Chapter 2

Duality, Bilinearity

In this chapter, the phrase "let ... be a linear space" will be used as a shorthand for "let ... be a finite-dimensional linear space over $\mathbb{R}$". (Actually, many definitions remain meaningful and many results remain valid when the given spaces are infinite-dimenisonal or when $\mathbb{R}$ is replaced by an arbitrary field. The interested reader will be able to decide for himself when this is the case.)

21 Dual Spaces, Transposition, Annihilators

Let $\mathcal{V}$ be a linear space. We write

$$\mathcal{V}^* := \mathrm{Lin}(\mathcal{V}, \mathbb{R})$$

and call the linear space $\mathcal{V}^*$ the **dual space** of the space $\mathcal{V}$. The elements of $\mathcal{V}^*$ are often called **linear forms** or *covectors*, depending on context. It is evident from Prop.7 of Sect. 17 that

$$\dim \mathcal{V}^* = \dim \mathcal{V}. \tag{21.1}$$

Let $\mathcal{V}$ and $\mathcal{W}$ be linear spaces and let $\mathbf{L} \in \mathrm{Lin}(\mathcal{V}, \mathcal{W})$ be given. It follows from Props.1 and 2 of Sect. 14 that the mapping $(\boldsymbol{\mu} \mapsto \boldsymbol{\mu}\mathbf{L}) : \mathcal{W}^* \to \mathcal{V}^*$ is linear. We call this mapping the **transpose** of $\mathbf{L}$ and denote it by $\mathbf{L}^\top$, so that

$$\mathbf{L}^\top \in \mathrm{Lin}(\mathcal{W}^*, \mathcal{V}^*) \quad \text{if} \quad \mathbf{L} \in \mathrm{Lin}(\mathcal{V}, \mathcal{W}) \tag{21.2}$$

and

$$\mathbf{L}^\top \boldsymbol{\mu} = \boldsymbol{\mu}\mathbf{L} \quad \text{for all} \quad \boldsymbol{\mu} \in \mathcal{W}^*. \tag{21.3}$$

It is an immediate consequence of Prop.3 of Sect. 14 that the mapping

$$(\mathbf{L} \mapsto \mathbf{L}^\top) : \mathrm{Lin}(\mathcal{V}, \mathcal{W}) \to \mathrm{Lin}(\mathcal{W}^*, \mathcal{V}^*) \tag{21.4}$$

is linear. This mapping is called the **transposition** on $\mathrm{Lin}(\mathcal{V}, \mathcal{W})$. The following rules follow directly from (21.3):

Proposition 1: *For every linear space $\mathcal{V}$, we have*

$$(\mathbf{1}_\mathcal{V})^\top = \mathbf{1}_{\mathcal{V}^*}. \tag{21.5}$$

Let $\mathcal{V}, \mathcal{V}', \mathcal{V}''$ be linear spaces. For all $\mathbf{L} \in \mathrm{Lin}(\mathcal{V}, \mathcal{V}')$ and all $\mathbf{M} \in \mathrm{Lin}(\mathcal{V}', \mathcal{V}'')$, we have

$$(\mathbf{ML})^\top = \mathbf{L}^\top \mathbf{M}^\top. \tag{21.6}$$

If $\mathbf{L} \in \mathrm{Lin}(\mathcal{V}, \mathcal{V}')$ is invertible, so is $\mathbf{L}^\top \in \mathrm{Lin}(\mathcal{V}'^*, \mathcal{V}^*)$, and

$$(\mathbf{L}^\top)^{-1} = (\mathbf{L}^{-1})^\top. \tag{21.7}$$

Definition: *Let $\mathcal{V}$ be a linear space and let $\mathcal{S}$ be a subset of $\mathcal{V}$. We say that a linear form $\boldsymbol{\lambda} \in \mathcal{V}^*$* **annihilates** *$\mathcal{S}$ if $\boldsymbol{\lambda}_>(\mathcal{S}) \subset \{0\}$, or, equivalently, $\boldsymbol{\lambda}|_\mathcal{S} = 0$, or, equivalently, $\mathcal{S} \subset \mathrm{Null}\, \boldsymbol{\lambda}$. The set of all linear forms that annihilate $\mathcal{S}$ is called the* **annihilator** *of $\mathcal{S}$ and is denoted by*

$$\mathcal{S}^\perp := \{\boldsymbol{\lambda} \in \mathcal{V}^* \mid \boldsymbol{\lambda}_>(\mathcal{S}) \subset \{0\}\}. \tag{21.8}$$

The following facts are immediate consequences of the definition.

Proposition 2: *$\emptyset^\perp = \{\mathbf{0}\}^\perp = \mathcal{V}^*$ and $\mathcal{V}^\perp = \{\mathbf{0}\}$. If $\mathcal{S}_1$ and $\mathcal{S}_2$ are subsets of $\mathcal{V}$ then*

$$\mathcal{S}_1 \subset \mathcal{S}_2 \implies \mathcal{S}_2^\perp \subset \mathcal{S}_1^\perp.$$

Proposition 3: *For every subset $\mathcal{S}$ of $\mathcal{V}$, $\mathcal{S}^\perp$ is a subspace of $\mathcal{V}^*$ and $(\mathrm{Lsp}\mathcal{S})^\perp = \mathcal{S}^\perp$.*

Proposition 4: *If $(\mathcal{S}_i \mid i \in I)$ is a family of subsets of $\mathcal{V}$, then*

$$(\bigcup_{i\in I} \mathcal{S}_i)^\perp = \bigcap_{i\in I} \mathcal{S}_i^\perp. \tag{21.9}$$

Combining Prop.3 and Prop.4, and using Prop.2 of Sect. 12, we obtain the following:

Proposition 5: *If $(\mathcal{U}_i \mid i \in I)$ is a family of subspaces of $\mathcal{V}$, then*

$$(\sum_{i\in I} \mathcal{U}_i)^\perp = \bigcap_{i\in I} \mathcal{U}_i^\perp. \tag{21.10}$$

In particular, if $\mathcal{U}_1, \mathcal{U}_2$ are subspaces of $\mathcal{V}$, then

$$(\mathcal{U}_1 + \mathcal{U}_2)^{\perp} = \mathcal{U}_1^{\perp} \cap \mathcal{U}_2^{\perp}. \tag{21.11}$$

The following result relates the annihilator of a subspace to the annihilator of the image of this subspace under a linear mapping.

Theorem on Annihilators and Transposes: *Let $\mathcal{V}, \mathcal{W}$ be linear spaces and let $\mathbf{L} \in \mathrm{Lin}(\mathcal{V}, \mathcal{W})$ be given. For every subspace $\mathcal{U}$ of $\mathcal{V}$, we then have*

$$(\mathbf{L}_{>}(\mathcal{U}))^{\perp} = (\mathbf{L}^{\top})^{<}(\mathcal{U}^{\perp}). \tag{21.12}$$

In particular, we have

$$(\mathrm{Rng}\,\mathbf{L})^{\perp} = \mathrm{Null}\ \mathbf{L}^{\top}. \tag{21.13}$$

Proof: Let $\boldsymbol{\mu} \in \mathcal{W}^*$ be given. Then, by (21.8) and (21.3),

$$\begin{aligned} \boldsymbol{\mu} \in (\mathbf{L}_{>}(\mathcal{U}))^{\perp} \quad &\Longleftrightarrow \quad \{0\} = \boldsymbol{\mu}_{>}(\mathbf{L}_{>}(\mathcal{U})) = (\boldsymbol{\mu}\mathbf{L})_{>}(\mathcal{U}) = (\mathbf{L}^{\top}\boldsymbol{\mu})_{>}(\mathcal{U}) \\ &\Longleftrightarrow \quad \mathbf{L}^{\top}\boldsymbol{\mu} \in \mathcal{U}^{\perp} \iff \boldsymbol{\mu} \in (\mathbf{L}^{\top})^{<}(\mathcal{U}^{\perp}). \end{aligned}$$

Since $\boldsymbol{\mu} \in \mathcal{W}^*$ was arbitrary, (21.12) follows. Putting $\mathcal{U} := \mathcal{V}$ in (21.12) yields (21.13). ∎

The following result states, among other things, that every linear form on a subspace of $\mathcal{V}$ can be extended to a linear form on all of $\mathcal{V}$.

Proposition 6: *Let $\mathcal{U}$ be a subspace of $\mathcal{V}$. The mapping*

$$(\boldsymbol{\lambda} \mapsto \boldsymbol{\lambda}|_{\mathcal{U}}) : \mathcal{V}^* \to \mathcal{U}^* \tag{21.14}$$

is linear and surjective, and its nullspace is $\mathcal{U}^{\perp}$.

Proof: It is evident that the mapping (21.14) is linear and that its nullspace is $\mathcal{U}^{\perp}$. By Prop.3 of Sect. 17, we may choose a supplement $\mathcal{U}'$ of $\mathcal{U}$ in $\mathcal{V}$. Now let $\boldsymbol{\mu} \in \mathcal{U}^*$ be given. By Prop.5 of Sect. 19 there is a $\boldsymbol{\lambda} \in \mathcal{V}^*$ such that $\boldsymbol{\lambda}|_{\mathcal{U}} = \boldsymbol{\mu}$ (and $\boldsymbol{\lambda}|_{\mathcal{U}'} = 0$, say). Since $\boldsymbol{\mu} \in \mathcal{U}^*$ was arbitrary, it follows that the mapping (21.14) is surjective. ∎

Using (21.1) and the Theorem on Dimensions of Range and Nullspace (Sect. 17), we see that Prop.6 has the following consequence:

Formula for Dimension of Annihilators: *For every subspace $\mathcal{U}$ of a given linear space $\mathcal{V}$ we have*

$$\dim \mathcal{U}^{\perp} = \dim \mathcal{V} - \dim \mathcal{U}. \tag{21.15}$$

Notes 21

(1) The notations $\mathcal{V}'$ or $\tilde{\mathcal{V}}$ are sometimes used for the dual $\mathcal{V}^*$ of a linear space $\mathcal{V}$.

(2) Some people use the term "linear functional" instead of "linear form". I prefer to reserve "linear functional" for the case when the domain is infinite-dimensional.

(3) The terms "adjoint" or "dual" are often used in place of the "transpose" of a linear mapping $\mathbf{L}$. Other notations for our $\mathbf{L}^\top$ are $\mathbf{L}^*$, $\mathbf{L}^t$, ${}^t\mathbf{L}$, and $\tilde{\mathbf{L}}$.

(4) The notation $\mathcal{S}^0$ instead of $\mathcal{S}^\perp$ is sometimes used for the annihilator of the set $\mathcal{S}$.

22 The Second Dual Space

In view of (21.1) and Corollary 2 of the Characterization of Dimension (Sect. 17), it follows that there exist linear isomorphisms from a given linear space $\mathcal{V}$ to its dual $\mathcal{V}^*$. However, if no structure on $\mathcal{V}$ other than its structure as a linear space is given, none of these isomorphisms is natural (see the Remark at the end of Sect.23). The specification of any one such isomorphism requires some capricious choice, such as the choice of a basis. By contrast, one can associate with each linear space $\mathcal{V}$ a natural isomorphism from $\mathcal{V}$ to its *second dual*, i.e. to the dual $\mathcal{V}^{**}$ of the dual $\mathcal{V}^*$ of $\mathcal{V}$. This isomorphism is an evaluation mapping as described in Sect.04.

Proposition 1: *Let $\mathcal{V}$ be a linear space. For each $\mathbf{v} \in \mathcal{V}$, the evaluation $\mathrm{ev}(\mathbf{v}) : \mathcal{V}^* \to \mathbb{R}$, defined by*

$$\mathrm{ev}(\mathbf{v})(\boldsymbol{\lambda}) := \boldsymbol{\lambda}\mathbf{v} \quad \text{for all} \quad \boldsymbol{\lambda} \in \mathcal{V}^*, \tag{22.1}$$

is linear and hence a member of $\mathcal{V}^{**}$. The evaluation mapping $\mathrm{ev} : \mathcal{V} \to \mathcal{V}^{**}$ defined in this way is a linear isomorphism.

Proof: The linearity of $\mathrm{ev}(\mathbf{v}) : \mathcal{V}^* \to \mathbb{R}$ merely reflects the fact that the linear-space operations in $\mathcal{V}^*$ are defined by value-wise applications of the operations in $\mathbb{R}$. The linearity of $\mathrm{ev} : \mathcal{V} \to \mathcal{V}^{**}$ follows from the fact that each $\boldsymbol{\lambda} \in \mathcal{V}^*$ is linear.

Put $\mathcal{U} := \mathrm{Null}\,(\mathrm{ev})$ and let $\mathbf{v} \in \mathcal{V}$ be given. By (22.1), we have

$$\mathbf{v} \in \mathcal{U} \quad \Longleftrightarrow \quad (\boldsymbol{\lambda}\mathbf{v} = 0 \quad \text{for all} \quad \boldsymbol{\lambda} \in \mathcal{V}^*),$$

which means, by the definition of annihilator (see Sect.21), that $\mathcal{V}^*$ coincides with the annihilator $\mathcal{U}^\perp$ of $\mathcal{U}$. Since $\dim \mathcal{V}^* = \dim \mathcal{V}$, we conclude from the Formula (21.15) for Dimension of Annihilators that $\dim \mathcal{U} = 0$ and hence $\mathrm{Null}\,(\mathrm{ev}) = \mathcal{U} = \{\mathbf{0}\}$. Since $\dim \mathcal{V}^{**} = \dim \mathcal{V}$, we can use the Pigeonhole

Principle for Linear Mappings (Sect. 17) to conclude that ev is invertible. ∎

We use the natural isomorphism described by Prop.1 to identify $\mathcal{V}^{**}$ with $\mathcal{V}$:

$$\mathcal{V}^{**} \cong \mathcal{V},$$

and hence we use the same symbol for an element of $\mathcal{V}$ and the corresponding element of $\mathcal{V}^{**}$. Thus, (22.1) reduces to

$$\mathbf{v}\boldsymbol{\lambda} = \boldsymbol{\lambda}\mathbf{v} \quad \text{for all} \quad \boldsymbol{\lambda} \in \mathcal{V}^*, \mathbf{v} \in \mathcal{V}, \tag{22.2}$$

where the $\mathbf{v}$ on the left side is interpreted as an element of $\mathcal{V}^{**}$.

The identification $\mathcal{V}^{**} \cong \mathcal{V}$ induces identifications such as

$$\operatorname{Lin}(\mathcal{V}^{**}, \mathcal{W}^{**}) \cong \operatorname{Lin}(\mathcal{V}, \mathcal{W})$$

when $\mathcal{V}$ and $\mathcal{W}$ are given linear spaces. In particular, if $\mathbf{H} \in \operatorname{Lin}(\mathcal{W}^*, \mathcal{V}^*)$, we will interpret $\mathbf{H}^\top$ as an element of $\operatorname{Lin}(\mathcal{V}, \mathcal{W})$. In view of (22.2) and (21.3), $\mathbf{H}^\top$ is characterized by

$$\boldsymbol{\mu}\mathbf{H}^\top\mathbf{v} = (\mathbf{H}\boldsymbol{\mu})\mathbf{v} \quad \text{for all} \quad \mathbf{v} \in \mathcal{V},\ \boldsymbol{\mu} \in \mathcal{W}^*. \tag{22.3}$$

Using (21.3) and (22.3), we immediately obtain the following:

Proposition 2: *Let $\mathcal{V}$, $\mathcal{W}$ be linear spaces. Then the transposition $(\mathbf{H} \mapsto \mathbf{H}^\top) : \operatorname{Lin}(\mathcal{W}^*, \mathcal{V}^*) \to \operatorname{Lin}(\mathcal{V}, \mathcal{W})$ is the inverse of the transposition $(\mathbf{L} \mapsto \mathbf{L}^\top) : \operatorname{Lin}(\mathcal{V}, \mathcal{W}) \to \operatorname{Lin}(\mathcal{W}^*, \mathcal{V}^*)$, so that*

$$(\mathbf{L}^\top)^\top = \mathbf{L} \quad \text{for all} \quad \mathbf{L} \in \operatorname{Lin}(\mathcal{V}, \mathcal{W}) \tag{22.4}$$

The identification $\mathcal{V}^{**} \cong \mathcal{V}$ also permits us to interpret the annihilator $\mathcal{H}^\perp$ of a subset $\mathcal{H}$ of $\mathcal{V}^*$ as a subset of $\mathcal{V}$.

Proposition 3: *For every subspace $\mathcal{U}$ of a given linear space $\mathcal{V}$, we have*

$$(\mathcal{U}^\perp)^\perp = \mathcal{U} \tag{22.5}$$

Proof: Let $\mathbf{u} \in \mathcal{U}$ be given. By definition of $\mathcal{U}^\perp$, we have $\boldsymbol{\lambda}\mathbf{u} = 0$ for all $\boldsymbol{\lambda} \in \mathcal{U}^\perp$. If we interpret $\mathbf{u}$ as an element of $\mathcal{V}^{**}$ and use (22.2), this means that $\mathbf{u}\boldsymbol{\lambda} = 0$ for all $\boldsymbol{\lambda} \in \mathcal{U}^\perp$ and hence $\mathbf{u}_>(\mathcal{U}^\perp) = \{0\}$. Therefore, we have $\mathbf{u} \in (\mathcal{U}^\perp)^\perp$. Since $\mathbf{u} \in \mathcal{U}$ was arbitrary, it follows that $\mathcal{U} \subset (\mathcal{U}^\perp)^\perp$. Applying the Formula (21.15) for Dimension of Annihilators to $\mathcal{U}$ and $\mathcal{U}^\perp$ and recalling $\dim\mathcal{V} = \dim\mathcal{V}^*$, we find $\dim\mathcal{U} = \dim(\mathcal{U}^\perp)^\perp$. By Prop.2 of Sect.17, this is possible only when (22.5) holds. ∎

Using (22.5) and (22.4) one obtains the following consequences of Props.3, 2, and 5 of Sect.21.

Proposition 4: *For every subset $\mathcal{S}$ of $\mathcal{V}$, we have*

$$\mathrm{Lsp}\mathcal{S} = (\mathcal{S}^{\perp})^{\perp}. \tag{22.6}$$

If $\mathcal{U}_1$ and $\mathcal{U}_2$ are subspaces of $\mathcal{V}$, then

$$\mathcal{U}_1^{\perp} \subset \mathcal{U}_2^{\perp} \implies \mathcal{U}_2 \subset \mathcal{U}_1$$

and

$$\mathcal{U}_1^{\perp} + \mathcal{U}_2^{\perp} = (\mathcal{U}_1 \cap \mathcal{U}_2)^{\perp}. \tag{22.7}$$

Using (22.5) and (22.4) one also obtains the following consequence of the Theorem on Annihilators and Transposes (Sect.21).

Proposition 5: *Let $\mathcal{V}$ and $\mathcal{W}$ be linear spaces and let $\mathbf{L} \in \mathrm{Lin}(\mathcal{V}, \mathcal{W})$ be given. For every subspace $\mathcal{H}$ of $\mathcal{W}^*$, we then have*

$$\mathbf{L}^{\top}_{>}(\mathcal{H}) = (\mathbf{L}^{<}(\mathcal{H}^{\perp}))^{\perp}. \tag{22.8}$$

In particular, we have

$$\mathrm{Rng}\,\mathbf{L}^{\top} = (\mathrm{Null}\;\mathbf{L})^{\perp}. \tag{22.9}$$

Notes 22

(1) Our notation $\boldsymbol{\lambda}\mathbf{v}$ for the value of $\boldsymbol{\lambda} \in \mathcal{V}^* := \mathrm{Lin}(\mathcal{V}, \mathbb{R})$ at $\mathbf{v} \in \mathcal{V}$, as in (22.1), is in accord with the general notation for the values of linear mappings (see Sect.13). Very often, a more complicated notation, such as $\langle \mathbf{v}, \boldsymbol{\lambda} \rangle$ or $[\mathbf{v}, \boldsymbol{\lambda}]$, is used. I disagree with the claim of one author that this complication clarifies matters later on; I believe that it obscures them.

23 Dual Bases

We now consider the space $\mathbb{R}^I$ of all families of real numbers indexed on a given finite set I. Let $\varepsilon := (\mathrm{ev}_i \mid i \in I)$ be the evaluation family associated with $\mathbb{R}^I$ (see Sect.04). As we already remarked in Sect.14, the evaluations $\mathrm{ev}_i : \mathbb{R}^I \to \mathbb{R}$, $i \in I$, are linear, i.e. they are members of $(\mathbb{R}^I)^*$. Thus, ε is a family in $(\mathbb{R}^I)^*$. The linear combination mapping $\mathrm{lnc}_{\varepsilon} : \mathbb{R}^I \to (\mathbb{R}^I)^*$ is defined by

$$\mathrm{lnc}_{\varepsilon}\lambda := \sum_{i \in I} \lambda_i \mathrm{ev}_i \quad \text{for all} \quad \lambda \in \mathbb{R}^I \tag{23.1}$$

(see Def.1 in Sect.15). It follows from (23.1) that

$$(\mathrm{lnc}_\varepsilon\lambda)\mu = \sum_{i\in I}\lambda_i \mathrm{ev}_i(\mu) = \sum_{i\in I}\lambda_i\mu_i \tag{23.2}$$

for all $\lambda, \mu \in \mathbb{R}^I$.

Let $\delta^I := (\delta^I_k \mid k \in I)$ be the standard basis of $\mathbb{R}^I$ (see Sect.16). Writing (23.2) with the choice $\mu := \delta^I_k$, $k \in I$, we obtain

$$(\mathrm{lnc}_\varepsilon\lambda)\delta^I_k = \lambda_k = \mathrm{ev}_k\lambda \tag{23.3}$$

for all $\lambda \in \mathbb{R}^I$ and all $k \in I$. Given $\alpha \in (\mathbb{R}^I)^*$, it easily follows from (23.3) that $\lambda := (\alpha\delta^I_i \mid i \in I)$ is the unique solution of the equation

$$?\lambda \in \mathbb{R}^I, \quad \mathrm{lnc}_\varepsilon\lambda = \alpha.$$

Since $\alpha \in (\mathbb{R}^I)^*$ was arbitrary, we can conclude that lnc_ε is invertible and hence a linear isomorphism.

It is evident from (23.2) that

$$(\mathrm{lnc}_\varepsilon\lambda)\mu = (\mathrm{lnc}_\varepsilon\mu)\lambda \quad \text{for all} \quad \lambda, \mu \in \mathbb{R}^I. \tag{23.4}$$

Comparing this result with (22.3) and using the identification $(\mathbb{R}^I)^{**} \cong \mathbb{R}^I$, we obtain $\mathrm{lnc}_\varepsilon^\top = \mathrm{lnc}_\varepsilon$, where $\mathrm{lnc}_\varepsilon^\top \in \mathrm{Lin}((\mathbb{R}^I)^{**}, (\mathbb{R}^I)^*)$ is identified with the corresponding element of $\mathrm{Lin}(\mathbb{R}^I, (\mathbb{R}^I)^*)$. In other words, the isomorphism $(\mathrm{lnc}_\varepsilon^\top)^{-1}\mathrm{lnc}_\varepsilon : \mathbb{R}^I \to (\mathbb{R}^I)^{**}$ coincides with the identification $\mathbb{R}^I \cong (\mathbb{R}^I)^{**}$ obtained from Prop.1 of Sect.22. Therefore, there is no conflict if we use lnc_ε to identify $(\mathbb{R}^I)^*$ with $\mathbb{R}^I$.

From now on we shall use lnc_ε to identify $(\mathbb{R}^I)^*$ with $\mathbb{R}^I$:

$$(\mathbb{R}^I)^* \cong \mathbb{R}^I,$$

except that, given $\lambda \in \mathbb{R}^I$, we shall write $\lambda\cdot := \mathrm{lnc}_\varepsilon\lambda$ rather than merely λ for the corresponding element in $(\mathbb{R}^I)^*$. Thus, (23.2) and (23.4) reduce to

$$\mu \cdot \lambda = \lambda \cdot \mu = \sum_{i\in I}\lambda_i\mu_i \quad \text{for all} \quad \lambda, \mu \in \mathbb{R}^I. \tag{23.5}$$

The equations (23.5) and (23.3) yield

$$\delta^I_k \cdot \lambda = \lambda \cdot \delta^I_k = \lambda_k = \mathrm{ev}_k\lambda \tag{23.6}$$

for all $\lambda \in \mathbb{R}^I$ and all $k \in I$. It follows that $\mathrm{ev}_k = \delta^I_k\cdot$ for all $k \in I$, i.e. that the standard basis δ^I becomes identified with the evaluation family ε. Since

$$\mathrm{ev}_k(\delta^I_i) = \delta_{k,i} \quad \text{for all} \quad i, k \in I$$

we see that

$$\delta_k^I \cdot \delta_i^I = \delta_{k,i} := \left\{ \begin{array}{ll} 1 & \text{if} \quad k = i \\ 0 & \text{if} \quad k \neq i \end{array} \right\} \tag{23.7}$$

holds for all $i, k \in I$.

The following result is an easy consequence of (22.3), (23.5), and (16.2). It shows that the transpose of a matrix as defined in Sect.02 is the same as the transpose of the linear mapping identified with this matrix, and hence that there is no notational clash.

Proposition 1: *Let* I *and* J *be finite index sets and let* $M \in \text{Lin}(\mathbb{R}^I, \mathbb{R}^J) \cong \mathbb{R}^{J \times I}$ *be given. Then*

$$M^\top \in \text{Lin}((\mathbb{R}^J)^*, (\mathbb{R}^I)^*) \cong \text{Lin}(\mathbb{R}^J, \mathbb{R}^I) \cong \mathbb{R}^{I \times J}$$

satisfies

$$(M^\top \mu) \cdot \lambda = \mu \cdot M\lambda \quad \textit{for all} \quad \mu \in \mathbb{R}^J,\ \lambda \in \mathbb{R}^I \tag{23.8}$$

and

$$(M^\top)_{i,j} = M_{j,i} \quad \textit{for all} \quad i \in I,\ j \in J. \tag{23.9}$$

Let $\mathcal{V}$ be a linear space, let $\mathbf{b} := (\mathbf{b}_i \mid i \in I)$ be a basis of $\mathcal{V}$, and let $\text{lnc}_{\mathbf{b}} : \mathbb{R}^I \to \mathcal{V}$ be the (invertible) linear combination mapping for $\mathbf{b}$ (see Sect.15). Using the identification $(\mathbb{R}^I)^* \cong \mathbb{R}^I$, we can regard $(\text{lnc}_{\mathbf{b}}^{-1})^\top$ as a mapping from $\mathbb{R}^I$ to $\mathcal{V}^*$. Using the standard basis δ^I of $\mathbb{R}^I$, we define

$$\mathbf{b}_i^* := (\text{lnc}_{\mathbf{b}}^{-1})^\top \delta_i^I \quad \text{for all} \quad i \in I \tag{23.10}$$

and call the family $\mathbf{b}^* := (\mathbf{b}_i^* \mid i \in I)$ in $\mathcal{V}^*$ the **dual** of the given basis $\mathbf{b}$. Since $(\text{lnc}_{\mathbf{b}}^{-1})^\top$ is invertible, it follows from Prop.2 of Sect.16 that the dual $\mathbf{b}^*$ is a basis of $\mathcal{V}^*$.

Using (23.10) and (21.7), we find that

$$\text{lnc}_{\mathbf{b}}^\top \mathbf{b}_i^* = \delta_i^I = \text{lnc}_{\mathbf{b}}^{-1} \mathbf{b}_i \quad \text{for all} \quad i \in I. \tag{23.11}$$

The dual basis $\mathbf{b}^*$ can be used to evaluate the family of components of a given $\mathbf{v} \in \mathcal{V}$ relative to the basis $\mathbf{b}$:

Proposition 2: *Let* $\mathbf{b} := (\mathbf{b}_i \mid i \in I)$ *be a basis of* $\mathcal{V}$ *and let* $\mathbf{b}^*$ *be its dual. For every* $\mathbf{v} \in \mathcal{V}$*, we then have*

$$(\text{lnc}_{\mathbf{b}}^{-1} \mathbf{v})_i = \mathbf{b}_i^* \mathbf{v} \quad \text{for all} \quad i \in I \tag{23.12}$$

and hence

$$\mathbf{v} = \sum_{i \in I} (\mathbf{b}_i^* \mathbf{v}) \mathbf{b}_i. \tag{23.13}$$

Proof: It follows from (23.10), (21.3) and (23.6) that

$$\mathbf{b}_i^*\mathbf{v} = ((\mathrm{lnc}_{\mathbf{b}}^{-1})^\top \delta_i^I)\mathbf{v} = (\mathrm{lnc}_b e^{-1}\mathbf{v}) \cdot \delta_i^I = (\mathrm{lnc}_{\mathbf{b}}^{-1}\mathbf{v})_i$$

for all $i \in I$. ■

Using Prop.2 and the formula (16.11) one easily obtains the following:

Proposition 3: *Let $\mathcal{V}$ and $\mathcal{W}$ be linear spaces and let $\mathbf{b} := (\mathbf{b}_i \mid i \in I)$ and $\mathbf{c} := (\mathbf{c}_j \mid j \in J)$ be bases of $\mathcal{V}$ and $\mathcal{W}$, respectively. Then the matrix $M \in \mathbb{R}^{J\times I}$ of a given $\mathbf{L} \in \mathrm{Lin}(\mathcal{V}, \mathcal{W})$ relative to $\mathbf{b}$ and $\mathbf{c}$ can be obtained by the formula*

$$M_{j,i} = \mathbf{c}_j^*\mathbf{L}\mathbf{b}_i \quad \text{for all} \quad i \in I, j \in J. \tag{23.14}$$

The following result gives the most useful characterization of the dual of a basis.

Proposition 4: *Let $\mathbf{b} := (\mathbf{b}_i \mid i \in I)$ be a basis of $\mathcal{V}$ and let $\boldsymbol{\beta} := (\boldsymbol{\beta}_i \mid i \in I)$ be a family in $\mathcal{V}^*$. Then*

$$\boldsymbol{\beta}_k\mathbf{b}_i = \delta_{k,i} \tag{23.15}$$

holds for all $i, k \in I$ if and only if $\boldsymbol{\beta}$ coincides with the dual $\mathbf{b}^*$ of $\mathbf{b}$.

Proof: In view of (23.7) and (21.7), the relation (23.15) is valid if and only if

$$\boldsymbol{\beta}_k\mathbf{b}_i = \delta_k^I \cdot \delta_i^I = \delta_k^I \cdot (\mathrm{lnc}_{\mathbf{b}}^{-1}\mathbf{b}_i) = ((\mathrm{lnc}_{\mathbf{b}}^{-1})^\top \delta_k^I)\mathbf{b}_i.$$

Therefore, since $\mathbf{b}$ is a basis, it follows from the uniqueness assertion of Prop.2 of Sect.16 that (23.15) holds for all $i \in I$ if and only if $\boldsymbol{\beta}_k = (\mathrm{lnc}_{\mathbf{b}}^{-1})^\top \delta_k^\top$. The assertion now follows from the definition (23.10) of the dual basis. ■

The following result furnishes a useful criterion for linear independence.

Proposition 5: *Let $\mathbf{f} := (\mathbf{f}_j \mid j \in J)$ be a family in $\mathcal{V}$ and $\boldsymbol{\varphi} := (\boldsymbol{\varphi}_j \mid j \in J)$ a family in $\mathcal{V}^*$ such that*

$$\boldsymbol{\varphi}_k\mathbf{f}_j = \delta_{k,j} \quad \text{for all} \quad j, k \in J. \tag{23.16}$$

Then $\mathbf{f}$ and $\boldsymbol{\varphi}$ are both linearly independent.

Proof: By the definition (15.1) of $\mathrm{lnc}_{\mathbf{f}}$ it follows from (23.16) that

$$\boldsymbol{\varphi}_k(\mathrm{lnc}_{\mathbf{f}}\lambda) = \sum_{j\in J} \lambda_j(\boldsymbol{\varphi}_k\mathbf{f}_j) = \lambda_k$$

for all $\lambda \in \mathbb{R}^{(J)}$ and all $k \in J$. Hence we can have $\lambda \in \mathrm{Null}\ \mathrm{lnc}_{\mathbf{f}}$ i.e. $\mathrm{lnc}_{\mathbf{f}}\lambda = 0$, only if $\lambda_k = 0$ for all $k \in J$, i.e. only if $\lambda = 0$. It follows that $\mathrm{Null}\ \mathrm{lnc}_{\mathbf{f}} = \{\mathbf{0}\}$, which implies the linear independence of $\mathbf{f}$ by Prop.1 of

Sect.15. The linear independence of $\boldsymbol{\varphi}$ follows by using the identification $\mathcal{V}^{**} \cong \mathcal{V}$ and by interchanging the roles of $\mathbf{f}$ and $\boldsymbol{\varphi}$. ∎

If we apply Prop.4 to the case when $\mathcal{V}$ is replaced by $\mathcal{V}^*$ and $\mathbf{b}$ by $\mathbf{b}^*$ and use (22.2), we obtain:

Proposition 6: *The dual* $\mathbf{b}^{**}$ *of the dual* $\mathbf{b}^*$ *of a basis* $\mathbf{b}$ *of* $\mathcal{V}$ *is identified with* $\mathbf{b}$ *itself by the identification* $\mathcal{V}^{**} \cong \mathcal{V}$*, i.e. we have*

$$\mathbf{b}^{**} = \mathbf{b}. \tag{23.17}$$

Using this result and Prop.3 we obtain:

Proposition 7: *Let* $\mathcal{V}$*,* $\mathcal{W}$*,* $\mathbf{b}$ *and* $\mathbf{c}$ *be given as in Prop.3. If* $M \in \mathbb{R}^{J\times I}$ *is the matrix of a given* $\mathbf{L} \in \mathrm{Lin}(\mathcal{V}, \mathcal{W})$ *relative to* $\mathbf{b}$ *and* $\mathbf{c}$*, then* $M^\top \in \mathbb{R}^{I\times J}$ *is the matrix of* $\mathbf{L}^\top \in \mathrm{Lin}(\mathcal{W}^*, \mathcal{V}^*)$ *relative to* $\mathbf{c}^*$ *and* $\mathbf{b}^*$*, i.e.*

$$M^\top_{i,j} = \mathbf{b}_i\mathbf{L}^\top\mathbf{c}^*_j = \mathbf{c}_j\mathbf{L}\mathbf{b}_i = M_{j,i} \quad \text{for all} \quad i \in I,\ j \in J \tag{23.18}$$

Let $\boldsymbol{\beta} := (\boldsymbol{\beta}_i \mid i \in I)$ be a family in $\mathcal{V}^*$, so that $\boldsymbol{\beta} \in \mathcal{V}^{*I}$. Using the identification $\mathcal{V}^{*I} = (\mathrm{Lin}(\mathcal{V}, \mathbb{R}))^I \cong \mathrm{Lin}(\mathcal{V}, \mathbb{R}^I)$ defined by termwise evaluation (see Sect.14), and the identification $(\mathbb{R}^I)^* \cong \mathbb{R}^I$ characterized by (23.5), we easily see that $\boldsymbol{\beta}^\top \in \mathrm{Lin}((\mathbb{R}^I)^*, \mathcal{V}^*) \cong \mathrm{Lin}(\mathbb{R}^I, \mathcal{V}^*)$ is given by

$$\boldsymbol{\beta}^\top = \mathrm{lnc}_{\boldsymbol{\beta}}. \tag{23.19}$$

Remark: Let $\mathbf{b}$ be a basis of $\mathcal{V}$. The mapping $(\mathrm{lnc}_{\mathbf{b}}^{-1})^\top \mathrm{lnc}_{\mathbf{b}}^{-1} : \mathcal{V} \to \mathcal{V}^*$ is a linear isomorphism. In fact, by (23.10) and (23.11) it is the (unique) linear isomorphism that maps the basis $\mathbf{b}$ termwise onto the dual basis $\mathbf{b}^*$. This isomorphism is not a natural isomorphism because it depends on the capricious choice of the basis $\mathbf{b}$. The natural isomorphism that is used for the identification $(\mathbb{R}^I)^* \cong \mathbb{R}^I$ exists because $\mathbb{R}^I$ has a natural basis, namely the standard basis. This basis gives $\mathbb{R}^I$ a structure beyond the mere linear-space structure. To use the metaphor mentioned in the Pitfall at the end of Sect.15, the bases in $\mathbb{R}^I$ do not form a "democracy", as is the case in linear spaces without additional structure. Rather, they form a "monarchy" with the standard basis as king. ∎

24 Bilinear Mappings

Definition 1: *Let* $\mathcal{V}_1$*,* $\mathcal{V}_2$*, and* $\mathcal{W}$ *be linear spaces. We say that the mapping* $\mathbf{B} : \mathcal{V}_1 \times \mathcal{V}_2 \to \mathcal{W}$ *is* **bilinear** *if* $\mathbf{B}(\mathbf{v}_1, \cdot) := (\mathbf{v}_2 \mapsto \mathbf{B}(\mathbf{v}_1, \mathbf{v}_2)) \ : \ \mathcal{V}_2 \to \mathcal{W}$ *is linear for all* $\mathbf{v}_1 \in \mathcal{V}_1$ *and* $\mathbf{B}(\cdot, \mathbf{v}_2) := (\mathbf{v}_1 \mapsto \mathbf{B}(\mathbf{v}_1, \mathbf{v}_2)) : \mathcal{V}_1 \to \mathcal{W}$ *is* *linear for all* $\mathbf{v}_2 \in \mathcal{V}_2$*. The set of all bilinear mappings from* $\mathcal{V}_1 \times \mathcal{V}_2$ *to* $\mathcal{W}$ *is denoted by* $\mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$*.*

Briefly, to say that $\mathbf{B}$ is bilinear means that $\mathbf{B}(\mathbf{v}_1, \cdot) \in \mathrm{Lin}(\mathcal{V}_2, \mathcal{W})$ for all $\mathbf{v}_1 \in \mathcal{V}_1$ and $\mathbf{B}(\cdot, \mathbf{v}_2) \in \mathrm{Lin}(\mathcal{V}_1, \mathcal{W})$ for all $\mathbf{v}_2 \in \mathcal{V}_2$.

Proposition 1: $\mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$ *is a subspace of* $\mathrm{Map}\,(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$*.*

Proof: $\mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$ is not empty because the zero-mapping belongs to it. To show that $\mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$ is stable under addition, consider two of its members $\mathbf{B}$ and $\mathbf{C}$. Using the definition of $\mathbf{B} + \mathbf{C}$ (see Sect.14) we have

$$(\mathbf{B} + \mathbf{C})(\mathbf{v}_1, \cdot) = \mathbf{B}(\mathbf{v}_1, \cdot) + \mathbf{C}(\mathbf{v}_1, \cdot)$$

for all $\mathbf{v}_1 \in \mathcal{V}_1$. Since $\mathrm{Lin}(\mathcal{V}_2, \mathcal{W})$ is stable under addition, it follows that $(\mathbf{B} + \mathbf{C})(\mathbf{v}_1, \cdot) \in \mathrm{Lin}(\mathcal{V}_2, \mathcal{W})$ for all $\mathbf{v}_1 \in \mathcal{V}_1$. A similar argument shows that $(\mathbf{B} + \mathbf{C})(\cdot, \mathbf{v}_2) \in \mathrm{Lin}(\mathcal{V}_1, \mathcal{W})$ for all $\mathbf{v}_2 \in \mathcal{V}_2$ and hence that $\mathbf{B} + \mathbf{C} \in \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$. It is even easier to show that $\mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$ is stable under scalar multiplication. ■

Pitfall: The space $\mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$ of bilinear mappings has little connection with the space $\mathrm{Lin}(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$ of linear mappings from the product space $\mathcal{V}_1 \times \mathcal{V}_2$ to $\mathcal{W}$. In fact, it is easily seen that as subspaces of $\mathrm{Map}\,(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$ the two are disjunct, i.e. they have only the zero mapping in common. ■

Proposition 2: *Let* $\mathcal{V}_1$*,* $\mathcal{V}_2$*, and* $\mathcal{W}$ *be linear spaces. For each* $\mathbf{B} \in \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$*, the mapping*

$$(\mathbf{v}_1 \mapsto \mathbf{B}(\mathbf{v}_1, \cdot)) : \mathcal{V}_1 \to \mathrm{Lin}(\mathcal{V}_2, \mathcal{W})$$

is linear. Moreover, the mapping

$$(\mathbf{B} \mapsto (\mathbf{v}_1 \mapsto \mathbf{B}(\mathbf{v}_1, \cdot))) : \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W}) \to \mathrm{Lin}(\mathcal{V}_1, \mathrm{Lin}(\mathcal{V}_2, \mathcal{W}))$$

is a linear isomorphism.

Proof: The first assertion follows from the definition of the linear-space operations in $\mathrm{Lin}(\mathcal{V}_2, \mathcal{W})$ and the linearity of $\mathbf{B}(\cdot, \mathbf{v}_2)$ for all $\mathbf{v}_2 \in \mathcal{V}_2$. The second assertion is an immediate consequence of the definitions of the spaces involved and of the definition of the linear-space operations in these spaces. ■

We use the natural isomorphism described in Prop.2 to identify:

$$\mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W}) \cong \mathrm{Lin}(\mathcal{V}_1, \mathrm{Lin}(\mathcal{V}_2, \mathcal{W})). \tag{24.1}$$

This identification is expressed by

$$(\mathbf{B}\mathbf{v}_1)\mathbf{v}_2 = \mathbf{B}(\mathbf{v}_1, \mathbf{v}_2) \quad \text{for all} \quad \mathbf{v}_1 \in \mathcal{V}_1, \mathbf{v}_2 \in \mathcal{V}_2, \tag{24.2}$$

where on the left side $\mathbf{B}$ is interpreted as an element of $\mathrm{Lin}(\mathcal{V}_1, \mathrm{Lin}(\mathcal{V}_2, \mathcal{W}))$ and on the right side $\mathbf{B}$ is interpreted as an element of $\mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$. The identification given by (24.1) and (24.2) is consistent with the identification described by (04.28) and (04.29).

Using Prop.7 of Sect.17 and Prop.2 above, we obtain the following formula for the dimension of spaces of bilinear mappings:

$$\dim \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W}) = (\dim \mathcal{V}_1)(\dim \mathcal{V}_2)(\dim \mathcal{W}). \tag{24.3}$$

Examples:

1. The scalar multiplication $\mathrm{sm} : \mathbb{R} \times \mathcal{V} \to \mathcal{V}$ of a linear space $\mathcal{V}$ is bilinear, i.e. an element of

$$\mathrm{Lin}_2(\mathbb{R} \times \mathcal{V}, \mathcal{V}) \cong \mathrm{Lin}(\mathbb{R}, \mathrm{Lin}(\mathcal{V}, \mathcal{V})) = \mathrm{Lin}(\mathbb{R}, \mathrm{Lin}\mathcal{V}).$$

The corresponding linear mapping $\mathrm{sm} \in \mathrm{Lin}(\mathbb{R}, \mathrm{Lin}\mathcal{V})$ is given by

$$\mathrm{sm} = (\xi \mapsto \xi \mathbf{1}_{\mathcal{V}}) : \ \mathbb{R} \ \to \ \mathrm{Lin}\mathcal{V}.$$

It is not only linear, but it also preserves products, i.e. $\mathrm{sm}(\xi\eta) = (\mathrm{sm}\xi)(\mathrm{sm}\eta)$ holds for all $\xi, \eta \in \mathbb{R}$. In fact, sm is an injective algebra-homomorphism from $\mathbb{R}$ to the algebra of lineons on $\mathcal{V}$ (see Sect.18).

2. Let $\mathcal{V}$ and $\mathcal{W}$ be linear spaces. Then

$$((\mathbf{L}, \mathbf{v}) \mapsto \mathbf{L}\mathbf{v}) : \mathrm{Lin}(\mathcal{V}, \mathcal{W}) \times \mathcal{V} \ \to \ \mathcal{W}$$

is bilinear. The corresponding linear mapping is simply $\mathbf{1}_{\mathrm{Lin}(\mathcal{V},\mathcal{W})}$. In the special case $\mathcal{W} := \mathbb{R}$, we obtain the bilinear mapping

$$((\boldsymbol{\lambda}, \mathbf{v}) \mapsto \boldsymbol{\lambda}\mathbf{v}) : \mathcal{V}^* \times \mathcal{V} \ \to \ \mathbb{R}.$$

The corresponding linear mapping is $\mathbf{1}_{\mathcal{V}^*}$.

3. Let S be a set and let $\mathcal{V}$ and $\mathcal{V}'$ be linear spaces. It then follows from Props.1 and 3 of Sect.14 that

$$((\mathbf{L}, \mathbf{f}) \mapsto \mathbf{L}\mathbf{f}) : \mathrm{Lin}(\mathcal{V}, \mathcal{V}') \times \mathrm{Map}\,(S, \mathcal{V}) \ \to \ \mathrm{Map}\,(S, \mathcal{V}')$$

is bilinear.

4. Let $\mathcal{V}$, $\mathcal{V}'$, $\mathcal{V}''$ be linear spaces. Then

$$((\mathbf{M},\mathbf{L}) \mapsto \mathbf{ML}) : \mathrm{Lin}(\mathcal{V}',\mathcal{V}'') \times \mathrm{Lin}(\mathcal{V},\mathcal{V}') \;\rightarrow\; \mathrm{Lin}(\mathcal{V},\mathcal{V}'')$$

defines a bilinear mapping. This follows from Prop.1 of Sect.13 and the result stated in the preceding example. ■

The following two results are immediate consequences of the definitions and Prop.1 of Sect.13.

Proposition 3: *The composite of a bilinear mapping with a linear mapping is again bilinear. More precisely, if $\mathcal{V}_1$, $\mathcal{V}_2$, $\mathcal{W}$, and $\mathcal{W}'$ are linear spaces and if $\mathbf{B} \in \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$ and $\mathbf{L} \in \mathrm{Lin}(\mathcal{W},\mathcal{W}')$, then $\mathbf{LB} \in \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W}')$.*

Proposition 4: *The composite of the cross-product of a pair of linear mappings with a bilinear mapping is again bilinear. More precisely, if $\mathcal{V}_1$, $\mathcal{V}_2$, $\mathcal{V}_1'$, $\mathcal{V}_2'$, and $\mathcal{W}$ are linear spaces and if $\mathbf{L}_1 \in \mathrm{Lin}(\mathcal{V}_1,\mathcal{V}_1')$, $\mathbf{L}_2 \in \mathrm{Lin}(\mathcal{V}_2,\mathcal{V}_2')$ and $\mathbf{B} \in \mathrm{Lin}_2(\mathcal{V}_1' \times \mathcal{V}_2', \mathcal{W})$, then $\mathbf{B} \circ (\mathbf{L}_1 \times \mathbf{L}_2) \in \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$.*

With every $\mathbf{B} \in \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$ we can associate a bilinear mapping $\mathbf{B}^\sim \in \mathrm{Lin}_2(\mathcal{V}_2 \times \mathcal{V}_1, \mathcal{W})$ defined by

$$\mathbf{B}^\sim(\mathbf{v}_2,\mathbf{v}_1) := \mathbf{B}(\mathbf{v}_1,\mathbf{v}_2) \quad \text{for all} \quad \mathbf{v}_1 \in \mathcal{V}_1,\ \mathbf{v}_2 \in \mathcal{V}_2 \tag{24.4}$$

We call $\mathbf{B}^\sim$ the **switch** of $\mathbf{B}$. It is evident that

$$(\mathbf{B}^\sim)^\sim = \mathbf{B} \tag{24.5}$$

holds for all bilinear mappings $\mathbf{B}$ and that the **switching**, defined by

$$(\mathbf{B} \mapsto \mathbf{B}^\sim) : \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W}) \rightarrow \mathrm{Lin}_2(\mathcal{V}_2 \times \mathcal{V}_1, \mathcal{W})$$

is a linear isomorphism.

Definition 2: *Let $\mathcal{V}$ and $\mathcal{W}$ be linear spaces. We say that a bilinear mapping $\mathbf{B} \in \mathrm{Lin}_2(\mathcal{V}^2, \mathcal{W})$ is* **symmetric** *if $\mathbf{B}^\sim = \mathbf{B}$, i.e. if*

$$\mathbf{B}(\mathbf{u},\mathbf{v}) = \mathbf{B}(\mathbf{v},\mathbf{u}) \quad \text{for all} \quad \mathbf{u},\mathbf{v} \in \mathcal{V}; \tag{24.6}$$

we say that it is **skew** *if $\mathbf{B}^\sim = -\mathbf{B}$, i.e. if*

$$\mathbf{B}(\mathbf{u},\mathbf{v}) = -\mathbf{B}(\mathbf{v},\mathbf{u}) \quad \text{for all} \quad \mathbf{u},\mathbf{v} \in \mathcal{V}. \tag{24.7}$$

We use the notations

$$\begin{aligned} \mathrm{Sym}_2(\mathcal{V}^2,\mathcal{W}) &:= \{\mathbf{S} \in \mathrm{Lin}_2(\mathcal{V}^2,\mathcal{W}) \mid \mathbf{S}^\sim = \mathbf{S}\}, \\ \mathrm{Skew}_2(\mathcal{V}^2,\mathcal{W}) &:= \{\mathbf{A} \in \mathrm{Lin}_2(\mathcal{V}^2,\mathcal{W}) \mid \mathbf{A}^\sim = -\mathbf{A}\}. \end{aligned}$$

Proposition 5: *A bilinear mapping* $\mathbf{A} \in \mathrm{Lin}_2(\mathcal{V}^2, \mathcal{W})$ *is skew if and only if*

$$\mathbf{A}(\mathbf{u}, \mathbf{u}) = 0 \quad \text{for all} \quad \mathbf{u} \in \mathcal{V}. \tag{24.8}$$

Proof: If $\mathbf{A}$ is skew, then, by (24.7), we have $\mathbf{A}(\mathbf{u}, \mathbf{u}) = -\mathbf{A}(\mathbf{u}, \mathbf{u})$ and hence $\mathbf{A}(\mathbf{u}, \mathbf{u}) = 0$ for all $\mathbf{u} \in \mathcal{V}$. If (24.8) holds, then

$$\begin{aligned} 0 \;\; = \mathbf{A}(\mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}) \;\; &= \mathbf{A}(\mathbf{u}, \mathbf{u}) + \mathbf{A}(\mathbf{u}, \mathbf{v}) + \mathbf{A}(\mathbf{v}, \mathbf{u}) + \mathbf{A}(\mathbf{v}, \mathbf{v}) \\ &= \mathbf{A}(\mathbf{u}, \mathbf{v}) + \mathbf{A}(\mathbf{v}, \mathbf{u}) \end{aligned}$$

and hence $\mathbf{A}(\mathbf{u}, \mathbf{v}) = -\mathbf{A}(\mathbf{v}, \mathbf{u})$ for all $\mathbf{u}, \mathbf{v} \in \mathcal{V}$. ∎.

Proposition 6: *To every* $\mathbf{B} \in \mathrm{Lin}_2(\mathcal{V}^2, \mathcal{W})$ *corresponds a unique pair* $(\mathbf{S}, \mathbf{A})$ *with* $\mathbf{S} \in \mathrm{Sym}_2(\mathcal{V}^2, \mathcal{W})$*,* $\mathbf{A} \in \mathrm{Skew}_2(\mathcal{V}^2, \mathcal{W})$ *such that* $\mathbf{B} = \mathbf{S} + \mathbf{A}$*. In fact,* $\mathbf{S}$ *and* $\mathbf{A}$ *are given by*

$$\mathbf{S} = \frac{1}{2}(\mathbf{B} + \mathbf{B}^{\sim}), \quad \mathbf{A} = \frac{1}{2}(\mathbf{B} - \mathbf{B}^{\sim}). \tag{24.9}$$

Proof: Assume that $\mathbf{S} \in \mathrm{Sym}_2(\mathcal{V}^2, \mathcal{W})$ and $\mathbf{A} \in \mathrm{Skew}_2(\mathcal{V}^2, \mathcal{W})$ are given such that $\mathbf{B} = \mathbf{S} + \mathbf{A}$. Since $\mathbf{B} \mapsto \mathbf{B}^{\sim}$ is linear, it follows that $\mathbf{B}^{\sim} = \mathbf{S}^{\sim} + \mathbf{A}^{\sim} = \mathbf{S} - \mathbf{A}$. Therefore, we have $\mathbf{B} + \mathbf{B}^{\sim} = (\mathbf{S}+\mathbf{A}) + (\mathbf{S}-\mathbf{A}) = 2\mathbf{S}$ and $\mathbf{B} - \mathbf{B}^{\sim} = (\mathbf{S} + \mathbf{A}) - (\mathbf{S} - \mathbf{A}) = 2\mathbf{A}$, which shows that $\mathbf{S}$ and $\mathbf{A}$ must be given by (24.9) and hence are uniquely determined by $\mathbf{B}$. On the other hand, if we *define* $\mathbf{S}$ and $\mathbf{A}$ by (24.9), we can verify immediately that $\mathbf{S}$ is symmetric, that $\mathbf{A}$ is skew, and that $\mathbf{B} = \mathbf{S} + \mathbf{A}$. ∎

In view of Prop.4 of Sect.12, Prop.6 has the following immediate consequence:

Proposition 7: $\mathrm{Sym}_2(\mathcal{V}^2, \mathcal{W})$ *and* $\mathrm{Skew}_2(\mathcal{V}^2, \mathcal{W})$ *are supplementary subspaces of* $\mathrm{Lin}_2(\mathcal{V}^2, \mathcal{W})$*.*

Let $\mathcal{V}$ and $\mathcal{W}$ be linear spaces. We consider the identifications

$$\mathrm{Lin}_2(\mathcal{V} \times \mathcal{W}^*, \mathbb{R}) \cong \mathrm{Lin}(\mathcal{V}, \mathcal{W}^{**}) \cong \mathrm{Lin}(\mathcal{V}, \mathcal{W})$$

and

$$\mathrm{Lin}_2(\mathcal{W}^* \times \mathcal{V}, \mathbb{R}) \cong \mathrm{Lin}(\mathcal{W}^*, \mathcal{V}^*).$$

Hence if $\mathbf{L} \in \mathrm{Lin}(\mathcal{V}, \mathcal{W})$, we can not only form its transpose $\mathbf{L}^{\top} \in \mathrm{Lin}(\mathcal{W}^*, \mathcal{V}^*)$ but, by interpreting $\mathbf{L}$ as an element of $\mathrm{Lin}_2(\mathcal{V} \times \mathcal{W}^*, \mathbb{R})$, we can also form its switch $\mathbf{L}^{\sim} \in \mathrm{Lin}_2(\mathcal{W}^* \times \mathcal{V}, \mathbb{R})$. It is easily verified that $\mathbf{L}^{\top}$ and $\mathbf{L}^{\sim}$ correspond under the identification, i.e.

$$\mathbf{L}^{\sim} = \mathbf{L}^{\top} \quad \text{for all} \quad \mathbf{L} \in \mathrm{Lin}(\mathcal{V}, \mathcal{W}). \tag{24.10}$$

Pitfall: For a bilinear mapping $\mathbf{B}$ whose codomain is not $\mathbb{R}$, the linear mapping corresponding to the switch $\mathbf{B}^{\sim}$ is not the same as the transpose $\mathbf{B}^{\top}$ of the linear mapping corresponding to $\mathbf{B}$. ■

Let $\mathcal{V}_1$, $\mathcal{V}_2$, $\mathcal{W}_1$, and $\mathcal{W}_2$ be linear spaces. Let $\mathbf{L}_1 \in \mathrm{Lin}(\mathcal{V}_1, \mathcal{W}_1)$, $\mathbf{L}_2 \in \mathrm{Lin}(\mathcal{V}_2, \mathcal{W}_2)$ and $\mathbf{B} \in \mathrm{Lin}_2(\mathcal{W}_1 \times \mathcal{W}_2, \mathbb{R}) \cong \mathrm{Lin}(\mathcal{W}_1, \mathcal{W}_2^*)$ be given. By Prop.4, we then have

$$\mathbf{B} \circ (\mathbf{L}_1 \times \mathbf{L}_2) \in \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathbb{R}) \cong \mathrm{Lin}(\mathcal{V}_1, \mathcal{V}_2^*).$$

Using these identifications, it is easily seen that

$$\mathbf{B} \circ (\mathbf{L}_1 \times \mathbf{L}_2) = \mathbf{L}_2^{\top} \mathbf{B} \mathbf{L}_1. \tag{24.11}$$

Notes 24

(1) The terms “antisymmetric” and “skewsymmetric” are often used for what we call, simply, “skew”. A bilinear mapping that satisfies the condition (24.8) is often said to be “alternating”. In the case considered here, this term is synonymous with “skew”, but one obtains a different concept if one replaces $\mathbb{R}$ by a field of characteristic 2.

(2) The pair $(\mathbf{S}, \mathbf{A})$ associated with the bilinear mapping $\mathbf{B}$ according to Prop.6 is sometimes called the “Cartesian decomposition” of $\mathbf{B}$.

25 Tensor Products

For any linear space $\mathcal{V}$ there is a natural isomorphism from $\mathrm{Lin}(\mathbb{R}, \mathcal{V})$ onto $\mathcal{V}$, given by $\mathbf{h} \mapsto \mathbf{h}(1)$. The inverse isomorphism associates with $\mathbf{v} \in \mathcal{V}$ the mapping $\xi \mapsto \xi\mathbf{v}$ in $\mathrm{Lin}(\mathbb{R}, \mathcal{V})$. We denote this mapping by $\mathbf{v}\otimes$ (read “vee tensor”) so that

$$\mathbf{v} \otimes \xi := \xi\mathbf{v} \quad \text{for all} \quad \xi \in \mathbb{R}. \tag{25.1}$$

In particular, there is a natural isomorphism from $\mathbb{R}$ onto $\mathbb{R}^* = \mathrm{Lin}(\mathbb{R}, \mathbb{R})$. It associates with every number $\eta \in \mathbb{R}$ the operation of multiplication with that number, so that (25.1) reduces to $\eta \otimes \xi = \eta\xi$. We use this isomorphism to identify $\mathbb{R}^*$ with $\mathbb{R}$, i.e. we write $\eta = \eta\otimes$. However, when $\mathcal{V} \neq \mathbb{R}$, we do *not* identify $\mathrm{Lin}(\mathbb{R}, \mathcal{V})$ with $\mathcal{V}$ because such an identification would conflict with the identification $\mathcal{V} \cong \mathcal{V}^{**}$ and lead to confusion.

If $\boldsymbol{\lambda} \in \mathcal{V}^* = \mathrm{Lin}(\mathcal{V}, \mathbb{R})$, we can consider $\boldsymbol{\lambda}^{\top} \in \mathrm{Lin}(\mathbb{R}^*, \mathcal{V}^*) \cong \mathrm{Lin}(\mathbb{R}, \mathcal{V}^*)$. Using the identification $\mathbb{R}^* \cong \mathbb{R}$, it follows from (21.3) and (25.1) that $\boldsymbol{\lambda}^{\top}\xi = \xi\boldsymbol{\lambda} = \boldsymbol{\lambda} \otimes \xi$ for all $\xi \in \mathbb{R}$, i.e. that $\boldsymbol{\lambda}^{\top} = \boldsymbol{\lambda}\otimes$.

Definition 1: *Let $\mathcal{V}$ and $\mathcal{W}$ be linear spaces. For every $\mathbf{w} \in \mathcal{W}$ and every $\boldsymbol{\lambda} \in \mathcal{V}^*$ the* **tensor product** *of $\mathbf{w}$ and $\boldsymbol{\lambda}$ is defined to be $\mathbf{w} \otimes \boldsymbol{\lambda} \in \operatorname{Lin}(\mathcal{V}, \mathcal{W})$, i.e. the composite of $\boldsymbol{\lambda}$ with $\mathbf{w}\otimes \in \operatorname{Lin}(\mathbb{R}, \mathcal{W})$, so that*

$$(\mathbf{w} \otimes \boldsymbol{\lambda})\mathbf{v} = (\boldsymbol{\lambda}\mathbf{v})\mathbf{w} \quad \text{for all} \quad \mathbf{v} \in \mathcal{V}. \tag{25.2}$$

In view of Example 4 of Sect. 24, it is clear that the mapping

$$((\mathbf{w}, \boldsymbol{\lambda}) \mapsto \mathbf{w} \otimes \boldsymbol{\lambda}) \ : \ \mathcal{W} \times \mathcal{V}^* \longrightarrow \operatorname{Lin}(\mathcal{V}, \mathcal{W}) \tag{25.3}$$

is bilinear.

In the special case when $\mathcal{V} := \mathbb{R}^I \cong (\mathbb{R}^I)^*$ and $\mathcal{W} := \mathbb{R}^J$ for finite index sets I and J, (16.4) and (25.2) show that the tensor product $\mu \otimes \lambda \in \operatorname{Lin}(\mathbb{R}^I, \mathbb{R}^J) \cong \mathbb{R}^{J \times I}$ of $\mu \in \mathbb{R}^J$ and $\lambda \in \mathbb{R}^I$ has the components

$$(\mu \otimes \lambda)_{j,i} = \mu_j \lambda_i \quad \text{for all} \quad (j, i) \in J \times I. \tag{25.4}$$

Using the identifications $\mathcal{V}^{**} \cong \mathcal{V}$ and $\mathcal{W}^{**} \cong \mathcal{W}$, we can form tensor products

$$\mathbf{w} \otimes \mathbf{v} \in \operatorname{Lin}(\mathcal{V}^*, \mathcal{W}), \quad \boldsymbol{\mu} \otimes \mathbf{v} \in \operatorname{Lin}(\mathcal{V}^*, \mathcal{W}^*),$$

$$\mathbf{w} \otimes \boldsymbol{\lambda} \in \operatorname{Lin}(\mathcal{V}, \mathcal{W}), \quad \boldsymbol{\mu} \otimes \boldsymbol{\lambda} \in \operatorname{Lin}(\mathcal{V}, \mathcal{W}^*)$$

for all $\mathbf{v} \in \mathcal{V}$, $\mathbf{w} \in \mathcal{W}$, $\boldsymbol{\lambda} \in \mathcal{V}^*$, $\boldsymbol{\mu} \in \mathcal{W}^*$. Also, using identifications such as

$$\operatorname{Lin}(\mathcal{V}, \mathcal{W}) \cong \operatorname{Lin}(\mathcal{V}, \mathcal{W}^{**}) \cong \operatorname{Lin}_2(\mathcal{V} \times \mathcal{W}^*, \mathbb{R})$$

we can interpret any tensor product as a bilinear mapping to $\mathbb{R}$. For example, if $\mathbf{w} \in \mathcal{W}$ and $\boldsymbol{\lambda} \in \mathcal{V}^*$ we have

$$(\mathbf{w} \otimes \boldsymbol{\lambda})(\mathbf{v}, \boldsymbol{\mu}) = (\boldsymbol{\lambda}\mathbf{v})(\mathbf{w}\boldsymbol{\mu}) \quad \text{for all} \quad \mathbf{v} \in \mathcal{V}, \ \boldsymbol{\mu} \in \mathcal{W}^* \tag{25.5}$$

The following is an immediate consequence of (25.5), the definition (24.4) of a switch, and (24.10).

Proposition 1: *For all $\mathbf{w} \in \mathcal{W}$ and $\boldsymbol{\lambda} \in \mathcal{V}^*$, we have*

$$(\mathbf{w} \otimes \boldsymbol{\lambda})^\sim = (\mathbf{w} \otimes \boldsymbol{\lambda})^\top = \boldsymbol{\lambda} \otimes \mathbf{w}. \tag{25.6}$$

The following facts follow immediately from the definition of a tensor product.

Proposition 2: *If $\mathbf{w} \neq \mathbf{0}$ then*

$$\operatorname{Null}\ (\mathbf{w} \otimes \boldsymbol{\lambda}) = \operatorname{Null}\ \boldsymbol{\lambda}; \tag{25.7}$$

if $\boldsymbol{\lambda} \neq \mathbf{0}$ *then*

$$\operatorname{Rng}(\mathbf{w} \otimes \boldsymbol{\lambda}) = \mathbb{R}\mathbf{w}. \tag{25.8}$$

Proposition 3: *Let* $\mathcal{V}$, $\mathcal{V}'$, $\mathcal{W}$, $\mathcal{W}'$ *be linear spaces. If* $\boldsymbol{\lambda} \in \mathcal{V}^*$, $\mathbf{w} \in \mathcal{W}$, $\mathbf{L} \in \operatorname{Lin}(\mathcal{W}, \mathcal{W}')$, *and* $\mathbf{M} \in \operatorname{Lin}(\mathcal{V}', \mathcal{V})$ *then*

$$\mathbf{L}(\mathbf{w} \otimes \boldsymbol{\lambda}) = (\mathbf{L}\mathbf{w}) \otimes \boldsymbol{\lambda} \tag{25.9}$$

and

$$(\mathbf{w} \otimes \boldsymbol{\lambda})\mathbf{M} = \mathbf{w} \otimes (\boldsymbol{\lambda}\mathbf{M}) = \mathbf{w} \otimes (\mathbf{M}^\top \boldsymbol{\lambda}). \tag{25.10}$$

If $\mathbf{v} \in \mathcal{V}$ *and* $\boldsymbol{\mu} \in \mathcal{V}'^*$, *then*

$$(\mathbf{w} \otimes \boldsymbol{\lambda})(\mathbf{v} \otimes \boldsymbol{\mu}) = (\boldsymbol{\lambda}\mathbf{v})(\mathbf{w} \otimes \boldsymbol{\mu}). \tag{25.11}$$

Tensor products can be used to construct bases of spaces of linear mappings:

Proposition 4: *Let* $\mathcal{V}$ *and* $\mathcal{W}$ *be linear spaces, let* $\mathbf{b} := (\mathbf{b}_i \mid i \in I)$ *be a basis of* $\mathcal{V}$. *Let* $\mathbf{b}^* := (\mathbf{b}_i^* \mid i \in I)$ *be the basis of* $\mathcal{V}^*$ *dual to* $\mathbf{b}$ *and let* $\mathbf{c} := (\mathbf{c}_j \mid j \in J)$ *be a basis of* $\mathcal{W}$. *Then* $(\mathbf{c}_j \otimes \mathbf{b}_i^* \mid (j,i) \in J \times I)$ *is a basis of* $\operatorname{Lin}(\mathcal{V}, \mathcal{W})$, *and the matrix* $M \in \mathbb{R}^{J \times I}$ *of the components of* $\mathbf{L}$ *relative to this basis is the same as the matrix* $\operatorname{lnc}_{\mathbf{c}}^{-1}\mathbf{L}\operatorname{lnc}_{\mathbf{b}} \in \operatorname{Lin}(\mathbb{R}^I, \mathbb{R}^J) \cong \mathbb{R}^{J \times I}$, *i.e. the matrix of* $\mathbf{L}$ *relative to the bases* $\mathbf{b}$ *and* $\mathbf{c}$ (see Sect. 16).

Proof: It is sufficient to prove that

$$\mathbf{L} = \sum_{(j,i) \in J \times I} M_{j,i}(\mathbf{c}_j \otimes \mathbf{b}_i^*) \tag{25.12}$$

holds when M is the matrix of $\mathbf{L}$ relative to $\mathbf{b}$ and $\mathbf{c}$. It follows from (16.11) and from (25.2) and Prop.4 of Sect. 23 that the left and right sides of (25.12) give the same value when applied to the terms $\mathbf{b}_k$ of the basis $\mathbf{b}$. Using Prop.2 of Sect. 16, we conclude that (25.12) must hold. ∎

Using (18.6) and (18.8), we obtain the following special case of (25.12):

Proposition 5: *Let* $\mathcal{V}$ *be a linear space and let* $\mathbf{b} := (\mathbf{b}_i \mid i \in I)$ *be a basis of* $\mathcal{V}$. *For every lineon* $\mathbf{L} \in \operatorname{Lin}\mathcal{V}$, *we then have*

$$\mathbf{L} = \sum_{(j,i) \in I \times I} ([\mathbf{L}]_{\mathbf{b}})_{j,i}(\mathbf{b}_j \otimes \mathbf{b}_i^*). \tag{25.13}$$

In particular, we have

$$\mathbf{1}_{\mathcal{V}} = \sum_{i \in I} \mathbf{b}_i \otimes \mathbf{b}_i^*. \tag{25.14}$$

Prop.4 has the following corollary:

Proposition 6: *If $\mathcal{V}$ and $\mathcal{W}$ are linear spaces, then*

$$\mathrm{Lin}(\mathcal{V},\mathcal{W}) = \mathrm{Lsp}\{\mathbf{w}\otimes\boldsymbol{\lambda} \mid \mathbf{w}\in\mathcal{W},\ \boldsymbol{\lambda}\in\mathcal{V}^*\}. \tag{25.15}$$

Notes 25

(1) The term "dyadic product" and the notation $\mathbf{w}\boldsymbol{\lambda}$ is often used in the older literature for our "tensor product" $\mathbf{w}\otimes\boldsymbol{\lambda}$. We cannot use this older notation because it would lead to a clash with the evaluation notation described by (22.2).

(2) For other uses of the term "tensor product" see Note (1) to Sect. 26.

26 The Trace

Let $\mathcal{V}$ and $\mathcal{W}$ be linear spaces. Since the mapping (25.3) is bilinear, it follows from Prop.3 of Sect. 24 that for every linear form Ω on $\mathrm{Lin}(\mathcal{V},\mathcal{W})$

$$((\mathbf{w},\boldsymbol{\lambda})\mapsto\Omega(\mathbf{w}\otimes\boldsymbol{\lambda}))\ :\ \mathcal{W}\times\mathcal{V}^*\longrightarrow\ \mathbb{R}$$

is a bilinear mapping Using the identifications $\mathrm{Lin}_2(\mathcal{W}\times\mathcal{V}^*,\mathbb{R})\cong\mathrm{Lin}(\mathcal{W},\mathcal{V}^{**})\cong\mathrm{Lin}(\mathcal{W},\mathcal{V})$ (see Sects.24 and 22) we see that there is a mapping

$$\tau\ :\ (\mathrm{Lin}(\mathcal{V},\mathcal{W}))^*\ \longrightarrow\mathrm{Lin}(\mathcal{W},\mathcal{V}) \tag{26.1}$$

defined by

$$\boldsymbol{\lambda}(\tau(\Omega)\mathbf{w}) = \Omega(\mathbf{w}\otimes\boldsymbol{\lambda})\quad\text{for all}\quad\Omega\in(\mathrm{Lin}(\mathcal{V},\mathcal{W}))^*,\ \ \mathbf{w}\in\mathcal{W},\ \ \boldsymbol{\lambda}\in\mathcal{V}^*. \tag{26.2}$$

Lemma: *The mapping (26.1) defined by (26.2) is a linear isomorphism.*

Proof: The linearity of τ follows from the fact that every member of $\mathrm{Lin}(\mathcal{V},\mathcal{W})$, and in particular $\mathbf{w}\otimes\boldsymbol{\lambda}$, can be identified with an element of $(\mathrm{Lin}(\mathcal{V},\mathcal{W}))^{**}$, i.e. a linear form on $(\mathrm{Lin}(\mathcal{V},\mathcal{W}))^*$.

If $\Omega\in\mathrm{Null}\ \tau$ then $\tau(\Omega)=\mathbf{0}$ and hence, by (26.2), $\Omega(\mathbf{w}\otimes\boldsymbol{\lambda})=0$ for all $\mathbf{w}\in\mathcal{W}$ and all $\boldsymbol{\lambda}\in\mathcal{V}^*$. By Prop.6 of Sect. 25 this is possible only when $\Omega=\mathbf{0}$. We conclude that $\mathrm{Null}\ \tau=\{\mathbf{0}\}$. Since $\dim(\mathrm{Lin}(\mathcal{V},\mathcal{W}))^* = (\dim\mathcal{V})(\dim\mathcal{W}) = \dim(\mathrm{Lin}(\mathcal{W},\mathcal{V}))$ (see (21.1) and (17.7)) it follows from the Pigeonhole Principle for Linear Mappings (Sect. 17) that τ is invertible. ∎

The following result shows that the algebra of lineons admits a natural linear form.

Characterization of the Trace: *Let $\mathcal{V}$ be a linear space. There is exactly one linear form* $\mathrm{tr}_{\mathcal{V}}$ *on* $\mathrm{Lin}\mathcal{V}$ *that satisfies*

$$\mathrm{tr}_{\mathcal{V}}(\mathbf{v}\otimes\boldsymbol{\lambda}) = \boldsymbol{\lambda}\mathbf{v} \quad \text{for all} \quad \mathbf{v}\in\mathcal{V},\ \boldsymbol{\lambda}\in\mathcal{V}^*. \tag{26.3}$$

This linear form $\mathrm{tr}_{\mathcal{V}}$ *is called the* **trace** *for* $\mathcal{V}$.

Proof: Assume that $\mathrm{tr}_{\mathcal{V}} \in (\mathrm{Lin}\mathcal{V})^*$ satisfies (26.3). Using (26.2) with $\mathcal{W} := \mathcal{V}$ and the choice $\Omega := \mathrm{tr}_{\mathcal{V}}$ we see that we must have $\tau(\mathrm{tr}_{\mathcal{V}}) = \mathbf{1}_{\mathcal{V}}$. By the Lemma, it follows that $\mathrm{tr}_{\mathcal{V}}$ is uniquely determined as $\mathrm{tr}_{\mathcal{V}} = \tau^{-1}(\mathbf{1}_{\mathcal{V}})$. On the other hand, if we define $\mathrm{tr}_{\mathcal{V}} := \tau^{-1}(\mathbf{1}_{\mathcal{V}})$ then (26.3) follows from (26.2). ■

If the context makes clear what $\mathcal{V}$ is, we often write tr for $\mathrm{tr}_{\mathcal{V}}$.

Using (26.3), (25.9), and (25.10), it is easily seen that the definition (26.2) of the mapping τ is equivalent to the statement that

$$\Omega\mathbf{L} = \mathrm{tr}_{\mathcal{V}}(\tau(\Omega)\mathbf{L}) = \mathrm{tr}_{\mathcal{W}}(\mathbf{L}\tau(\Omega)) \tag{26.4}$$

holds for all $\Omega \in (\mathrm{Lin}(\mathcal{V},\mathcal{W}))^*$ and all $\mathbf{L}\in\mathrm{Lin}(\mathcal{V},\mathcal{W})$ that are tensor products, i.e. of the form $\mathbf{L} = \mathbf{w}\otimes\boldsymbol{\lambda}$ for some $\mathbf{w}\in\mathcal{W}$, $\boldsymbol{\lambda}\in\mathcal{V}^*$. Since these tensor products span $\mathrm{Lin}(\mathcal{V},\mathcal{W})$ (Prop.6 of Sect. 25), it follows that the mapping τ can be characterized by the statement that (26.4) holds for all $\mathbf{L}\in\mathrm{Lin}(\mathcal{V},\mathcal{W})$, whether $\mathbf{L}$ is a tensor product or not. Using this fact and the Lemma we obtain the following two results.

Representation Theorem for Linear Forms on a Space of Linear Mappings: *Let $\mathcal{V}$ and $\mathcal{W}$ be linear spaces. Every linear form Ω on* $\mathrm{Lin}(\mathcal{V},\mathcal{W})$ *is represented by a unique* $\mathbf{M}\in\mathrm{Lin}(\mathcal{W},\mathcal{V})$ *in the sense that*

$$\Omega\mathbf{L} = \mathrm{tr}_{\mathcal{V}}(\mathbf{M}\mathbf{L}) \quad \text{for all} \quad \mathbf{L}\in\mathrm{Lin}(\mathcal{V},\mathcal{W}). \tag{26.5}$$

Proposition 1: *For every* $\mathbf{L}\in\mathrm{Lin}(\mathcal{V},\mathcal{W})$ *and every* $\mathbf{M}\in\mathrm{Lin}(\mathcal{W},\mathcal{V})$ *we have*

$$\mathrm{tr}_{\mathcal{V}}(\mathbf{M}\mathbf{L}) = \mathrm{tr}_{\mathcal{W}}(\mathbf{L}\mathbf{M}). \tag{26.6}$$

Now let a linear space $\mathcal{V}$ be given.

Proposition 2: *For every* $\mathbf{L}\in\mathrm{Lin}\mathcal{V}$ *we have*

$$\mathrm{tr}_{\mathcal{V}^*}\mathbf{L}^\top = \mathrm{tr}_{\mathcal{V}}\mathbf{L}. \tag{26.7}$$

Proof: By the Theorem on Characterization of Trace it suffices to show that the mapping $\mathbf{L}\mapsto\mathrm{tr}_{\mathcal{V}^*}\mathbf{L}^\top$ from $\mathrm{Lin}\mathcal{V}$ into $\mathbb{R}$ is linear and has the value $\boldsymbol{\lambda}\mathbf{v}$ when $\mathbf{L} = \mathbf{v}\otimes\boldsymbol{\lambda}$. The first is immediate and the second follows from (25.6) and (26.3), applied to the case when $\mathcal{V}$ is replaced by $\mathcal{V}^*$. ■

Proposition 3: *Let* $\mathbf{b} := (\mathbf{b}_i \mid i \in I)$ *be a basis of* $\mathcal{V}$*. For every lineon* $\mathbf{L} \in \mathrm{Lin}\mathcal{V}$*, we have*

$$\mathrm{tr}_{\mathcal{V}}\mathbf{L} = \sum_{i\in I}([\mathbf{L}]_{\mathbf{b}})_{i,i}, \tag{26.8}$$

where $[\mathbf{L}]_{\mathbf{b}} := \mathrm{lnc}_{\mathbf{b}}^{-1}\mathbf{L}\mathrm{lnc}_{\mathbf{b}}$ *is the matrix of* $\mathbf{L}$ *relative to* $\mathbf{b}$*.*

Proof: It follows from (26.3) and Prop.4 of Sect. 23 that

$$\mathrm{tr}_{\mathcal{V}}(\mathbf{b}_j \otimes \mathbf{b}_i^*) = \mathbf{b}_i^*\mathbf{b}_j = \delta_{i,j} \quad \text{for all} \quad i,j \in I.$$

Therefore, (26.8) is a consequence of (25.13) and the linearity of $\mathrm{tr}_{\mathcal{V}}$. ■

If we apply (26.8) to the case when $\mathbf{L} := \mathbf{1}_{\mathcal{V}}$ and hence $([\mathbf{L}])_{\mathbf{b}})_{i,i} = 1$ for all $i \in I$ (see (18.8)), we obtain

$$\mathrm{tr}_{\mathcal{V}}\mathbf{1}_{\mathcal{V}} = \dim \mathcal{V}. \tag{26.9}$$

Proposition 4: *Let* $\mathcal{U}$ *be a subspace of* $\mathcal{V}$ *and let* $\mathbf{P} : \mathcal{V} \to \mathcal{U}$ *be a projection* (see Sect. 19). *Then*

$$\mathrm{tr}_{\mathcal{U}}(\mathbf{K}) = \mathrm{tr}_{\mathcal{V}}(\mathbf{1}_{\mathcal{U}\subset\mathcal{V}}\mathbf{K}\mathbf{P}) \tag{26.10}$$

for all $\mathbf{K} \in \mathrm{Lin}\mathcal{U}$*.*

Proof: It follows from Prop.1 that

$$\mathrm{tr}_{\mathcal{V}}((\mathbf{1}_{\mathcal{U}\subset\mathcal{V}}\mathbf{K})\mathbf{P}) = \mathrm{tr}_{\mathcal{U}}(\mathbf{P}(\mathbf{1}_{\mathcal{U}\subset\mathcal{V}}\mathbf{K})) = \mathrm{tr}_{\mathcal{U}}(\mathbf{P}|_{\mathcal{U}}\mathbf{K})$$

for all $\mathbf{K} \in \mathrm{Lin}\mathcal{U}$. By the definition of a projection (Sect. 19) we have $\mathbf{P}|_{\mathcal{U}} = \mathbf{1}_{\mathcal{U}}$ and hence (26.10) holds. ■

Proposition 5: *The trace of an idempotent lineon* $\mathbf{E}$ *on* $\mathcal{V}$ *is given by*

$$\mathrm{tr}_{\mathcal{V}}\mathbf{E} = \dim \mathrm{Rng}\,\mathbf{E} \tag{26.11}$$

Proof: Put $\mathcal{U} := \mathrm{Rng}\,\mathbf{E}$ and $\mathbf{P} := \mathbf{E}|^{\mathcal{U}}$. We then have $\mathbf{E} = \mathbf{1}_{\mathcal{U}\subset\mathcal{V}}\mathbf{P}$ and, by Prop.1 of Sect. 19, $\mathbf{P} : \mathcal{V} \to \mathcal{U}$ is a projection. Applying (26.10) to the case when $\mathbf{K} := \mathbf{1}_{\mathcal{U}}$ we obtain $\mathrm{tr}_{\mathcal{U}}(\mathbf{1}_{\mathcal{U}}) = \mathrm{tr}_{\mathcal{V}}(\mathbf{E})$. The desired formula (26.11) now follows from (26.9). ■

Proposition 6: *Let* $\mathcal{V}_1$*,* $\mathcal{V}_2$*, and* $\mathcal{W}$ *be linear spaces. There is exactly one mapping*

$$\Lambda : \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W}) \longrightarrow \mathrm{Lin}(\mathrm{Lin}(\mathcal{V}_2^*, \mathcal{V}_1), \mathcal{W})$$

such that

$$\mathbf{B}(\mathbf{v}_1, \mathbf{v}_2) = \Lambda(\mathbf{B})(\mathbf{v}_1 \otimes \mathbf{v}_2) \quad \text{for all} \quad \mathbf{v}_1 \in \mathcal{V}_1, \ \mathbf{v}_2 \in \mathcal{V}_2 \tag{26.12}$$

and all $\mathbf{B} \in \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathcal{W})$*. Moreover,* Λ *is a linear isomorphism.*

Proof: For every $\mathbf{B} \in \mathrm{Lin}_2(\mathcal{V} \times \mathcal{V}_2, \mathcal{W})$ and $\boldsymbol{\mu} \in \mathcal{W}^*$ we have $\boldsymbol{\mu}\mathbf{B} \in \mathrm{Lin}_2(\mathcal{V}_1 \times \mathcal{V}_2, \mathbb{R}) \cong \mathrm{Lin}(\mathcal{V}_1, \mathcal{V}_2^*)$ (see Prop.3 of Sect. 24). Using the identification $\mathcal{V}_2^{**} \cong \mathcal{V}_2$, we see that (26.12) holds if and only if

$$\mathbf{v}_2((\boldsymbol{\mu}\mathbf{B})\mathbf{v}_1) = (\boldsymbol{\mu}\mathbf{B})(\mathbf{v}_1, \mathbf{v}_2) = (\boldsymbol{\mu}\Lambda(\mathbf{B}))(\mathbf{v}_1 \otimes \mathbf{v}_2) \tag{26.13}$$

for all $\mathbf{v}_1 \in \mathcal{V}_1, \mathbf{v}_2 \in \mathcal{V}_2$ and all $\boldsymbol{\mu} \in \mathcal{W}^*$. Comparing (26.13) with (26.2), we conclude that (26.12) is equivalent to

$$\tau(\boldsymbol{\mu}\Lambda(\mathbf{B})) = \boldsymbol{\mu}\mathbf{B} \quad \text{for all} \quad \boldsymbol{\mu} \in \mathcal{W}^*,$$

where τ is defined according to (26.1) and (26.2), in which $\mathcal{V}$ and $\mathcal{W}$ must be replaced by $\mathcal{V}_1$ and $\mathcal{V}_2^*$, respectively. Since τ is invertible by the Lemma, we conclude that (26.12) is equivalent to

$$\boldsymbol{\mu}\Lambda(\mathbf{B}) = \tau^{-1}(\boldsymbol{\mu}\mathbf{B}) \quad \text{for all} \quad \boldsymbol{\mu} \in \mathcal{W}^*. \tag{26.14}$$

Since $\mathcal{W}$ is isomorphic to $\mathcal{W}^{**}$, the uniqueness and existence of Λ follow from the equivalence of (26.14) with (26.12). The linearity of Λ follows from the linearity of τ^{-1}. Also it is clear from (26.14) that Λ has an inverse Λ, which is characterized by

$$\boldsymbol{\mu}\Lambda^{-1}(\Phi) = \tau(\boldsymbol{\mu}\Phi) \quad \text{for all} \quad \boldsymbol{\mu} \in \mathcal{W}^* \tag{26.15}$$

and all $\Phi \in \mathrm{Lin}(\mathrm{Lin}(\mathcal{V}_2^*, \mathcal{V}_1), \mathcal{W})$. ∎

Remark: Prop.6 shows that every bilinear mapping $\mathbf{B}$ on $\mathcal{V}_1 \times \mathcal{V}_2$ is the composite of the tensor product mapping from $\mathcal{V}_1 \times \mathcal{V}_2$ into $\mathrm{Lin}(\mathcal{V}_2^*, \mathcal{V}_1)$ with a linear mapping $\Lambda(\mathbf{B})$. In a setting more abstract than the one used here, the term *tensor product mapping* is often employed for *any* bilinear mapping $\otimes$ on $\mathcal{V}_1 \times \mathcal{V}_2$ with the following universal factorization property: Every bilinear mapping on $\mathcal{V}_1 \times \mathcal{V}_2$ is the composite of $\otimes$ with a unique linear mapping. The codomain of $\otimes$ is then called a *tensor product space* and is denoted by $\mathcal{V}_1 \otimes \mathcal{V}_2$. For the specific tensor product used here, we have $\mathcal{V}_1 \otimes \mathcal{V}_2 := \mathrm{Lin}(\mathcal{V}_2^*, \mathcal{V}_1)$. ∎

Notes 26

(1) There are many ways of constructing a tensor-product space in the sense of the Remark above from given linear spaces $\mathcal{V}_1$ and $\mathcal{V}_2$. The notation $\mathcal{V}_1 \otimes \mathcal{V}_2$ for such a space is therefore ambiguous and we will not use it. One can associate with the construction of tensor-product spaces a concept of "tensor product" of linear mappings which generalizes the concept of tensor product introduced in Sect. 25. Tensor-product spaces are of little practical value in the present context, even though they give important insights in abstract algebra.

27 Bilinear Forms and Quadratic Forms

We assume that a linear space $\mathcal{V}$ is given. The bilinear mappings from $\mathcal{V}^2$ to $\mathbb{R}$, i.e. the members of $\mathrm{Lin}_2(\mathcal{V}^2, \mathbb{R})$ are called **bilinear forms** on $\mathcal{V}$. The identification $\mathrm{Lin}_2(\mathcal{V}^2, \mathbb{R}) \cong \mathrm{Lin}(\mathcal{V}, \mathcal{V}^*)$—a special case of (24.1)—enables us to interpret bilinear forms as mappings from $\mathcal{V}$ to its dual $\mathcal{V}^*$. We use the notation $\mathrm{Sym}(\mathcal{V}, \mathcal{V}^*)$ and $\mathrm{Skew}(\mathcal{V}, \mathcal{V}^*)$ for the subspaces of $\mathrm{Lin}(\mathcal{V}, \mathcal{V}^*)$ that correspond to the subspaces $\mathrm{Sym}_2(\mathcal{V}^2, \mathbb{R})$ and $\mathrm{Skew}_2(\mathcal{V}^2, \mathbb{R})$ of $\mathrm{Lin}_2(\mathcal{V}^2, \mathbb{R})$ (see Def.2 in Sect. 24). In view of (24.10), we have

$$\mathrm{Sym}(\mathcal{V}, \mathcal{V}^*) = \{\mathbf{S} \in \mathrm{Lin}(\mathcal{V}, \mathcal{V}^*) \mid \mathbf{S}^\top = \mathbf{S}\}, \tag{27.1}$$

$$\mathrm{Skew}(\mathcal{V}, \mathcal{V}^*) = \{\mathbf{A} \in \mathrm{Lin}(\mathcal{V}, \mathcal{V}^*) \mid \mathbf{A}^\top = -\mathbf{A}\}. \tag{27.2}$$

By Prop.7 of Sect. 24, $\mathrm{Sym}(\mathcal{V}, \mathcal{V}^*)$ and $\mathrm{Skew}(\mathcal{V}, \mathcal{V}^*)$ are supplementary subspaces of $\mathrm{Lin}(\mathcal{V}, \mathcal{V}^*)$.

Let $\mathbf{b} := (\mathbf{b}_i \mid i \in I)$ be a basis of $\mathcal{V}$ and let $\mathbf{B} \in \mathrm{Lin}_2(\mathcal{V}^2, \mathbb{R}) \cong \mathrm{Lin}(\mathcal{V}, \mathcal{V}^*)$ be a bilinear form. The matrix $M \in \mathbb{R}^{I \times I}$ defined by

$$M_{j,i} := \mathbf{B}(\mathbf{b}_i, \mathbf{b}_j) \quad \text{for all} \quad i, j \in I \tag{27.3}$$

is called the **matrix of B relative to b**. It is easily seen that M coincides with the matrix of $\mathbf{B}$ when regarded as an element of $\mathrm{Lin}(\mathcal{V}, \mathcal{V}^*)$, relative to the bases $\mathbf{b}$ and $\mathbf{b}^*$ in the sense of the definition in Sect. 16. Using the fact that $\mathbf{b}^{**} = \mathbf{b}$, (see (23.17)) it follows from Prop.4 of Sect. 25 that M is also the matrix of the components of $\mathbf{B}$ relative to the basis $(\mathbf{b}_i^* \otimes \mathbf{b}_j^* \mid (i, j) \in I \times I)$ of $\mathrm{Lin}_2(\mathcal{V}^2, \mathbb{R}) \cong \mathrm{Lin}(\mathcal{V}, \mathcal{V}^*)$, i.e. that

$$\mathbf{B} = \sum_{(i,j) \in I \times I} M_{i,j} (\mathbf{b}_i^* \otimes \mathbf{b}_j^*) \tag{27.4}$$

holds if and only if M is given by (27.3). It is clear from (27.3) that $\mathbf{B}$ is symmetric, i.e. $\mathbf{B}^\top = \mathbf{B}^\sim = \mathbf{B}$, if and only if

$$M_{i,j} = M_{j,i} \quad \text{for all} \quad i, j \in I, \tag{27.5}$$

and that $\mathbf{B}$ is skew, i.e. $\mathbf{B}^\top = \mathbf{B}^\sim = -\mathbf{B}$, if and only if

$$M_{i,j} = -M_{j,i} \quad \text{for all} \quad i, j \in I \tag{27.6}$$

Proposition 1: *If $n := \dim \mathcal{V}$ then*

$$\dim \mathrm{Sym}_2(\mathcal{V}^2, \mathbb{R}) \;=\; \dim \mathrm{Sym}(\mathcal{V}, \mathcal{V}^*) = \frac{n(n+1)}{2} \tag{27.7}$$

and

$$\dim \mathrm{Skew}_2(\mathcal{V}^2, \mathbb{R}) \quad = \quad \dim \mathrm{Skew}(\mathcal{V}, \mathcal{V}^*) = \frac{n(n-1)}{2}. \tag{27.8}$$

Proof: We choose a list basis $\mathbf{b} := (\mathbf{b}_i \mid i \in n^{]})$ of $\mathcal{V}$. Let $\mathbf{A} \in \mathrm{Skew}_2(\mathcal{V}^2, \mathbb{R})$ be given. We claim that

$$\mathbf{A} = \sum (K_{i,j}(\mathbf{b}_i^* \otimes \mathbf{b}_j^* - \mathbf{b}_j^* \otimes \mathbf{b}_i^* \mid (i,j) \in n^{]} \times n^{]},\ i < j) \tag{27.9}$$

holds if and only if

$$K_{i,j} = \mathbf{A}(\mathbf{b}_i, \mathbf{b}_j) \quad \text{for all} \quad (i,j) \in n^{]} \times n^{]} \quad \text{with} \quad i < j. \tag{27.10}$$

Indeed, since $(\mathbf{b}_i^* \otimes \mathbf{b}_j^* \mid (i,j) \in n^{]} \times n^{]})$ is a basis of $\mathrm{Lin}_2(\mathcal{V}^2, \mathbb{R})$, we find, by comparing (27.9) with (27.4) and (27.10) with (27.3), that (27.9) can hold only if (27.10) holds. On the other hand, if we *define* $(K_{i,j} | (i,j) \in n^{]} \times n^{]},\ i < j)$ by (27.10), it follows from (27.6) that the matrix M of $\mathbf{A}$ is given by

$$M_{i,j} = \left\{ \begin{array}{ccc} K_{i,j} & \text{if} & i < j \\ -K_{j,i} & \text{if} & j < i \\ 0 & \text{if} & i = j \end{array} \right\} \quad \text{for all} \quad i, j \in I.$$

Therefore, (27.4) reduces to (27.9).

It follows from (25.6) that the terms of the family

$$\mathbf{b}^* \wedge \mathbf{b}^* := (\mathbf{b}_i^* \otimes \mathbf{b}_j^* - \mathbf{b}_j^* \otimes \mathbf{b}_i^* \mid (i,j) \in n^{]} \times n^{]},\ i < j) \tag{27.11}$$

all belong to $\mathrm{Skew}_2(\mathcal{V}^2, \mathbb{R})$. Therefore, the equivalence of (27.9) and (27.10) proves that the family (27.11) is a basis of $\mathrm{Skew}_2(\mathcal{V}^2, \mathbb{R})$. This basis has $\sharp\{(i,j) \in n^{]} \times n^{]} \mid i < j\} = \frac{n(n-1)}{2}$ terms, and hence (27.8) holds.

Since $\mathrm{Sym}_2(\mathcal{V}^2, \mathbb{R})$ is a supplement of $\mathrm{Skew}_2(\mathcal{V}^2, \mathbb{R})$ in $\mathrm{Lin}_2(\mathcal{V}^2, \mathbb{R})$ and since $\dim \mathrm{Lin}_2(\mathcal{V}^2, \mathbb{R}) = n^2$ by (24.3), we can use Prop.5 of Sect. 17 and (27.8) to obtain

$$\begin{aligned} \dim \mathrm{Sym}_2(\mathcal{V}^2, \mathbb{R}) \quad &= \quad \dim \mathrm{Lin}_2(\mathcal{V}^2, \mathbb{R}) - \dim \mathrm{Skew}_2(\mathcal{V}^2, \mathbb{R}) \\ &= \quad n^2 - \frac{n(n-1)}{2} = \frac{n(n+1)}{2}. \end{aligned}$$

■

We note that if $\mathbf{S} \in \mathrm{Sym}_2(\mathcal{V}^2, \mathbb{R})$, then $\mathbf{S} \circ (\mathbf{1}_\mathcal{V}, \mathbf{1}_\mathcal{V}) = (\mathbf{u} \mapsto \mathbf{S}(\mathbf{u}, \mathbf{u}))$ is a real-valued function on $\mathcal{V}$ (see Sect. 04). Thus, one can consider the mapping

$$(\mathbf{S} \mapsto \mathbf{S} \circ (\mathbf{1}_\mathcal{V}, \mathbf{1}_\mathcal{V}))\ :\ \mathrm{Sym}_2(\mathcal{V}^2, \mathbb{R})\ \rightarrow\ \mathrm{Map}\,(\mathcal{V}, \mathbb{R}). \tag{27.12}$$

Proposition 2: *The mapping* (27.12) *is linear and injective.*

Proof: The linearity follows from Prop.1 of Sect. 14. To show that (27.12) is injective, it suffices to show that its nullspace contains only the zero of $\mathrm{Sym}_2(\mathcal{V}^2, \mathbb{R})$. If $\mathbf{S}$ belongs to this nullspace, then $\mathbf{S} \circ (1_{\mathcal{V}}, 1_{\mathcal{V}}) = 0$, i.e. $\mathbf{S}(\mathbf{u}, \mathbf{u}) = 0$ for all $\mathbf{u} \in \mathcal{V}$. By Props.5 and 7 of Sect. 24 we conclude that $\mathbf{S}$ must also be skew and hence that $\mathbf{S} = \mathbf{0}$. ■

Definition 1: *The range space of the mapping* (27.12) *will be denoted by* $\mathrm{Qu}(\mathcal{V})$. *Its members are called* **quadratic forms** *on* $\mathcal{V}$. *We write*

$$\overset{\frown}{\mathbf{S}} := \mathbf{S} \circ (\mathbf{1}_{\mathcal{V}}, \mathbf{1}_{\mathcal{V}}) \quad \textit{when} \quad \mathbf{S} \in \mathrm{Sym}_2(\mathcal{V}^2, \mathbb{R}), \tag{27.13}$$

and the inverse of the linear isomorphism

$$(\mathbf{S} \mapsto \overset{\frown}{\mathbf{S}}) \ : \ \mathrm{Sym}_2(\mathcal{V}^2, \mathbb{R}) \longrightarrow \mathrm{Qu}(\mathcal{V})$$

will be denoted by

$$(\mathbf{Q} \mapsto \overset{\smile}{\mathbf{Q}}) \ : \ \mathrm{Qu}(\mathcal{V}) \to \mathrm{Sym}_2(\mathcal{V}^2, \mathbb{R}) \cong \mathrm{Sym}(\mathcal{V}, \mathcal{V}^*),$$

so that $\mathbf{Q} = \overset{\smile}{\mathbf{Q}} \circ (\mathbf{1}_{\mathcal{V}}, \mathbf{1}_{\mathcal{V}})$, *i.e.,*

$$\mathbf{Q}(\mathbf{u}) = \overset{\smile}{\mathbf{Q}}(\mathbf{u}, \mathbf{u}) = (\overset{\smile}{\mathbf{Q}}\mathbf{u})\mathbf{u} \quad \textit{for all} \quad \mathbf{u} \in \mathcal{V}. \tag{27.14}$$

We say that $\mathbf{Q} \in \mathrm{Qu}(\mathcal{V})$ *is* **non-degenerate** *if* $\overset{\smile}{\mathbf{Q}} \in \mathrm{Sym}(\mathcal{V}, \mathcal{V}^*)$ *is injective,* **positive** [**negative**] *if* $\mathrm{Rng}\,\mathbf{Q} \subset \mathbb{P}$ $[\mathrm{Rng}\,\mathbf{Q} \subset -\mathbb{P}]$, **strictly positive** [**strictly negative**] *if* $\mathbf{Q}_>(\mathcal{V}^\times) \subset \mathbb{P}^\times$ $[\mathbf{Q}_>(\mathcal{V}^\times) \subset -\mathbb{P}^\times]$. *We say that* $\mathbf{Q}$ *is* **single-signed** *if it is either positive or negative,* **double-signed** *otherwise. The same adjectives will be used for the corresponding bilinear form* $\overset{\smile}{\mathbf{Q}}$.

Example: Let $\mathcal{V}$ be a linear space. The mapping

$$(\mathbf{L} \mapsto \mathrm{tr}_{\mathcal{V}}(\mathbf{L}^2)) \ : \ \mathrm{Lin}\mathcal{V} \longrightarrow \mathbb{R} \tag{27.15}$$

is a quadratic form on the algebra of lineons $\mathrm{Lin}\mathcal{V}$, i.e. a member of $\mathrm{Qu}(\mathrm{Lin}\mathcal{V})$. The corresponding element of $\mathrm{Sym}_2((\mathrm{Lin}\mathcal{V})^2, \mathbb{R})$ is given by $(\mathbf{L}, \mathbf{M}) \mapsto \mathrm{tr}_{\mathcal{V}}(\mathbf{LM})$. It follows from the Representation Theorem for Linear Forms on $\mathrm{Lin}\mathcal{V}$ (Sect. 26) that the quadratic form (27.15) is non-degenerate. ■

Using the symmetry and bilinearity of $\overset{\smile}{\mathbf{Q}}$, one obtains the formulas

$$\mathbf{Q}(\xi\mathbf{u}) = \xi^2\mathbf{Q}(\mathbf{u}) \quad \text{for all} \quad \xi \in \mathbb{R}, \ \mathbf{u} \in \mathcal{V}, \tag{27.16}$$

$$\mathbf{Q}(\mathbf{u} + \mathbf{v}) = \mathbf{Q}(\mathbf{u}) + 2\overset{\smile}{\mathbf{Q}}(\mathbf{u}, \mathbf{v}) + \mathbf{Q}(\mathbf{v}) \quad \text{for all} \quad \mathbf{u}, \mathbf{v} \in \mathcal{V}. \tag{27.17}$$

As a consequence of (27.17), we find

$$\begin{aligned}\overline{\mathbf{Q}}(\mathbf{u},\mathbf{v}) &= \frac{1}{2}(\mathbf{Q}(\mathbf{u}+\mathbf{v})-\mathbf{Q}(\mathbf{u})-\mathbf{Q}(\mathbf{v}))\\ &= \frac{1}{2}(\mathbf{Q}(\mathbf{u})+\mathbf{Q}(\mathbf{v})-\mathbf{Q}(\mathbf{u}-\mathbf{v})) \qquad (27.18)\\ &= \frac{1}{4}(\mathbf{Q}(\mathbf{u}+\mathbf{v})-\mathbf{Q}(\mathbf{u}-\mathbf{v})),\end{aligned}$$

which give various ways of expressing $\overline{\mathbf{Q}}$ explicitly in terms of $\mathbf{Q}$.

The following result follows from Prop.4 of Sect. 24, from (24.11), and from Prop.1 of Sect. 14.

Proposition 3: *Let $\mathcal{V}$ and $\mathcal{W}$ be linear spaces and let $\mathbf{L}\in \mathrm{Lin}(\mathcal{V},\mathcal{W})$. For every $\mathbf{Q}\in \mathrm{Qu}(\mathcal{W})$ we have $\mathbf{Q}\circ\mathbf{L}\in\mathrm{Qu}(\mathcal{V})$ and*

$$\overline{\mathbf{Q}\circ\mathbf{L}} = \overline{\mathbf{Q}}\circ(\mathbf{L}\times\mathbf{L}) = \mathbf{L}^{\top}\overline{\mathbf{Q}}\,\mathbf{L}. \qquad (27.19)$$

The mapping

$$(\mathbf{Q}\mapsto\mathbf{Q}\circ\mathbf{L}) \;:\; \mathrm{Qu}(\mathcal{W}) \;\to\; \mathrm{Qu}(\mathcal{V}) \qquad (27.20)$$

is linear; it is invertible if and only if $\mathbf{L}$ is invertible.

If $\mathcal{U}$ is a subspace of $\mathcal{V}$ and $\mathbf{Q}\in\mathrm{Qu}(\mathcal{V})$, then $\mathbf{Q}|_{\mathcal{U}}\in\mathrm{Qu}(\mathcal{U})$ and $\overline{\mathbf{Q}|_{\mathcal{U}}} = \overline{\mathbf{Q}}|_{\mathcal{U}\times\mathcal{U}}$. If $\mathbf{Q}$ is positive, strictly positive, negative, or strictly negative, so is $\mathbf{Q}|_{\mathcal{U}}$. However, $\mathbf{Q}|_{\mathcal{U}}$ need not be non-degenerate even if $\mathbf{Q}$ is.

Proposition 4: *A positive [negative] quadratic form is strictly positive [strictly negative] if and only if it is non-degenerate.*

Proof: Let $\mathbf{Q}\in\mathrm{Qu}(\mathcal{V})$ and $\mathbf{u}\in\mathrm{Null}\,\overline{\mathbf{Q}}$, so that $\overline{\mathbf{Q}}\,\mathbf{u}=\mathbf{0}$. Then

$$0 = (\overline{\mathbf{Q}}\,\mathbf{u})\mathbf{u} = \overline{\mathbf{Q}}(\mathbf{u},\mathbf{u}) = \mathbf{Q}(\mathbf{u}). \qquad (27.21)$$

Now, if $\mathbf{Q}$ is strictly positive then (27.21) can hold only if $\mathbf{u}=\mathbf{0}$, which implies that $\mathrm{Null}\,\overline{\mathbf{Q}}=\{\mathbf{0}\}$ and hence that $\overline{\mathbf{Q}}$ is injective, i.e. that $\mathbf{Q}$ is positive and non-degenerate. Assume that $\mathbf{Q}$ is positive and non-degenerate and that $\mathbf{Q}(\mathbf{v})=0$ for a given $\mathbf{v}\in\mathcal{V}$. Using $(27.18)_2$ we find that

$$\xi\,\overline{\mathbf{Q}}(\mathbf{u},\mathbf{v}) = \overline{\mathbf{Q}}(\mathbf{u},\xi\mathbf{v}) = \frac{1}{2}(\mathbf{Q}(\mathbf{u})+\mathbf{Q}(\xi\mathbf{v})-\mathbf{Q}(\mathbf{u}-\xi\mathbf{v})),$$

and hence, since $\mathbf{Q}(\mathbf{u}-\xi\mathbf{v})\geq 0$ and $\mathbf{Q}(\xi\mathbf{v})=\xi^2\mathbf{Q}(\mathbf{v})=0$,

$$\xi\,\overline{\mathbf{Q}}(\mathbf{u},\mathbf{v}) \leq \frac{1}{2}\mathbf{Q}(\mathbf{u})$$

for all $\mathbf{v}\in\mathcal{V}$ and all $\xi\in\mathbb{R}^{\times}$. It is easily seen that this is possible only when $0=\overline{\mathbf{Q}}(\mathbf{u},\mathbf{v})=(\overline{\mathbf{Q}}\,\mathbf{u})\mathbf{v}$ for all $\mathbf{v}\in\mathcal{V}$, i.e. if $\overline{\mathbf{Q}}\mathbf{u}=\mathbf{0}$. Since $\overline{\mathbf{Q}}$ is injective, it follows that $\mathbf{u} = \mathbf{0}$. We have shown that $\mathbf{Q}(\mathbf{u}) = 0$ implies $\mathbf{u} = \mathbf{0}$, and hence, since $\mathbf{Q}$ was assumed to be positive, that $\mathbf{Q}$ is strictly positive. ■

Notes 27

(1) Many people use the clumsy terms "positive semidefinite" or "non-negative" when we speak of a "positive" quadratic or bilinear form. They then use "positive definite" or "positive" when we use "strictly positive".

(2) The terms "single-signed" and "double-signed" for quadratic and bilinear forms are used here for the first time. They are clearer than the terms "definite" and "indefinite" found in the literature, sometimes with somewhat different meanings.

28 Problems for Chapter 2

1. Let $\mathcal{V}$ be a linear space and let $\mathbf{L} \in \mathrm{Lin}\mathcal{V}$ and $\lambda, \mu \in \mathbb{R}$ with $\lambda \neq \mu$ be given. Prove that

$$\text{Null }(\mathbf{L}^\top - \mu \mathbf{1}_{\mathcal{V}^*}) \subset (\text{Null }(\mathbf{L} - \lambda \mathbf{1}_{\mathcal{V}}))^\perp. \tag{P2.1}$$

2. Let $n \in \mathbb{N}$ be given and let the linear space $\mathcal{P}_n$ be defined as in Problem 4 in Chap.1. For each $k \in n^[$, let $\beta_k : \mathcal{P}_n \to \mathbb{R}$ be defined by

$$\beta_k(f) := f^{(k)}(0) \quad \text{for all} \quad f \in \mathcal{P}_n, \tag{P2.2}$$

where $f^{(k)}$ denotes the k'th derivative of f (see Sect. 08). Note that, for each $k \in n^[$, β_k is linear, so that $\beta_k \in {\mathcal{P}_n}^*$.

(a) Determine a list $h := (h_k \mid k \in n^[)$ in $\mathcal{P}_n$ such that

$$\beta_k h_j = \delta_{k,j} \quad \text{for all} \quad k, j \in n^[. \tag{P2.3}$$

(b) Show that the list h determined in (a) is a basis of $\mathcal{P}_n$ and that the dual of h is given by

$$h^* = \beta := (\beta_k \mid k \in n^[). \tag{P2.4}$$

(Hint: Apply Props.4 and 5 of Sect. 23).

For each $t \in \mathbb{R}$, let the evaluation $\text{ev}_t : \mathcal{P}_n \to \mathbb{R}$ be defined by

$$\text{ev}_t(f) := f(t) \quad \text{for all} \quad f \in \mathcal{P}_n \tag{P2.5}$$

(see Sect. 04) and note that ev_t is linear, so that $\text{ev}_t \in \mathcal{P}_n{}^*$.

(c) Let $t \in \mathbb{R}$ be given. Determine $\lambda := \text{lnc}_\beta{}^{-1}\text{ev}_t \in \mathbb{R}^{n^[}$, where β is the basis of $\mathcal{P}_n{}^*$ defined by $(P2.4)$ and $(P2.2)$, so that

$$\text{ev}_t = \text{lnc}_\beta \lambda = \textstyle\sum_{k \in n^[} \lambda_k \beta_k. \tag{P2.6}$$

3. Let $n \in \mathbb{N}$ be given and let the linear space $\mathcal{P}_n$ be defined as in Problem 4 of Chap.1. Also, let a subset F of $\mathbb{R}$ with n elements be given, so that $n = \sharp F$.

   (a) Determine a family $g := (g_s \,|\, s \in F)$ in $\mathcal{P}_n$ such that

   $$\text{ev}_t(g_s) = \delta_{t,s} \quad \text{for all} \quad t, s \in F, \tag{P2.7}$$

   where $\text{ev}_t \in \mathcal{P}_n{}^*$ is the evaluation given by (P2.5).

   (b) Show that the family g determined in (a) is a basis of $\mathcal{P}_n$ and that the dual of g is given by

   $$g^* = (\text{ev}_t \,|\, t \in F). \tag{P2.8}$$

   (Hint: Apply Props.4 and 5 of Sect. 23).

4. Let $\mathcal{V}$ be a linear space. Let the mappings Λ and Υ from $\text{Lin}\mathcal{V} \times \text{Lin}\mathcal{V}$ to $\text{Lin}\mathcal{V}$ be defined by

   $$\Lambda(\mathbf{L}, \mathbf{M}) := \mathbf{LM} - \mathbf{ML} \tag{P2.9}$$

   and

   $$\Upsilon(\mathbf{L}, \mathbf{M}) := \mathbf{LM} + \mathbf{ML} \tag{P2.10}$$

   for all $\mathbf{L}, \mathbf{M} \in \text{Lin}\mathcal{V}$.

   (a) Show that Λ is bilinear and skew, so that $\Lambda \in \text{Skew}_2((\text{Lin}\mathcal{V})^2, \text{Lin}\mathcal{V})$. Also, show that

$$\sum_{\gamma \in C_3} \Lambda(\Lambda(\mathbf{L}_{\gamma(1)}, \mathbf{L}_{\gamma(2)}), \mathbf{L}_{\gamma(3)}) = 0 \tag{P2.11}$$

for every triple $(\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3)$ in $\mathrm{Lin}\mathcal{V}$, where C_3 denotes the set of cyclic permutations of $3^]$, i.e.

$$C_3 := \left\{ 1_{3^]}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}.$$

(b) Show that Υ is bilinear and symmetric, so that $\Upsilon \in \mathrm{Sym}_2((\mathrm{Lin}\mathcal{V})^2, \mathrm{Lin}\mathcal{V})$. Also, show that

$$\Upsilon(\mathbf{L}, \Upsilon(\Upsilon(\mathbf{L}, \mathbf{L}), \mathbf{M})) = \Upsilon(\Upsilon(\mathbf{L}, \mathbf{L}), \Upsilon(\mathbf{L}, \mathbf{M})) \tag{P2.12}$$

for all $\mathbf{L}, \mathbf{M} \in \mathrm{Lin}\mathcal{V}$.

Remark: The mapping Λ endows $\mathrm{Lin}\mathcal{V}$ with the structure of a "Lie-algebra" and the mapping Υ endows $\mathrm{Lin}\mathcal{V}$ with the structure of a "Jordan-algebra". The theory of these is an important topic in abstract algebra. ■

5. Let $D, M \in \mathrm{Lin}\,\mathrm{C}^\infty(\mathbb{R})$ and $\mathcal{P}_n$, with $n \in \mathbb{N}$, be defined as in Problems 4 and 5 of Chapt.1.

   (a) Show that $\mathcal{P}_n$ is D-subspace, a (MD)-subspace, and a (DM)-subspace of $\mathrm{C}^\infty(\mathbb{R})$. Calculate $\mathrm{tr} D_{|\mathcal{P}_n}, \mathrm{tr}(MD)_{|\mathcal{P}_n}$, and $\mathrm{tr}(DM)_{|\mathcal{P}_n}$.

   (b) Prove: If $\mathcal{V}$ is a linear space of finite and non-zero dimension, there do not exist $\mathbf{L}, \mathbf{K} \in \mathrm{Lin}\mathcal{V}$ such that $\mathbf{LK} - \mathbf{KL} = \mathbf{1}_\mathcal{V}$. (Compare this assertion with the one of Problem 5 of Chap.1.)

6. Let $\mathcal{V}$ be a linear space.

   (a) Prove that $\Omega \in (\mathrm{Lin}\mathcal{V})^*$ satisfies

$$\Omega(\mathbf{LM}) = \Omega(\mathbf{ML}) \quad \text{for all} \quad \mathbf{L}, \mathbf{M} \in \mathrm{Lin}\mathcal{V} \tag{P2.13}$$

   if and only if $\Omega = \xi \mathrm{tr}_\mathcal{V}$ for some $\xi \in \mathbb{R}$.

   (Hint: Consider the case when $\mathbf{L}$ and $\mathbf{M}$ in $(P2.13)$ are tensor products.)

   (b) Consider the left-multiplication mapping $\mathrm{Le}_\mathbf{L} \in \mathrm{Lin}(\mathrm{Lin}\mathcal{V})$ defined by $(P1.1)$ for each $\mathbf{L} \in \mathrm{Lin}\mathcal{V}$. Prove that

$$\mathrm{tr}_{\mathrm{Lin}\mathcal{V}}(\mathrm{Le}_\mathbf{L}) = (\dim \mathcal{V})\mathrm{tr}_\mathcal{V}\mathbf{L} \quad \text{for all} \quad \mathbf{L} \in \mathrm{Lin}\mathcal{V} \tag{P2.14}$$

(Hint: Use part (a)).

7. Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be supplementary subspaces of the given linear space $\mathcal{V}$ and let $\mathbf{P}_1, \mathbf{P}_2$ be the projections associated with $\mathcal{U}_1, \mathcal{U}_2$ according to Prop.4 of Sect. 19. Prove that, for every $\mathbf{L} \in \mathrm{Lin}\mathcal{V}$, we have

$$\mathrm{tr}_{\mathcal{V}}\mathbf{L} = \mathrm{tr}_{\mathcal{U}_1}(\mathbf{P}_1\mathbf{L}|_{\mathcal{U}_1}) + \mathrm{tr}_{\mathcal{U}_2}(\mathbf{P}_2\mathbf{L}|_{\mathcal{U}_2}). \tag{P2.15}$$

(Hint: Apply Prop.1 of Sect. 26 and Prop.5 of Sect. 19.)

8. Let $\mathcal{V}$ be a linear space and put $n := \dim\mathcal{V}$.

   (a) Given a list basis $\mathbf{b} := (\mathbf{b}_i \,|\, i \in n^]) $ of $\mathcal{V}$ construct a basis of $\mathrm{Sym}_2(\mathcal{V}^2, \mathbb{R}) \cong \mathrm{Sym}(\mathcal{V}, \mathcal{V}^*)$ by a procedure analogous to the one described in the proof of Prop.1 of Sect. 27.

   (b) Show that, for every $\boldsymbol{\lambda} \in \mathcal{V}^*$, the function $\boldsymbol{\lambda}^2 : \mathcal{V} \to \mathbb{R}$ obtained from $\boldsymbol{\lambda}$ by value-wise squaring (i.e. by $\boldsymbol{\lambda}^2(\mathbf{v}) := (\boldsymbol{\lambda}\mathbf{v})^2$ for all $\mathbf{v} \in \mathcal{V}$), is a quadratic form on $\mathcal{V}$, i.e. that $\boldsymbol{\lambda}^2 \in \mathrm{Qu}(\mathcal{V})$. (See Def.1 in Sect. 27)

   (c) Prove that

$$\mathrm{Lsp}\{\boldsymbol{\lambda}^2 \,|\, \boldsymbol{\lambda} \in \mathcal{V}^*\} = \mathrm{Qu}(\mathcal{V}). \tag{P2.16}$$