .

- (b) Two vectors  $v, w \in M^4$  are orthogonal if  $v \cdot w = 0$ ; thus a light-like vector is orthogonal to itself.
- (c) A basis  $\mathcal{B}$  for  $\mathbb{R}^4$  is *orthonormal* if the action of  $\eta$  with respect to  $\mathcal{B}$  is diagonal, that is if and only if the matrix  $\eta^{\mathcal{B}}$  has the form

$$\eta^{\mathcal{B}} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We simply denote  $\eta_{\mu\nu} = (\eta^{\mathcal{B}})_{\mu\nu}$  with  $\mathcal{B}$  orthonormal.

- (d) A matrix  $A \in \mathbb{R}^{4,4}$  is a *Lorentz matrix* if its columns yield an orthonormal basis for  $M^4$ .

We omit the proof of the following results, which generalise to  $M^4$  analogous results valid in  $E^n$ .

**Proposition 13.3.3** Let  $\mathcal{B} = (e_0, e_1, e_2, e_3)$  be an orthonormal basis for  $M^4$ , with  $A \in \mathbb{R}^{4,4}$  and  $\phi \in \text{End}(M^4)$ .

- (a) The matrix  $A$  is a Lorentz matrix if and only if  $A \eta A = \eta$ .
- (b) It holds that  $\phi(v) \cdot \phi(w) = v \cdot w$  for any  $v, w \in M^4$  if and only if  $M_{\phi}^{\mathcal{B}, \mathcal{B}}$  is a Lorentz matrix.
- (c) The system  $\mathcal{B}' = (\phi(e_0), \dots, \phi(e_3))$  is an orthonormal basis for  $M^4$  if and only if for any  $v, w \in M^4$  one has  $\phi(v) \cdot \phi(w) = v \cdot w$ , that is if and only if

$$\phi(e_{\mu}) \cdot \phi(e_{\nu}) = e_{\mu} \cdot e_{\nu} = \eta_{\mu\nu}.$$

As an immediate consequence of such proposition, one proves that, if  $u \in M^4$  is a space-like vector, there exists an orthonormal basis  $\mathcal{B}'$  for  $M^4$  with respect to which  $u = (0, u'_1, u'_2, u'_3)_{\mathcal{B}'}$ . Analogously, if  $u$  is a time-like vector, there exists a basis  $\mathcal{B}''$  with respect to which  $u = (u''_0, 0, 0, 0)_{\mathcal{B}''}$ .

Indeed it is straightforward to prove that the set of Lorentz matrices form a group, for matrix multiplication, denoted by  $O(3, 1)$  and called the *Lorentz group*. If the endomorphism  $\phi$  is represented, with respect to an orthonormal basis for  $M^4$  by a

Lorentz matrix, then  $\phi$  is said to be a Lorentz transformation. This means that the set of Lorentz transformations is a group isomorphic to the Lorentz group.

*Example 13.3.4* In the special theory of relativity the position of a point mass at a given time  $t$  is represented with a vector  $x = (x_0 = ct, x_1, x_2, x_3)_{\mathcal{B}}$  in  $M^4$  with respect to an orthonormal basis  $\mathcal{B}$ , with  $(x_1, x_2, x_3)$  giving the so called *spatial* components of  $x$  and  $c$  denoting the speed of light. Such a vector  $x$  is also called an *event*. The linear map

$$\begin{pmatrix} x'_0 \\ x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

with

$$\beta = v/c \quad \text{and} \quad \gamma = (1 - \beta^2)^{-1/2},$$

yields the components of the vector  $x$  with respect to an orthonormal basis  $\mathcal{B}'$  corresponding to an *inertial reference system*, (an inertial observer) which is moving with constant spatial velocity  $v$  along the direction  $e_1$ . Notice that, being  $c$  a limit value for the velocity, we have  $|\beta| < 1$  and then  $\gamma \geq 1$ . It is easy to see that this map is a Lorentz transformation, and that the matrix gives the change of basis  $M^{\mathcal{B}', \mathcal{B}}$  in  $M^4$ .

From the identity  ${}^tA \eta A = \eta$  one gets  $\det A = \pm 1$  for a Lorentz matrix  $A$ . The set of Lorentz matrices whose determinant is positive is the (sub)group  $\text{SO}(3, 1)$  of *proper* Lorentz matrices.

If  $A_{\mu\nu}$  denotes the entries of a Lorentz matrix  $A$ , then from the same identity we can write that

$$-A_{00}^2 + \sum_{k=1}^3 A_{k0}^2 = -1 \quad \text{and} \quad -A_{00}^2 + \sum_{k=1}^3 A_{0k}^2 = -1,$$

thus proving that  $A_{00}^2 \geq 1$ . Lorentz matrices with  $A_{00} > 1$  are called *ortochronous*. We omit the proof that the set of ortochronous Lorentz matrices form a group as well. Proper *and* ortochronous Lorentz matrices form therefore a group, that we denote by

$$\text{SO}(3, 1)^\uparrow = \{A \in \text{O}(3, 1) : \det A = 1, A_{00} > 1\}.$$

Notice that the Lorentz matrix given in Example 13.3.4 is proper and ortochronous. Given the physical interpretation of the components of a vector in  $M^4$  mentioned before, it is natural to call the endomorphisms represented by the Lorentz matrices

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

the (*spatial*) *parity* and the *time reversal*. The matrix  $P$  is improper and orthochronous, while  $T$  is improper and antichronous.

We can generalise the final remark from Example 11.3.1 to the Lorentz group case. If  $A$  is an improper orthochronous Lorentz matrix, then it is given by the product  $PA'$  with  $A' \in \text{SO}(3, 1)^\uparrow$ . If  $A$  is an improper antichronous Lorentz matrix, then it is given by the product  $TA'$  with  $A' \in \text{SO}(3, 1)^\uparrow$ . If  $A$  is the product  $PTA'$  with  $A' \in \text{SO}(3, 1)^\uparrow$ , it is called a proper antichronous Lorentz matrix.

Let us describe the structure of the group  $\text{SO}(3, 1)^\uparrow$  in more details. Firstly, notice that if  $R \in \text{SO}(3)$  then all matrices of the form

$$A_R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & R & & \\ 0 & & & \end{pmatrix}$$

are elements in  $\text{SO}(3, 1)^\uparrow$ . The set of such matrices  $A$  is clearly isomorphic to the group  $\text{SO}(3)$ , so we can refer to  $\text{SO}(3)$  as the subgroup of *spatial rotations* within the Lorentz group.

The Lorentz matrix in the Example 13.3.4 is not such a rotation. From the Exercise 11.2.3 we write

$$e^{uS_1} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{with} \quad S_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

with  $\sinh u = \beta\gamma$  and  $\cosh u = \gamma$  so that  $\text{tgh } u = v/c$ .

We therefore have a closer look at the exponential of symmetric matrices of the form

$$S(u) = \begin{pmatrix} 0 & u_1 & u_2 & u_3 \\ u_1 & 0 & 0 & 0 \\ u_2 & 0 & 0 & 0 \\ u_3 & 0 & 0 & 0 \end{pmatrix} = u_1S_1 + u_2S_2 + u_3S_3, \tag{13.2}$$

with  $u = (u_1, u_2, u_3)$  a triple of real parameters. If the matrix  $R = (R_{ij})$  represents a spatial rotation, a direct computation shows that

$$A_{R^{-1}} S A_R = \begin{pmatrix} 0 & \sum_{k=1}^3 R_{k1}u_k & \sum_{k=1}^3 R_{k2}u_k & \sum_{k=1}^3 R_{k3}u_k \\ \sum_{k=1}^3 R_{k1}u_k & 0 & 0 & 0 \\ \sum_{k=1}^3 R_{k2}u_k & 0 & 0 & 0 \\ \sum_{k=1}^3 R_{k3}u_k & 0 & 0 & 0 \end{pmatrix} :$$

We see that  $u = (u_1, u_2, u_3)$  transforms like a vector in a three dimensional euclidean space, and therefore we write the identity above as

$$S(R^{-1}u) = A_{R^{-1}}S(u)A_R.$$

This identity allows us to write (see the Proposition 11.2.2)

$$e^{S(u)} = A_{R^{-1}}e^{S(Ru)}A_R.$$

If  $R$  is a proper rotation mapping  $u \mapsto (\|u\|_E, 0, 0)$ , with  $\|u\|_E^2 = u_1^2 + u_2^2 + u_3^2$  the square of the euclidean three-norm, we get

$$e^{S(u)} = A_{R^{-1}}e^{(\|u\|_E S_1)}A_R.$$

Alternatively, one shows by direct computations that

$$S^2(u) = \begin{pmatrix} \|u\|_E^2 & 0 & 0 & 0 \\ 0 & & & \\ 0 & Q & & \\ 0 & & & \end{pmatrix} \quad \text{and} \quad S^3(u) = \|u\|_E^2 S(u)$$

$$\Rightarrow S^{2k}(u) = \|u\|_E^{2(k-1)} S(u), \quad S^{2k+1}(u) = \|u\|_E^{2k} S(u),$$

where  $Q \in \mathbb{R}^{3,3}$  has entries  $Q_{ij} = u_i u_j$ , so that  $Q^2 = \|u\|_E^2 Q$ . These identities give then

$$e^{S(u)} = 1 + \frac{1}{\|u\|_E^2} (\cosh \|u\|_E^2 - 1) S^2(u) + \frac{1}{\|u\|_E} \sinh \|u\|_E S(u).$$

It is easy to show that  $e^{S(u)} \in \text{SO}(3, 1)^\uparrow$ . Such transformations are called *Lorentz boosts*, or hyperbolic rotations. They give the matrices of change of bases  $M^{\mathcal{B}, \mathcal{B}'}$  where  $\mathcal{B}'$  is the orthonormal basis corresponding to an inertial reference system moving with constant velocity  $v = (v_1, v_2, v_3)$  in the physical euclidean three dimensional space with respect to the reference system represented by  $\mathcal{B}$ , by identifying for the velocity,

$$c (\text{tgh } \|u\|_E) = \|v\|_E.$$

From the properties of the group  $\text{SO}(3)$  we know that each proper spatial rotation is the exponential of a suitable antisymmetric matrix, that is  $A_R = e^{\tilde{L}}$  where  $\tilde{L}$  is an element in the three dimensional vector space spanned by the matrices  $\tilde{L}_j \subset \mathbb{R}^{4,4}$  of the form

$$\tilde{L}_j = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & & & \\ 0 & L_j & & \\ 0 & & & \end{pmatrix}$$

with the antisymmetric matrices  $L_j$ ,  $j = 1, 2, 3$ , those of the Exercise 11.1.10, the generators of the Lie algebra  $\mathfrak{so}(3)$ . With the symmetric matrices  $S_j$  in (13.2), we compute the commutators to be

$$\begin{aligned}
[\tilde{L}_i, \tilde{L}_j] &= \sum_{i,j=1}^3 \varepsilon_{ijk} \tilde{L}_k \\
[S_i, S_j] &= - \sum_{i,j=1}^3 \varepsilon_{ijk} \tilde{L}_k \\
[S_i, \tilde{L}_j] &= \sum_{i,j=1}^3 \varepsilon_{ijk} S_k,
\end{aligned}$$

thus proving that the six dimensional vector space  $\mathcal{L}(\tilde{L}_1, \tilde{L}_2, \tilde{L}_3, S_1, S_2, S_3)$  is a matrix Lie algebra (see the Definition 11.1.7) which is denoted  $\mathfrak{so}(3, 1)$ . What we have discussed gives the proof of the first part of the following proposition, which is analogous of the Proposition 11.2.6.

**Proposition 13.3.5** *If  $M$  is a matrix in  $\mathfrak{so}(3, 1)$ , then  $e^M \in \text{SO}(3, 1)^\uparrow$ . When restricted to  $\mathfrak{so}(3, 1)$ , the exponential map is surjective onto  $\text{SO}(3, 1)^\uparrow$ .*

This means that the group of proper and orthochronous Lorentz matrices is given by spatial rotations, hyperbolic rotations (that is boosts) and their products.

## 13.4 Electro-Magnetism

By recalling the framework of Sect. 1.4, in the standard euclidean formulation on the space  $E^3$  representing the physical space  $\mathcal{S}$  (and with an orthonormal basis) one describes the three dimensional electric  $\mathbf{E}(t, \mathbf{x})$  field and the magnetic field  $\mathbf{B}(t, \mathbf{x})$  as depending on both the three dimensional position vector  $\mathbf{x} = (x_1, x_2, x_3)$  and the time coordinate  $t$ . In this section we show that the Maxwell equations for electro-magnetism can be naturally formulated in terms of the geometry of the Minkowski space  $M^4$ .

*Example 13.4.1* The Maxwell equations in vacuum for the pair  $(\mathbf{E}(t, \mathbf{x}), \mathbf{B}(t, \mathbf{x}))$  are written as

$$\begin{aligned}
\text{div } \mathbf{B} &= 0, & \text{rot } \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0 \\
\text{div } \mathbf{E} &= \frac{\rho}{\varepsilon_0}, & \text{rot } \mathbf{B} &= \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}
\end{aligned}$$

where  $\varepsilon_0$  and  $\mu_0$  are the vacuum permittivity and permeability, with  $c^2 \varepsilon_0 \mu_0 = 1$ . The *sources* of the fields are the electric charge density  $\rho$  (a scalar field) and the current density  $\mathbf{J}$  (a vector field).

For the homogeneous Maxwell equations (the first two) the vector fields  $\mathbf{E}$  and  $\mathbf{B}$  can be written in terms of a vector potential  $\mathbf{A}(t, \mathbf{x}) = (A_1(t, \mathbf{x}), A_2(t, \mathbf{x}), A_3(t, \mathbf{x}))$  and a scalar potential  $\phi(t, \mathbf{x})$ , as

$$\mathbf{B} = \text{rot } \mathbf{A}, \quad \mathbf{E} = -\text{grad } \phi - \frac{\partial \mathbf{A}}{\partial t},$$

that makes the homogeneous equations automatically satisfied from the identity  $\text{div}(\text{rot}) = 0$  and  $\text{rot}(\text{grad}) = 0$ , in Exercise 1.4.1. If the potentials satisfy the so called *Lorentz gauge condition*

$$\text{div } \mathbf{A} + \frac{1}{c^2} \frac{\partial \phi}{\partial t} = 0,$$

the two Maxwell equations depending on the sources can be written as

$$\begin{aligned} \nabla^2 A_j - \frac{1}{c^2} \frac{\partial^2 A_j}{\partial t^2} &= -\mu_0 J_j, \quad \text{for } j = 1, 2, 3, \\ \nabla^2 \phi - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} &= -\frac{\rho}{\varepsilon_0} \end{aligned}$$

where  $\nabla^2 = \sum_{k=1}^3 \partial_k^2$  is the spatial Laplacian operator with  $\partial_k = \partial/\partial x_k$ .

If we define the four-potential as  $A = (A_0 = -\frac{\phi}{c}, \mathbf{A})$ , then the Lorentz gauge condition is written as (recall the Definition 13.3.2 for the metric  $\eta_{\mu\nu}$ )

$$\sum_{\mu, \nu=0}^3 \eta_{\mu\nu} \partial_\mu A_\nu = 0,$$

where we also define  $\partial_0 = \partial/\partial x_0 = \partial/c\partial t$ . In terms of the four-current  $J = (J_0 = -\rho/c\varepsilon_0, \mu_0 \mathbf{J})$ , the inhomogeneous Maxwell equations are written as

$$\sum_{\mu, \nu}^3 \eta_{\mu\nu} \partial_\mu \partial_\nu A_\rho = -J_\rho.$$

Using the four-dimensional ‘nabla’ operator  $\nabla = (\partial_0, \partial_1, \partial_2, \partial_3)$  we can then write the Lorentz gauge condition as

$$\nabla \cdot A = \sum_{\mu, \nu=0}^3 \eta_{\mu\nu} \partial_\mu A_\nu = 0,$$

and the inhomogeneous Maxwell equations as

$$\nabla^2 A_\rho = \sum_{\mu, \nu=0}^3 \eta_{\mu\nu} \partial_\mu \partial_\nu A_\rho = -J_\rho, \quad \text{for } \rho = 0, 1, 2, 3,$$

thus generalising to the Minkowski spacetime the analogous operations written for the euclidean space  $E^3$  in Sect. 1.4.

*Example 13.4.2* From the relations defining the vector fields  $\mathbf{E}$  and  $\mathbf{B}$  in terms of the four-potential vector  $A$ , we can write for their components in the physical space

$$B_a = \sum_{b,c=1}^3 \varepsilon_{abc} \partial_b A_c$$

$$E_a = c(\partial_a A_0 - \partial_0 A_a)$$

for  $a = 1, 2, 3$ . This shows that the quantity

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

with  $\mu, \nu \in \{0, \dots, 3\}$ , defines the entries of the antisymmetric *field strength matrix* (or more precisely field strength ‘tensor’)  $F$  given by

$$F = (F_{\mu\nu}) = \begin{pmatrix} 0 & -E_1/c & -E_2/c & -E_3/c \\ E_1/c & 0 & B_3 & -B_2 \\ E_2/c & -B_3 & 0 & B_1 \\ E_3/c & B_2 & -B_1 & 0 \end{pmatrix}.$$

Merging the definition of  $F$  with the Lorentz gauge condition we have

$$\begin{aligned} \sum_{\mu,\nu}^3 \eta_{\mu\nu} \partial_\mu \partial_\nu A_\rho &= \sum_{\mu,\nu=0}^3 \eta_{\mu\nu} \partial_\mu (F_{\nu\rho} + \partial_\rho A_\nu) \\ &= \sum_{\mu,\nu=0}^3 \eta_{\mu\nu} \partial_\mu F_{\nu\rho} + \partial_\rho \left( \sum_{\mu,\nu=0}^3 \eta_{\mu\nu} \partial_\mu A_\nu \right) = \sum_{\mu,\nu=0}^3 \eta_{\mu\nu} \partial_\mu F_{\nu\rho} \end{aligned}$$

so we can write the inhomogeneous Maxwell equations as

$$\sum_{\mu,\nu=0}^3 \eta_{\mu\nu} \partial_\mu F_{\nu\rho} = -J_\rho \quad \text{for } \rho = 0, 1, 2, 3.$$

The homogeneous Maxwell equation can be written in a similar way by means of another useful quantity, the *dual field strength matrix* (or tensor)  $\tilde{F}_{\mu\nu}$ . For this one needs the (four dimensional) *totally antisymmetric* symbol  $\varepsilon_{a_1 a_2 a_3 a_4}$  with indices  $a_j = 0, 1, 2, 3$  and defined by

$$\varepsilon_{a_1 a_2 a_3 a_4} = \begin{cases} +1 & \text{if } (a_1, a_2, a_3, a_4) \text{ is an even permutation of } (0,1,2,3) \\ -1 & \text{if } (a_1, a_2, a_3, a_4) \text{ is an odd permutation of } (0,1,2,3) \\ 0 & \text{if any two indices are equal} \end{cases}.$$

Also, let  $\eta^{-1} = (\eta^{\mu\nu})$  be the inverse of the matrix  $\eta = (\eta_{\mu\nu})$ . The dual field strength matrix is the antisymmetric matrix defined by

$$\tilde{F} = (\tilde{F}_{\mu\nu}), \quad \tilde{F}_{\mu\nu} = \frac{1}{2} \sum_{\alpha, \beta, \gamma, \delta=0}^3 \varepsilon_{\mu\nu\gamma\delta} \eta^{\gamma\alpha} \eta^{\delta\beta} F_{\alpha\beta} = \begin{pmatrix} 0 & B_1 & B_2 & B_3 \\ -B_1 & 0 & -E_3/c & E_2/c \\ -B_2 & E_3/c & 0 & -E_1/c \\ -B_3 & -E_2/c & E_1/c & 0 \end{pmatrix}.$$

Notice that the elements of  $\tilde{F}$  are obtained from those of  $F$  by the exchange  $\mathbf{E} \leftrightarrow -c\mathbf{B}$ .

A straightforward computation shows that the homogeneous Maxwell equations can be written as

$$\sum_{\mu, \nu=0}^3 \eta_{\mu\nu} \partial_\mu \tilde{F}_{\nu\rho} = 0, \quad \text{for } \rho = 0, 1, 2, 3.$$

In terms of  $F_{\mu\nu}$  rather than  $\tilde{F}_{\mu\nu}$ , these homogeneous equations are the four equations

$$\partial_\rho F_{\mu\nu} + \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} = 0$$

for  $\mu, \nu, \rho$  any three of the integers 0, 1, 2, 3.

We now have a glimpse at the geometric nature of the four-potential  $A$  and of the antisymmetric matrix  $F$ , that is we study how they transform under a change of basis from  $\mathcal{B}$  to  $\mathcal{B}'$  for  $M^4$ . If two inertial observers (for the orthonormal bases  $\mathcal{B}$  and  $\mathcal{B}'$  for  $M^4$ ) relate their spacetime components as in the Example (13.3.4), we know from physics that for the transformed electric and magnetic fields  $\mathbf{E}'$  and  $\mathbf{B}'$  one has

$$\begin{aligned} E'_1 &= E_1, & B'_1 &= B_1 \\ E'_2 &= \gamma(E_2 - vB_3), & B'_2 &= \gamma(B_2 + (v/c^2)E_3) \\ E'_3 &= \gamma(E_3 + vB_2), & B'_3 &= \gamma(B_3 - (v/c^2)E_2) \end{aligned}$$

For the transformed potential  $A' = (A'_\rho)$  and matrix  $F' = (F'_{\mu\nu})$  with  $F'_{\mu\nu} = \partial'_\mu A'_\nu - \partial'_\nu A'_\mu$  (where  $\partial'_a = \partial/\partial x'_a$ ), one then finds

$$A' = M^{\mathcal{B}', \mathcal{B}} A$$

and

$$F' = {}^t(M^{\mathcal{B}, \mathcal{B}'}) F M^{\mathcal{B}, \mathcal{B}'}$$

It is indeed possible to check that such identities are valid for any proper and orthochronous Lorentz matrix giving the change of orthonormal basis  $\mathcal{B} \rightarrow \mathcal{B}'$ .

If we denote by  $M^{4*}$  the space dual to  $(\mathbb{R}^4, \eta)$  with  $\{\epsilon_0, \epsilon_1, \epsilon_2, \epsilon_3\}$  the basis dual to  $\mathcal{B} = (e_0, \dots, e_3)$ , the definition

$$\eta(\epsilon_a, \epsilon_b) = \eta(e_a, e_b)$$

clearly defines a Minkowski quadratic form on  $\mathbb{R}^{4*}$ , making then the space  $M^{4*}$ . Also, if  $\mathcal{B}$  is orthonormal, then  $\mathcal{B}^*$  is orthonormal as well.

Recall now the results described in Sect. 8.1 on the dual of a vector space. The previous relations, when compared with the Example 13.3.4, show that the vectors  $A = (A_0, \mathbf{A})$  is indeed an element in the dual space  $M^{4*}$  to  $M^4$  with respect to the dual basis  $\mathcal{B}^*$  to  $\mathcal{B}$ . From the Proposition 13.1.2 we see also that the matrix elements  $F$  transform as the entries of a quadratic form in  $M^{4*}$  (although  $F$  is antisymmetric). All this means that the components of the electro-magnetic fields  $\mathbf{E}$ ,  $\mathbf{B}$  are the entries of an antisymmetric matrix  $F$  which transform as a ‘contravariant’ quadratic form under (proper and orthochronous) Lorentz transformations.